Administrator's and User's Guide
Note!

Before using this information and the product it supports, read the information in Appendix D, “Notices,” on page 307.

This edition of IBM Tivoli System Automation for Multiplatforms Administrator’s and User’s Guide applies to Version 3, Release 1, Modification 0 of IBM Tivoli System Automation for Multiplatforms, program number 5724-M00, and to all subsequent releases and modifications of this product until otherwise indicated in new editions.

This edition replaces SC33-8415-01.

IBM welcomes your comments. A form for readers’ comments may be provided at the back of this publication, or you may address your comments to the following address:

IBM® Deutschland Research & Development GmbH
Department 3248
Schoenaicher Strasse 220
71032 Boeblingen
Germany

FAX (Germany): 07031 16-3456
FAX (Other Countries): +49 7031 16-3456

Internet e-mail: eservdoc@de.ibm.com

If you would like a reply, be sure to include your name, address, telephone number, or FAX number.

Make sure to include the following in your comment or note:
• Title and order number of this book
• Page number or topic related to your comment

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.
Chapter 10. Controlling and administering System Automation for Multiplatforms

Controlling System Automation for Multiplatforms

TimeOut and RetryCount
Automation
ExcludedNodes
ResourceRestartTimeout
Examples
Managing automation policies for System Automation for Multiplatforms domains
Using the sampolicy command to manage policies
Using XML for automation policies
Using policy templates
Starting and stopping resource groups and group members
Scope of start and stop requests
Changing the priority of start and stop requests
Canceling start and stop requests
Displaying request lists
Example scenario
Using requests to lock and unlock resource groups and resources
Scope of lock requests
Changing the priority of requests
Example scenario
Prioritization of requests
Sources and default priorities of requests
How requests are prioritized
Moving resource groups with the rggreq command
Scope of a move
Processing of a move request
Move and relationships
Using shadow resources
Defining shadow resources
Setting up security for System Automation for Multiplatforms clusters
Limitations of the non-root security setup
Diagnosing System Automation for Multiplatforms Resources
Using the TEC event interface with System Automation for Multiplatforms
What is the Tivoli Enterprise Console?
What are events?
Sending events to the TEC
Enabling the TEC publisher function
Setting a new language locale for the TEC event messages

Chapter 11. Automating System Automation for Multiplatforms resources

Using the Global Resource Manager
What is the IBM.Application resource class?
What is the IBM.ServiceIP resource class?
Using the Test Resource Manager
What is the IBM.Test resource class?
Example: Create a test resource and manipulate its OpState

Chapter 12. Automating RSCT Storage resource manager resources

Supported storage devices
Automation support for single-path and multipath storage devices of the IBM TotalStorage DS4000 family
Support for single-path and multipath storage devices on other IBM TotalStorage disk system families
Support for non-DS4000 single-path storage devices
Support for network file systems
Creating user-defined IBM.AgFileSystem resources
Automating user-defined IBM.AgFileSystem resources in an LVM environment
Attributes used by IBM.AgFileSystem
Attribute details

Chapter 13. System Automation for Multiplatforms on Windows:

Particularities and restrictions
Using System Automation for Multiplatforms on Windows systems
Contents of the SA for Multiplatforms
Working with the IBM Tivoli System Automation Shell
Resource instrumentation with resource classes
Startup procedure
Checking that System Resource Controller (SRC) is running
Restarting SRC and RMC
Dead-Man-Switch (DMS) to protect critical resources
## Figures

1. Symbols used in this guide .................................................. xii
2. Overview of the System Automation for Multiplatforms architecture ......................................... 9
3. Sample configuration setup .................................................................. 99
4. Operational state transitions of automated resources .......................................................... 100
5. State cycle of a resource group .......................................................... 113
6. Overview of the operations console ........................................................................ 121
7. Main panel of the operations console ........................................................................ 122
8. Layout of the resources section ........................................................................ 125
9. Main panel of the policy editor ........................................................................... 143
10. Quorum – majority of nodes ........................................................................ 161
11. Two-node cluster system logs ........................................................................ 178
12. Priority ranking of requests ........................................................................... 196
13. Scenario 1: DependsOn relationship, without shadow resources ........................................ 199
14. DependsOn relationship, with shadow resources .................................................. 200
15. Syntax format and sample of the publisher configuration file ........................................ 210
16. Syntax format and sample of the TEC configuration file ........................................ 212
17. Aggregate resource and constituent resources .................................................. 222
18. Configuring a supporting resource for an IBM.Application resource .............................. 235
19. Windows Computer Management Console showing all services ................................ 262
20. Setting up the operating system to restart after BSOD ............................................... 264
21. Problems planning a highly available network .................................................. 267
22. Two nodes, one interface ........................................................................... 270
23. Two nodes, one interface – interface failure .................................................. 271
24. Two nodes, two interfaces, two physically separated networks ...................................... 272
25. Two nodes, two interfaces, one physical network .................................................. 273
26. Two physically separated networks, dynamic routing and VIPA ................................ 275
27. Network interfaces bonded together to one logical network device .............................. 276
### Tables

1. System Automation for Multiplatforms commands used with managed resources ........................................ 29
2. System Automation for Multiplatforms commands used with resource groups ........................................ 35
3. MoveStatus values ................................................................................................................................ 45
4. System Automation for Multiplatforms commands used with equivalencies ............................................ 51
5. System Automation for Multiplatforms commands used with managed relationships ............................. 57
7. System Automation actions regarding OpState changes of resource Res2 .............................................. 101
8. Resource group OpState determination ..................................................................................................... 102
9. System Automation actions when a StartCommand still running ............................................................ 104
10. System Automation actions after a StartCommand completed successfully ........................................... 105
11. System Automation actions when a StartCommand completed with an error or timed out .................... 106
12. System Automation actions and StopCommand still under execution .................................................... 109
13. System Automation actions after StopCommand successfully finished ................................................ 110
14. System Automation actions after StopCommand finished with an error or timed out ............................ 110
15. System Automation actions after the MonitorCommand reports an OpState change for the resource on another node ............................................................................................................. 111
17. Icons used for the elements of the topology tree ...................................................................................... 124
18. Icons in the Status column of the topology tree ..................................................................................... 124
19. Resource icons in the resource column .................................................................................................... 128
20. Operator request icons in the information area .......................................................................................... 136
21. Types of tie breakers .................................................................................................................................. 163
22. Comparison of network-based and disk-based tie breakers .................................................................... 176
23. Sources and priority levels of start and stop requests ............................................................................. 195
25. Operational quorum protection methods ................................................................................................. 263
About this book

This book explains how to implement and exploit the policy-based self-healing capabilities provided by IBM Tivoli System Automation for Multiplatforms (System Automation for Multiplatforms).

System Automation for Multiplatforms provides high-availability for resources on AIX® clusters (on IBM System p), Solaris clusters (on SPARC), Linux clusters (on IBM System x, System z, System i, and System p), and Windows clusters (on IBM System x).

Who should use this book

This book is intended for system administrators and operators who want to use the automation and failover capabilities of System Automation for Multiplatforms.

How to use this book

This book contains the information you need to understand and use IBM Tivoli System Automation for Multiplatforms (System Automation for Multiplatforms).

- Chapter 1 provides an introduction to System Automation for Multiplatforms. It gives an overview of System Automation for Multiplatforms, introduces the components, and explains the technical terms used in this manual.
- Chapter 2 is a quick guide to the tasks you need to complete to work with System Automation for Multiplatforms. It describes how clusters and nodes are administered and how an automation policy is defined.
- Chapter 3 describes the common attributes of System Automation for Multiplatforms.
- Chapter 4 describes how resource groups are created and used.
- Chapter 5 describes how equivalencies are created and used.
- Chapter 6 describes how managed relationships are created and used.
- Chapter 7 describes how System Automation for Multiplatforms processes system information.
- Chapter 8 describes the operations console, the policy editor, and the other commands and applications available in Integrated Solutions Console.
- Chapter 9 describes how System Automation for Multiplatforms protects your resources.
- Chapter 10 describes how System Automation for Multiplatforms is controlled and administered.
- Chapter 11 describes the resource managers provided by System Automation for Multiplatforms.
- Chapter 12 describes how to automate RSCT Storage resource manager resources.
- Chapter 13 describes how System Automation for Multiplatforms on Windows differs from System Automation for Multiplatforms on AIX, Solaris, and Linux systems.
- Appendix A describes how a highly available network is set up.
- Appendix B provides useful information for troubleshooting.
- Appendix C describes how to use IBM Support Assistant.
Where to find more information

In addition to this manual, the IBM Tivoli System Automation for Multiplatforms library comprises the following books:

- **IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide**, SC33-8416
- **IBM Tivoli System Automation for Multiplatforms Reference**, SC33-8417


The System Automation for Multiplatforms home page offers up-to-date information and services, and other items of interest to System Automation for Multiplatforms users.


Conventions

The following highlighting conventions are used in this book:

<table>
<thead>
<tr>
<th>Conventions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bold</strong></td>
<td>Identifies commands, subroutines, keywords, files, structures, directories, and other items whose names are predefined by the system. Also identifies graphical objects such as buttons, labels, and icons that the user selects.</td>
</tr>
<tr>
<td><strong>Italic</strong></td>
<td>Identifies parameters whose actual names or values are to be supplied by the user.</td>
</tr>
<tr>
<td><strong>monospace</strong></td>
<td>Identifies examples of specific data values, examples of text similar to what you might see displayed, examples of portions of program code similar to what you might write as a programmer, messages from the system, or information you should actually type.</td>
</tr>
</tbody>
</table>

This manual uses symbols to show resources, resource groups, equivalencies, and relationships. The symbols used are as follows:

- Resource Group
- Equivalency
- fixed Resource
- floating Resource
- Relationship

*Figure 1. Symbols used in this guide*

ISO 9000

ISO 9000 registered quality systems were used in the development and manufacturing of this product.
Related information

The following IBM Reliable Scalable Cluster Technology (RSCT) publications are available on the System Automation for Multiplatforms CD:

- RSCT Administration Guide
- RSCT for AIX 5L: Technical Reference
- RSCT for Multiplatforms: Technical Reference
- RSCT Messages
- RSCT Diagnosis Guide

RSCT publications can also be found at the following Web site:
www.ibm.com/servers/eserver/clusters/library/

You might also need to refer to the following IBM Redpaper:
- Linux on IBM zSeries® and S/390®: High Availability for z/VM® and Linux

It can be found at the following Web site:

How to obtain publications

The System Automation for Multiplatforms publications are also available (valid at the time of release) at these Web sites:
- .ibm.com/servers/eserver/clusters/library/
- .ibm.com/servers/eserver/zseries/software/sa/
- .ibm.com/software/sysmgmt/products/support/

How to reach us by e-mail

If you would like to contact us by e-mail, send your comments to eservdoc@de.ibm.com
Summary of changes

What's new in release 3.1

The following new features and enhancements are provided in Version 3 Release 1:

- System Automation for Multiplatforms can now be installed on:
  - AIX 5.3 and 6.1, including the RSCT prerequisite
  - Solaris 10 on SPARC
  - Windows Server 2008
- The minimum Red Hat Enterprise Linux version is now 4.6
- The configuration tool can now be used in silent mode without the need for a window environment (such as X-Window)
- System Automation for Multiplatforms on Windows can now be added to a System Automation Application Manager domain
- Graphical policy editor
- Running the product with the English language is now supported in Turkish locales
- Volume group mirroring on AIX is supported
- You can now use the operations console to move resources from one node to another, see “Moving resources” on page 137
Chapter 1. Introduction

Overview

System Automation for Multiplatforms manages the availability of applications running on Linux, AIX, Solaris, and Windows clusters. For detailed support platform information, see IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide System Automation for Multiplatforms provides the following features:

High availability and resource monitoring

System Automation for Multiplatforms provides a high-availability environment, in which systems are continuously available and whose self-healing infrastructure prevents downtime caused by system problems. The self-healing infrastructure detects improper operations of systems, transactions, and processes, and initiates corrective actions without disrupting users.

System Automation for Multiplatforms offers mainframe-like high availability by utilizing fast outage detection and sophisticated knowledge about application components and their relationships. It ensures quick and consistent recovery of failed resources and complete business applications, either on the same or on a different system within a Linux, AIX, Solaris, or Windows cluster, without requiring operator intervention, which reduces the risk of operator errors by relieving them from manual monitoring and from having to remember application components and relationships.

Policy-based automation

System Automation for Multiplatforms enables you to configure high-availability systems through the use of automation policies in which the relationships among the various components are defined. These policies can be applied to existing applications with minor modifications. Once the relationships are established, System Automation for Multiplatforms will assume responsibility for managing the applications on the specified nodes as configured. This reduces implementation time and the need for complex coding of applications. In addition, systems can be added without modifying scripts, and resources can be easily added, too.

Sample automation policies can be downloaded from the IBM Tivoli Open Process Automation Library at:

http://catalog.lotus.com/wps/portal/topal

To access the samples, click Tivoli products in the Browse by panel and then select Tivoli System Automation from the drop-down list.

Automatic recovery

System Automation for Multiplatforms quickly and consistently restarts failed resources and complete business applications, either on the same or on a different system within a Linux, AIX, Solaris, or Windows cluster. This greatly reduces the risk of application and service outages.

Automatic movement of applications

System Automation for Multiplatforms manages cluster-wide relationships among resources for which it is responsible. If applications need to be moved among
nodes, the start and stop relationships, node requirements and any preliminary or follow-up actions are automatically handled by System Automation for Multiplatforms. This again relieves the operator from manual command entry and reduces the risk of operator errors.

**Resource grouping**

In System Automation for Multiplatforms, resources can be grouped together. Once grouped, all relationships among the members of the group can be established; for example, location relationships and start and stop relationships. After the configuration is complete, operations can be performed against the group as a single unit, which once again relieves operators from having to remember application components and relationships and reduces the risk of operator errors.
System Automation for Multiplatforms terms

This section gives an overview of the terms this manual uses when describing System Automation for Multiplatforms.

Clusters and peer domains

Throughout this guide, the terms cluster and peer domain are used interchangeably to refer to a group of host systems that is configured for high availability and on which System Automation for Multiplatforms is used for automating and managing resources. A cluster can consist of a single or multiple nodes (the terms node and system are used interchangeably as well). System Automation for Multiplatforms supports up to 32 nodes within a cluster.

Resource

A resource is any piece of hardware or software that can be defined to System Automation for Multiplatforms. Users can define resources to System Automation for Multiplatforms in a number of ways:

- From the command line with the command mkrsrc ("make resource")
- In an XML automation policy file.
- In the graphical policy editor.
- Additionally, some resources are automatically defined through the resource harvesting function of the cluster infrastructure.

The resources are controlled by the appropriate resource managers (see "Resource managers provided by RSCT and System Automation for Multiplatforms" on page 2). Resources have characteristics, or attributes, which can be defined. For example, the attributes of a resource that represents an IP address would include the IP address itself and the net mask.

There are two types of resources:

Fixed resource

A resource of which there is only a single instance within the cluster. It represents one entity that is defined for a single node, and this is the only node on which it runs.

Floating resource

A resource that can run on several nodes in the cluster but only on one node at a time (see "ResourceType attribute" on page 31).

Resource attributes

A resource attribute describes some characteristics of a resource. There are two types of resource attributes:

Persistent attributes

The attributes of an IP address (the IP address itself and the net mask) are examples of persistent attributes – they describe enduring characteristics of a resource. Although you could change the IP address and net mask, these characteristics are, in general, stable and unchanging.

Dynamic attributes

Dynamic attributes, on the other hand, represent changing characteristics of the resource. Dynamic attributes of an IP address, for example, would identify such things as its operational state.
Resource class

A resource class is a collection of resources of the same type. For example, if an application is a resource, then all applications defined in the cluster would comprise a resource class. Resource classes allow you to define the common characteristics among the resources in its class. In the case of applications, the resource class can define identifying characteristics such as the name of the application, and varying characteristics such as whether or not the application is running. So each resource in the class can then be noted by its characteristics at any given time. Resource classes are managed by various resource managers – see “Resource managers provided by RSCT and System Automation for Multiplatforms” on page 7.

Resource group

Resource groups are logical containers for a collection of resources. This container allows you to control multiple resources as a single logical entity. Resource groups are the primary mechanism for operations within System Automation for Multiplatforms. Resource groups can also be nested, meaning that applications can be split into several resource groups which themselves are part of another higher level resource group. Also, resource groups can be defined in such a way that their members can be located on different systems in the cluster.

Managed resources

Managed resources are resources that have been defined to System Automation for Multiplatforms. Resources become managed resources by being added to a resource group or an equivalency or by being defined as the target of a managed relationship.

Nominal state

The nominal state of a resource group indicates to System Automation for Multiplatforms whether the resources with the group should be online or offline at this point in time. So setting the nominal state to “Offline” indicates that you wish for System Automation for Multiplatforms to stop the resources in the group, and setting the nominal state to “Online” is an indication you wish to start the resources in the resource group. You can change the value of the NominalState resource group attribute, but you cannot set the nominal state of a resource directly. See “NominalState attribute” on page 41.

Equivalency

An equivalency is a collection of resources that provides the same functionality. For example, equivalencies are used for selecting network adapters that should host an IP address. If one network adapter goes offline, System Automation for Multiplatforms selects another network adapter to host the IP address.

Relationships

System Automation for Multiplatforms allows the definition of relationships between resources in a cluster. There are two different relationship types:

- Start-/stop relationships
  Relationships are used to define start and stop dependencies between resources. You can use the StartAfter, StopAfter, DependsOn, DependsOnAny, and ForcedDownBy relationships to achieve this. For example, if a resource must only be started after another resource has been started, you can define this by using the policy element StartAfter relationship.
Location relationships

Location relationships are applied when resources must, or should if possible, be started on the same or a different node in the cluster. System Automation for Multiplatforms provides the following location relationships: Collocation, AntiCollocation, Affinity, AntiAffinity, and IsStartable. A simple example is that a webserver and its corresponding service IP address, which could be started on any node in the cluster, should always be kept together. In the past, this behavior had to be defined by writing complex scripts. Now System Automation for Multiplatforms allows the usage of a location relationship, which simplifies the policy definition for the administrator.

Relationships provide the following additional features:
- The possibility to define relationships between resource groups, resources, and equivalencies.
- The possibility to define relationships between resources running on different systems in the cluster.

Quorum

The main goal of quorum operations is to keep data consistent and to protect critical resources. Quorum can be seen as the number of nodes in a cluster that are required to modify the cluster definition or perform certain cluster operations. There are two types of quorum:

Configuration quorum

Configuration quorum determines when configuration changes in the cluster will be accepted. Operations affecting the configuration of the cluster or resources are only allowed when the absolute majority of nodes is online. See “Configuration quorum” on page 162 for a detailed description.

Operational quorum

Operational quorum is used to decide whether resources can be safely activated without creating conflicts with other resources. In case of a cluster splitting, resources can only be started in the subcluster which has a majority of nodes or obtained a tie breaker. See “Operational quorum” on page 162 for a detailed description.

Tie breaker

To resolve tie situations in which a cluster with an equal number of nodes is partitioned into sub-clusters that each have exactly half of the defined nodes, a tie breaker is used to determine which subcluster has operational quorum.

Operations console

In this guide, the terms operations console and SA operations console are used interchangeably to refer to the browser-based graphical user interface of System Automation for Multiplatforms.

The operations console has the following characteristics:
- It is shipped as a part of System Automation for Multiplatforms and can be installed separately.
- The operations console and the end-to-end automation adapter, which establishes and manages the communication between the clusters whose resources are automated by System Automation for Multiplatforms (automation domains) and the SA operations console, are configured for direct-access mode.
Introduction

End-to-end automation adapter

The end-to-end automation adapter of System Automation for Multiplatforms is required to establish and manage the communication between the automation domains and the SA operations console (or, if installed, IBM Tivoli System Automation Application Manager). The adapter is installed automatically with System Automation for Multiplatforms.
RSCT as an integral part of System Automation for Multiplatforms

Reliable Scalable Cluster Technology (RSCT) is a product that is fully integrated into System Automation for Multiplatforms. RSCT is a set of software products that together provide a comprehensive clustering environment for AIX, Linux, Solaris, and Windows. System Automation for Multiplatforms uses the RSCT infrastructure to provide clusters with improved system availability, scalability, and ease of use.

RSCT includes the following basic subsystems:

- **RMC** (Resource Monitoring and Control subsystem): RMC is used for single-system monitoring and for monitoring nodes in a cluster. In a cluster, RMC provides global access to subsystems and resources throughout the cluster, thus providing a single monitoring and management infrastructure for clusters.

- **HAGS** (High Availability Group Services subsystem) HAGS is a distributed coordination, messaging, and synchronization service.

- **HATS** (High Availability Topology Services subsystem) HATS provides a scalable heartbeat for adapter and node failure detection, and a reliable messaging service in a peer domain.

- **SRC** (System Resource Controller) The System Resource Controller (SRC) is the bootstrap process that gets started first and brings up all other daemon processes required by RSCT and System Automation for Multiplatforms to run. It is also responsible for restarting these daemon processes on failure. The System Resource Controller provides a set of commands and subroutines to make it easier for system managers to create and control subsystems.
  
  On AIX, and Linux, the System Resource Controller is started from the `inittab`. On Solaris, the System Resource Controller is started from the Service Management Facility (SMF). On Windows, the System Resource Controller runs as a Windows Service.

Resource managers provided by RSCT and System Automation for Multiplatforms

Resource classes are managed by so-called resource managers (RM). Which resource manager is responsible for managing a particular resource depends on the resource class. A resource manager is a software layer between a resource and RMC. A resource manager runs as a daemon which is controlled by the system resource controller (SRC).

RSCT provides a core set of resource managers for managing System Automation for Multiplatforms resources on single systems and across clusters. The set of RSCT resource managers is extended by the following resource managers specific to System Automation for Multiplatforms:

**Recovery resource manager (Recovery RM)**

The recovery RM (IBM.RecoveryRM) serves as the decision engine for System Automation for Multiplatforms. Once an automation policy is provided in which the resource availabilities and relationships are defined, the policy information is supplied to the Recovery RM. A Recovery RM runs on every node in the cluster, with exactly one Recovery RM designated as the master. The master Recovery RM is responsible for evaluating the monitoring information supplied by the various...
resource managers. When a situation develops that requires intervention, the Recovery RM drives the decisions that result in start or stop operations on the resources as needed.

**Global Resource RM**
The Global Resource RM (IBM.GblResRM) supports two resource classes:

**IBM.Application**
The IBM.Application resource class defines the behavior for general application resources. This class can be used to start, stop, and monitor processes. As a generic class, it is very flexible and can be used to monitor and control various kind of resources. Most of the applications that you will automate will be of this class. For more information, refer to "Using the Global Resource Manager" on page 221.

**IBM.ServiceIP**
This application class defines the behavior of Internet Protocol (IP) address resources. It allows you to assign IP addresses to an adapter. In effect, it allows IP addresses to ‘float’ among nodes. For more information, refer to "What is the IBM.ServiceIP resource class?" on page 236.

**Storage resource manager (Storage RM)**
The Storage resource manager provides monitoring and control for storage resources within an automation domain. The Storage resource manager provides the interface between RMC and the physical and logical storage entities within the peer domain by mapping these entities to instances of the resource classes it provides.

Running as a daemon process on each node in the automation domain, the Storage resource manager collects information about locally-attached physical disks (and related storage entities) and maps these to resource class instances. These separate views of the storage resources from each individual node are then collected together to provide the Storage resource manager with a global view of the storage resources within an automation domain.

In a shared storage environment (in which multiple nodes within the automation domain have access to the same disk subsystem and therefore access to the same disks contained within the disk subsystem), the Storage resource manager’s global view of storage resources enables it to identify such shared disks and provide serial access through hardware reserves.

**Configuration resource manager (Configuration RM)**
The Configuration RM (IBM.ConfigRM) is used in cluster definition. In addition, quorum support, which is a means of insuring data integrity when portions of a cluster lose communication, is provided.

**Test resource manager (Test RM)**
The Test RM (IBM.TestRM) manages test resources and can be used to manipulate their operational state. The Test RM does not control actual resources.

The Test RM is operational in a peer domain mode only and provides the resource class IBM.Test. A detailed description of the Test RM is given in Chapter 11, "Automating System Automation for Multiplatforms resources," on page 221.

Figure 2 on page 9 shows a diagram of the previously described components.
Introduction

Figure 2. Overview of the System Automation for Multiplatforms architecture
Using System Automation for Multiplatforms on Windows


You can use System Automation for Multiplatforms on Windows to manage and automate standard Win32 applications and services.

To interact with the product, operators and administrators may use either the operations console (see “Using the SA operations console” on page 121) or the “Tivoli System Automation Shell”. This shell is a specialized “Korn Shell” process. For the rest of this guide it is assumed that any user has the basic skills required to work with UNIX-like command shells. This is especially also true for System Automation for Multiplatforms running on Windows servers. All users should be able to navigate in the standard UNIX file systems, view and edit text files, change access rights, and understand utilities such as “ps” and “grep”. All commands that are described in the remainder of this document must be issued from a UNIX shell. They cannot be invoked from within a Windows command shell.
Chapter 2. Getting started

Note: Before you begin to use System Automation for Multiplatforms on Windows, refer to the information in Chapter 13, “System Automation for Multiplatforms on Windows: Particularities and restrictions,” on page 255.

This chapter lists and describes the steps shown below that you have to perform to start working with System Automation for Multiplatforms:

Step 1: Defining and administering a cluster (see page 12)
This step shows how you can create and remove a cluster, add nodes to and remove them from a cluster, and how you can check the status of System Automation for Multiplatforms.

Step 2: Defining RSCT resources (see page 19)
This step shows how you can create a resource like a web server and how you can create an equivalency relationship.

Step 3: Defining the automation policy from the command line (see page 23)
This step shows how you can define the relationships among the components created in Step 1 and Step 2. This is called defining the automation policy.

For detailed information about the commands that are mentioned in this chapter, refer to the following documentation:

IBM Reliable Scalable Cluster Technology (RSCT) commands
For a detailed description of the RSCT commands, refer to the appropriate man pages and to the RSCT manuals listed in “Related information” on page xiii.

The RSCT manuals are available on the System Automation for Multiplatforms CD. They can also be downloaded from the following website:
www.ibm.com/servers/eserver/clusters/library/

The Cluster information center is available at:
http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp

System Automation for Multiplatforms commands
For a detailed description of these commands, refer to IBM Tivoli System Automation for Multiplatforms Reference. You will use some System Automation for Multiplatforms commands in Step 3.
Getting started

Step 1: Defining and administering a cluster

In the scenarios below you learn how you create a cluster, set up a tie breaker, add nodes to the cluster, and how you can check the status of the System Automation for Multiplatforms Recovery RM daemon (IBM.RecoveryRM).

Command overview

The following list gives you an overview of the Reliable Scalable Cluster Technology (RSCT) commands you will use when you work with cluster definitions. You will use some of these commands in Step 1. For a detailed description of the commands, refer to the RSCT manuals listed on page [11].

preprpnode This command prepares the security settings for the node to be included in a cluster. When issued, public keys are exchanged among the nodes, and the RMC access control list (ACL) is modified to enable access to cluster resources by all the nodes of the cluster.

mkrpdomain This command creates a new cluster definition. It is used to specify the name of the cluster, and the list of nodes to be added to the cluster.

lsrpdomain This command lists information about the cluster to which the node belongs where the command runs.

startrpdomain / stoprpdomain These commands are used to bring the cluster online and offline, respectively.

addrpnode Once a cluster has been defined and is operational, this command is used to add new nodes to the cluster.

startrpnode / stoprpnode These commands are used to bring individual nodes online and offline to the cluster. They are often used when performing maintenance to a particular system. The node is stopped, repairs or maintenance is performed, then the node is restarted, at which time it rejoins the cluster.

lsrpnode This command is used to view the list of nodes defined to a cluster, and the operating state (OpState) of each node. Note that this command is useful only on nodes that are Online in the cluster, otherwise it will not display the list of nodes.

lssrc This command lists the status of subsystems.

rmpdomain This command removes a defined cluster.

rmrpnodel This command removes one or more nodes from a cluster definition.

startsrc This command starts an individual subsystem.

stopsrc This command stops an individual subsystem.
Creating a two-node cluster

Before you begin to create a cluster, make sure that your network setup is correct:

- IP, netmask, and broadcast addresses must be consistent on each cluster node.
- Make sure that the name resolution is correct, DNS entries are consistent, or the entries in your local /etc/hosts files are identical on all nodes.
- Do not define more than one network interface on a node to the same subnet.
- All firewalls, including local firewalls, must allow ICMP pings and IP-traffic to flow through from 12347/udp for cthats, 12348/udp for ctags, rmc 657/udp and tcp.

For more information, refer to Appendix A, “Setting up a highly available network,” on page 267.

To create a two-node cluster, you need to perform these steps:

1. Log on to each node in the cluster:
   - **AIX/Solaris/Linux**:
     - Access a console on each node and log on as root.
     - Set the environment variable CT_MANAGEMENT_SCOPE to 2 on each node. To permanently set the environment variable to this value, you can add it to /etc/profile.
     - Remove extraneous entries that refer to the same host name from the /etc/hosts table (Example: 127.0.0.2 my_hostname). These entries are created during the installation of the operating system but conflict with the cluster infrastructure.
   - **Windows**:
     Perform these steps:
     a. Log in on each node with the user account that was used to install System Automation for Multiplatforms.
     b. Open a Windows Remote console.
     c. Open an "IBM Tivoli System Automation Shell". Typically, you find it in the folder Start > All Programs > Tivoli SA MP Base. If you have enabled the telnet service on the Windows nodes, you can open a telnet session.
     d. Ensure that /etc/hosts, /etc/services, and /etc/protocols are symbolic links to the corresponding files in the <WINDOWS>\system32\drivers\ etc directory.

2. Issue the preprpnode command on all nodes to allow communication between the cluster nodes.
   
   preprpnode node01 node02

3. You can now create a cluster with the name **SA_Domain** running on node01 and node02. The following command can be issued from any node.
   
   mkrpdomain SA_Domain node01 node02

   Note that when creating RSCT peer domains (clusters) using mkrpdomain, the characters used for the peer domain name are limited to the following ASCII characters: A-Z, a–z, 0-9, . (period), and _(underscore).

4. To look up the status of **SA_Domain**, issue the lsrpdomain command:
   
   lsrpdomain

   Output:

   Name OpState RSCTActiveVersion MixedVersions TSPort GSPort
   SA_Domain Offline 2.5.5.1 No 12347 12348
The cluster is defined but offline.

5. Issue the `startrpdomain` command to bring the cluster online.

```
startrpdomain SA_Domain
```

When you run the `lsrpdomain` command again, you see that the cluster is still in the process of starting up, the OpState is Pending Online.

```
Name     OpState        RSCTActiveVersion MixedVersions TSPort GSPort
SA_Domain Pending online 2.5.5.1            No    12347 12348
```

After a short time the cluster will be started, so when you issue the `lsrpdomain` command again, you will see that the cluster is now online:

```
Name     OpState        RSCTActiveVersion MixedVersions TSPort GSPort
SA_Domain Online      2.5.5.1            No    12347 12348
```

Notes:

1. You may receive an error message like:
   
   **2632-044** The domain cannot be created due to the following errors that were detected while harvesting information from the target nodes:
   
   node1: **2632-068** This node has the same internal identifier as node2 and cannot be included in the domain definition.

   This error most often occurs if you have cloned Linux images. Something went wrong with the cluster and the entire configuration should be reset. Solve such problems by running the following command on the node that is specified in the error message in order to reset the node ID:
   
   `/usr/sbin/rsct/install/bin/recfgct`

   Continue with the `preprpnod`e command.

2. You may also get an error message like:
   
   **2632-044** The domain cannot be created due to the following errors that were detected while harvesting information from the target nodes:
   
   node1: **2610-418** Permission is denied to access the resources or resource class specified in this command.

   Check your host name resolution. Make sure that all entries for each node of the cluster in your local `/etc/hosts` files on all nodes and the nameserver entries are identical.

3. You may get an error message like:
   
   **2612-022** A session could not be established with the RMC daemon on `<node_name>`.

   This error most often occurs if a firewall is running on one of the cluster nodes. To resolve the problem, enable the following ports in the firewall software:
   
   `rmc` port `657` TCP IN/OUT Source Port Range: ephemeral port
   `rmc` port `657` UDP IN/OUT Source Port Range: ephemeral port
   `cthats` port `12347` UDP IN/OUT Source Port Range: `1024 - 65535`
   `cthags` port `12348` UDP IN/OUT Source Port Range: `1024 - 65535`

   Note that all firewall rules should allow BROADCAST packets to go through the `cthats` port. For more information, refer to *IBM RSCT Administration Guide* (Appendix A. RSCT network considerations).

### Setting up a tie breaker

A tie breaker must be set up for all clusters that consist of an even number of nodes.

For two-node clusters, a tie breaker is required for the following reasons:
Getting started

- Without a tie breaker, System Automation for Multiplatforms automates resources only when more than half of the nodes of a cluster are online. For a two-node cluster without a tie breaker this means that both nodes must be online for automation to be performed.
- A tie breaker ensures that the cluster remains operational even when it is split in such a way that the communication between the two nodes fails.

A tie breaker should be set up for all clusters with an even number of nodes. The tie breaker ensures that situations can be resolved in which only half of the nodes of the cluster are online or when the cluster is split.

For detailed information about setting up, administering, and using tie breakers, refer to [Chapter 9, “Protecting your resources – quorum support,” on page 161](#).

**Adding a node to an existing cluster**

After having created the two-node cluster, you might want to add a third node to **SA_Domain**. In order to do this, you need to:

1. **Issue the `lsrpdomain` command to see if your cluster is online:**

   ```
   Name OpState RSCTActiveVersion MixedVersions TSPort GSPort
   SA_Domain Online 2.5.5.1 No 12347 12348
   ```

   Issue the `lsrpnode` command to see which nodes are online:

   ```
   Name OpState RSCT Version
   node02 Online 2.5.5.1
   node01 Online 2.5.5.1
   ```

2. **Issue the following `preprpnode` commands to allow communication between the existing nodes and the new node.**
   - Log on to node03 and enter:
     ```
     preprpnode node01 node02
     ```
   - Log on to node02 and enter:
     ```
     preprpnode node03
     ```
   - Log on to node01 and enter:
     ```
     preprpnode node03
     ```
   - It is strongly recommended that you issue a `preprpnode` command on each node for all nodes.

3. **In order to add node03 to the cluster definition, issue the `addrpnode` command on node01 or node02, which are already online on the cluster.**

   ```
   addrpnode node03
   ```

   Again issue the `lsrpnode` command to see the status of all nodes:

   ```
   Name OpState RSCT Version
   node02 Online 2.5.5.1
   node03 Offline 2.5.5.1
   node01 Online 2.5.5.1
   ```

4. **Start node03 from an online node:**

   ```
   startrpnode node03
   ```

   After a short time, node03 should be online as well.
Taking an entire cluster or individual nodes offline

In order to perform node maintenance or make application upgrades, you might want to take an entire cluster or individual nodes of a cluster offline:

- In order to perform maintenance on cluster **SA-Domain**, you might wish to take it offline. Use the **stoprpdomain** command from any online node in the cluster to do this.

```
stoprpdomain SA_Domain
```

Issue the **lsrpdomain** command to check the status of cluster **SA-Domain**:

```
Name    OpState    RSCTActiveVersion MixedVersions TSPort GSPort
SA_Domain Offline 2.5.5.1 No 12347 12348
```

As stopping a cluster does not remove the cluster definition, the cluster can be brought back online with the **startrpdomain** command.

- To take one or more cluster **nodes** offline, you use the **stoprpnode** command. You might need to take a cluster node offline, for example:
  - to apply application upgrades
  - to perform maintenance on a node
  - before removing the node from the cluster
  - Since a node may be defined in multiple clusters, but only online in one cluster at a time, you might also need to take a node offline in one cluster to be able to bring it online in another cluster.

To take a node offline, issue the **stoprpnode** command from any online node in the cluster, and pass to it the cluster node name of the node to take offline. For example, do this to stop node03:

```
stoprpnode node03
```

**Note**: Be careful when stopping multiple nodes of a cluster. You will lose quorum if less than half of the nodes are online, which may lead to outages if there are resources running on the online nodes of the cluster. See Chapter 9, “Protecting your resources – quorum support,” on page 161 for more information.

Issue the **lsrpnode** command to see if node03 has gone offline:

```
lsrpnode node03
Name    OpState    RSCT Version
node03  Offline  2.5.5.1
```

Removing nodes from a cluster, or removing a complete cluster

When upgrading hardware or otherwise reorganizing your overall cluster configuration, you may need to remove individual nodes from a cluster, or remove an entire cluster definition.

- To remove a node from a cluster, you use the **rmrpnode** command. In order to remove a node from a cluster, the node must be offline. If the node you wish to remove is not currently offline, you must use the **stoprpnode** command to take it offline. You can also remove multiple nodes from the cluster, using the **rmrpnode** command. In order to see which node is offline, issue the **lsrpnode** command from any online node in the cluster.

```
lsrpnode
Name    OpState    RSCT Version
node02  Online  2.5.5.1
node03  Offline  2.5.5.1
node01  Online  2.5.5.1
```
Then issue the `rmrpnode` command from any online node in the cluster to remove node03.

```
rmrpnode node03
```

Issue the `lsrpnode` command again to see if node03 has been removed.

```
lsrpnode
```

```
Name   OpState RSCT Version
node02 Online 2.5.5.1
node01 Online 2.5.5.1
```

- To remove a complete cluster definition, you use the `rmrpdomain` command. Removing a cluster involves removing the cluster definition from each node on the cluster. To do this efficiently, all nodes in the cluster should be online. You can bring individual nodes online using the `startrpnode` command, or you can bring all offline nodes in the cluster online using the `startrpdomain` command. The `rmrpdomain` command removes the cluster definition on all of the nodes that are reachable from the node where the command was issued. If the command is issued from an online node in a cluster, and all the nodes are online, then the command will attempt to remove all of their cluster definition files. If a node is not reachable from the node where the `rmrpdomain` command is run (for example, the node is offline or inoperative), the `rmrpdomain` command will not be able to remove the cluster definition on that node. In case the cluster cannot be brought online, you can use the force option `-f` to remove nodes or the cluster.

Issue the `startrpdomain` command to bring all nodes of cluster `SA_Domain` online:

```
startrpdomain SA_Domain
```

Then issue the `rmrpdomain` command to remove cluster `SA_Domain`:

```
rmrpdomain SA_Domain
```

**Administering the recovery resource manager**

On each node that is online in a cluster, an System Automation for Multiplatforms Recovery RM daemon (IBM.RecoveryRM) is running. The daemon is started automatically when the cluster is brought online.

To retrieve the status and the process ID of the daemon, use this command:

```
1ssrc -s IBM.RecoveryRM
```

The output is similar to the following:

```
Subsystem    Group    PID   Status
IBM.RecoveryRM rsct_rm 18283 active
```

If necessary, you can manually stop the daemon with the following command:

```
stopsrc -s IBM.RecoveryRM
```

To start the daemon you use:

```
startsrc -s IBM.RecoveryRM
```

One of the daemons is the so-called *master daemon*. This daemon is responsible for driving all the necessary decisions. To find out on which node the master daemon is located, use this command:

```
1ssrc -ls IBM.RecoveryRM | grep Master
```

You will get the following output:
Master Node Name : node03 (node number = 3)

In the example, the master daemon runs on a node called node03.

The other daemons are called ‘peer daemons’. These peer daemons are a hot standby if the master daemon or the node the master daemon is located on runs into problems. In this case, one of the peer daemons becomes the master. Of course, the takeover between the daemons is done without interruption of the automation functionality of System Automation for Multiplatforms.
Step 2: Defining RSCT resources

Command overview

The following gives you an overview of the Reliable Scalable Cluster Technology (RSCT) commands you will use when you define RSCT resources. You will need some of these commands in Step 2. For a detailed description of the commands, refer to the RSCT manuals listed on page [11].

- **chrsrc**: This command changes persistent attribute values of a resource within a specified resource class.
- **lsrsr**: This command lists resources of a resource class.
- **mkrsr**: This command creates resources of a specified resource class.
- **resetrsr**: This command resets resources of a specified resource class.
- **rmrsr**: This command removes resources of a specified resource class.
- **runact**: This command runs an action on a resource class.
- **startrsrc**: This command brings an individual resource online.
- **stoprsr**: This command brings an individual resource offline.

Defining RSCT resources for a Web server

The following example shows how to define a highly available Web server on the three nodes of the cluster SA_Domain. ["Step 1: Defining and administering a cluster" on page 12] describes how this cluster and its nodes (node01, node02, and node03) are defined.

The following requirements have to be met for the highly available Web server:

- The Web server should be startable on any node in the cluster but will only run on one node at any point in time.
- In case of a failure, the Web server should automatically be restarted on the same or on a different node in the cluster. This mechanism can also be utilized for planned outages of nodes for service and maintenance.
- The Web server should be addressable with the same IP address, regardless of which node it currently runs on. Thus the location of the Web server is transparent outside the cluster where no adaptation has to be performed when the Web server is moved to another node.

As the basis for automation, the components involved must first be described in a set of RSCT defined resources. Due to often uncommon characteristics of resources, there are various RSCT resource classes to accommodate the differences. In this example we will need to define three RSCT resources from different classes:

1. An application resource named **apache1**, which represents the Web server daemon. The resource is from a class called IBM.Application. apache1 will be a floating resource, since the Web server is not tied to a particular node in the cluster.
Getting started

2. An IP address named **apache1IP**, which is used to represent the Web server’s IP address. apache1IP is from a class called IBM.ServiceIP. apache1IP will be a floating resource, since it can move around in the cluster following the location of the Web server.

3. An representation named **netequ** for the network interface cards that can be used for the apache1IP address. This is called an equivalency and belongs to a class IBM.Equivalency. The characteristic “floating” or “fixed” is not meaningful for this class.
Creating the application resource apache1

As part of the definition of application resource apache1, commands or scripts for starting, stopping and querying the Web server have to be specified. These commands and/or scripts can be different ones, but it is often convenient to gather these functions in a single script, which has a command line parameter to select start/stop/status actions. These scripts will often be user-written. See Chapter 11, "Automating System Automation for Multiplatforms resources," on page 221 for details on the requirements of such scripts.

In this example, we use a script

```
/cluster/scripts/apache
```

which has the following content for a Linux system:

```
#!/bin/bash

OPSTATE_ONLINE=1
OPSTATE_OFFLINE=2

Action=${1}

case ${Action} in
  start)
    /usr/sbin/apachectl start >/dev/null 2>&1
    logger -i -t "SAM-apache" "Apache started"
    RC=0
    ;;
  stop)
    /usr/sbin/apachectl stop >/dev/null 2>&1
    logger -i -t "SAM-apache" "Apache stopped"
    RC=0
    ;;
  status)
    ps -ax |grep -v "grep"|grep "/usr/sbin/httpd">/dev/null
    if [ $? == 0 ]
      then
        RC=${OPSTATE_ONLINE}
      else
        RC=${OPSTATE_OFFLINE}
      fi
    ;;
  esac

exit $RC
```

Make sure to make the script accessible on all nodes with the same directory path.

RSCT resource definitions are created with the `mkrsrc` command. All resource characteristics can be provided in command line parameters, but the `mkrsrc` command also accepts a definition file in plain text. We will use the second approach with a definition file named apache1.def, which may look like the following:

```
PersistentResourceAttributes::
  Name="apache1"
  StartCommand="/cluster/scripts/apache start"
  StopCommand="/cluster/scripts/apache stop"
  MonitorCommand="/cluster/scripts/apache status"
  MonitorCommandPeriod=5
  MonitorCommandTimeout=5
  NodeNameList="node01","node02","node03"
  StartCommandTimeout=10
  StopCommandTimeout=10
  UserName="root"
  ResourceType=1
```
The resource definition can now be created with the `mkrsrc` command using the definition file.
```
mkrsrc -f apache.def IBM.Application
```

**Creating the IP address resource apache1IP**

The web server's IP address apache1IP is a separate IP address in the cluster and does not match any IP address assigned to the network adapters on each cluster node, that are made in system definitions outside of System Automation for Multiplatforms. The address for apache1IP is in contrast created by System Automation for Multiplatforms and is an additional alias address on an appropriate network adapter on the node where the web server resides. When the web server moves to a new location, the alias address is removed from the former node and recreated on the new node, where the web server is about to be restarted.

In this example apache1IP has the following attributes:

- IP 9.152.172.11
- Netmask 255.255.255.0
- The IP address may be created on any node in the cluster.

This time, we use command line parameters to the `mkrsrc` command to create the apache1IP resource:
```
mkrsrc IBM.ServiceIP \  
  NodeNameList="{'node01','node02','node03'}" \  
  Name="apache1IP" \  
  NetMask=255.255.255.0 \  
  IPAddress=9.152.172.11
```

Note that the command shown is split onto separate lines for readability only. In fact, the apache1IP has more attributes than the ones specified in the command shown. We leave the rest of them to their default values, such as the ResourceType attribute, which marks the resource as “floating” by default.

Also note that the managed resources are not started/stopped by a third party like, for example, the Linux run level or manually by the operator.
Step 3: Defining the automation policy from the command line

Before defining the automation policy there are some important things to get familiar with:

- By adding a resource into a resource group, this resource is automated and controlled by System Automation for Multiplatforms. This means that System Automation for Multiplatforms will manage to change the operational state (OpState) of the resource to the defined desired/nominal state (NominalState) of the resource group. If any resource automated by System Automation for Multiplatforms is manually started or stopped by an operator, System Automation for Multiplatforms will recognize this by monitoring this resource, and finally start or stop the resource again to bring it to the state defined in the automation policy.

If a resource is defined in the automation policy, then System Automation for Multiplatforms will control this resource, and every manually operated action against this resource will usually trigger a reverse action of System Automation for Multiplatforms.

This is also true for all resources on nodes that are excluded from automation (see the samctrl command described in IBM Tivoli System Automation for Multiplatforms Reference).

If at a certain point in time System Automation for Multiplatforms should not control the resources, then automation can be manually set with the following command:

```
# samctrl -M T
```

To return the control over the resources back to System Automation for Multiplatforms it has to be set in automatic mode again, using the following command:

```
# samctrl -M F
```

See the samctrl and lssamctrl commands described in IBM Tivoli System Automation for Multiplatforms Reference for further information.

- Resources (applications) that are controlled by System Automation for Multiplatforms can only be safely started and stopped by using the appropriate System Automation for Multiplatforms commands (chrg, rgreq or rgmbrreq).

Only these commands will reflect the change of the desired state of a resource to System Automation for Multiplatforms. If a resource is started or stopped using its ‘native’ command, then System Automation for Multiplatforms still has the ‘old’ desired state of this resource in the automation policy, and finally will trigger a start or stop command against this resource to bring it again into the desired state.

- Special maintenance considerations apply to nodes that are running in a System Automation for Multiplatforms domain. This is especially important for the commands used for stopping nodes (AIX/Solaris/Linux: halt and reboot, Windows: shutdown, reboot, hibernate, freeze).

Before a node is stopped, it should be added to the automation exclude list, which will trigger System Automation for Multiplatforms to stop all resources running on that node and to move these resources to another node. After all resources are offline on that node, the node can be safely stopped or rebooted. Failure to do so will result in an immediate reboot, if any ‘critical resource’ is running on the stopped/rebooted node (see Chapter 9, “Protecting your resources – quorum support,” on page 161). The following command adds a node to the exclude list:

```
# samctrl -u a <node-name>
```
Getting started

Now the node can be safely stopped. After the node is online again in the domain, it must be deleted from the exclude list with the following command:

```
# samctrl -u d <node-name>
```

- System Automation for Multiplatforms will not start resources on nodes that are on the exclude list, even if the other nodes are offline, so a node should only stay on the exclude list for the time of the maintenance. The exclude list can be checked using the following command:

```
# lssamctrl
```

- System Automation for Multiplatforms domains with two nodes (or an even number of nodes) require a tie breaker to be set up. Without a tie breaker, the second node will not take over resources that were running on a node that crashed or was rebooted. See Chapter 9, “Protecting your resources – quorum support,” on page 161 for tie breaker types and setup.

The following examples show how the resources `apache1` and `apache1IP` are turned into System Automation for Multiplatforms managed resources (see “What is a managed resource?” on page 30) providing high availability for a web server in a cluster environment. See “Step 2: Defining RSCT resources” on page 19 how the resources `apache1` and `apache1IP` were defined.

To turn resources "apache1" and "apache1IP" into managed resources, they have to be added to a resource group. When this has been done, System Automation for Multiplatforms starts controlling the resource group and its member resources.

In most cases, it is not enough to automate a managed resource on its own, because the resources are often related to each other. For instance, both resources "apache1" and "apache1IP" from our example must be made available on the same node. Such dependencies between managed resources must be described and defined through managed relationships (see “What is a managed relationship?” on page 57).

At last, an automation goal must be specified, that is: should a managed resource be available/started in the cluster or should it be offline.

**Command overview**

The following list provides an overview of the commands you use for defining policies. For a detailed description of the commands, refer to [IBM Tivoli System Automation for Multiplatforms Reference](#).

- `mkequ` Makes an equivalency resource
- `chequ` Changes a resource equivalency
- `lsequ` Lists equivalencies and their attributes
- `rmequ` Removes one or more resource equivalencies
- `mkrg` Makes a resource group
- `chrg` Changes persistent attribute values of a resource group (including starting and stopping a resource group)
- `lsrg` Lists persistent attribute values of a resource group or its resource group members
- `rmrg` Removes a resource group
- `mkrel` Makes a managed relationship between resources
Creating an equivalency definition for network adapters

When a node in the cluster has multiple network attachments, not all of them might be equally suited to host the apache1IP address as an alias. An equivalency definition will specify the network adapters that can be used to carry the apache1IP address. Equivalency means that each of the adapters in the equivalency can provide the same required function regardless of its own unique characteristics. Since the web server should be startable on each node in the cluster, at least one of the adapters on each node has to appear in the equivalency.

An equivalency groups together a set of resources from another class. Network adapters belong to a class named IBM.NetworkInterface. There is no need to provide resource definitions for all the network adapters on the cluster nodes, since RSCT has a harvesting function which automatically creates appropriate resource definitions for many system defined resources.

The following command (Linux sample) creates a so-called static equivalency named netequ, which contains a network adapter from each node of the cluster:

```
mkequ netequ IBM.NetworkInterface:eth0:node01,eth0:node02,eth0:node03
```

In a static equivalency, such as the one created with the command above, the members of the equivalency are specified at the creation time of the equivalency.
The advantage of this equivalency type is that an order can be specified for the resources contained in the equivalency. A disadvantage is that network interfaces that are deleted from the cluster definition are also removed from the equivalency and will not be added automatically again if they are harvested later, thus resulting in a reduced member set.

A network interface is deleted from the cluster definition if its IP address is removed, or if the network interface is detached. If the network interface is activated again later, it will be harvested by System Automation and again included in the cluster definition, but it will not be automatically added to the equivalency definition again.

To overcome this limitation, you can create an equivalency with a dynamic selection string using the following command (Linux sample):

```
mkequ -D 'Name like "eth0" ' netequ IBM.NetworkInterface
```

This definition includes all ‘eth0’ network interfaces available in the cluster. The selection string is evaluated each time System Automation has to find out on which node to start resources. The list of available ‘eth0’ resources may change as System Automation harvests new network interfaces in the cluster periodically, or when network interfaces that are undefined (IP address is removed or detached) are deleted from the cluster definition. However, equivalencies with a dynamic selection string cannot be ordered.

### Creating a resource group for apache1 and apache1IP

A resource group is created with the `mkrg` command. The following command creates a resource group named "apacherg":

```
mkrg apacherg
```

Both resources, "apache1" and "apache1IP", will be added to the resource group "apacherg". This is done with the `addrgmbr` command. Adding the resources to the resource group turns them into managed resources:

```
addrgmbr -g apacherg IBM.Application:apache1
addrgmbr -g apacherg IBM.ServiceIP:apache1IP
```

### Defining relationships between the Web server resources

There are two conditions that relate resources apache1 and apache1IP to one another:

- Both resources must be started/available on the same node in the cluster. This is called a collocated relationship (see "Collocated relationship" on page 79).
- It is futile to start the Web server apache1 on a node, on which the IP address apache1IP has not been established yet. That is, apache1IP must be available before apache1 can be started.

System Automation for Multiplatforms provides a relationship type called DependsOn (see "DependsOn relationship" on page 68) that gathers both required conditions. A managed relationship is defined with the `mkrel` command. The following command creates a managed relationship named `apache1_dependson_ip1` that establishes the dependency of resource apache1 on the IP address apache1IP:

```
mkrel -p DependsOn -S IBM.Application:apache1 -G IBM.ServiceIP:apache1IP apache1_dependson_ip1
```

Our example needs a second relationship. In "Step 2: Defining RSCT resources" on page 19 we have created an equivalency of those network adapters, that can be
Getting started

used for aliasing the apache1IP address. This will be described in a second relationship called `apache1IP_dependson_netequ`, that ties apache1IP and the equivalency netequ together:

```
mkrel -p DependsOn -S IBM.ServiceIP:apache1IP -G IBM.Equivalency:netequ apache1IP_dependson_netequ
```

**Bringing the Web server resource group online**

When resources are added to resource groups, they become managed resources with a default automation goal of offline. This can be changed at the level of the resource group with the `chrg` command. To bring resource group apacherg online use the command:

```
chrg -o online apacherg
```

After your clusters and nodes have been created and configured, you can begin to use System Automation for Multiplatforms commands to:

- make, remove, change, list, and check the online status of resource groups. See Chapter 4, “Using resource groups,” on page 35 for details.
- make, remove, change, list, and check the online status of the resource group member resources. See Chapter 3, “Using resources,” on page 29 for details.
- make, remove, change, list, and check status details of equivalency resources. See Chapter 5, “Using equivalencies,” on page 51 for details.
- make, remove, change, list, and check status details of managed relationship resources. See Chapter 6, “Using managed relationships,” on page 57 for details.

The complete list of System Automation for Multiplatforms commands is provided in *IBM Tivoli System Automation for Multiplatforms Reference*. 
Chapter 3. Using resources

The following main sections of this chapter describe how you use resources:

- “What is a resource?”
- “What is a managed resource?”
- “Attributes used by resources”

These are the System Automation for Multiplatforms commands that you use for managed resources:

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>addrgmbr</td>
<td>Adds one or more resources to a resource group</td>
<td>For details, see the command description in the IBM Tivoli System Automation for Multiplatforms Reference.</td>
</tr>
<tr>
<td>chrgmbr</td>
<td>Changes persistent attribute values of a managed resource in a resource group</td>
<td></td>
</tr>
<tr>
<td>lsr</td>
<td>Lists an already-defined resource group or its member resources</td>
<td></td>
</tr>
<tr>
<td>rmrgmbr</td>
<td>Removes one or more resources from a resource group</td>
<td></td>
</tr>
</tbody>
</table>

What is a resource?

A resource is any piece of hardware or software that has been defined to RMC in any of the ways described in “Resource”.

As described in “RSCT as an integral part of System Automation for Multiplatforms” on page 7, System Automation for Multiplatforms uses as its basis the functionality of RMC. A resource is therefore sometimes referred to as an RMC resource.

Resources (adapter, program, services, disk, and so on) are controlled by a Resource Manager (abbreviated to RM).

What is a resource class?

A resource class is a set of resources of the same type. For example, while a resource might be a particular file system or particular host machine, a resource class would be the set of file systems, or the set of host machines. A resource class defines the common characteristics that instances of the resource class can have (for example, all file systems will have identifying characteristics (such as a name), as well as changing characteristics (such as whether or not it is mounted). Each individual resource instance of the resource class will then define what its particular characteristic values are (for example, this file system is named “/var”, and it is currently a mounted file system).

What are resource attributes?

A resource attribute describes some characteristics of a resource. If the resource represents a host machine, its attributes would identify such information as the host name, size of its physical memory, machine type, and so on.
There are two types of resource attributes – persistent attributes and dynamic attributes.

**Persistent attributes**

The attributes of a host machine that were mentioned above (host name, size of physical memory, and machine type) are examples of persistent attributes – they describe enduring characteristics of a resource. Although you can change the host name or increase the size of its physical memory, these characteristics are in general stable and unchanging.

**Dynamic attributes**

Dynamic attributes represent changing characteristics of the resource. Dynamic attributes of a host resource, for example, would identify such things as the average number of processes that are waiting in the run queue, processor idle time, the number of users currently logged on, and so on.

**What is a managed resource?**

A resource becomes an *System Automation for Multiplatforms managed resource* (referred to simply as a *managed resource*) as soon as the resource has been inserted in a System Automation for Multiplatforms resource group. This is done using the System Automation for Multiplatforms `addrgmbr` command (described in [IBM Tivoli System Automation for Multiplatforms Reference](#)). From this point onwards, the resource can be managed using the System Automation for Multiplatforms commands and programs.

Managed resources are supplied in System Automation for Multiplatforms resource class IBM.ManagedResource.

**Working with resources**

Note that System Automation for Multiplatforms does not allow you to start or stop resources directly.

Resource starting and stopping is based on the setting of the NominalState attribute of a resource group. For example, setting the NominalState attribute of a resource group to "Online" indicates that you wish to start the resources in the resource group. Setting the NominalState attribute of a resource group to "Offline" indicates that you wish to stop the resources in the resource group. See the description in "Starting and stopping a resource group" on page 49 and "NominalState attribute" on page 41.

**Attributes used by resources**

Resources can have the following attributes:

**NodeNameList**

Indicates on which nodes the resource is allowed to run. NodeNameList is a persistent attribute of an RSCT resource.

**SelectFromPolicy**

Contains the list of nodes on which the resource is available. SelectFromPolicy is a persistent attribute of a managed resource.
Attributes used by resources

**ResourceType**  Indicates whether the resource is allowed to run on multiple nodes or on a single node only. ResourceType is a persistent attribute of an RSCT resource.

**OpState**  Indicates the operational state of a resource or a resource group. OpState is a dynamic attribute of an RSCT resource.

**NodeNameList attribute**

The NodeNameList persistent attribute represents the set of nodes on which the resource can run.

The nodes resources started on by System Automation for Multiplatforms are controlled by the order of the nodes in the NodeNameList of the resources. By default, each floating resource will be placed on the first node in its node list where it can be started with respect of all its relationships to other resources. This behavior can be modified by using the SelectFromPolicy attribute described below. Floating resources which should be collocated are placed before independent floating resources (these are resources which do not have relationships to other resources) and anticollocated floating resources.

If there are collocated resources with different node lists, the resources are placed on the node which is chosen by the majority of resources. In a tie situation, one node is chosen randomly.

**SelectFromPolicy attribute**

As described above, the NodeNameList persistent attribute defines the list of nodes the resource is available on. This list is ordered, either in the sequence as specified by the user or in the sequence the resource manager owning the resource put them in. The SelectFromPolicy attribute gives the user more flexibility. It allows to tell System Automation for Multiplatforms which algorithm to use when selecting a node from the list. This can be either ordered, meaning that System Automation for Multiplatforms always starts from the beginning of the list when determining the next available node, or any, meaning that System Automation for Multiplatforms picks a node randomly.

**ResourceType attribute**

The ResourceType attribute is either defined by the resource manager or during creation of the resource. The ResourceType persistent attribute specifies whether a resource is:

- **Serial fixed** (its NodeNameList attribute contains a single node entry).

  **Fixed resource**

  A fixed resource is a resource of which there is only a single instance within the cluster. It is defined upon a single node, and that is where it runs. It represents one entity such as a process, a mount point, or a network adapter.

  If you want to define an application which runs on more than one node in the cluster at a time, you must define one fixed resource for each node, specifying the same attributes (for example, name, start/stop/monitor commands) but a different node name for each fixed resource.

- **Serial floating** (its NodeNameList attribute contains one or multiple entries).

  Although multiple nodes are defined for possible use, only one instance of the resource may be active at any time. For example, an IP address that can be
Attributes used by resources

moved from one machine to another is a floating resource; although multiple machines may have use of the IP address at some point, only one machine at a time will use it.

Floating resource

A floating resource is a resource which can run on several nodes in the cluster. A floating resource is represented in RMC in the following way:
You have one aggregate resource and one constituent resource on each node belonging to the aggregate resource.

The aggregate resource has a ResourceType attribute value of 1. The set of nodes where the resource should be able to run is defined in the NodeNameList attribute of the aggregate resource. The other attributes of this resource are defined by the resource manager and its class definition.

If you create a floating resource you create the aggregate resource. The resource manager responsible for this type of resource will create constituent resources on each node the resource is supposed to run on. The constituent resources have their own values of the attributes. The ResourceType of a constituent resource is 0 (a fixed resource), and the NodeNameList contains one node only. At time of creation the other attributes have identical values as the aggregate resource.

The following happens if you change attributes of a floating resource:
- A change of the NodeNameList of the aggregate resource causes deletion or creation of constituent resources.
- If you change an attribute of the aggregate resource this changes the according attributes of all constituent resources.
- If you change an attribute of a constituent resource this affects the constituent resource only and is not conveyed to other constituent resources or the aggregate resource.

A floating resource represents an automatable entity such as an application or a service ip address which can run on several nodes.

Note: A floating resource is labelled as move group in the operations console.

OpState attribute

RMC uses the OpState dynamic attribute to specify the operational state of a resource. It is mandatory for resources added to a resource group.
Attributes used by resources

These are the possible values that the OpState attribute can have:

**Offline** The resource is not started.

**Pending Online** The resource has been started, but is not yet ready for work

**Online** The resource is ready for work.

**Pending Offline** The resource is in the process of being stopped

Some of the operational states indicate problems:

**Failed Offline** The resource is broken and cannot be used. You have to reset the resource when you have fixed it.

**Stuck Online** The resource was being started, but did not become ready for work within the expected time interval and cannot be brought offline. Another possibility is that the resource was online, but an offline request could not bring it offline.

**Unknown** System Automation for Multiprocessing is unable to obtain reliable state information from the RMC managing the resource.

**Note:** You might have to reset a resource that has the Failed Offline state. To do so, use the RMC command `resetrsrc`. For details, refer to the man page for this command, or the RSCT documentation (see “Related information” on page xiii).

When a node of a resource is Offline, the resource is considered to be Failed Offline, even though its operation state at that point is Unknown. System Automation for Multiprocessing can do this because it has separate state data for the resources node.

**NominalState can only be set for resource groups**

The NominalState attribute can only be set for resource groups. The value of the NominalState of a resource group (Online or Offline) determines whether all of its members are automated to be online or offline. The fact that individual resources do not have a NominalState attribute is one of the reasons why all resources must be members of a resource group.
Attributes used by resources
Chapter 4. Using resource groups

This chapter describes how you use resource groups, in these main sections:

- "What is a resource group?"
- "Attributes used by resource groups" on page 38
- "Making (creating) a resource group" on page 47
- "Changing attributes of a resource group" on page 49
- "Removing a resource group" on page 50
- "Listing resource groups and their members" on page 48
- "Adding a member resource to a resource group" on page 47
- "Removing a member resource from a resource group" on page 49
- "Changing the attributes of resource group members" on page 49

Related Section:
- "Events that might allow a resource group to become Online" on page 94

Table 2 lists the System Automation for Multiplatforms commands that you use together with resource groups. For details, see the command description in IBM Tivoli System Automation for Multiplatforms Reference.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mkrg</td>
<td>Make a resource group</td>
</tr>
<tr>
<td>rmrg</td>
<td>Remove a resource group</td>
</tr>
<tr>
<td>chrg</td>
<td>Change persistent attributes of a resource group (including starting and stopping a resource group)</td>
</tr>
<tr>
<td>lsrg</td>
<td>List one or more resource groups</td>
</tr>
<tr>
<td>addrgmbr</td>
<td>Add member resources to a resource group</td>
</tr>
<tr>
<td>rmrgmbr</td>
<td>Remove member resources from a resource group</td>
</tr>
<tr>
<td>chrgmbr</td>
<td>Change attributes of the member resources of the resource group</td>
</tr>
<tr>
<td>rgreq</td>
<td>Start, stop, cancel, or move a resource group</td>
</tr>
<tr>
<td>rgmbreq</td>
<td>Request a managed resource to be started, stopped, or canceled</td>
</tr>
<tr>
<td>lsrgreq</td>
<td>List the outstanding requests applied against resource groups or managed resources</td>
</tr>
</tbody>
</table>

What is a resource group?

Resource groups are the central unit in System Automation for Multiplatforms. They are logical containers for a collection of resources that can be treated as one logical instance:

- You can use resource groups to control all of their members collectively. For example, if you set a resource group’s NominalState to Online all members are started and kept online. If you set the NominalState to Offline, all members are stopped and kept offline.
- You can monitor their OpState, which provides a consolidation of the OpStates of the individual resource group members.

Members of a resource group can be of type:
Resource Groups

- Serial fixed.
- Serial floating.
- And even resource groups itself which means that nested groups can be defined.

An example for a resource group containing fixed resources is a resource group RG_Fix which contains serial fixed resources. These are a web server FixWebServer which can only run on node1, and a database resource FixDB2 located on node2.

In order to start both resources FixWebServer and FixDB2, set the NominalState of RG_Fix to Online. This example also shows that System Automation for Multiplatforms can handle resource group members that are distributed on different nodes in a cluster.

An example for floating resource group members is the following: A web server apache1 could run either on node1, node2, or node3. The resource group RG_WebApp would look very similar except that the web server could be started on either of the three nodes.

This example shows that resource groups can contain a mixture of members of different resource types.

The concept of resource groups is very powerful as it allows defining resource groups as members of other resource groups. An example is resource group RG_A which has as members resource A, which is a fixed resource, and RG_WebApp, the
Resource Groups

resource group from the previous example. Nested resource groups allow structuring complex environments in several layers. The nesting level is 50.

Another flexibility of the resource groups functionality is that all kinds of relationships like start/stop relationships and location constraint relationships can be defined with resource groups as source or target resource. Furthermore it is allowed that resource group members can be part of such relationships as source or target resource.

Resource groups are defined in System Automation for Multiplatforms resource class IBM.ResourceGroup.

Rules for using resource groups

These are the rules for using resource groups:
1. A resource group cannot contain an equivalency or vice versa.
2. A resource can only be in one group.
3. A member cannot be in a group and in an equivalency.
4. The nesting level of a resource group is limited to 50.
5. The number of resources linked by groups or relationships is limited to 100.
Resource Groups

Attributes used by resource groups

A resource group provides the following persistent RSCT attributes which can be defined by the user:

AllowedNode limits the nodes on which the resource group members are allowed to be started.

MemberLocation defines if all members of a resource group have to be collocated or not. Collocated means that all members have to run on the same node. None means that the members are not dependent from each other and can arbitrarily run on the nodes on which they are defined.

Name defines a unique name for a resource group.

NominalState is the desired state of the resource group. System Automation for Multiplatforms tries to bring up and keep the resource group in this state.

Priority defines the importance of a resource group in a conflict situation.

ExcludedList defines a list of nodes which are temporarily excluded from the node list of the group.

ActivePeerDomain is the name of the peer domain the group is defined to. This cannot be set by the user.

Description can contain descriptive text about the resource group.

InfoLink can be used to specify a URL of an HTML page where the operator can find additional information about the resource.

Owner provides information about the owner of the resource group, for example, the name and telephone number of a responsible person.

Subscription can contain subscription information from end-to-end management.

Note: The persistent attributes described in this section can only be modified if the resource group containing them is Offline. An exception are the NominalState attribute and the informational attributes Description, InfoLink, and Owner, which can also be modified when the resource group is Online.

A resource group provides the following dynamic attributes:

OpState Specifies the aggregate operational state of the collection of managed resources.

TopGroup Shows the name of the top level resource group of a resource group.

AutomationDetails Shows System Automation for Multiplatforms internal states of the resource group.

MoveStatus Shows the progress of a move of a resource group initiated by an rgreq command.

ConfigValidity Shows if a policy has become invalid.
**AllowedNode attribute**

You use the AllowedNode attribute to define a set of nodes in a cluster on which the members of a resource group are limited to run.

You can choose between the following parameters:

- **All** is the default value. It means that no limitation is made by the resource group. It can run on all nodes in the cluster.

- **One node** defines a specific node on which all resource group members have to run on. If the specified node is removed from the cluster at a later time, then AllowedNode will default to All.

- **Equivalency of nodes** contains a set of nodes where the resource group members are limited to run on. Only static equivalencies are allowed. See also Chapter 5, “Using equivalencies,” on page 51.

**Node limitation aspect of AllowedNode parameter**

In certain cases it may be necessary to limit a member of a resource group to run on a set of nodes. For example, when a floating resource is defined a NodeNameList has to be specified which is in general independent from the System Automation for Multiplatforms usage. The NodeNameList of the floating resource is used by the resource managers (for instance GblResRM) which own a floating resource. For them the NodeNameList defines on which nodes the constituents of a floating resource could potentially run. Whereas the AllowedNode attribute belongs to a resource group parameter that has different resource group members which are floating resources. Here the AllowedNode attribute allows to limit the nodes on which the resource group members are allowed to run. This means for resource group members of type:

- **Fixed resource**, that the NodeNameList which contains only one node on which the fixed resource is located has to be part of the AllowedNode parameter.

- **Floating resource**, that the intersection of the NodeNameList and the AllowedNode parameter defines the set of nodes on which the floating resource is started and controlled by System Automation for Multiplatforms.

- **Resource Group**, that an intersection of the AllowedNode parameter of the inner and the outer group derives a resulting list of allowed nodes for the inner group. The resulting list defines the nodes on which the members of the inner resource group are allowed to run on.

The following example explains the behavior of the AllowedNode parameter on the node limitation: Given an outer resource group RG_A with a member FixA with a NodeNameList = {node1}, a floating member FloatB with a NodeNameList = {node1, node2, node3} and a resource group RG_B with AllowedNode = {node2, node3, node4}. The AllowedNode list of RG_A defines {node1, node2, node 4}. RG_B contains FloatC with a NodeNameList of {node1, node2, node3, node4} and FloatD with NodeNameList of {node3, node4}.
The result of this scenario is that:
- FixA can only be started by System Automation for Multiplatforms on node1.
- FloatB can only be started by System Automation for Multiplatforms on node1 and node2.
- RG_B’s members are limited to run on node2 and node4.
- FloatC can only be started by System Automation for Multiplatforms on node2 and node4.
- FloatD can only be started by System Automation for Multiplatforms on node4.

The AllowedNode list of an inner group is not affected when its outer group has no node limitation due to its intersection of the own AllowedNode parameter and the intersection with the AllowedNode list of its outer group.

MemberLocation attribute

You use the MemberLocation persistent attribute to specify the default location between resources in the resource group. Valid values are either:

Collocated (the default)
Collocated means that all members have to run on the same node. If you specify Collocated, you cannot apply the Affinity, AntiAffinity, and AntiCollocated managed relationships between the resources in the resource group to specify on which nodes the member resources should be located. See Chapter 6, “Using managed relationships,” on page 57 for a detailed explanation of relationships.

None
None means that the resource group implies no restriction regarding the location of its members.

The MemberLocation attribute is applied to all member resources of the resource group.

If resource groups are nested, the value of the MemberLocation attribute of the outer resource groups must allow the inner-group(s) specification. Therefore, if a resource group’s MemberLocation attribute is collocated, only collocated resource groups are allowed as members.
**Name attribute**

The Name of a resource group has to be unique in a cluster.

**NominalState attribute**

You use the NominalState persistent attribute to control a resource group. By setting the NominalState of a resource group to Online all of its members are started and kept online.

The NominalState attribute can be either Online or Offline. A NominalState of Offline causes that all resource group members are stopped and kept offline. There are exceptions:

1. If resource groups are nested, a NominalState of online of an outer group overrules a NominalState of offline of an inner group and starts these inner groups. You cannot stop such an inner group separately.

2. If resource groups are nested, a NominalState of offline of an outer group will not overrule a NominalState of online of an inner group. A NominalState of online of an inner group will start this group and groups contained in this inner group. The outer groups’ OpState will change, but the other group members remain untouched.

3. Cross-group dependencies might force individual group members to be started. See "Relationships for start / stop behavior" on page 60.

The default value for the NominalState of a resource group is Offline.

This attribute can be modified at any time.

**Note:** You can modify this attribute with the `chrg -o` command (see the description of the `chrg` command in [IBM Tivoli System Automation for Multiplatforms Reference](#)). The numeric value of this attribute is

1   Offline
0   Online

**Priority attribute**

You use the Priority persistent attribute to specify the relative priority of this resource group in relationship to other resource groups.

The Priority attribute is used to resolve conflicts when resource groups are being started that have conflicting managed relationships (described in Chapter 6, “Using managed relationships,” on page 57) to other started or online resources. These conflicts may be between the resource group or a managed resource contained in that group and any other started or running resource.

If a group has members in state Unbound or Offline, there is no adjustment value of +10. For resource groups that have the same actual priority, the number of non-mandatory members determines the order. The group with the lowest number of non-mandatory members is sacrificed first. If the groups have the same actual priority and the same number of non-mandatory members, the algorithm decides incidentally which group it will be sacrificed.

For instance, you have a cluster with only one node online, and two resource groups with an AntiCollocated relationship against each other. This means that the resource groups must never be started on the same node at a time. System Automation for Multiplatforms now uses the value of the Priority attribute of the
resource groups to find out which of the resource groups should be online, and which cannot be online because of the conflicting AntiCollocated relationship.

If a lower priority active resource group prevents the activation of a higher priority resource group because of conflicting relationships, the lower priority resource group is stopped in order to allow the activation of the resource group with higher priority to proceed.

If resource groups are nested, the outer resource group must have a higher or equal priority than any of the inner resource groups.

The default value of the Priority attribute is zero, which is the lowest value. The maximum value is 200.

The actual priority of a resource group is computed from the value of the priority attribute and an adjustment value which is determined by the current observed state of the group (see “AutomationDetails attribute” on page 44). The adjustment value for groups that are online is +10, to make sacrificing started resources harder, and -10 for groups that are failed offline. The actual priority value can assume values from -10 to +210. (For more information on the meaning of the priority, see “Location relationship resolution: Binding algorithm” on page 91.)

Hint: System Automation for Multiplatforms also uses the Mandatory attribute of the managed resources (described in “Attributes used for resource group members” on page 46) to determine which resources will be started in a conflict situation. In case of nested resource groups, the resource groups that are non-mandatory members should have a lower priority than mandatory members. Otherwise Mandatory members may be discarded.

**ExcludedList attribute**

You use the ExcludedList attribute to temporarily exclude one or a list of nodes from the node list of the group. When excluding a node, the resources residing on the node being excluded are not automatically forced offline. The move must be triggered by means of the `rgreq` command.

This means that placing a node in the exclude list causes System Automation for Multiplatforms to not consider the node as a potential candidate for hosting the resource. It can be used to gradually and non-disruptively move resources away from a node in preparation for an EXCLUDE (via the `samctrl` command) at a later point in time.

The following rules apply:
1. Excluding a node means for a fixed resource group member that the resource cannot be started anymore.
2. Excluding a node means for a floating resource group member that its constituent on that node cannot be started any more.

**ActivePeerDomain**

This attribute shows the name of the RSCT Peer Domain the group is defined in.

**Description**

This attribute can contain descriptive text about the resource group. This attribute only serves informational purposes and does not affect automation functions.
InfoLink
This attribute can be used to specify a URL of an HTML page where the operator can find additional information about the resource. This attribute only serves informational purposes and does not affect automation functions.

Owner
This attribute provides information about the owner of the resource group, for example, the name and telephone number of a responsible person. This attribute only serves informational purposes and does not affect automation functions.

Subscription
This attribute contains subscription information from end-to-end management.

OpState attribute
System Automation for Multiplatforms uses the OpState dynamic attribute to specify the aggregate operational state of the collection (of managed resources). It is determined from the individual operational states of the member resources of the resource group. These are the possible values that the OpState attribute can have:

<table>
<thead>
<tr>
<th>Status</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td>1</td>
<td>Specifies that all of the Mandatory member resources are Online. The Non-Mandatory member resources are ignored.</td>
</tr>
<tr>
<td>Offline</td>
<td>2</td>
<td>Specifies that all of the member resources are Offline.</td>
</tr>
<tr>
<td>Failed Offline</td>
<td>3</td>
<td>Specifies that one or more member resources contained in the resource group, are FailedOffline. In this case, all resources contained in the resource group will be set to Offline. A resource group can reach the status Failed Offline as a result of a binder run if a resource group’s members could not be placed (see also “Location relationship resolution: Binding algorithm” on page 91). If this is the case, the automation details show a BindingState of Sacrificed (see also “AutomationDetails attribute” on page 44).</td>
</tr>
<tr>
<td>Stuck Online</td>
<td>4</td>
<td>Specifies that a member resource is stuck online.</td>
</tr>
<tr>
<td>Pending Online</td>
<td>5</td>
<td>Specifies that a Start command is executed (the resource group’s NominalState attribute is set to Online). The resource group must begin processing an Online action.</td>
</tr>
<tr>
<td>Pending Offline</td>
<td>6</td>
<td>Specifies that an Offline action has been initiated.</td>
</tr>
</tbody>
</table>

TopGroup attribute
This attribute shows the top level resource group of a resource group.

In System Automation for Multiplatforms, resource groups can be members of another resource group, meaning that resource groups can be nested. The TopGroup attribute shows the name of the top level resource group of the current group. The attribute can be displayed using the lsrg command as described in IBM Tivoli System Automation for Multiplatforms Reference and shown below.
Resource Groups

Note: When using the lsrg -g command to query a resource group, the NominalState of the top group is shown in the output for user convenience together with the TopGroup attribute. The output looks similar to:

```
TopGroup = apacherg
TopGroupNominalState = Offline
```

**AutomationDetails attribute**

This attribute shows System Automation internal states of the group. These states include:

**CompoundState**
Overall status of the resource group including group dependencies. An example is "Satisfactory" – resource/group has reached the requested user status.

**DesiredState**
User requested status of the resource group. An example is "online" – user requested that the resource group should be online.

**ObservedState**
Real status of the resource group from an automation point of view. An example is "online" – the resource group is currently online.

**BindingState**
Status indicating if the resource group is bound to a specific system. An example is "bound" – the resource group is currently bound to a specific system.

**AutomationState**
Status indicating if the resource group is currently being automated. An example is "Idle" – System Automation for Multiplatforms is currently not trying to start or stop the resource group.

**ControlState**
Status indicating if the resource group can be controlled by automation. An example is "startable" - it is currently possible to start this resource group.

**HealthState**
currently not used, reserved for future releases.

To show the automation details as indicated above you use the lsrg command with the -A d and the -V option. For example, to show the automation details of a resource group named "apacherg", you use the command:

```
lsrg -A d -V -g apacherg
```

**MoveStatus attribute**

This dynamic attribute shows the progress of a move of a resource group that was initiated through a rgreq command.

To retrieve the move status, you use the lsrg command with the options -A d and -V.

**Example:**

To retrieve the move status of the resource group "apacherg", you use the following command:

```
lsrg -A d -V -g apacherg
```
Table 3. MoveStatus values

<table>
<thead>
<tr>
<th>MoveStatus values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undefined</td>
<td>State during initialization</td>
</tr>
<tr>
<td>Not Supported</td>
<td>Move is not supported</td>
</tr>
<tr>
<td>None</td>
<td>Not in the process of moving</td>
</tr>
<tr>
<td>Preparing</td>
<td>Move operations has been started</td>
</tr>
<tr>
<td>Stopping</td>
<td>Offline Phase 1: resources are being stopped</td>
</tr>
<tr>
<td>Canceling</td>
<td>Move is cancelled (not available)</td>
</tr>
<tr>
<td>Starting</td>
<td>Online Phase: resources are restarted</td>
</tr>
<tr>
<td>Backing Out</td>
<td>Move was not successful, restart groups on original nodes</td>
</tr>
<tr>
<td>Completing</td>
<td>Move is completing</td>
</tr>
</tbody>
</table>

ConfigValidity

After a policy was established, several things can happen to make the policy invalid. This attribute shows the reason for invalidity.

For example, if a node is removed from the peer domain which was the only common node for the members of a collocated resource group, the resources can no longer be started. This will be indicated by the ConfigValidity attribute.
**Attributes used for resource group members**

In addition for each resource group member a user has to define persistent attribute:

**Mandatory**
Defined for each resource group member and specifies whether it is mandatory for the group. Alternatively a member can also be non-mandatory.

**MemberOf**
The name of the resource group of which the resources are members.

**SelectFromPolicy**
Used to tell System Automation for Multiplatforms where floating resources should be preferably started.

**ConfigValidity**
This attribute is reserved for future use.

**Mandatory attribute**
You use the Mandatory persistent attribute to specify whether a managed resource is Mandatory or Non-Mandatory.

When a resource group is started, all managed resources within that group that are Mandatory must also be started. Managed resources that are Non-Mandatory (whose Mandatory attribute is set to False) might not be started when a conflict exists. If a managed resource that is Mandatory fails, the entire resource group is stopped and started on another node, but if a non-mandatory member of a resource group fails, the resource group stays Online on that node.

Resources that are members of a resource group are implicitly mandatory unless this attribute value is explicitly set to False.

Member resources whose Mandatory managed resource attribute is False may be sacrificed in order to activate the resource group.

**MemberOf attribute**
Indicates that the resource is contained in a resource group resource. The MemberOf persistent attribute is generated implicitly when resources (including nested resource groups) are added to a resource group. The MemberOf attribute is used to determine the set of resources to be started and stopped when the resource group is activated or deactivated (either explicitly through a stop order, or implicitly through a non-recoverable member resource failure). A resource can be a member of one and only one resource group.
Defining and administering a resource group

Making (creating) a resource group

To make (create) a resource group you use the `mkrg` command, in which you define to System Automation for Multiplatforms:

- Where the resource group is allowed to run.
- The relative importance of the resource group in relation to other resource groups (using the `Priority` attribute, as explained in "Priority attribute" on page 41).
- The `Location` relationship among the member resources of the resource group (explained in "Location relationships" on page 77).

Newly-created resource groups will default to an NominalState of Offline. This allows a user or administrator to fully configure the resource group and its resources.

For example, to define a new resource group called `apacherg2` with location relationship “None” and allowed node name “node03”, you would enter:

```
mkrg -l None -n node03 apacherg2
```

For further details, see either the `mkrg` man page, or the command description in IBM Tivoli System Automation for Multiplatforms Reference.

To establish the node list for a resource group, you can use either:

- `mkrg` command, to create a new resource group.
- `chrg` command, for an already-existing resource group.

To specify the node list to be used with the above two commands, you can either:

- Enter a node name when using the `mkrg/chrg` command.
- Establish as an equivalency, the set of nodes on which the resource group can be activated. This must be done before the node list is established. You then use the `mkrg/chrg` command, and the required equivalency will be attached to the resource group. (For details of equivalencies, see Chapter 5, “Using equivalencies,” on page 51).

Adding a member resource to a resource group

To add one or more new member resources to a resource group, you use the `addrgmbr` command.

Notes:

1. A member resource cannot be included in more than one resource group at the same time.
2. A member resource cannot be in a resource group and in an equivalency at the same time.

For example, to add member resource `apache1`, belonging to resource class `IBM.Application`, to a resource group `apacherg2`, you would enter:

```
addrgmbr -g apacherg2 IBM.Application:apache1
```

For further details, see either the `addrgmbr` man page, or the command description in IBM Tivoli System Automation for Multiplatforms Reference.
Resource Groups

Listing resource groups and their members

Use the following commands to list resource groups and their members:

**Issam**  Lists resource groups and their members in tree format. For detailed information about the command, see the **Issam** man page or the command description in [IBM Tivoli System Automation for Multiplatforms Reference](#). On Windows, **Issam** can be invoked from the folder **Start > All Programs > Tivoli SA MP Base > List Defined Resource Groups - Issam**.

**lsrg**  Lists resource groups or the members of a resource group. This section gives you an overview of how the command is used and provides some usage examples.

For further details, see either the **lsrg** man page, or the command description in [IBM Tivoli System Automation for Multiplatforms Reference](#).

To list a resource group or its members, you can use the **lsrg** command. If the resource group name is omitted, all of the resource groups are listed. If the resource group name is specified with the **-m** option, the member resource names and resource classes are listed.

If the **Attr** parameter is specified, then the attributes specified for the resource group are listed.

Three resource groups are defined in the following examples:

**Example 1:** If command **lsrg** is entered, this type of information is displayed:

- **Resource Group Names:**
  - apacherg2
  - apacherg3
  - apacherg4

**Example 2:** To list all members of all resource groups, enter:

```
lsrg -m
```

This information is then displayed:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM.Application:apache1</td>
<td>True</td>
<td></td>
<td>Offline</td>
<td>Nominal</td>
<td>node03</td>
</tr>
</tbody>
</table>

**Example 3:** **apacherg2** contains one resource. To list the members of **apacherg2**, this command is entered:

```
lsrg -m -g apacherg2
```

This information is then displayed:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM.Application:apache1</td>
<td>True</td>
<td>apacherg2</td>
<td>Offline</td>
</tr>
</tbody>
</table>

**Example 4:** To list the attributes of a resource group **apacherg2**, this command is entered:

```
lsrg -g apacherg2
```

This information is then displayed:
Resource Group 1:
- Name: apacherg2
- MemberLocation: None
- Priority: 0
- AllowedNode: node03
- NominalState: Offline
- OpState: Offline

Starting and stopping a resource group
To start or stop a resource group, you set the NominalState attribute of the resource group to online or offline, respectively. Use the `chrg` command to do this.

For example, to start a resource group called apacherg2, you would enter:
```
chrg -o online apacherg2
```

For further details, see either the `chrg` man page, or the command description in `IBM Tivoli System Automation for Multiplatforms Reference`.

Changing attributes of a resource group
To change the persistent attribute values of one or more resource groups, you use the `chrg` command. The name of a resource group can also be changed with this command, using the `-c` option.

Example 1: To change the location relationship of a group apacherg2 to collocated, you would enter:
```
chrg -l collocated apacherg2
```

Example 2: To change the name of a group apacherg3 to apacherg4, you would enter:
```
chrg -c apacherg4 apacherg3
```

For further details, see either the `chrg` man page, or the command description in `IBM Tivoli System Automation for Multiplatforms Reference`.

Changing the attributes of resource group members
To change the attributes of resource group members, you use the `chrgmbr` command.

This command also allows you to specify changes to the Mandatory attribute of a managed resource by using the `-m` option, and to change the resource group to which the resource belongs by using the `-c` option.

For example, to change the resource group to which member resource apache2 of resource class IBM.Application belongs, from the current resource group apacherg2 to resource group apacherg3, you would enter:
```
chrgmbr -c apacherg3 -g apacherg2 IBM.Application:apache2
```

For further details, see either the `chrgmbr` man page, or the command description in `IBM Tivoli System Automation for Multiplatforms Reference`.

Removing a member resource from a resource group
You use the `rmrgmbr` command to remove:
- all the member resources of a specified resource group.
- only the specified member resources of the specified resource group.
Resource Groups

- the member resources that match a selection string.

System Automation for Multiplatforms also ensures that any associated managed relationship or equivalency is also updated.

For example, to remove member resource apache2, which belongs to resource class IBM.Application, from resource group apacherg3, you would enter:

```
rmrgmbr -g apacherg3 IBM.Application:apache2
```

For further details, see either the `rmrgmbr` man page, or the command description in `IBM Tivoli System Automation for Multiplatforms Reference`.

Removing a resource group

To remove one or more resource groups, you use the `rmrg` command. The resource groups to be removed are specified either using the `Resource_group` parameter, or by matching a selection string. The member resources associated with the removed resource groups are also removed by System Automation for Multiplatforms. If the resource group to be removed is still online, the resource group is not removed. Note that relationships in which the deleted resource groups are the source are also deleted.

This implies that all resource groups that are nested within the resource group to be removed are also removed. If you want to prevent contained resource groups from being deleted recursively, proceed as follows:

1. Remove these resource groups as members from the resource group to be removed using the `rmrgmbr` command.
2. Remove the containing resource group.

For example, to remove resource groups called apacherg2 and apacherg3, you would enter:

```
rmrg apacherg2 apacherg3
```

For further details, see either the `rmrg` man page, or the command description in `IBM Tivoli System Automation for Multiplatforms Reference`.
Chapter 5. Using equivalencies

This chapter describes equivalencies in these main sections:

- "What is an equivalency?"
- "Attributes used by equivalencies" on page 53
- "Rules for using equivalencies" on page 52
- "Making (creating) an equivalency" on page 54
- "Changing an equivalency" on page 55
- "Removing an equivalency" on page 55
- "List one or more equivalencies" on page 54

Table 4 lists the System Automation for Multiplatforms commands that you use for equivalencies. For details, see the command descriptions in IBM Tivoli System Automation for Multiplatforms Reference.

Table 4. System Automation for Multiplatforms commands used with equivalencies

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mkequ</td>
<td>Make an equivalency resource</td>
</tr>
<tr>
<td>rmequ</td>
<td>Remove an equivalency resource</td>
</tr>
<tr>
<td>chequ</td>
<td>Change an equivalency resource</td>
</tr>
<tr>
<td>lsequ</td>
<td>List equivalency resources</td>
</tr>
</tbody>
</table>

What is an equivalency?

An equivalency is a collection of resources that provide the same functionality. An equivalency consists of a set of fixed resources from the same resource class.

For example, network adapters might be defined as equivalencies. If one network adapter goes offline, another network adapter can take over the processing from the offline adapter.

Equivalencies are also used for establishing a resource group’s node list. From this equivalency, one or more resources can be selected to satisfy a managed relationship. But only one member is started on a node to satisfy a managed relationship. For details of managed relationships, see Chapter 6, “Using managed relationships,” on page 57.

There are two types of equivalency:

1. Equivalencies with a static membership list. An equivalency of this type contains a fixed set of resources which a user has explicitly added to the equivalency.

2. Equivalencies with a SelectString list. An equivalency if this type dynamically determines at run-time which resources are contained within the equivalency. If RMC resources are created that match the dynamic select string, these are automatically contained within the equivalency. Specifying a policy is not reasonable for this type of equivalency because the resources are not ordered.

The following types of equivalency can be used to effect a particular automation behavior:
### Equivalencies

#### Failback
Can be used for ordered equivalencies only. Equivalencies of this type show the following behavior:

System Automation for Multiplatforms ensures that all resources of the equivalency run on the node hosting the first member of the equivalency. If this member goes Offline, System Automation for Multiplatforms starts the resources on the node hosting the second member of the equivalency. When the first member has come back Online, System Automation for Multiplatforms stops all dependent resources and restarts them again on the node hosting the first member of the equivalency.

Note that a move request against a resource group that has a relationship to an equivalency of type Failback does not result in the expected behavior: The move will take place, however, immediately after the move is complete, the resource group is failed back to the original node!

#### NoFailure
Equivalencies of this type show the following behavior:

Resources that have a dependency to the equivalency will not be set to a Failed Offline state if they do not reach the Online operational state within the timeout interval. This is useful for resources that take a long time to start. The meaning is similar as if the StartCommandTimeout of a resource were set to infinite, but differs in that these resources can be automated again after the Online timeout is reached, while resources with very high StartCommandTimeout values cannot be automated until the StartCommandTimeout is reached because they are in a Waiting state.

### Shadow Resources/Shadow equivalencies
A shadow resource monitors the OpState of another resource. System Automation for Multiplatforms only evaluates the OpState of shadow resources but neither starts nor stops them. Shadow resources are used to define relationships from a floating resource to one of two or more fixed resources running on different nodes. For a detailed description, see "Using shadow resources" on page 199.

Equivalencies are supplied in System Automation for Multiplatforms resource class IBM.Equivalency.

### Rules for using equivalencies
These are the rules for using Equivalencies:
1. A resource can be a member of either an equivalency or a resource group, but not both.
2. A resource may be in more than one equivalency.
3. The specified resources must all be from the same resource class.
4. Equivalencies cannot be members of an equivalency.
5. Resource groups cannot be members of an equivalency.
6. Equivalencies cannot be members of a resource group.
7. An equivalency that satisfies a relationship for an active resource cannot be modified.
8. An equivalency can only be the target of a managed relationship (it cannot be the source of a managed relationship).
9. The members of an equivalency must be fixed resources. Floating resources are not allowed.
Attributes used by equivalencies

MemberClass attribute
System Automation for Multiplatforms uses the MemberClass persistent attribute to determine the resource class of all the member resources.

Membership attribute
System Automation for Multiplatforms uses the Membership persistent attribute to determine the set of resources contained within the equivalency. If a Membership attribute is specified, no SelectString attribute is allowed.

SelectString attribute
System Automation for Multiplatforms uses the SelectString persistent attribute to dynamically determine the resources contained within an equivalency. If resources matching the selection string are inserted in, or removed from, the system, then System Automation for Multiplatforms automatically modifies the equivalency. If a SelectString attribute is specified, no Membership attribute is allowed.

SelectFromPolicy attribute
System Automation for Multiplatforms uses the SelectFromPolicy persistent attribute to determine the policy to be used in making a selection from the equivalency. You can modify this attribute when the resources referencing the equivalency are offline. Possible policies:

- **Any (the default)**
  In this case, if a resource contained within an equivalency fails, System Automation for Multiplatforms chooses any resource without referring to a pre-specified order of selection.

- **Ordered**
  In this case, if a resource contained within an equivalency fails, System Automation for Multiplatforms always starts from the beginning of the selection list.

  **Note:** An Ordered policy may not be used when a dynamic SelectString is used.

There are additional settings for the SelectFromPolicy persistent attribute, which can be combined with the Any or Ordered setting to effect a particular automation behavior:

- **Failback**
  The Failback option can only be specified for an Ordered policy. If the option is specified, System Automation for Multiplatforms will fail back the resources to the node hosting the first member of the equivalency when the first member becomes available again.

  Note that this feature has an impact on move requests against resource groups (see also "Moving resource groups with the rgreq command" on page 197): If an equivalency with a Failback policy is involved in a move request, the move of the resource group is performed, but immediately after the resources are Online on the target node, the Failback option of the equivalency causes the resource group to be moved back to the original node.

- **NoFailure**
  If the NoFailure option is specified, resources that have a DependsOn...
Equivalencies

relationship to the equivalency will not be set to Failed Offline if an Online request against these resources does not succeed within the Online timeout interval. This implies that these resources will be failed over to another node only if the MonitorCommand for these resources returns Failed Offline.

NoControl
The NoControl option identifies the members of the equivalency as shadow resources. In this case, System Automation for Multiplatforms will not start or stop the members of the equivalency but will only react to changes of the OpState of the equivalency resource members. By default, System Automation for Multiplatforms will start and stop controllable resources within an equivalency, for example, of class IBM.Application.

Attributes used by members of equivalencies

- A resource that is to be added to an equivalency must have this attribute:
  - OpState
- A resource that is to be added to an equivalency may have these attributes:
  - NodeNameList
  - ResourceType

For details of these attributes, refer to "Attributes used by resources" on page 30.

Defining and administering equivalencies

Making (creating) an equivalency
To make an equivalency among resources, you use the mkequ command.

For further details, see either the mkequ man page, or see the description of the mkequ command in IBM Tivoli System Automation for Multiplatforms Reference.

For example, to create a static equivalency called NetworkInterfaces of two ethernet interfaces eth0 located on Linux systems node01 and node02 of the resource class IBM.NetworkInterface, you would enter:
mkequ NetworkInterfaces IBM.NetworkInterface:eth0:node01,eth0:node02

To create a dynamic equivalency called NetworkInterfacesDynamic containing all available ethernet interfaces in a cluster of Linux systems, you would enter:
mkequ -D "Name like 'eth\%'" NetworkInterfacesDynamic IBM.NetworkInterface

In a cluster of AIX systems, you would enter:
mkequ -D "Name like 'en\%'" NetworkInterfacesDynamic IBM.NetworkInterface

In a cluster of Solaris systems, you would enter:
mkequ -D "Name like 'bge\%'" NetworkInterfacesDynamic IBM.NetworkInterface

List one or more equivalencies
To list one or more equivalencies, you use the lsequ command.

For further details, see either the lsequ man page, or the description of the lsequ command in IBM Tivoli System Automation for Multiplatforms Reference.
Equivalencies

If you omit an equivalency name, all of the defined equivalencies will be listed. If you specify an equivalency, the persistent attributes of the this equivalency will be listed. If you specify the attribute name as operand, the attributes specified for the equivalency will be listed.

For example, to list the persistent attributes of the equivalency NetworkInterfaces, you would enter:
```
1sequ -A p -e NetworkInterfaces
```

Changing an equivalency

To add, remove, or totally replace the resources contained in an equivalency, you use the `chequ` command.

For further details, see either the `chequ` man page or the command description in IBM Tivoli System Automation for Multiplatforms Reference.

You can also use this command to change the name of the equivalency.

For example, to add the resource eth1 located on the Linux system node01 that belongs to the resource class IBM.NetworkInterface, to an equivalency called NetworkInterfaces, you would enter:
```
chequ -u a NetworkInterfaces IBM.NetworkInterface:eth1:node01
```

Removing an equivalency

To remove one or more equivalencies, you use the `rmequ` command.

For further details, see either the `rmequ` man page, or the description in IBM Tivoli System Automation for Multiplatforms Reference.

You specify one or more equivalencies using either the Equivalency name as operand, or selection string.

For example, to remove an equivalency called NetworkInterfaces, you would enter:
```
rmequ NetworkInterfaces
```
Equivalencies
Chapter 6. Using managed relationships

This chapter describes how you use managed relationships in these main sections:
- “What is a managed relationship?”
- “Attributes used by managed relationships” on page 58
- “Relationships for start / stop behavior” on page 60
- “Location relationships” on page 77
- “Creating and administering relationships” on page 88

Note: Note that System Automation decisions are based on a various factors, such as combinations of relationships between groups and resources, on attribute settings, and on the nominal states of groups. Each of these factors may have an impact on the automation behavior. The examples described below only illustrate the typical automation behavior that occurs when relationships are used, and the actual behavior may differ from the descriptions provided here.

Table 5 lists the System Automation for Multiplatforms commands that you use for managed relationships. For details, see the command description in IBM Tivoli System Automation for Multiplatforms Reference.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mkrel</td>
<td>Make a managed relationship</td>
</tr>
<tr>
<td>lsrel</td>
<td>List managed relationships</td>
</tr>
<tr>
<td>rmrel</td>
<td>Remove a managed relationship</td>
</tr>
<tr>
<td>chrel</td>
<td>Change a managed relationship</td>
</tr>
</tbody>
</table>

What is a managed relationship?

A managed relationship exists between a source resource, and one or more target resources. As a first example, during a resource group startup the source resource must be started after the target resource has become online: this example of a managed relationship uses a StartAfter value of the Relationship attribute.

As shown in the example, relationships always indicate a direction. Relationships can cross node boundaries.

As a second example, a source resource should if possible be started on the same node as the target resource: this example of a managed relationship uses an Affinity value of the Relationship attribute (described in “Relationship attribute” on page 59).

Relationship attributes may be used with Condition attributes (described in “Condition attribute” on page 59). As a third example, a source resource should, if possible, be started on the same node as the target resource, but only if the target
Managed Relationships

resource is online. This example of a managed relationship uses an Affinity value of the Relationship attribute, together with an IfOnline value of the Condition attribute.

The Relationship attribute is described in "Relationship attribute" on page 59, the Condition attribute is described in "Condition attribute" on page 59.

By using combinations of managed relationships, complex automation scenarios can be defined.

As mentioned above, a relationship is defined between a source and one or more target resources. If you remove the source resource of a relationship (can be a resource group member or the underlying RMC resource), the relationship is deleted. If you remove the last target resource from the list of target resources (can be a resource group member or the underlying RMC resource), the relationship is not deleted. You must remove this relationship with the `rmrel` command.

The reason for this behavior is that no relationships should accidentally be deleted. There was a reason why, for example, a DependsOn was defined from a source resource to a target resource. The source resource cannot function properly without the target resource. So the relationship should not be deleted automatically unless you tell System Automation for Multiplatforms to do so.

Managed relationships are supplied in System Automation for Multiplatforms resource class IBM.ManagedRelationship.

Attributes used by managed relationships

Related Sections:
- "Attributes used by resources" on page 30
- "Attributes used by resource groups" on page 38
- "Attributes used by equivalencies" on page 53

The following picture shows another example of a managed relationship:

A managed relationship has the attributes described in the sections below.

Name attribute
You use the Name persistent attribute to specify the name you wish to use for the managed relationship. This attribute is optional. It makes it easier to change or delete relationships.

Source attribute
You use the Source persistent attribute to specify the source resource of the managed relationship.

Target attribute
You use the Target persistent attribute to specify the list of target resources of the managed relationship.
Managed Relationships

Relationship attribute

You use the Relationship persistent attribute to specify the relationship that is to be applied between source and target resources. There are two types of relationships, start / stop dependencies and location dependencies:

**Start / Stop dependencies:**
- StartAfter
- StopAfter
- DependsOn
- DependsOnAny
- ForcedDownBy

Start / Stop dependencies are used to define a start / stop behavior.

**Location dependencies**
- Collocated
- AntiCollocated
- Affinity
- AntiAffinity
- IsStartable

Location dependencies are used for locating resources on nodes.

Condition attribute

The Condition persistent attribute specifies a condition to be used together with all Location relationships (described in "Location relationships" on page 77), except for the IsStartable managed relationship. The IfPossible condition can only be used in combination with the StartAfter relationship.

The Condition persistent attribute defines when the relationship is considered applicable. These are the conditions that can be applied:
- IfOnline
- IfNotOnline
- IfOffline
- IfNotOffline
- IfPossible
- None
Managed Relationships

Relationships for start / stop behavior

System Automation for Multiplatforms provides the following relationships which can be used to define a start/stop behavior:

- **StartAfter**
- **StopAfter**
- **DependsOn**
- **DependsOnAny**
- **ForcedDownBy**

The source of a start/stop relationship is either a member of a resource group or a resource group. See “What is a resource group?” on page 35 for more information about resource groups.

The target of a start/stop relationship is either:
- a member of a resource group or a resource group.
- an equivalency.
- an RSCT resource (which is not a managed resource) which has to provide an OpState attribute.

Note that in case of a DependsOn relationship and source or target resources or both being groups, these groups must have a member location of collocated.

A start command cannot be issued against a resource directly. Therefore you start a resource by setting the nominal state of the resource group of which resource is a member to online.

**StartAfter relationship**

Use the StartAfter relationship to ensure that the source resource is only started when the target resource(s) are online.

The StartAfter relationship provides the following behavior scheme:

- With the start behavior StartAfter defines a start sequencing for resources A and B:
  - When source resource A has to be started, then the target resource B is started first. After resource B has become online, resource A is started.
  - Note that resource A and resource B can be started on different nodes.

Using the IfPossible condition with StartAfter relationship between two resource groups induces a variant of this behavior, because the target resource group B is bypassed if it cannot be bound and ends up in a sacrificed binding state (see “Location relationship resolution: Binding algorithm” on page 91). In this case, the source of the relationship is started ignoring the relationship.

The StartAfter relationship does not provide a force down behavior (see “DependsOn relationship” on page 68).
Details on the start behavior of the StartAfter relationship

The start behavior is controlled via the operational state (OpState) of the target resource. At the time when the operational state of resource B has become online, resource A is started.

In many cases resources A and B are members of the same resource group.

```
  A   B
   RG_AB
     StartAfter
```

Setting their resource group’s nominal state to online causes that both members A and B are started. Due to the StartAfter relationship from A to B, resource B is started first. When the operational state of resource B is online, resource A is started.

If resource A is a member of resource group RG_A, and resource B is a member of resource group RG_BC, and a StartAfter relationship is defined between A and B. Then the start behavior of the StartAfter relationship is triggered by setting the nominal state of RG_A to online.

```
  A   B   C
   RG_A
     StartAfter
   RG_BC
```

Due to the start sequence of the StartAfter relationship, resource B has to be started first. In case RG_BC’s nominal state is set to offline, the following conflict exists: RG_BC wants resource B to be offline whereas the StartAfter relationship forces B to be started. System Automation for Multiplatforms resolves this conflict in such a way that the online request is always more important than the offline request. Therefore resource B is started even though other possible group members of RG_BC will not be started since the nominal state of their group is offline. After resource B is online, System Automation for Multiplatforms will start resource A. Resource C is not started.

If resource A was started and resource B is started due to a StartAfter relationship, but the nominal state of the comprising group remains Offline, resource B is recovered, which means that B is restarted in case of failure. This enhanced behavior is introduced with release 2.3 of System Automation for Multiplatforms. In earlier releases, B would not have been restarted in case of failure.
Managed Relationships

The start order only acts in the forward direction of the relationship. In case resource A and resource B are part of different resource groups (A belongs to RG_A and B belongs to RG_B), then setting the nominal state of RG_B to online does not cause any action on resource A since resource B has no forward relationship to resource A.

When RG_A’s nominal state is set to online, the resource A can be started right away since resource B is already online.

In another scenario it also might be the case that resource A has a StartAfter relationship to resource B and resource C.

In this case starting A requires that both resources B and C are online before System Automation for Multiplatforms can start resource A. For instance A, B, and C are members of the resource group RG_ABC. Setting the nominal state of RG_ABC to online causes that resources B and C are started in parallel first. When the operational state of both resources is online, then resource A is started.

It is also possible that resource A is member of resource group RG_A, resource B is member of resource group RG_B, and resource C is member of resource group RG_C.
A has a StartAfter relationship to both B and C. Setting RG_A’s nominal state to online causes that due to the StartAfter relationship resource C and resource B are started. After both resources B and C are online, A is started.

The example depicted below shows the start behavior that occurs when two resource groups (“RG1” and “RG2”) are connected through a StartAfter/IfPossible relationship and the target resource group (“RG2”) cannot be bound:

- The source of the relationship (“RG1”) is started.
- The relationship to the target resource group (“RG2”) is not honored and its bindingState is set to Sacrificed. Note that a different behavior occurs if both resource groups are members of another group.

Using the IfPossible condition with StartAfter relationships: You can specify the IfPossible condition with StartAfter relationships. The condition specifies that the target resource group may be bypassed if it cannot be bound. In such a case, the target resource group ends up in state Sacrificed and the StartAfter relationship is ignored. StartAfter/IfPossible should only be used between groups; the results are undefined if managed resources are specified as supporting resources.

Details on the stop behavior of the StartAfter relationship

Target resource B cannot be stopped while source resource A is online. If the NominalState attribute of source resource A is changed to Offline, target resource B automatically stops. Both resources can be simultaneously stopped.
In many cases source resource A and target resource B are members of the same resource group. So their NominalState values are identical.

Set the NominalState attribute of the RG_AB to offline to stop both members A and B. Since the StartAfter relationship does not require a stop sequence, resources A and B can be simultaneously stopped.

Provided that RG_B has a NominalState of Online, you can start and stop RG_A without affecting resource group RG_B. It remains online.

If you set the NominalState of RG_B to Offline and set the NominalState of RG_A to Online, target resource B will start before source resource A.

If you set the NominalState of RG_A to Offline, then resources A and B are simultaneously stopped.

Consider the following stop behavior:

Provided the NominalState of RG_A is Online and the NominalState of RG_B is Offline, resource A and resource B are online. Now set the NominalState of RG_A to Offline. Resource A and resource B simultaneously stop. The reason for this behavior is that resource B started due to the start request on resource group RG_A which was passed on via the StartAfter relationship. Setting RG_A to Offline removes the start request, and the NominalState of Offline from resource group RG_B causes resource B to be stopped.

The StartAfter relationship causes the typical stop behavior: Resource A and B can be simultaneously stopped.

Resources A, B, and C are members of individual resource groups RG_A, RG_B, and RG_C.
Managed Relationships

Resource C must be online to support both resources A and B. As long as the NominalState of either or both RG_A and RG_B is online, resource C must be kept Online, even if RG_C’s NominalState is Offline. Only when both RG_A and RG_B have a NominalState of Offline, resource C can be stopped. This will be the case if RG_C’s NominalState is Offline, too.

Rules for using the StartAfter relationship

1. The StartAfter relationship must not conflict with an existing DependsOn relationship.
2. The StartAfter relationship does not assume that a Location relationship exists between managed resources. If you wish to define a Location relationship (see "Location relationships" on page 77), you must create an additional relationship for this purpose.
3. If System Automation for Multiplatforms is requested to start the source resource, it will however always attempt to first start the target resource.
4. If the target resource fails, this does not mean that the source resource will then be stopped.
Managed Relationships

StopAfter relationship

Use the StopAfter relationship to ensure that the source resource can only be stopped when the target resource has been already stopped.

The StopAfter relationship provides the following behavior scheme:

- Resource A will not be stopped unless the target resource has been brought Offline before (including Failed Offline).

The StopAfter relationship does not provide a start and a force down behaviour (see “StartAfter relationship” on page 60 and “DependsOn relationship” on page 68).

Details on the stop behavior of the StopAfter relationship

Source resource A cannot be stopped while target resource B is Online. If the OpState attribute of target resource B changes to Offline or Failed Offline, source resource A automatically stops.

In many cases source resource A and target resource B are members of the same resource group. Set the NominalState attribute of the RG_AB to Online to start both members A and B. Since the StopAfter relationship does not require a start sequence, resources A and B can be simultaneously started. Setting their resource group’s NominalState attribute to Offline causes that members are stopped. Due to the relationship from A to B, resource B is stopped first. When the operational state of resource B is Offline, resource A is stopped.

In case resource A and B are part of different resource groups (A belongs to RG_A, and B belongs to RG_B) and RG_B has a NominalState of Offline, you can start and stop RG_A without any dependency to resource group RG_B. If you set the NominalState of RG_B to Online and set the NominalState of RG_A to Offline, source resource A cannot stop as long as target resource B is Online.

If the NominalState of RG_A is Offline, you can start or stop RG_B without any dependency to resource A.

It is also possible that resource A is a member of resource group RG_A, resource B
is a member of resource group RG_B, and resource C is a member of resource group RG_C. A has a StopAfter relationship to both B and C.

If the NominalState of RG_A is Online and you want to stop it, RG_A cannot be stopped as long as the NominalState of both RG_B and RG_C is Online. Only when both RG_B and RG_C have a NominalState of Offline or Failed Offline, resource A can be stopped.
Managed Relationships

**DependsOn relationship**

System Automation for Multiplatforms uses the DependsOn relationship to ensure that the source resource can only be started when the target resource(s) is online. It is used in a similar way to the StartAfter relationship, except:

- A DependsOn relationship also includes an implicit collocation (explained in “Collocated relationship” on page 79) between the source and target resources.
- If a target resource fails, the source resource will also be stopped.

The DependsOn relationship provides the following three behavior schemes:

1. With the start behavior DependsOn defines a start sequencing for resources A and B with an implicit collocation: When resource A (source) has to be started then the target resource B is started first. After resource B has become online, resource A (source) is started on the same node.
2. With the stop behavior DependsOn defines a stop sequence for resource A and B: When resource B (target) has to be stopped, then source resource A is stopped first. After resource A has become offline, resource B (target) is stopped
3. Force down behavior in case the target resource fails: When target resource B has failed resource A is also stopped. Then a restart is triggered according to the start behavior described in 1.

**Details on the start behavior of the DependsOn relationship**

The start sequencing of the DependsOn relationship is controlled via the operational state (OpState) of the target resource. At the time when the operational state of resource B has become online, resource A is started. In addition to the start sequence, DependsOn provides a collocated constraint which causes that resource A has to be started on the same node where resource B was started. Therefore resource B is already started on a node where resource A can be started afterwards. The collocated constraint which is part of the DependsOn relationship corresponds to the behavior of the collocated relationship. For further details on this behavior see “Collocated relationship” on page 79.

In many cases resource A and resource B are members of the same resource group.

Setting their resource group’s nominal state to online causes that both members A and B are started. Due to the DependsOn relationship from A to B resource B is started first. When resource B’s operational state is online, resource A is started on the same node.
If resource A is a member of resource group RG_A, and resource B is a member of resource group RG_BC, and a DependsOn relationship is defined from A to B, then the start behavior of the DependsOn relationship is triggered by setting the nominal state of RG_A to online.

Due to the start sequence of the DependsOn relationship, resource B has to be started first. In case RG_BC’s nominal state is set to offline there is the following conflict: RG_BC wants resource B to be offline whereas the DependsOn relationship forces B to be started. System Automation for Multiplatforms resolves the conflict in such a way that the online request is always more important than the offline request. Therefore resource B is started even though other possible group members of RG_BC will not be started as their group’s nominal state is offline. After resource B is online, System Automation for Multiplatforms will start resource A. Of course the resources A and B are started on the same node. Resource C is not started.

The start order only takes effect in the forward direction of the relationship. In case resource A and resource B are part of different resource groups (A belongs to RG_A and B belongs to RG_B),

then setting the nominal state of RG_B to online does not cause any action on resource A as resource B has no forward relationship to resource A. When RG_A’s nominal state is then also set to online, the resource A can be started right away on the same node as resource B is already online.

In another scenario it also might be the case that resource A has a DependsOn relationship to resource B and resource C.

In this case starting A requires that both resources B and C are online before System Automation for Multiplatforms can start resource A. For instance A, B, and C are members of the resource group RG_ABC. Setting the nominal state of
Managed Relationships

RG_ABC to online causes that resources B and C are started in parallel first. When both resources’ operational state is online then resource A is started. All three resources are started on the same node as A has to be started on the same node where B and C are running.

It is also possible that resource A is member of resource group RG_A, resource B is member of resource group RG_B and resource C is member of resource group RG_C.

A has a DependsOn relationship to both B and C. Setting RG_A’s nominal state to online causes resource B and resource C to be started. After both resources B and C are online, A is started on the same node.

**Details on the stop behavior of the DependsOn relationship**
You can control the stop sequencing of the DependsOn relationship via the operational state (OpState) of the source resource.

When the resource A’s OpState has become offline, then resource B can be stopped.

In many cases resource A and resource B are members of the same resource group.

Set the nominal state attribute of the resource group to offline to stop both members A and B. Due to the DependsOn relationship resource A is stopped first. When resource A is offline resource B is stopped.

Resource A is a member of resource group RG_A, and resource B is a member of resource group RG_B, and a DependsOn relationship is defined from A to B. You can trigger the stop behavior of the DependsOn relationship by setting the nominal state of RG_B to offline (stopping resources directly is not possible in System Automation for Multiplatforms). Due to the DependsOn relationship resource A should stop first. There is a conflict if the nominal state of resource group A is set to online: RG_A wants resource A to be online whereas the DependsOn relationship causes it to be stopped.
This conflict is resolved in such a way that in System Automation for Multiplatforms always the online request is more important than the offline request. Therefore resource A is kept online and resource B cannot be stopped. Only if the nominal state of RG_A is set to offline, resource A can be stopped. When resource A is offline, resource B is stopped afterwards.

There is also an implicit stop behavior to consider:

When the nominal state of RG_A is online, and the nominal state of resource group RG_B is set to offline, then as described above in the start scenario resource A and B are online. Now the nominal state of RG_A is set to offline. This causes resource A to stop. In addition, resource B will be stopped. The reason for this is that it was started due to the start request on resource group RG_A which was propagated via the DependsOn relationship to resource B. Since this resource group RG_A is set to offline, the start request is removed and the nominal state of offline from resource group RG_B causes B to be stopped. The DependsOn relationship causes the typical stop behavior: Resource B cannot be stopped before resource A is stopped. Therefore resource A is stopped first. When A is offline, then resource B is stopped.

Another scenario is that resource A and B have a DependsOn relationship to C. Stopping resource C requires that both resources A and B are brought offline first.

For instance A, B, and C are members of the same resource group RG_ABC. Setting the nominal state of RG_ABC to offline causes that resources A and B are stopped first. When the operational state of both resources is offline, then resource C is stopped. An alternative example is that resources A, B, and C are members of individual resource groups RG_A, RG_B, and RG_C, respectively.
Setting the nominal state of RG_C to offline triggers the stop behavior of the DependsOn relationship. Here the nominal state of the resource groups RG_A and RG_B might overrule the stop behavior. As long as the nominal state of RG_A or RG_B is online, resource C cannot be stopped. The reason for this is that in conflict situations an online request always overrules an offline request. Therefore the stop behavior of the DependsOn relationship is deferred until the nominal state of RG_A and RG_B is set to offline. When their members A and B are offline, then resource C is also stopped.

Details on the force down behavior of the DependsOn relationship

The basic principle of the DependsOn relationship is that source A depends on the functionality of the target resource B. When the target resource B fails then the source resource A cannot function anymore. Therefore, it is not sufficient to restart B. Due to a failure of B resource A will also be forced down. And then both resources will be restarted according to the start behavior: First B, then A.

As example one can define a resource A which has a DependsOn relationship to resource B.

Both resources are online. In case resource B fails, resource A stops and then the normal start behavior occurs. Resource B will be restarted, and then resource A will be started.

Resource A and resource B are members of the same resource group RG_AB.

In addition, the relationship resource A DependsOn resource B is defined. When RG_AB is set to online, resource B is started first, then resource A is started. In case resource B fails or goes offline, resource A is also stopped. Afterwards a normal restart is performed with the start sequence of DependsOn: B is started before A is started.

It could also be the case that resource A is member of resource group RG_A, and resource B is member of resource group RG_B, and A has a DependsOn relationship with resource B.
When resource group RG_A is set to online and RG_B’s nominal state is offline, then resource B is started first and resource A is started afterwards. The force down behavior of the DependsOn relationship is triggered by a failure of resource B. It causes that resource A will also be stopped. This will occur even though RG_A’s nominal state is online. In System Automation for Multiplatforms such a conflict is always resolved in such a way that a force down behavior is always more important than the online request of a resource group.

The force down behavior is propagated through chains of DependsOn relationship. Given the following scenario: resource A is member of resource group RG_A, Resource B is member of resource group RG_B, and resource C is member of resource group RG_C with the relationships A DependsOn B and B DependsOn C.

Let’s assume that resource group RG_A is set to online which causes that the three resources C, B, and A were sequentially started and are in online state. Now resource C fails. This causes resources A and B to be forced down: First, resource A is stopped, then resource B is stopped. The reason is that the force down behavior has a higher importance than a normal online request.

**Rule for using the DependsOn relationship**

There is one rule for using the DependsOn relationship:

1. If the source or target resource is a group, all members of the group must be collocated.
**Managed Relationships**

**DependsOnAny relationship**

The behavior of the DependsOnAny relationship is identical to the DependsOn relationship except that it does not provide the collocated constraint for the start sequence. Therefore source and target resources can be started either on the same node or on different nodes.

The DependsOnAny relationship provides the following three behavior schemes:

1. **With the start behavior** DependsOnAny defines a start sequencing for resource A and B without a location relationship:
   - When resource A (source) has to be started, then the target resource B is started first. After resource B has become online, resource A (source) is started. Note that the only difference to the DependsOn relationship is that Resource A and resource B can be started on different nodes.

2. **With the stop behavior** DependsOnAny defines a stop sequence for resource A and B:
   - When resource B (target) has to be stopped, then source resource A is stopped first. After resource A has become offline, resource B (target) is stopped.

3. **Force down behavior in case the target resource fails**: When target resource B has failed resource A is also stopped. Then a restart is triggered according to the start behavior described in 1.


**Note:** The scenario $A \rightarrow DependsOn \rightarrow B$ corresponds to the scenario $A \rightarrow DependsOnAny \rightarrow B$ and $A \rightarrow Collocated \rightarrow B$

![Diagram showing the relationships between A and B with and without Collocated constraint](image_url)
ForcedDownBy relationship

Use the ForcedDownBy relationship to ensure that the source resource will be brought down if the target resource comes offline.

The ForcedDownBy relationship provides the following behavior scheme:

- Resource A must be forced Offline in the event that either the target resource goes Offline unexpectedly or the target resource itself is forced Offline. The stop of resources A and B can happen in parallel. The force down of resource A will be triggered when resource B enters any of the regular down states (Offline) after having previously been in an Online state or in any of the terminal down states (Failed Offline), regardless of its former state.

The ForcedDownBy relationship does not provide a start and a stop behaviour (see “StartAfter relationship” on page 60, “StopAfter relationship” on page 66, and “DependsOn relationship” on page 68).

Details on the force down behavior of the ForcedDownBy relationship

The basic principle of the ForcedDownBy relationship is that source A must be forced Offline when target resource B goes Offline or fails.

As example one can define a resource A which has a ForcedDownBy relationship to resource B.

Both resources are Online. In case resource B is stopped or fails, resource A will be forced down.

It could also be the case that resource A is member of resource group RG_A, and resource B is member of resource group RG_B, and A has a ForcedDownBy relationship with resource B.

When the NominalState attribute of the resource groups RG_A and RG_B is set to Online, then resource A and B are started without any dependencies on each other. The force down behavior of the ForcedDownBy relationship is triggered either:

1. By a failure of resource B. It causes that resource A will be stopped, too. This will occur even though RG_A’s nominal state is Online. But since the nominal state of RG_A is still Online in this case, resource A will be restarted by System Automation for Multiplatforms.

2. Or by a stopping of resource B.
Managed Relationships

Setting the nominal state of RG_B to Offline causes that resource A will be stopped, too. This will occur even though RG_A’s nominal state is Online. But since the nominal state of RG_A is still Online in this case, resource A will be restarted by System Automation for Multiplatforms.
Managed Relationships

Location relationships

System Automation for Multiplatforms provides the following relationships which can be used to define location relationships:

- Collocated
- AntiCollocated
- Affinity
- AntiAffinity
- IsStartable

For example, resources A and B are floating resources which can be started on node1, node2, and node3:

The idea behind these relationships is to define location constraints between resources. Resource types like floating resources, and groups provide a list of nodes on which these can be started. Resource A and resource B are floating resources which can be started on node1, node2, and node3.

A requirement could be that resource A must always be started on the node where resource B is already running or supposed to be running. This behavior can be specified by defining a Collocated relationship from A to B.

The opposite behavior which requires that resource A must not be started on the node where resource B is already running can be specified by defining the AntiCollocated relationship.

In case of the requirement that resource A should - if possible - be started on the node where resource B is running, otherwise it can be started elsewhere, the Affinity relationship is used. Compared to the Collocated relationship, the Affinity relationship has ‘soft’ location relationships.

The AntiAffinity relationship is used to define that resource A should not be started - if possible - where B is already running. Only if this requirement cannot be satisfied, process A can be started on the node where B is located. Like the Affinity relationship also the AntiAffinity relationship has ‘soft’ location constraints compared to the AntiCollocated relationship.

The IsStartable relationship defines that source resource A can only be placed on a node where target resource B is startable. This relationship is only considered if the source and target resources have nominal state online. When one of the resources (source or target) does not have a nominal state of online, the IsStartable relationship will be discarded together with the resources that have a nominal state of Offline.

Conditions IfOnline, IfOffline, IfNotOnline, and IfNotOffline

You can specify the following conditions together with all location relationships except IsStartable. These conditions are:

**IfOnline**
IfOnline defines that a location relationship is only evaluated when the target resource’s OpState is Online. Otherwise the location is ignored. IfOnline does not include states such as Pending Online, and Pending Offline.
**Managed Relationships**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IfOffline</strong></td>
<td>IfOffline means that a location relationship is only evaluated when the target resource’s OpState is either Offline or Failed Offline or Unknown. Otherwise the location relationship is ignored.</td>
</tr>
<tr>
<td><strong>IfNotOnline</strong></td>
<td>IfNotOnline means that a location relationship is only evaluated when the target resource is <strong>not</strong> in an Online state. IfNotOnline includes states such as Pending Online and Stuck Online. Otherwise the location relationship is ignored.</td>
</tr>
<tr>
<td><strong>IfNotOffline</strong></td>
<td>IfNotOffline means that a location relationship is only evaluated when the target resource is <strong>not</strong> in an Offline or Failed Offline or Unknown state. Otherwise the location relationship is ignored.</td>
</tr>
</tbody>
</table>

**Rules for using location relationships**

1. The source of a location relationship is either a member of a resource group or a resource group. See “What is a resource group?” on page 35 for more information about resource groups.
2. The target of a location relationship is either
   - a member of a resource group or a resource group.
   - an RMC resource (which is not a managed resource) which has to provide a start /stop method and an OpState attribute.
3. If the source or target resource is a group, all members of the group must be collocated.
Collocated relationship

System Automation for Multiplatforms uses the Collocated relationship to ensure that the source resource and its target resource are located on the same node. The Collocated relationship provides the following behavior scheme:

- The Collocated relationship defines that on start of resource A it can only be started on the node where resource B is already running.

The Collocated relationship can be used together with a Condition attribute as described on page 80.

Details on the principle behavior of the Collocated relationship

The following describes in detail four states the Collocated relationship can take:

Case I:

On start of resource A place it on the same node where resource B is already running. ‘Running’ means that the OpState of resource B is either Online, Pending Online, Stuck Online, or Pending Offline.

This behavior represents the standard situation.

The collocated relationship tries to optimize the node selection based on predictions for future situations. Here the following cases are possible:

Case II:

Resource B is started and resource A is in an Offline, Failed Offline or Unknown state.

Generally you would expect that the node selection for resource B is independent of resource A. But when System Automation for Multiplatforms selects a node for resource B, then a node is selected on which resource A could also be started in future. The reason for this prediction approach is that it later simplifies the start behavior for resource A: if no error situation occurs it is ensured that after resource B was started, resource A can be started on the same node where resource B runs.

Case III:

Resource A is started and resource B is in an offline state.

Theoretically resource A could now be placed on any of the nodes of its node list since A cannot be bound to a node on which B is running. Here again the prediction approach tries to find a node location for resource A where resource B
Managed Relationships

could also be started in future. Therefore, System Automation for Multiplatforms determines the same node location for both resources A and B even though it will only start resource A. The internal System Automation for Multiplatforms behavior works as follows: When resource A has to be started System Automation for Multiplatforms determines a node location for both resources A and B, and then starts resource A. (Note: the start of resource B is not driven by the collocated relationship. This is done by another start/stop relationship or a group behavior).

A summary of the prediction approach is: If either resource A or resource B is started, and the other resource is in an offline state, then System Automation for Multiplatforms determines a node location where both resources A and B are logically bound before one of them is started.

Note that the optimization for the node location is just a prediction based on the current circumstances. The prerequisites the decision of the node selection was based upon can change over time.

A scenario of an erroneous prediction of a node selection could be the following: Resources A and B are floating resources and can be located on node 1, 2, 3. The relationship A -- Collocated ---> B is defined. Now resource B has to be started. Due to the collocated relationship System Automation for Multiplatforms could select node1 for resources A and B. Then resource B is started. After a while an administrator usage error causes that resource A cannot be started on node 1 anymore. The OpState of the resource on node 1 is FailedOffline. Then a request causes that resource A has to be started. Since resource A cannot be started on node 1 anymore, a conflict situation occurs which has to be solved as described later.

Case IV:
Another possible state could be that resource A is already in a running state (OpState is either Online, PendingOnline, StuckOnline, or PendingOffline) when resource B is started.

At the time when resource A was started resource B already got the same node selected. If no error occurred resource B can be started there. If there was a problem which prevented resource B from starting on the previous selected node, the resource gets unbound and at start time of resource B a new node location has to be found. This means that resource B can be started on another node.

The following relationships with conditions can be defined:

- **Collocated/IfOnline**
  The relationship A --> Collocated/IfOnline --> B means that the location relationship is only considered when resource B is in a online state. Otherwise the location relationship is ignored. IfOnline does not include states such as Pending Online, and Pending Offline.

- **Collocated/IfOffline**
  The relationship A --> Collocated/IfOffline --> B means that the location relationship is only valid when resource B is in an Offline, Failed Offline or Unknown state.

- **Collocated/IfNotOnline**
The relationship A ---> Collocated/IfNotOnline ----> B means that the location relationship is only valid, when resource B is not in an online state.

- Collocated/IfNotOffline
  The relationship A ---> Collocated/IfNotOffline ----> B means that the location relationship is only valid, when resource B is not in an Offline, Failed Offline or Unknown state.
AntiCollocated relationship

System Automation for Multiplatforms uses the AntiCollocated relationship to ensure that the source resource and its target resource are located on different nodes. The AntiCollocated relationship provides the following behavior scheme:

- The AntiCollocated relationship defines that on start of resource A it can only be started on a different node where resource B is already running.

The AntiCollocated relationship can be used together with the Condition attribute as described on page 83.

Details on the principle behavior of the AntiCollocated relationship

The following describes in detail four states the AntiCollocated relationship can take:

Case I:
On start of resource A place it on a different node than the one where resource B is currently running. ‘Running’ means that the OpState of resource B is either Online, Pending Online, Stuck Online, or Pending Offline.

This behavior represents the standard situation.

The AntiCollocated relationship tries to optimize the node selection based on predictions for future situations. Here the following cases are possible:

Case II:
Resource B is started and resource A is in an Offline, Failed Offline or Unknown state.

Generally you would expect that the node selection for resource B is independent from resource A. But when System Automation for Multiplatforms selects a node for resource B, then a node is selected which allows that resource A can be started on another node in the future. The reason for this prediction approach is that it later simplifies the start behavior for resource A: if no error situation occurs it is ensured that after resource B was started, resource A can be started on another node where resource B is not running. This corresponds to the description of Case I.

Case III:
Resource A is started and resource B is in an offline state (Offline, Failed Offline).

Theoretically resource A could now be placed on any of the nodes of its node list.
Managed Relationships

Here again the prediction approach tries to find a node location for resource A that allows resource B to be started on another node in the future. Therefore System Automation for Multiplatforms determines a node location for resource B even though it will only start resource A.

Summary of the prediction approach is:
If resource A is in an offline state and either resource A or resource B is started (see Case II and Case III), then System Automation for Multiplatforms determines a different node location for both resources A and B before one of them is started

As already mentioned in the description of the Collocated relationship it may happen that the prediction based on the current circumstances might be wrong over time. Nevertheless the prediction approach will simplify the automation behavior in most cases.

Case IV:
Resource A is already in a running state (OpState is either Online, Pending Online, Stuck Online, or Pending Offline) when resource B is started.

At the time when resource A was started (see Case III) resource B already got another node selected. If no error occurred, resource B can be started there. If there was a problem which caused that resource B cannot be started anymore on its previous selected node, at start time of resource B a new node location is found. This means that resource B can be started anywhere, even where resource A is already running.

The following relationships with conditions can be defined:
• AntiCollocated/IfOnline
  The relationship A --> AntiCollocated/IfOnline -----> B means that the location relationship is only valid, when resource B is in a online state. Otherwise the location relationship is ignored. IfOnline does not include states such as Pending Online, and Pending Offline.
• AntiCollocated/IfOffline
  The relationship A --> AntiCollocated/IfOffline -----> B means that the location relationship is only valid when resource B is in an Offline, Failed Offline or Unknown state.
• AntiCollocated/IfNotOnline
  The relationship A --> AntiCollocated/IfNotOnline -----> B means that the location relationship is only valid, when resource B is not in an online state.
• AntiCollocated/IfNotOffline
  The relationship A --> AntiCollocated/IfNotOffline -----> B means that the location relationship is only valid, when resource B is not in an Offline, Failed Offline or Unknown state.
Managed Relationships

Affinity relationship

The Affinity relationship provides the following behavior scheme:

- The Affinity relationship defines that on start of resource A the same node is chosen where resource B is already running, if possible. If other location relationships are inhibiting this, resource A can also run on another node.

The Affinity relationship is very similar to the Collocated relationship. Therefore the Affinity relationship defines a soft location relationship whereas the Collocated relationship is a hard location relationship.

The Affinity relationship can be used together with the Condition attribute (described in "Condition attribute" on page 59).

The following relationships with conditions can be defined:

- **Affinity/IfOnline**
  The relationship A ----> Affinity/IfOnline -----> B means that the location relationship may only be considered when resource B is in an online state. Otherwise the location relationship is ignored. IfOnline does not include states such as Pending Online, and Pending Offline.

- **Affinity/IfOffline**
  The relationship A ----> Affinity/IfOffline -----> B means that the location relationship may only be valid when resource B is in an Offline, Failed Offline or Unknown state.

- **Affinity/IfNotOnline**
  The relationship A ----> Affinity/IfNotOnline -----> B means that the location relationship may only be valid, when resource B is not in an online state.

- **Affinity/IfNotOffline**
  The relationship A ----> Affinity/IfNotOffline -----> B means that the location relationship may only be valid, when resource B is not in an Offline, Failed Offline or Unknown state.
AntiAffinity relationship

The AntiAffinity relationship provides the following behavior scheme:

- The AntiAffinity relationship defines that on start of resource A a different node than the one where resource B is already running is chosen, if possible. If other location relationships are inhibiting this, resource A can also run on the same node.

The AntiAffinity relationship is very similar to the AntiCollocated relationship. Therefore the AntiAffinity relationship defines a soft location relationship whereas the AntiCollocated relationship is a hard location relationship.

The AntiAffinity relationship can be used together with the Condition attribute (described in “Condition attribute” on page 59).

See also “Location relationships” on page 77.

The following relationships with conditions can be defined:

- **AntiAffinity/IfOnline**

  The relationship A ---> AntiAffinity/IfOnline -----> B means that the location relationship may only be valid, when resource B is in a online state. Otherwise the location relationship is ignored. IfOnline does not include states such as Pending Online, and Pending Offline.

- **AntiAffinity/IfOffline**

  The relationship A ---> AntiAffinity/IfOffline -----> B means that the location relationship may only be valid when resource B is in an Offline, Failed Offline or Unknown state.

- **AntiAffinity/IfNotOnline**

  The relationship A ---> AntiAffinity/IfNotOnline -----> B means that the location relationship may only be valid, when resource B is not in an online state.

- **AntiAffinity/IfNotOffline**

  The relationship A ---> AntiAffinity/IfNotOffline -----> B means that the location relationship may only be valid, when resource B is not in an Offline, Failed Offline or Unknown state.
**IsStartable relationship**

The IsStartable relationship provides the following behavior scheme:

![Diagram showing the IsStartable relationship]

- The IsStartable relationship defines that resource A can only be placed on a node where resource B is startable when the resources A and B have a nominal state of online.

IsStartable does not imply that the target resource will actually be startable at a later time. This is because resource failures may prevent all of its relationships from being resolved at that later time.

See also "Location relationships" on page 77.

**Details on the principle behavior of the IsStartable relationship:**

The IsStartable relationship causes the following behavior:

- The IsStartable relationship defines that the source resource can only be placed on a node where the target resource is startable. This relationship is only considered if the source and target resources have nominal state online. When one of the resources (source or target) does not have a nominal state of online, the IsStartable relationship will be discarded together with the resources that have a nominal state of Offline.

The following example explains the behavior of the IsStartable relationship:

Resource A and resource B are floating resources and members of the same resource group RG_A. Resource A can run on node1 and node2, and resource B on node2 and node3. An IsStartable relationship is defined from resource A to resource B.

Both members are started when the nominal state of the resource group is set to online. Based on the IsStartable relationship resource A and resource B are started on node2, as this node is the intersecting node for both resources. When resource B is in failed offline state on node 2, starting of the resource group RG_A does not start resource A, as no node exists where both resource A and resource B can be started.

The following example provides more information about the IsStartable relationship. In this scenario resource A can run on node1, node2, and node3, and is member of the resource group RG_A. Resource B can run on node1 and node2, and is member of resource group RG_B. An IsStartable relationship from resource A to resource B is defined.

The following describes the possible states in this example:

- RG_A’s nominal state is set to online while RG_B is offline. Since the IsStartable relationship is only taken into consideration if the source and target resources have a nominal state of online (here RG_A and RG_B), and RG_B’s nominal state
is offline in this case, the relationship will be ignored. Therefore resource A can either start on node1, or node2, or on node3.

- RG_A’s nominal state is set to online while RG_B is already online. In this case the IsStartable relationship is taken into account and System Automation for Multiplatforms starts resource A on a node where resource B can start (node1 or node2).

- Due to a problem resource B cannot start on node1 and node2, and the nominal state of RG_B is online. Starting resource group RG_A causes that resource A cannot become online because resource B is not startable on the intersecting nodes node1 and node2.

- Due to a problem resource B cannot start on node1 and node2, and the nominal state of the resource group RG_B is offline. When resource group RG_A’s nominal state is set to online, System Automation for Multiplatforms discards resource B, and the IsStartable relationship is ignored because of the desired state offline of resource group RG_B.
Creating and administering relationships

Creating a relationship

To create a relationship between a source resource and one or more target resources, you use the `mkrel` command.

The source resource must be a member of a resource group. The target resource does not have to be in a resource group.

For example, to define an AntiCollocated relationship for a source resource `FloatWebServerA` of class IBM.Application to target resource `FloatWebServerB` of class IBM.Application with condition 'IfOnline' and name 'Rel1', you would enter:

```
mkrel -p anticollocated -o ifonline -S IBM.Application:FloatWebServerA -G IBM.Application:FloatWebServerB Rel1
```

For further details, see either the `mkrel` man page, or the description of the `mkrel` command in the `IBM Tivoli System Automation for Multiplatforms Reference`.

Listing a relationship

To list a relationship, you use the `lsrel` command.

If you do not enter a relationship name, all relationships currently defined will be listed:

```
lsrel
```

Displaying Managed Relations:

<table>
<thead>
<tr>
<th>Name</th>
<th>Class:Resource:Node[Source]</th>
<th>ResourceGroup[Source]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rel1</td>
<td>IBM.Application:FloatWebServerA</td>
<td>RG_WebApp</td>
</tr>
</tbody>
</table>

If you specify a relationship name with the `-M` option, the persistent attributes of the specified relationship will be listed. For example, to list the attributes of the relationship Rel1 you would enter:

```
lsrel -M Rel1
```

Displaying Managed Relationship Information:

for Managed Relationship "Rel1".

Managed Relationship 1:

- **Name**: Rel1
- **Class:Resource:Node[Source]**: IBM.Application:FloatWebServerA
- **Class:Resource:Node[Target]**: {IBM.Application:FloatWebServerB}
- **Relationship**: AntiCollocated
- **Conditional**: IfOnline
- **ResourceGroup[Source]**: RG_WebApp

You may get a similar output if you list all relationships where IBM.Application:FloatWebServerA is the source of (-S option):

```
lsrel -S IBM.Application:FloatWebServerA
```

Displaying Managed Relationship Information:

Managed Relationship 1:

- **Name**: Rel1
- **Class:Resource:Node[Source]**: IBM.Application:FloatWebServerA
- **Class:Resource:Node[Target]**: {IBM.Application:FloatWebServerB}
Managed Relationships

<table>
<thead>
<tr>
<th>Relationship</th>
<th>= AntiCollocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional</td>
<td>= IfOnline</td>
</tr>
<tr>
<td>ResourceGroup[Source]</td>
<td>= RG_WebApp</td>
</tr>
</tbody>
</table>

For further details, see either the `lsrel` man page, or the description of the `lsrel` command in [IBM Tivoli System Automation for Multiplatforms Reference](#).

**Changing a relationship**

To change a relationship, you use the `chrel` command.

For example, to change a relationship named Rel1 (created above) to AntiAffinity, you would enter:

```
chrel -p antiaffinity Rel1
```

For further details, see either the `chrel` man page, or the description of the `chrel` command in [IBM Tivoli System Automation for Multiplatforms Reference](#).

**Removing a relationship**

To remove a relationship between source and target resources, you use the `rmrel` command.

For example, to remove a relationship for a source resource `FloatWebServerA` of class IBM.Application, you would enter:

```
rrel -S IBM.Application:FloatWebServerA
```

For further details, see either the `rmrel` man page, or the description of the `rmrel` command in [IBM Tivoli System Automation for Multiplatforms Reference](#).
Chapter 7. How System Automation for Multiplatforms processes the system information

This chapter describes in its first part the binding algorithm, also called binder. This is a System Automation for Multiplatforms internal function responsible for the node placements of all resources. The second part of this chapter discusses events which allow a resource group to become online. The third part of this chapter is about behavior patterns of System Automation for Multiplatforms.

Location relationship resolution: Binding algorithm

The binder is invoked whenever a resource has to be started for which System Automation for Multiplatforms has not yet assigned a node placement. Resources which have a node location assigned are also called bound. An example is a floating resource A which could potentially run on several nodes. Here, the binding algorithm has to determine (bind) a node location for the floating resource considering all of its location relationships. Based on previous binding algorithm runs which already bound resources, a new solution has to be found. Binding solutions do not necessarily have to be non-ambiguous. Many constellations allow several alternative solutions where it is arbitrary which one is chosen by System Automation for Multiplatforms.

An example for an ambiguous scenario is a resource group with a collocated location relationship containing two floating resources A and B which can run on node1 and node2. When the group is started two alternative solutions are possible: either A and B are bound to node1 or both are bound to node2.

If the binding algorithm can find a solution for the node placement of all involved resources, the resource(s) are started. It is, however, obvious that location relationships can lead to conflicts that must be resolved.

Example: The two floating resources "A" and "B" could be located on "node1" and "node2" but due to performance constraints they cannot run concurrently on the same node. This is achieved by specifying AntiCollocated relationships from "A" to "B" and from "B" to "A". It is assumed that resource "A" is already running on node 1 when node 2 fails. If a user now starts resource "B", a location relationship conflict results because the resources cannot be started on the same node. In this situation, there is no perfect solution to this problem that would allow both resources to run. To resolve such conflicts, System Automation for Multiplatforms performs a so-called "discarding step".

It is possible that resources that are already online are part of the problem. These resources get an additional priority bonus of 10 to the priority set by the resource group (for more information about the Priority attribute and its effects, see Attributes used by resource groups "Priority attribute" on page 41).
System Automation for Multiplatforms logic

The following section describes in detail System Automation for Multiplatforms’s solution finding for location relationships and its conflict resolution handling. This whole process is called the binding algorithm.

The binding algorithm consists of several steps:

1. **Discovery step: Determining configuration subsets for which the location relationships can be independently solved**

   The discovery algorithm consists of several substeps:

   a. **Step 1a: Find all involved resources (configuration subset)**

   Location relationships might separate a customer configuration into several configuration subsets which can be solved independently. The reason for this is that location relationships often affect only a subset of resources of the configuration. An example is the configuration with A --> Collocated --> B, B--> Collocated --> C, and D --> Collocated --> E. Here the location relationships for A, B, and C can be independently solved from D and E. For those two subsets all following steps are separately made.

   b. **Step 1b: Ignore all resources with OpState = Failed Offline**

   It is obvious that all instances which have an OpState of Failed Offline cannot contribute to a binding solution. Those instances are removed from the configuration subset which is used to find a binding solution. An example for this is a Resource Group R1 containing two floating resources A and B which can run on node1 and node2. The Group has a collocated parameter set which means that resource A and B have to be started on the same node. Assume that node2 is broken down which causes that the constituents of the floating resources A and B on node2 are in a Failed Offline state. Therefore those are removed from the configuration subset since instances on node2 will not help to solve the binding problem.

   c. **Step 1c: Cleaning up resource groups which cannot be started**

   In case mandatory resource group members are in a Failed Offline state, the resource group cannot be started according to the resource group behavior. Therefore all other resource group members of such a resource group have to be stopped.

   An example is resource group R1 with floating resource A and B as described above. If floating resource A cannot be started on either of the nodes due to an application error, and if it is a mandatory resource group member, the floating resource B is also stopped (see “Attributes used for resource group members” on page 46).

2. **Perfect solution step: Try to find a ‘perfect’ solution**

   At first, the recovery resource manager tries to find a perfect solution of all involved location relationships for a configuration subset. In this step it tries to find bindings as described in “Location relationships” on page 77. Since in this first step the goal is to find a perfect solution, all Affinity and AntiAffinity relationships are treated as if they were pure Collocated and AntiCollocated relationships. In addition, even resources which are Offline and are not intended to start are also tried to be bound if necessary. If no location relationship conflict occurs, the necessary resources are bound and the binding algorithm is done. As a next step System Automation for Multiplatforms can start those resources which have to be started.

   There are situations in which this binding step gets into a conflict situation with contradicting constraints that cannot be overcome. To resolve this System Automation for Multiplatforms provides a discarding step consisting of several substeps as described below.

3. **Discarding step: Resolve situations with conflicting location relationships**
The discarding step consists of a number of substeps:

a. **Step 3a: Ignore all Affinity and AntiAffinity relationships**

   The first approach to overcome the conflicting situation is to ignore all Affinity and AntiAffinity relationships since those are ‘soft’ location relationships. Based on the previous bindings System Automation for Multiplatforms tries to find a solution for the resources which have to be bound. Since all Affinity and AntiAffinity relationships are ignored, the location relationships are simplified and the probability that a binding solution can be found is increased. In case a solution can be found, the discarding step is left. But there is still the chance that the conflicting situation cannot be overcome. Then the next level of the discarding step is reached.

b. **Step 3b: Ignore all resources with OpState = Offline and which do not have to be started**

   If ignoring all Affinity and AntiAffinity relationships did not help to find a solution for the binding problem (see step 3a) then the next level is to ignore all resources from the binding evaluation which are Offline and which are currently not intended to be started. This increases the possibility that a binding solution can be found.

   In case there is a binding solution available, then the discarding step is left. Otherwise the next step of the discarding process is reached.

   An example is ResourceGroup R1 containing floating resource A, and resource group R2 containing floating resource B, and a relationship A AntiCollocated B. Floating resource A and B can run on node1 and node2, but node2 is broken down.

   ![Diagram](https://example.com/diagram.png)

   Now R1’s nominal state is set to online which causes that resource A has to be bound before it can be started. At first System Automation for Multiplatforms tries to find a perfect solution. Therefore it tries to bind A and B. But here a solution cannot be found. Then System Automation for Multiplatforms ignores all Affinity and AntiAffinity relationships, which does not provide a solution either. Then it ignores all resources with an Offline state and which do not have to be started. This causes that resource B is ignored for the evaluation. Now it is possible to bind resource A to node1.

   The non-mandatory resources are removed from the binding set in the following order:

   1) Resources that are not startable are removed first.
   2) If no resource could be removed, the resources that are currently offline are considered next.
   3) If still no resource could be removed, all non-mandatory resources are removed.
   4) If the resource group itself is a non-mandatory member, all resources, including mandatory ones, are removed.

   c. **Step 3c: Stopping the least important resource group members**

   The next level of the discarding step is to stop resource group members and ignore those members in the binding evaluation. Since each resource group has a priority value assigned, resource group members of the group(s) with the lowest priority value are stopped first, and then it is attempted to find a
binding solution without them. In case this does not satisfy the binding constraints, resource group members with the next group priority level are chosen.

In addition to the priority schema, two substeps are performed to stop and remove resource groups:

1) All non-mandatory members of group priority level are stopped and ignored for the binding solution. If a group is a non-mandatory member of another group, mandatory members are discarded explicitly. This additional step helps to organize more complex policies like the SAP policy and allows dealing with situations where resources must be stopped before a solution for the binding problem can be found, for example, when an SAP enqueue server failed.

2) If the conflict still exists, the members of the group with the next lowest priority are stopped as described above.

These steps are performed iteratively until a binding solution can be found or the binding problem is regarded as unsolvable. In the latter case, all resources participating in the binding problem are stopped and a new solving attempt will be made when these resources are down.

**Hints:**

- Outer groups must have the same or a higher priority than the inner groups. Otherwise the outer groups would be discarded before the inner groups. But if the outer groups are discarded, the inner groups are stopped automatically.
- Resource groups that are non-mandatory members of another resource group should have a lower priority than resource groups that are mandatory members of the same resource group. Otherwise mandatory members may be discarded.
- For more information about priority handling, see Attributes used by resource groups, “Priority attribute” on page 41.

**Events that might allow a resource group to become Online**

All top-level resource groups whose NominalState attribute is Online will be automated: this means, an attempt will be made to start such top-level resource groups and the managed resources within these groups, providing the managed relationships of the managed resources within the resource groups can be satisfied.

If a resource group or a managed resource cannot be brought Online (when one or more of its member resources fail completely to reach the Online required state), the resource group is in an Offline state. The resource group remains Offline until an event occurs that informs System Automation for Multiplatforms that it should again attempt to start the resource group.

These are the possible events which might cause an Offline resource group to become Online:

- Changing the **AllowedNodes attribute** (explained in “AllowedNode attribute” on page 39) of the resource group, for example to include an additional node where the resource group can be started. For details, see the description of the **chrg** command in IBM Tivoli System Automation for Multiplatforms Reference.
- Removing a managed resource from a resource group. As a result, the other member resources might then be startable because a resource which cannot be started is removed. For details, see the description of the **rmrgmbr** command in IBM Tivoli System Automation for Multiplatforms Reference.
System Automation for Multiplatforms logic

- Adding a resource to a resource group which is a target resource of a managed relationship, but which is currently not a member of a resource group. It will be automated by System Automation for Multiplatforms then. As a result, the other member resources might then be startable because a managed relationship can be satisfied. For details, see the description of the `adrgmbr` command in IBM Tivoli System Automation for Multiplatforms Reference.

- Starting a resource which is not controllable by System Automation for Multiplatforms and which has a managed relationship to a resource group member. As a result, the resource goes Online and the managed relationship is satisfied. This might allow the resource group to also be brought online. To start the resource, you might use the RMC `startsrc` command (for details, see the man page for this command).

- Adding a constituent to an aggregate that will make the aggregate resource available on more nodes, and may result in System Automation for Multiplatforms being able to satisfy all of the managed-relationship constraints. If the constituent is a piece of hardware, this would require that you install the hardware, or define it correctly. If the resource is a floating resource, you add a constituent by adding a node name to the `NodeNameList` attribute. For details, refer to the RMC documentation and man pages.

- A new resource is found by an equivalency that uses a dynamic select string. As a result, this resource is added to the equivalency, and may resolve a managed relationship to this equivalency.

- Making a managed resource `NotMandatory`, which allows this resource to be sacrificed. As a result, the other managed resources can be started. For details, see the description of the `chrgmbr` command in IBM Tivoli System Automation for Multiplatforms Reference.

- Performing a Reset on an aggregate or one of its constituents after a failure has been fixed. As a result, the resource will be Offline, and can then be started by System Automation for Multiplatforms. For details, see the man page for the RMC `resetsrc` command.

- A node that was offline becomes online. As a result, System Automation for Multiplatforms may be able to bring the resource group online.

- Changing the priority attribute of a resource group (explained in “Priority attribute” on page 41). A resource group might not be startable because of a priority conflict with another resource group. In this case, increase the priority of the group you want to start or decrease the priority of the other group. The changes will take effect when you next start the group.

- Stopping a higher-priority resource group, which prevents a lower-priority resource group from starting. As a result, a managed relationship conflict is avoided. For details, see the description of the `chr` command in IBM Tivoli System Automation for Multiplatforms Reference.

If a resource group is currently at its NominalState value, the following events may cause additional automation actions:

1. The NominalState attribute value changes from Offline to Online.
2. The NominalState attribute value changes from Online to Offline.

Behavior Patterns of IBM Tivoli System Automation for Multiplatforms

This section describes how System Automation for Multiplatforms behaves and reacts in certain situations.
General considerations

The following sections describe issues relating to the StartCommand, MonitorCommand, and StopCommand.

StartCommand issues

IBM Tivoli System Automation for Multiplatforms uses the command that is specified in the resource's StartCommand attribute to bring the resource online.

It invokes the StartCommand in the following situations:

• Immediately after the NominalState attribute of a resource group changes to Online and all start dependencies of the resource are satisfied.
• Immediately after the OpState of a resource changes from Online to Offline because the resource failed.

The StartCommand is not invoked if the OpState change to Offline results from one of the following events:

• The NominalState of the resource group was changed to Offline.
• The resource was stopped or forced down by System Automation for Multiplatforms to satisfy a dependency on another resource.

When the StartCommand has run but the resource is still Offline when the Online timeout expires and the maximum number of StartCommand execution attempts (as defined in the RetryCount attribute) has not been reached, the Online timeout (in seconds) for a resource is calculated using the following formula:

$$\text{MAX(StartCommandTimeout, MonitorCommandPeriod, MonitorCommandTimeout)} + 5$$

Notes:

1. $+ 5$ is not an absolute value because System Automation for Multiplatforms does not use a real timer. The actual value may range from 5 to 8 seconds, depending on when the System Automation for Multiplatforms daemon is woken up again, which happens frequently.
2. This Online timeout is only evaluated if the resource's OpState did not change during the previous execution of the StartCommand. If the OpState did change (for example, to Pending Online or Online), the Online timeout timer is canceled and the StartCommand of the resource is invoked immediately after the OpState of the resource changes to Offline again.

By default, System Automation for Multiplatforms runs the StartCommand synchronously, which means that System Automation for Multiplatforms waits until the command completes and it obtains knowledge of any return code.

In addition to the Online timeout, each resource has a StartCommandTimeout attribute. The StartCommandTimeout attribute determines the maximum run time of the StartCommand. If the StartCommand does not return within the StartCommandTimeout period, System Automation for Multiplatforms kills the StartCommand with the SIGKILL command and logs a message to the system log of the node.

This may lead to problems if an application process that is started within the StartCommandTimeout period does not return control to System Automation for Multiplatforms. In such a case, the application process is killed whenever the StartCommandTimeout is reached, which is why System Automation for
Multiplatforms cannot start the application as a resource. To correct the problem, the application process must be detached from the calling StartCommand by using one of the following methods:

- Redirect all file handles to a file and start the application process in the background, for example:

  
  ```
  /usr/bin/application >/outputfile 2>&1 &
  ```

- Create a small wrapper application that uses the `setsid()` C-function to detach the application process from the calling StartCommand.

- If the above methods do not work or if they are inappropriate for an application, set the value of the resource's RunCommandsSync attribute to 0. If you do so, System Automation for Multiplatforms will ignore the resource's StartCommandTimeout attribute and the StartCommand and all of its child processes can stay indefinitely on the node.

  Note: If you use this method, System Automation for Multiplatforms will not wait for a return code of this StartCommand and such a resource will not be failed over even if the StartCommand fails. If the resource does not come Online during the Online timeout period, System Automation for Multiplatforms will reinvoke the StartCommand until the RetryCount is reached.

MonitorCommand issues

System Automation for Multiplatforms uses the MonitorCommand of an IBM.Application resource to determine the OpState of the resource on a node. System Automation for Multiplatforms starts monitoring a resource as soon as it is added to a resource group or equivalency. The resource is monitored on all nodes on which it is allowed to run, which are the nodes that specified in the resource’s NodeNameList attribute.

After the initial invocation, the MonitorCommand is invoked in the frequency defined in the MonitorCommandPeriod attribute. The resource is monitored indefinitely on each node on which it is defined until the resource is removed from the resource group or equivalency.

From System Automation for Multiplatforms release 1.2, the MonitorCommand is also invoked immediately after the StartCommand or StopCommand of a resource has completed (only for synchronous commands, if RunCommandsSync attribute for this resource is set to 1, which is the default). This was introduced to enhance the performance of the start or stop of an entire resource group, as now the OpState of the resource is immediately checked after the StartCommand or StopCommand has finished. After this execution of the MonitorCommand, the frequency of MonitorCommandPeriod seconds is honored again, meaning that the next MonitorCommand is executed after MonitorCommandPeriod seconds.

From System Automation for Multiplatforms release 2.2, the value of the MonitorCommandPeriod attribute can be lower than that of the MonitorCommandTimeout attribute for a given resource. This is possible because the MonitorCommandPeriod actually specifies the waiting period between the end of the last execution of the MonitorCommand and the start of the next. As a result, the MonitorCommandTimeout can now be specified independently from the MonitorCommandPeriod, and thus a frequent MonitorCommand execution can be realized even with long-running MonitorCommand scripts. However, it is still recommended that the MonitorCommand be kept as efficient as possible.

There are two issues regarding the MonitorCommand that should be kept in mind to avoid problems or a strange behavior:
System Automation for Multiplatforms logic

1. The MonitorCommand is executed on all nodes on which the resource is allowed to run. If a resource should be down (NominalState of the resource group is offline) and an operator starts the resource manually, System Automation for Multiplatforms will notice this through the resource's MonitorCommand and will execute the StopCommand for the resource to bring it offline again.

   If it is necessary to bring an individual resource group member online or offline, for example, to perform a backup, an System Automation for Multiplatforms request has to be used (the `rgreq` command). This will overrule the NominalState of the resource group and will allow a resource group member to be started even if the NominalState of the resource is offline.

2. If a running MonitorCommand does not complete before the timeout period that is specified in the MonitorCommandTimeout attribute expires, System Automation for Multiplatforms kills the MonitorCommand with the SIGKILL command, logs a message to the system log of the node, and sets the OpState of the resource to Unknown. As a result, this resource and any dependent resources will not be automated until a meaningful state can again be determined. If the message frequently appears in the system log, you should check the value of the MonitorCommandTimeout attribute and adjust it if necessary.

StopCommand issues

By default, the StopCommand is executed synchronously by System Automation for Multiplatforms, meaning that System Automation for Multiplatforms waits for the command to finish and obtain knowledge of any return code. Furthermore there is an attribute StopCommandTimeout for each resource which determines how long it takes at most to execute the StopCommand. If the StopCommand does not return within the StopCommandTimeout time period, the StopCommand is killed by System Automation for Multiplatforms using the SIGKILL command. If this happens a message is logged into the system log of that node.

If the StopCommand for a resource has been executed but the OpState of the resource does not change to Offline within the Offline timeout period, System Automation for Multiplatforms issues a reset operation against the resource. The Offline timeout period (in seconds) for a resource is calculated using the following formula:

\[ \text{MAX}(\text{StopCommandTimeout}, \text{MonitorCommandPeriod}, \text{MonitorCommandTimeout}) + 5 \]

Note that + 5 is not an absolute value as System Automation for Multiplatforms does not use a real timer. The System Automation for Multiplatforms daemons are woken up frequently, and this may result in the additional value, which may be in the range of 5 to 8 seconds.

Issuing a reset operation against a resource results in a second execution of the StopCommand, but this time a special environment variable (SA_RESET) is set to 1. This can be exploited within the StopCommand of a resource to achieve an enhanced force behavior, for example:

```
#!/bin/sh
# A sample stop/reset automation script for the lpd application
if [ $SA_RESET == 1 ]; then
  killall -9 lpd
  exit $?
else
  /etc/init.d/lpd stop
  exit $?
fi
```
If a second execution of the StopCommand is not desired, the script could just exit, for example:

```
#!/bin/sh
# A sample stop/reset automation script for the lpd application
if [ $SA_RESET == 1 ]; then
  exit 0
else
  /etc/init.d/lpd stop
  exit $?
fi
```

If the resource still does not stop after the second execution of the StopCommand, the OpState of the resource is set to Stuck Online, which means that manual intervention is now required to stop the resource. System Automation for Multiplatforms will resume automating the resource only after the problem has been resolved by an operator.

This behavior was introduced in IBM Tivoli System Automation for Multiplatforms, release 2.2. The following scenarios are affected by the change:

- A stop command fails: In 2.2 or later, a second stop command is issued
- Manual reset: A manual reset now causes a stop command to run for all specified resources. If the `resetrsrc` command is executed without specifying the `NodeNameList` attribute within the selection string, the stop command will be executed against all resources. If only the resource name is specified in the selection string for a floating resource, the stop command is executed for each constituent as well as for the aggregate resource, as the selection string comprises the aggregate resource and the constituent resources. As stop command handling for the aggregate resource itself is delegated to all of the constituent resources, all constituent resources will execute the stop command twice. This is why it is recommended that you specify the `NodeNameList` attribute whenever `resetrsrc` is executed within the selection string of the command.
- State handling while resources are started or stopped: For detailed information, see Chapter 7, “How System Automation for Multiplatforms processes the system information,” on page 91.

How System Automation for Multiplatforms reacts to the possible OpState changes of a resource that is online on a node

The following sample configuration is used for the discussion in the next section:

```
Figure 3. Sample configuration setup
```

Chapter 7. How System Automation for Multiplatforms processes the system information 99
The setup of this sample configuration is as follows:

- Cluster of 2 nodes.
- Disk TieBreaker
  - Node1: production system.
  - Node2: stand-by system.
- Resource group: RG1.
  - Floating resource: Res1
  - Floating resource: Res2
  - Relationship: Res1 DependsOn. Res2
- Resources are Online on Node1.

Figure 4 shows the typical OpState transitions for resources that are automated by System Automation for Multiplatforms:

Notes:

1. Offline (ForceFailed). Running the offline method stopsrc with the failed flag set causes a resource to be marked as permanently failed offline. This might happen after a failed start attempt. See Table 11 on page 106.

The OpState of a resource can have seven values. The OpState of a resource is determined by System Automation for Multiplatforms with the MonitorCommand, the actual OpState of a resource is provided to System Automation for Multiplatforms with the return code of the MonitorCommand. Note that it is sufficient for a MonitorCommand to return the OpState values Online and Offline to System Automation for Multiplatforms, the other OpState values a resource can have may be exploited optionally.
Some OpState values like Unknown or Failed Offline may be also set by System Automation for Multiplatforms, for instance, the OpState Unknown is set for a resource if the MonitorCommand for this resource timed out. System Automation for Multiplatforms therefore has no knowledge about the OpState of this resource anymore.

The following two tables illustrate how System Automation for Multiplatforms reacts to an OpState change of the resources Res1 and Res2 from the example above. Note that the tables in this chapter contain all possible OpState values of a resource, even if a particular OpState does not make sense in this situation. The columns which contain these unlikely OpState values are preceded by the word ‘unlikely’ in the following tables.

**OpState change of resource Res1**
The current status of Res1 and Res2 on node1 is Online. The following table shows the actions that System Automation for Multiplatforms performs depending on the return value of the MonitorCommand for Res1.

*Table 6. System Automation actions regarding OpState changes of resource Res1*

<table>
<thead>
<tr>
<th>MonitorCommand (OpState)</th>
<th>First action of System Automation</th>
<th>Second action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (Unknown) =&gt;</td>
<td>Nothing, wait for next MonitorCommand with RC&lt;&gt;0</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=1 (Online)</td>
<td>Nothing</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=2 (Offline)</td>
<td>Start Res1</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=3 (Failed Offline)</td>
<td>Stop Res2</td>
<td>After Res2 is Offline, start both resources on node2 in correct order</td>
</tr>
<tr>
<td>RC=4 (Stuck Online)</td>
<td>Nothing; wait for Operator action</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=5 (Pending Online)</td>
<td>Unlikely, wait for Online</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=6 (Pending Offline)</td>
<td>Wait for Offline</td>
<td>Nothing</td>
</tr>
</tbody>
</table>

**OpState change of resource Res2**
The current status of Res1 and Res2 on node1 is Online. The following table shows the actions that System Automation for Multiplatforms performs depending on the return value of the MonitorCommand for Res2.

*Table 7. System Automation actions regarding OpState changes of resource Res2*

<table>
<thead>
<tr>
<th>MonitorCommand (OpState)</th>
<th>First action of System Automation</th>
<th>Second action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (Unknown) =&gt;</td>
<td>Nothing, wait for next MonitorCommand with RC&lt;&gt;0</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=1 (Online)</td>
<td>Nothing</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=2 (Offline)</td>
<td>Force down Res1</td>
<td>After Res1 is Offline, start Res2, after Online of Res2 -&gt; start Res1</td>
</tr>
<tr>
<td>RC=3 (Failed Offline)</td>
<td>Force down Res1</td>
<td>Start both resources on node2 in correct order</td>
</tr>
</tbody>
</table>
System Automation for Multiplatforms logic

Table 7. System Automation actions regarding OpState changes of resource
Res2 (continued)

<table>
<thead>
<tr>
<th>MonitorCommand (OpState)</th>
<th>First action of System Automation</th>
<th>Second action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=4 (Stuck Online) =&gt;</td>
<td>Nothing: wait for Operator action</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=5 (Pending Online) =&gt;</td>
<td>Unlikely, wait for Online</td>
<td>Nothing</td>
</tr>
<tr>
<td>RC=6 (Pending Offline) =&gt;</td>
<td>Wait for Offline</td>
<td>Nothing</td>
</tr>
</tbody>
</table>

How System Automation composes the OpState of a resource group

IBM Tivoli System Automation for Multiplatforms is a policy-based automation product. The control point for the automation is the resource group level, meaning that an Operator usually starts or stops an entire resource group rather than starting or stopping single resources. This is done by changing the NominalState attribute of a resource group to Online or Offline. Immediately after changing this attribute, System Automation for Multiplatforms will decide which resources needs to be started or stopped to meet the rules of the changed policy.

The OpState (Operational State) attribute of a resource group is an aggregation of the OpState attributes of all resources contained in that resource group in relation to the NominalState value of the resource group. So if the NominalState of a resource group has been changed to Online, the OpState of this resource group is showing Pending Online until all of the resources in that resource group are Online. Finally, if all resources of that resource group have reached the value of the NominalState attribute of the resource group, the OpState of the resource group changes to Online, and this value of the OpState attribute of a resource group can now be used to monitor the status of the resources in that group.

Note: Member resources may have a different OpState than the containing group when requests were issued directly against the member resources. For more information, see "Starting and stopping resource groups and group members" on page 190.

The following table shows how System Automation for Multiplatforms composes the value of the OpState attribute of a resource group based on the OpState of the two contained resources Res1 and Res2 from the above example. Note that this picture becomes more complex the more resources are contained in a single resource group.

Table 8. Resource group OpState determination

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>=&gt; Unknown</td>
<td>Nothing</td>
</tr>
<tr>
<td>Offline</td>
<td>Offline</td>
<td>=&gt; Offline</td>
<td>Nothing</td>
</tr>
<tr>
<td>Pending Online</td>
<td>Offline</td>
<td>=&gt; Pending Online</td>
<td>Wait until Res2 is Online</td>
</tr>
<tr>
<td>Online</td>
<td>Offline</td>
<td>=&gt; Pending Online</td>
<td>Start Res1</td>
</tr>
</tbody>
</table>
System Automation for Multiplatforms logic

Table 8. Resource group OpState determination (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Online</td>
<td>Pending Online</td>
<td>=&gt; Pending Online</td>
<td>Wait until Res1 is Online</td>
</tr>
<tr>
<td>Online</td>
<td>Online</td>
<td>=&gt; Online</td>
<td>Nothing</td>
</tr>
<tr>
<td>Online</td>
<td>Pending Offline</td>
<td>=&gt; Pending Offline</td>
<td>Wait until Res1 is Offline</td>
</tr>
<tr>
<td>Online</td>
<td>Failed Offline</td>
<td>=&gt; Pending Offline</td>
<td>Stop Res2</td>
</tr>
<tr>
<td>Pending Offline</td>
<td>Offline</td>
<td>=&gt; Pending Offline</td>
<td>Wait until Res2 is Offline</td>
</tr>
<tr>
<td>Failed Offline</td>
<td>Offline</td>
<td>=&gt; Offline</td>
<td>Nothing</td>
</tr>
</tbody>
</table>

How System Automation reacts to OpState changes of a resource that is started or stopped

System Automation for Multiplatforms usually automates resources based on the DesiredState of the resource group and on the OpState value of the resource. The goal is to achieve and maintain a state in which a resource’s OpState and DesiredState are identical. System Automation for Multiplatforms also takes action when the OpState of a resource changes, for example, when a resource that was running is now reported as Offline by the MonitorCommand.

Another trigger for automation actions is the return code of the StartCommand. If the StartCommand returns an error (a non-zero return code) and the resource is not monitored Online, System Automation for Multiplatforms performs a failover of the resource to another eligible node.

The following sections describe the actions that are performed by System Automation for Multiplatforms if the OpState of a resource changes during or shortly after the execution of the StartCommand or StopCommand for a resource.

The enhancements introduced in System Automation for Multiplatforms V2R2 enforce a more compliant implementation of a resource’s MonitorCommand. This is especially important at the time when the StartCommand for a resource is executed. At that time, System Automation for Multiplatforms knows that the OpState of the resource is Pending Online, and so this OpState is set for the resource. The only exceptions are:

- the OpState reported by the MonitorCommand changes to Online (in which case the automation goal is reached)
- the OpState reported by the MonitorCommand changes to Failed Offline (which informs System Automation for Multiplatforms that the resource is broken and cannot be started on that node without prior manual intervention)

This applies to the execution of the StopCommand in a similar way, where the OpState of the resource is set to Pending Offline during that period. Possible exceptions are:

- the MonitorCommand returns Offline or Failed Offline (which indicate that the automation goal has already been reached)
- the MonitorCommand returns Stuck Online (which indicates that manual intervention is required because the resource cannot be stopped by System Automation for Multiplatforms)
**System Automation for Multiplatforms logic**

**StartCommand**

The following tables illustrate how System Automation for Multiplatforms reacts to OpState changes during the execution of a StartCommand. There is one table for each of the three possible situations in which the MonitorCommand can report an OpState change:

1. The StartCommand is still running (long-running StartCommand).
2. The StartCommand has completed successfully (this is the normal situation).
3. The StartCommand has completed with an error or has timed out.

**StartCommand is still running:**

*Table 9. System Automation actions when a StartCommand still running*

<table>
<thead>
<tr>
<th>StartCommand</th>
<th>MonitorCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>StartCommand started, but not complete.</td>
<td>RC=0 (Unknown)</td>
<td>OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=1 (Online)</td>
<td>Start other resources, if any</td>
</tr>
<tr>
<td></td>
<td>RC=2 (Offline)</td>
<td>OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=3 (Failed Offline)</td>
<td>The resource is failed over to another node; other dependent resources may be forced down</td>
</tr>
<tr>
<td></td>
<td>RC=4 (Stuck Online)</td>
<td>Unlikely OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=5 (Pending Online)</td>
<td>OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=6 (Pending Offline)</td>
<td>Unlikely OpState is set to Pending Online, Wait for Online</td>
</tr>
</tbody>
</table>

**Notes:**

1. When the MonitorCommand reports the resource as Online, System Automation for Multiplatforms ignores the return code of the StartCommand, because the goal to bring the resource Online has been achieved.
2. When the MonitorCommand reports the resource as Failed Offline, a failover is triggered. All other potential OpState values are mapped to Pending Online and System Automation waits for the OpState to change to Online.

**StartCommand completed successfully:** This table describes the typical automation actions that are performed when a StartCommand has completed successfully.
Table 10. System Automation actions after a StartCommand completed successfully

<table>
<thead>
<tr>
<th>StartCommand</th>
<th>MonitorCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (successful) and actual retry count &lt; RetryCount (samctrl)</td>
<td>RC=0 (Unknown)</td>
<td>OpState is set to Unknown until new state information is received</td>
</tr>
<tr>
<td></td>
<td>RC=1 (Online)</td>
<td>OpState is set to Online. Start command processing is considered complete.</td>
</tr>
<tr>
<td></td>
<td>RC=2 (Offline)</td>
<td>OpState is set to Pending Online, Wait for Online. An online timeout triggers these actions:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. A reset operation is performed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. If the reset is successful, a start attempt is made.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. The retry count is increased.</td>
</tr>
<tr>
<td></td>
<td>RC=3 (Failed Offline)</td>
<td>OpState is set to Failed Offline. Start command processing is considered complete.</td>
</tr>
<tr>
<td></td>
<td>RC=4 (Stuck Online)</td>
<td>Unlikely OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=5 (Pending Online)</td>
<td>OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td></td>
<td>RC=6 (Pending Offline)</td>
<td>Unlikely OpState is set to Pending Online, Wait for Online</td>
</tr>
<tr>
<td>RC=0 (successful) and actual retry count = RetryCount (samctrl) and after Online timeout</td>
<td></td>
<td>Resource is set to Failed Offline, stop command is issued against the resource, which is then failed over to another node; other dependent resources may be forced down</td>
</tr>
</tbody>
</table>

StartCommand completed with an error or timed out: The following table describes the typical automation actions that are performed when a resource’s StartCommand returns with an error or times out, depending on the OpState of the resource:
### System Automation for Multiplatforms logic

**Table 11. System Automation actions when a StartCommand completed with an error or timed out**

<table>
<thead>
<tr>
<th>MonitorCommand</th>
<th>StartCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (Unknown)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>When the StartCommand returns 1, the resource is immediately failed over to another node. The return code of the MonitorCommand is ignored.</td>
</tr>
<tr>
<td>RC=1 (Online)</td>
<td>RC=1 (not zero) not successful</td>
<td>When the MonitorCommand returns 1 (Online), the return code of the StartCommand is ignored. No further actions are taken. The resource stays Online.</td>
</tr>
</tbody>
</table>
### System Automation for Multiplatforms logic

<table>
<thead>
<tr>
<th>MonitorCommand</th>
<th>StartCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=1 (Online)</td>
<td>StartCommand timed out</td>
<td>If the MonitorCommand returns 1 (Online), a timeout of the StartCommand triggers neither a StopCommand execution nor any other automation action. However, System Automation kills the StartCommand when it times out, which may also kill the resource process itself if it cannot be successfully detached from the StartCommand script at the time. If this happens, the OpState of the resource changes to Offline again. As soon as the MonitorCommand detects that the resource is Offline, System Automation will restart the resource on the same node. Note that this behavior may result in a loop if the StartCommand again runs into a timeout. In such a case, the RetryCount has no effect, because the internal counter is reset when the OpState of the resource changes to Online. Hint: Usually the loop is a result of an error within the StartCommand. However, the loop can be prevented if the MonitorCommand of the resource returns Failed Offline, in case the application process has been killed unexpectedly. In this case, a take over would take place, and the resource would be started on another node. For additional information about how the StartCommand of an IBM.Application resource is handled, see &quot;Using the Global Resource Manager&quot; on page 221.</td>
</tr>
</tbody>
</table>

Table 11. System Automation actions when a StartCommand completed with an error or timed out (continued)
## System Automation for Multiplatforms logic

Table 11. System Automation actions when a StartCommand completed with an error or timed out (continued)

<table>
<thead>
<tr>
<th>MonitorCommand</th>
<th>StartCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=2 (Offline)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>Immediately after the StartCommand returns with RC=1, the resource is stopped by System Automation for Multiplatforms, and then a failover takes place. If the start command timed out, its process is killed with <code>killpg()</code> (see also “Using the Global Resource Manager” on page 221).</td>
</tr>
<tr>
<td>RC=3 (Failed Offline)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>If a MonitorCommand did return 3 (Failed Offline) previously, then a failover already takes place, and the unsuccessful execution of the StartCommand has no additional effect.</td>
</tr>
<tr>
<td>RC=4 (Stuck Online)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>Unlikely, immediately after the StartCommand returns with RC=1, the resource is stopped by System Automation for Multiplatforms, and then wait for Operator action.</td>
</tr>
<tr>
<td>RC=5 (Pending Online)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>Immediately after the StartCommand returns with RC=1, the resource is stopped by System Automation for Multiplatforms, and then a failover takes place after the resource is reported Offline.</td>
</tr>
<tr>
<td>RC=6 (Pending Offline)</td>
<td>RC=1 (not zero) not successful, or timed out</td>
<td>Unlikely, immediately after the StartCommand returns with RC=1, the resource is stopped by System Automation for Multiplatforms, and then a failover takes place after the resource is reported Offline.</td>
</tr>
</tbody>
</table>

**Note:** When the MonitorCommand reports the resource as Online, the return code of the StartCommand is ignored. In this case, the results of the two commands are inconsistent:

- The StartCommand indicates that the start of the resource has failed
- The MonitorCommand reports the resource as Online.
StopCommand
The tables in this section illustrate how System Automation for Multiplatforms reacts to OpState changes during the execution of the StopCommand. StopCommand refers to the System Automation methods stop and reset, that is, in both cases the StopCommand is executed. Similarly, both timeout values, the StopCommandTimeout as well as the ResetCommandTimeout, refer to the same timeout value.

There is one table for each of the three situations in which the MonitorCommand can report an OpState change:
1. The StopCommand is still under execution (long-running StopCommand).
2. The StopCommand has successfully finished (this is the normal situation).
3. The StopCommand has finished with an error or has timed out.

StopCommand is still under execution:

Table 12. System Automation actions and StopCommand still under execution

<table>
<thead>
<tr>
<th>StopCommand</th>
<th>MonitorCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>StopCommand started, but not finished.</td>
<td>RC=0 (Unknown)</td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td></td>
<td>RC=1 (Online)</td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td></td>
<td>RC=2 (Offline)</td>
<td>Go on stopping other resources</td>
</tr>
<tr>
<td></td>
<td>RC=3 (Failed Offline)</td>
<td>Go on stopping other resources</td>
</tr>
<tr>
<td></td>
<td>RC=4 (Stuck Online)</td>
<td>Wait for Operator action</td>
</tr>
<tr>
<td></td>
<td>RC=5 (Pending Online)</td>
<td>Unlikely OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td></td>
<td>RC=6 (Pending Offline)</td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
</tbody>
</table>

Note that once the MonitorCommand has reported the resource as Offline or Failed Offline, System Automation for Multiplatforms ignores the return code of the StopCommand because the goal to bring the resource Offline is already achieved.

If the OpState for the resource is reported as Stuck Online, System Automation for Multiplatforms waits for the intervention by an operator, because the resource cannot be stopped. All other potential OpState values are mapped to Pending Offline, and System Automation for Multiplatforms waits for the OpState to change to Offline.

StopCommand completed successfully: This table describes the typical behavior of System Automation for Multiplatforms:
System Automation for Multiplatforms logic

Table 13. System Automation actions after StopCommand successfully finished

<table>
<thead>
<tr>
<th>StopCommand</th>
<th>MonitorCommand</th>
<th>Action of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (successful) and (Offline time out is not reached) or (reset sent against the resource and reset timeout is not reached)</td>
<td>RC=0 (Unknown)</td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td>RC=1 (Online)</td>
<td></td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td>RC=2 (Offline)</td>
<td></td>
<td>OpState is set to Offline. StopCommand processing is considered complete.</td>
</tr>
<tr>
<td>RC=3 (Failed Offline)</td>
<td></td>
<td>OpState is set to Failed Offline. StopCommand processing is considered complete.</td>
</tr>
<tr>
<td>RC=4 (Stuck Online)</td>
<td></td>
<td>OpState is set to Stuck Online. StopCommand processing is considered complete, Wait for Operator action.</td>
</tr>
<tr>
<td>RC=5 (Pending Online)</td>
<td></td>
<td>Unlikely, OpState is set to Pending Offline, Wait for Offline</td>
</tr>
<tr>
<td>RC=6 (Pending Offline)</td>
<td></td>
<td>OpState is set to Pending Offline, Wait for Offline</td>
</tr>
</tbody>
</table>

StopCommand finished with an error or timed out: The table below lists the automation actions that occur when a StopCommand completes with an error or times out. The actions depend on whether the StopCommand was invoked for a stop or reset operation.

Table 14. System Automation actions after StopCommand finished with an error or timed out

<table>
<thead>
<tr>
<th>StopCommand</th>
<th>MonitorCommand</th>
<th>Automation actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=1 (not zero; not successful, or timed out) and (Offline time out is not reached) or (reset sent against the resource and reset timeout is not reached)</td>
<td>RC=0 (Unknown)</td>
<td>Send reset order immediately</td>
</tr>
<tr>
<td></td>
<td>RC=1 (Online)</td>
<td>Send reset order immediately</td>
</tr>
<tr>
<td></td>
<td>RC=2 (Offline)</td>
<td>Go on stopping other resources</td>
</tr>
<tr>
<td></td>
<td>RC=3 (Failed Offline)</td>
<td>Go on stopping other resources</td>
</tr>
<tr>
<td></td>
<td>RC=4 (Stuck Online)</td>
<td>Send reset order immediately</td>
</tr>
<tr>
<td></td>
<td>RC=5 (Pending Online)</td>
<td>Send reset order immediately</td>
</tr>
<tr>
<td></td>
<td>RC=6 (Pending Offline)</td>
<td>Send reset order immediately</td>
</tr>
</tbody>
</table>
System Automation for Multiplatforms logic

Table 14. System Automation actions after StopCommand finished with an error or timed out (continued)

<table>
<thead>
<tr>
<th>StopCommand</th>
<th>MonitorCommand</th>
<th>Automation actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=1 (unsuccessful) and after Offline timeout</td>
<td></td>
<td>OpState is set to Pending Offline, Reset sent against the resource</td>
</tr>
<tr>
<td>RC=1 (unsuccessful) and after reset timeout</td>
<td></td>
<td>OpState is set to Stuck Online, Wait for Operator action</td>
</tr>
</tbody>
</table>

Notes:
1. The OpState of the resource is set to Pending Offline until its OpState changes to Offline, Failed Offline, or Stuck Online. If the OpState does not change to Offline or Failed Offline within the Offline timeout period, System Automation for Multiplatforms issues a reset operation against the resource, which triggers an execution of the StopCommand. If the OpState still does not change to Offline, Failed Offline, or Stuck Online within the reset timeout period, the resource is finally set to Stuck Online to indicate that an operator must take action to stop the resource. If the return code of the StopCommand or reset operation is not 0 (zero), the resource will end in automation state 'Problem', which requires manual intervention. To get this resolved, the resource needs to be reset using resetrsrc; however, the return code of this reset operation needs to be 0 (zero) otherwise the resource will stay in automation state 'Problem'.
2. If the order was a reset, the resource is set to stuck online, instead of sending a reset order.

How System Automation reacts if a resource is Online on a certain node and the MonitorCommand reports an OpState for the resource on another node at the same time

The following table shows the actions System Automation for Multiplatforms will perform, if a resource is Online on a node, and the MonitorCommand returns a certain OpState for this resource on another node.

Table 15. System Automation actions after the MonitorCommand reports an OpState change for the resource on another node

<table>
<thead>
<tr>
<th>MonitorCommand on stand-by Node</th>
<th>Actions of System Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC=0 (Unknown)</td>
<td>No automation action possible for this resource until the MonitorCommand returns with RC&lt;&gt;0.</td>
</tr>
</tbody>
</table>
**System Automation for Multiplatforms logic**

**Table 15. System Automation actions after the MonitorCommand reports an OpState change for the resource on another node** (continued)

<table>
<thead>
<tr>
<th>MonitorCommand on stand-by Node</th>
<th>Actions of System Automation</th>
</tr>
</thead>
</table>
| RC=1 (Online)                   | 1. The resources on both nodes are stopped. This affects the following resources, which are stopped as well:  
   - dependent resources that are linked through dependsOn or forcedDownBy relationships  
   - other group members  
   The scope of resources that will be stopped can be determined by the scope of a binder run, that is, all resources that are connected through relationships, including group membership, will be stopped (see “Location relationship resolution: Binding algorithm” on page 91). |
| RC=2 (Offline)                  | No action, this is usual OpState of the resource on a stand-by node. |
| RC=3 (Failed Offline)           | No action, but no failover is possible to this node anymore. |
| RC=4 (Stuck Online)             | Unlikely, script error (requires Online OpState before ...) |
| RC=5 (Pending Online)           | Same as Online |
| RC=6 (Pending Offline)          | Unlikely, script error (requires Online OpState before ...) |

Most important to notice is the fact, that System Automation for Multiplatforms will stop the resource on both nodes, if it is monitored Online on more than one node at a time, and not just the resource on the stand-by node. This also implies all resources that have a dependency to this resource may also be stopped. For instance all resources that have a DependsOn relationship to this resource will be stopped too. It is therefore recommended not to start and stop applications and resources manually that are under control of System Automation for Multiplatforms.

**State cycle of a resource group**

Figure 5 on page 113 shows the state transitions that occur when System Automation successfully starts and stops a resource group.
At the beginning of the cycle (1), the resource group's DesiredState is Offline, its ObservedState is Offline, and the members of the resource group are unbound, which results in a BindingState of Unbound.

If the NominalState of the resource group is changed to Online (2), the DesiredState of the resource group also changes to Online. The BindingState is still Unbound (3).

In the binding step (4), the binder tries to find a placement for all members of the resource group that satisfies all dependencies defined in the automation policy. If the binding step is successful, the BindingState of the resource group changes to Bound (5), otherwise the BindingState of the resource group will end up in state Sacrificed and the resources will not be started.

When all resource group members are bound, System Automation knows on which node to start them. The logic deck sends out start orders for the members of the resource group (6) and ensures that all dependencies defined in the automation policy are fulfilled. While the resources are starting, the ObservedState of the resource group changes to Pending Online (7).

When all resources are Online, the ObservedState of the resource group changes to Online (8).
System Automation for Multiplatforms logic

When the NominalState of the resource group is changed to Offline (9), the DesiredState of the resource group changes to Offline (10). This will trigger System Automation to stop all members of the resource group. The logic deck sends out stop orders for the resources (11) and ensures that all dependencies defined in the automation policy are fulfilled. While the resources are stopping, the ObservedState changes to Pending Offline (12). When all resources are Offline, the ObservedState of the resource group changes to Offline (13).

In the unbinding step (14), the node placement of the resource group members is deleted from the internal tables and the BindingState of the resource group changes to Unbound (1).
Chapter 8. Using the Integrated Solutions Console

This chapter describes the tasks you can perform in the SA operations console, the policy editor, and the other System Automation tasks and applications available in Integrated Solutions Console.

Authorizing users to perform System Automation tasks from Integrated Solutions Console

The installation task creates one authorized user (default user ID: eezadmin) with which System Automation for Multiplatforms tasks can be performed from Integrated Solutions Console. To authorize additional users, you must assign the user IDs they have on the automation domain to System Automation for Multiplatforms specific user groups which were created and assigned to System Automation for Multiplatforms specific access roles during the installation of the SA operations console. Note that each user ID for the automation domain must consist of an alphanumeric string with characters that are part of the local code set.

The operations console users can store and change their user credentials for the domains in the credential vault of Integrated Solutions Console (refer to "Managing your user credentials" on page 141 for more information).

Creating and authorizing users to work with System Automation for Multiplatforms from Integrated Solutions Console

The following user IDs and user groups are created during the installation of the SA operations console. The user groups are associated with the corresponding user roles.

The WebSphere administrative role supressmonitor determines whether you have access to WebSphere specific administration tasks from Integrated Solutions Console. If you only have this role assigned (and no other role that would give you access to WebSphere administrative tasks), the WebSphere administrative tasks are suppressed in the user interface.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Default user IDs (created during installation)</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>EEZAdministratorGroup</td>
<td>eezadmin</td>
<td>EEZAdministrator,suppressmonitor</td>
</tr>
<tr>
<td>EEZOperatorGroup</td>
<td></td>
<td>EEZOperator,suppressmonitor</td>
</tr>
<tr>
<td>EEZConfiguratorGroup</td>
<td></td>
<td>EEZConfigurator,suppressmonitor</td>
</tr>
<tr>
<td>EEZMonitorGroup</td>
<td></td>
<td>EEZMonitorGroup,suppressmonitor</td>
</tr>
</tbody>
</table>

For detailed information on how to create and authorize users in Integrated Solutions Console, see the Websphere Application Server documentation.

Table 16 on page 116 describes the access roles that determine which System Automation for Multiplatforms tasks are available to a user in Integrated Solutions Console.
The access roles are created during the installation of the SA operations console and assigned to the user groups listed in the right column of the table. To assign access roles to individual users, you assign the users’ IDs to the corresponding user groups.

The EEZ* access roles only authorize users to access and work with System Automation for Multiplatforms tasks. Other administrative console tasks are only available to users who have the Websphere Application Server Administrator access role.

Table 16. Access roles for System Automation for Multiplatforms

<table>
<thead>
<tr>
<th>Role</th>
<th>Permissions</th>
<th>Group name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EEZMonitor</td>
<td>Grants minimum access rights. Users who have this role can perform query-type operations but cannot activate and deactivate automation policies or perform actions that modify the state of resources, for example, they cannot submit start requests. In the navigation tree, the following tasks are available to EEZMonitor users: • SA operations console • Stored domain credentials</td>
<td>EEZMonitorGroup</td>
</tr>
<tr>
<td>EEZOperator</td>
<td>In addition to the permissions granted by the EEZMonitor role, users who have this role have the ability to issue requests against resources. The role does not permit users to perform tasks that change the configuration, such as activating and deactivating policies. In the navigation tree, the following tasks are available to EEZOperator users: • SA operations console • Stored domain credentials</td>
<td>EEZOperatorGroup</td>
</tr>
<tr>
<td>EEZConfigurator</td>
<td>In addition to the permissions granted by the EEZMonitor role, users who have this role have the ability to perform tasks that change the configuration, such as activating and deactivating policies. Users who have only this role cannot submit requests against resources. The role is required to be able to work with policies. In the navigation tree, the following tasks are available to EEZConfigurator users: • SA operations console • Stored domain credentials • Activate an automation policy • Deactivate active automation policy</td>
<td>EEZConfiguratorGroup</td>
</tr>
</tbody>
</table>
### Configuring your Web browser for Integrated Solutions Console

The following minimum Web browsers are supported:
- Microsoft Internet Explorer V6.0 SP1
- Mozilla V1.7.8
- Firefox 1.5

To display Integrated Solutions Console in your Web browser, the following settings are required:
- JavaScript must be enabled in all Web browsers.
- For Microsoft Internet Explorer, the following settings are required:
  - Set the security level to medium.
    Do not set the security level to “high”. If high security is required, be sure to set the entry **ActiveX controls and plugins - Initialize and Script ActiveX controls not marked as safe** on the Security settings page to **Enable**. Otherwise, the information displayed on the console is not updated automatically.
  - Set **Scripting - Active Scripting** to **Enable** on the Security settings page. Otherwise, navigation problems may occur.

### Logging in to Integrated Solutions Console

To access Integrated Solutions Console, perform the following steps:

1. Open a Web browser window and type the address of Integrated Solutions Console in the **Address** field.
   The entry must have the following form:
   \[http://<hostname>:<port>/ibm/console\]

   where `<hostname>` is the name of the host on which Integrated Solutions Console is running and `<port>` is the port number of Integrated Solutions Console. The default port is 9060.
2. Wait for the console to load into the browser. A login page is displayed after the console starts.

3. Specify your user ID and password and click Log in.

The user interface for Integrated Solutions Console is displayed.

After you are logged in, be sure to use the Logout link in the console toolbar when you are finished using the console and to prevent unauthorized access. If there is no activity during this login session for an extended period of time, the session expires and you must log in again to access the console.

If the user ID that you provide is already logged in at a different location, you are prompted to choose between logging out from the other location or returning to the login page.

It is recommended that you do not use multiple browser windows on the same client system simultaneously to connect to the same Integrated Solutions Console, because browser types other than Microsoft Internet Explorer will share a single HTTP session between multiple browser instances if these instances are running on the same system and connect to the same Integrated Solutions Console. Working with multiple browser instances using the same HTTP session will cause unexpected results.

The same situation occurs if you open multiple Microsoft Internet Explorer browser windows using File > New Window (or Ctrl+N) from an existing Integrated Solutions Console session, because in this case the new browser window and the one from which it was opened will also share the same session.

### Integrated Solutions Console layout

This topic provides an overview the layout of the user interface of Integrated Solutions Console. For detailed information about using the user interface, see the Integrated Solutions Console online help, which is accessible through the Help link in the console toolbar.

**Banner**
Displays a common image across all Integrated Solutions Console installations. The banner includes a greeting to the user who is logged in and links to log out of the console and to open console help.

**Navigation tree**
Lists the tasks available in the console. Tasks are grouped into organizational nodes that represent categories of tasks. The organizational nodes can be nested in multiple levels. Which tasks are shown depends on your user role and on your current View settings. When you click a task in the navigation, a page is displayed in the work area containing one or more modules for completing the task.

Use the View selection list at the top of the navigation tree to modify the list of tasks according to your preferences. You can organize the tasks as follows:

**All tasks**
This shows all tasks in the console. Tasks are grouped into organizational nodes. The tasks that are specific to System Automation for Multiplatforms are available under the node IBM Tivoli System Automation for Multiplatforms.

**My tasks**
This shows only the tasks that you have added to the view. This list is initially empty, but provides a link to the My Tasks page.
Use My Tasks to add and remove from the My Tasks list in the navigation. For detailed information about using My Tasks, see the console help.

**Product selection**
Selecting a product name shows only the tasks for the particular product, for example, IBM Tivoli System Automation for Multiplatforms.

**Work area**
When you launch a page, the content of the page is displayed in the work area. If you have not launched any pages, the Integrated Solutions Console Welcome page is displayed in the work area. It displays the products that are installed that use the Integrated Solutions Console as common user interface. You can open the Welcome page for IBM Tivoli System Automation for Multiplatforms by clicking the product name in the list.

---

## System Automation for Multiplatforms tasks in the navigation tree

Which tasks you can see in the navigation tree depends on your System Automation for Multiplatforms user role and on your current View settings. If you have not customized the View task list, you expand the IBM Tivoli System Automation for Multiplatforms entry in the navigation tree to display the list of product-specific tasks you have access to. The following list gives an overview of all available Tivoli System Automation tasks.

### Welcome
Opens the Welcome page for Tivoli System Automation. To open the page, you can also click the **IBM Tivoli System Automation for Multiplatforms** link on the Welcome page for Integrated Solutions Console.

### SA operations console
Opens the SA operations console in the work area. You use the SA operations console for managing and monitoring automated resources and for performing administrative tasks, some of which are also available from the navigation tree (for example, policy activation and deactivation for automation domains that support the tasks).

### Operational tasks
The node contains the following tasks:

**Activate an automation policy**
Allows you to activate automation policies for automation domains that support policy activation. You can also perform the task from the SA operations console.

**Deactivate active automation policy**
Allows you to deactivate the currently active automation policy for automation domains that support policy deactivation. You can also perform the task from the SA operations console.

### Settings
The node contains the following tasks:

**Stored domain credentials**
Use this task to administer your user credentials for the first-level automation domains to which you have access. The user credentials for an automation domain are saved to the credential vault when you access the domain for the first time.
**Tivoli Enterprise Portal launch-in-context configuration**

If you are using both the SA operations console and Tivoli Enterprise Portal for resource monitoring and management, use this task to set up launch-in-context support for Tivoli Enterprise Portal. Launch-in-context enables users to launch Tivoli Enterprise Portal work spaces from the SA operations console with a single mouse click. The configuration of the launch-in-context function is described in *IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide* and in the SA operations console online help.

**SA Policy Management**

Enables you to create and save new policies using a graphical user interface. You can also browse and load existing policies for editing, rename existing policies, and delete existing policies. When you select any of the following tasks, a new tab page opens in the work area of Integrated Solutions Console:

**Edit an existing policy**

Loads an existing policy into the policy editor, where it can be modified and saved. You can load the currently active policy from a connected automation domain, select a policy stored in the policy pool of a domain, or select a policy saved on the local file system. While the policy is loading, it is checked for semantic and syntactical errors.

**Create a new policy**

Enables you to create a new policy. When you have finished, you can save the policy in the policy pool of a connected automation domain or on the local file system.

**Browse Policies in Policy Pool**

Enables you to manage policies stored in policy pools. After browsing to the policy pool of a connected automation domain, you can edit, rename, or delete a listed policy.
Using the SA operations console

This section gives you an overview of the SA operations console.

Note: The operations console is not available in the Solaris environment. However, an operations console installed on another operating system platform can be used to manage Solaris domains.

Overview of the SA operations console

The SA operations console is a browser-based graphical user interface that is displayed in IBM Integrated Solutions Console (ISC).

The following figure shows the setup of the SA operations console both for users of System Automation for Multiplatforms and for users of IBM Tivoli System Automation Application Manager.

System Automation for Multiplatforms deals with the left part of the figure which shows how the operations console is used in direct access mode without the end-to-end automation part for which IBM Tivoli System Automation Application Manager is required. Direct access mode means that you can only monitor and manage domains controlled by IBM Tivoli System Automation for Multiplatforms. System Automation Application Manager provides two more modes: the end-to-end automation mode and the first-level automation mode. Both modes are described in IBM Tivoli System Automation Application Manager Administrator’s and User’s Guide.
by the operations console (as shown on the left side of the figure) or by the end-to-end automation management part (as shown on the right side of the figure). Both cannot be done simultaneously.

**Understanding the layout of the SA operations console**

The main panel of the operations console is divided into several areas:

**Menu bar**

Use the entries in the Menu which is available on the menu bar to update the information displayed in the topology tree and the resource table, to change your user preferences, and to display information about the version of the operations console you are using.

**Title bar**

Use the controls on the title bar to display the online help page for the panel you are displaying and for minimizing and maximizing the panel.

**Information bar**

The information bar is not shown in the figure above. It is displayed below the menu bar when you have performed an action on an element in the operations console. It displays a message confirming that the request or command has been submitted for processing. The message on the information bar only confirms the initial action, it is not updated while the command or request is being processed. The results of the system actions that are performed due to the request or command are reflected in the operations console itself. There you can see, for example, that the status of a resource has changed.

The confirmation message is replaced with a new message whenever you perform an action against an element in the operations console. Clicking **Clear** on the information bar hides the information bar from view. It reappears with a new confirmation message when you perform an action on an element.
3 Information area
Use the pages in the information area to obtain information about the element you have selected in the topology tree or resource table, and to perform actions against the element. For more information about the information area, refer to “What you must know about the information area” on page 130.

4 Resources section
Use the areas of the resources section to work with resources:

View and Search
The View and Search functions allow you to limit the scope of the resource table. View allows you to display only resources that are in an error or warning state. Search allows you to display only resources that meet specific search criteria.

Resource table
Displays a list of resources and their states. You use it to select and work with resources. The resource table has two views:

Search results view
When you use Search to see only a specific set of resources in the resource table, the search results are displayed in the search results view. For more information, refer to “Search results view” on page 129.

Group hierarchy view
The group hierarchy view is displayed when you are not displaying the results of a search. For more information, refer to “Group hierarchy view” on page 126.

For more information about the resources section, refer to “What you must know about the resources section” on page 125.

5 Topology tree
The topology tree shows the automation domains and the nodes that belong to the domains. The topology tree displays state-related information, allows you to select and work with domains and nodes, and is used to control what is displayed in the resource table. For more information, refer to “What you must know about the topology tree.”

What you must know about the topology tree
The Topology tree has three columns:

- The Topology column shows the automation domains and the nodes that belong to a domain in a hierarchical view.
- The Status column shows the health status of the domain.
- The Located here column is used to identify by which domain a resource is hosted and on which node or nodes it is located.

Navigating the topology tree: You click the twistie in front of a domain icon to expand or collapse the nodes belonging to the domain.

Selecting an element in the topology tree: To select an element in the tree, click its name. When you select a domain or node, you influence what is displayed in the resource table and in the information area:

- The resource table shows the resources that are hosted by the selected domain or that are located on the selected node.
The pages in the information area show information about the element that is selected in the topology tree. Depending on which type of element you have selected, buttons are enabled on the pages that let you perform actions against the element.

Notes:
1. To select its resources in an automation domain, you must first expand the domain by clicking the adjacent blue square in the topology tree.
2. The selected resources for which a reference is created are displayed in the right pane of the window.

Limiting the scope of the topology tree: By default, all automation domains are displayed in the topology tree. When you are not interested in seeing all automation domains, you can hide domains from view. To limit the scope of the topology tree, you use the Visible domains page of the Preferences panel. To open the Preferences panel, click Menu in the menu bar and select Preferences.

What is displayed in the topology column: In the topology column you see the automation domains and the nodes that are managed by each automation domain.

The following icons are used to identify the elements of the topology tree:

<table>
<thead>
<tr>
<th>Icon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Automation Domain]</td>
<td>An automation domain. When the domain is not online or its state is unknown, the icon is grayed-out.</td>
</tr>
<tr>
<td>![Node]</td>
<td>A node that belongs to an automation domain. When a node is not online, the icon is grayed-out.</td>
</tr>
</tbody>
</table>

What you can see in the Status column: The Status column is used to inform you of the health status of a domain. When the domain is healthy, the column is empty.

A domain is considered healthy if none of the resources that you are using as domain health indicators has encountered a problem that may require your attention. By default, the top-level resources of a domain are used as domain health indicators. However, on the Visible domains page on Preferences panel (Menu → Preferences) you can define that a different set of resources is to be used for the purpose.

When a resource that is used as domain health indicator has a problem, one of the following icons appears in the Status column:

<table>
<thead>
<tr>
<th>Icon</th>
<th>The icon indicates ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Warning]</td>
<td>A warning has been issued. The problem may still be solved automatically, but the element should be monitored carefully.</td>
</tr>
<tr>
<td>![Error]</td>
<td>The red error icon indicates that an error has occurred. To resolve the error, operator intervention is required.</td>
</tr>
<tr>
<td>![Black Error]</td>
<td>The black error icon indicates that an unrecoverable error has occurred. To resolve the problem, urgent operator intervention is required.</td>
</tr>
</tbody>
</table>
Note: If you see a warning icon right beside the domain name, this may indicate that the user ID and password for this System Automation for Multiplatforms domain has not been entered. Double click on the domain name and you are prompted for entering this user ID and password. Your system administrator may help you in getting the user ID and password.

What you can see in the Located here column: You use the Located here column to find out to which domain or domains a resource or group of resources belongs and on which nodes they are located. To determine the location of a resource or resource group, you select the resource or resource group in the resource table. When you have made your selection, a check mark appears in the Located here column for the domain that hosts the resource. If you are displaying the node hierarchy, a check mark will identify the node or nodes on which the resource is located.

What you must know about the resources section

The following figure shows the layout of the resources section.

![Figure 8. Layout of the resources section](image-url)

The resources section has the following areas:

**Section header**: The section header displays the name of the domain or node that is currently selected in the topology tree.

**View and Search**: The View and Search functions allow you to limit the scope of the resource table:

**View** Select the Errors and warnings item from the View drop-down list to display only resources that are in an error or warning state. The view is always applied to the list of resources which is currently displayed in the resource table.
Search
Allows you to display only resources that meet specific search criteria.

Resource table views: The resources which you are monitoring and managing from the operations console are displayed in the resource table. The resource table has two views, which are described in the sections below. In both views, you can perform the following basic actions:

Select a resource
To select a resource, you click its name in the resource column.

Control the sort order of the resource table
You can sort the resource table on any column by clicking the sort arrow in the column header.

A solid sort arrow in a column header indicates that the table is currently sorted on the column. The direction in which the solid sort arrow is pointing indicates the current sort order (ascending or descending). By clicking on the solid sort arrow, you can toggle between ascending and descending sort order.

When you position the cursor over a sort arrow, a hover help text appears showing the current sort status of the column and the sort order that will result when the sort arrow is clicked.

Page through the resource table
The resource table may extend over multiple pages. To page through the table or to navigate to a specific page you use the controls that are available in the status line below the table.

Group hierarchy view: The group hierarchy view is displayed when you are not displaying the results of a search. In the following figure, the top-level resources of the automation domain “FEPLEX1”, which is selected in the topology tree, are displayed in the resource table.

Resources of FEPLEX1

<table>
<thead>
<tr>
<th>Resource name</th>
<th>Compound state</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUS CONTIN/APG</td>
<td>Error</td>
</tr>
<tr>
<td>DR2 BACKUP/APG/SYS5</td>
<td>OK</td>
</tr>
<tr>
<td>DR2 BACKUP/APG/SYS6</td>
<td>OK</td>
</tr>
<tr>
<td>IP ADAPTER/APG</td>
<td>OK</td>
</tr>
<tr>
<td>SY5/SYG/SYS5</td>
<td>OK</td>
</tr>
<tr>
<td>SY56/SYG/SYS6</td>
<td>OK</td>
</tr>
<tr>
<td>SY7/SYG/SYS7</td>
<td>OK</td>
</tr>
</tbody>
</table>

Resource column:
In this column, you see a list of resources and their states. Which resources are listed depends on which selection you have made:

- A domain or node is selected in the topology tree: The top-level resources of the domain or node are displayed.
A group is selected in the resource table: The members of the group are displayed.

To sort the resources alphabetically by name, click the sort arrow in the column header.

When you select a group in the resource table, the members of the group are displayed in the resource table. In the area above the table, a bread crumb trail appears. On the trail, the name of the group whose members are listed in the resource table is highlighted, indicating that the group is selected.

The bread crumb trail is useful for both navigation and orientation:

- When you drill down into the group hierarchy, an entry is added to the trail for each group you select.
- The last entry on the trail identifies the group whose members are currently displayed in the resource table. When the group name is highlighted, the group is selected and the group details are displayed in the information area.
- When you click **Top** on the bread crumb trail, the top-level resources of the automation domain or node that is selected in the topology tree are again displayed in the resource table and the bread crumb trail disappears.
- When the bread crumb trail starts to get deeper than three levels, an ellipsis symbol (...) replaces all but the last two entries on the trail.

The ellipsis symbol cannot be clicked. To navigate upward through the group hierarchy, click an available group name on the trail until the group you want to view appears again.
The resource icon to the left of the resource name in the resource column indicates both the resource type and its online status:

- When the resource is online, its icon is active, when the resource is offline, the icon is grayed out.
- When a resource is in a warning or error state, the resource icon is highlighted with a warning or error icon.
- An operator icon indicates that an operator request was submitted against the resource. The color of the operator icon changes while the request is being processed, yellow indicates that the request has been submitted, green indicates that the request was completed successfully.

The following types of resources are displayed in the resource column:

<table>
<thead>
<tr>
<th>Icon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>![icon]</td>
<td>A resource</td>
</tr>
<tr>
<td>![icon]</td>
<td>A resource group</td>
</tr>
<tr>
<td>![icon]</td>
<td>A move group, also called a floating resource.</td>
</tr>
</tbody>
</table>

The following table lists the warning and error icons that appear in the resource column when a resource is in an error or warning state.

<table>
<thead>
<tr>
<th>Icon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>![icon]</td>
<td>The yellow warning icon indicates that the resource is in warning state. The problem may be resolved automatically but the resource should be monitored carefully.</td>
</tr>
<tr>
<td>![icon]</td>
<td>The red error icon indicates that the resource is in an error state. Operator intervention is required.</td>
</tr>
<tr>
<td>![icon]</td>
<td>The black error icon indicates that the resource has encountered an unrecoverable error. Immediate operator intervention is required.</td>
</tr>
</tbody>
</table>

**Compound state column:**
The column shows the compound state of the resource. By sorting on this column, you can group the resources by state.

The compound state can have one of the following values:

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK</td>
<td>The resource is working as desired. Select the resource to display more information.</td>
</tr>
<tr>
<td>Warning</td>
<td>The resource is in warning state. Select the resource to display more information.</td>
</tr>
</tbody>
</table>
### State Description

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>The resource is in an error state. Select the resource to display more information.</td>
</tr>
<tr>
<td>Fatal</td>
<td>The resource has encountered an unrecoverable error. Select the resource to display more information.</td>
</tr>
</tbody>
</table>

**Search results view:** When you use **Search** to see only a specific set of resources in the resource table, the search results are displayed in the search results view. In the area above the resource table, the search criteria that were used for the search are displayed. In this view, the resource table has the following layout:

**Resources of FEPEX1**

<table>
<thead>
<tr>
<th>Resource name</th>
<th>Compound state</th>
<th>Member of</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP APP/APG</td>
<td>OK</td>
<td>BUS_CONTII/APG</td>
</tr>
<tr>
<td>SAP APP/APG/SYS5</td>
<td>OK</td>
<td>SAP_APP/APG</td>
</tr>
<tr>
<td>SAP APP/APG/SYS8</td>
<td>OK</td>
<td>SAP_APP/APG</td>
</tr>
<tr>
<td>SAP APP/APG/SYS7</td>
<td>OK</td>
<td>SAP_APP/APG</td>
</tr>
<tr>
<td>SAP DB2/APG</td>
<td>OK</td>
<td>BUS_CONTII/APG</td>
</tr>
<tr>
<td>SAP DB2/APG/SYS5</td>
<td>OK</td>
<td>SAP_DB2/APG</td>
</tr>
<tr>
<td>SAP DB2/APG/SYS6</td>
<td>OK</td>
<td>SAP_DB2/APG</td>
</tr>
<tr>
<td>SAP DB2/APG/SYS7</td>
<td>OK</td>
<td>SAP_DB2/APG</td>
</tr>
<tr>
<td>SAP ENQ/APG</td>
<td>OK</td>
<td>BUS_CONTII/APG</td>
</tr>
<tr>
<td>SAP APP/PLI/SYS3</td>
<td>OK</td>
<td>SAP_APP/APG/SYS3</td>
</tr>
</tbody>
</table>

To limit the scope of resources that are currently displayed in the resource table to those that are in an error or warning state, you can additionally apply the **Errors and warnings** view that is provided in the **View** field.

**Resource table columns:**

In the search results view, the resource table has three columns:

**Resource column**

In the column, the resources that match the search criteria are listed.

- To sort the resources alphabetically by name, click the sort arrow in the column header.
- If a resource is in a warning or error state, the resource icon is highlighted with a warning or error icon.
- If an operator request was submitted against the resource, an operator icon is displayed.
- Clicking a resource selects the resource and its details are displayed in the information area.
Note: When you select a group in the search results view, the group details will be displayed in the information area, but the resource table will not switch to the group hierarchy view to display the group members.

To display the group’s members in the group hierarchy view, you must select the group and click Clear results (see “Clearing the search results”).

**Compound state column**

The column shows the compound state of the resource. By sorting on this column, you can group the resources by state.

**Member of column**

If a resource is a member of a group, the name of the group is displayed in this column. When you sort the resource table on this column, the resources that are members of the same group are listed next to each other.

*Clearing the search results:*

When you click Clear results, the resource table switches back to the group hierarchy view. Which resources are then displayed in the group hierarchy view, depends on your selection in the search results view:

- No resource was selected: The top-level resources of the automation domain or node that is selected in the topology tree are displayed.
- A resource group was selected: The group members are displayed. On the bread crumb trail, the name of the group is highlighted, the group details are displayed in the information area.
- A resource that is a member of a group was selected: The group members are displayed, the group name is displayed on the bread crumb trail but is not highlighted, the name of the selected resource is highlighted in the resource list.

**What you must know about the information area**

This section gives a short overview about the pages in the information area.

In the information area you find detailed information about the element that is currently selected in the topology tree or in the resource table. On the pages in the information area, controls are available that let you perform actions on the selected resource, group, node, or automation domain. Which pages are displayed and what they contain depends on the type of element that is currently selected in the topology tree or resource table:

<table>
<thead>
<tr>
<th>When you select ...</th>
<th>...these pages are available</th>
</tr>
</thead>
</table>
| an automation domain in the topology tree | • General  
  • Policy  
  • Additional Info |
| a node in the topology tree | • General  
  • Additional Info (available only if additional information exists) |
<table>
<thead>
<tr>
<th>When you select ...</th>
<th>...these pages are available</th>
</tr>
</thead>
</table>
| a resource or a resource group in the resource table | • General  
• Relationships (available only if the resource has relationships)  
• Location info  
• Additional Info (available only if additional information exists) |

**General page:** The General page is always available when an element is selected in the topology tree or the resource table. Which information is provided on the General page and which actions are possible depends on the selection you have made in the topology tree or the resource table.

**General page for a domain**
Use the General page for a domain to get information about the state of the domain, its communication state, and the state of the resources it hosts and to display the log file for the domain.

**General page for a node**
Use the General page for a node to get detailed information about the node, for example, its name, class, and possibly, a description of the node. In addition, the page contains information about the observed state of the node. A button is provided that lets you exclude the node from automation or include it in automation.

**General page for a resource group**
Use the General page for a resource group:
• to get more detailed information about the resource group itself, for example, its class.
• to display detailed information about the different states of the group.
• to submit a start or stop request.
• to view the request stack of the resource.
• to find out if the resource group is a member of any other resource group. If this is the case, a link is provided that lets you jump to that resource group.

**General page for a resource**
Use the General page:
• to display detailed information about the resource, for example, its class and owner.
• to obtain detailed information about the different states of the resource.
• to view the operator requests if there are any.
• to submit a start or stop request.
• to find out whether the resource is a member of a group.

**Policy page:** The Policy page is available when a domain is selected in the topology tree. It allows you to check, activate, and deactivate automation policies that are stored in the policy pool.

**Location info:** The page is available for resource groups. You use this page to determine the location and state of the members of the resource group.
**Additional Info page**: An Additional Info page can be displayed for automation domains, nodes, resources, and resource groups when additional information is available.

**Relationships page**: A Relationships page is available when you have selected a resource or group in the resource table and backward or forward relationships have been defined for the resource.

## Working with automation policies

### Checking automation policies

Perform this task to check whether the automation policies in the policy pool are ready for activation or contain problems that need to be resolved prior to activation.

To perform the task, the following prerequisites must be met:

- You must have at least EEZConfigurator privileges.
- The policy pool directory is configured for the automation domain.
- The automation policies you want to check are stored in the policy pool directory.

Perform these steps:

1. Open the "Select an automation policy" page in one of these ways:
   - In the console navigation tree, click **Tivoli System Automation for Multiplatforms > Operational Tasks > Activate an automation policy**. On the "Activate an automation policy" page, select the appropriate automation domain and click Next.
   - Open the SA operations console, select the appropriate automation domain in the topology tree, open the domain’s Policy page, and click **Activate new policy**.

2. If a warning or error icon appears in the right column of the policy table, warnings, errors, or both were issued during the validity check that was performed when you opened the page.

   To view the list of problems for a policy, select the policy and click the **View warnings** or **View errors** button that appears below the **Description** field. If errors where found in the file, you must correct them before the policy can be activated. Although warnings do not prevent the policy from being activated, you should check if they cannot be avoided.

3. Click **Cancel** to close the policy list. Repeat the procedure until all problems in the policy are resolved and it is ready for activation.

### Policy activation check when the end-to-end automation adapter is automated:

In environments in which the end-to-end automation adapter has been automated, ensure that a policy contains the resource definitions of the samadapter resource group before you activate the policy.

To check that the new policy contains resource definitions of the samadapter resource group:

- Browse the XML policy file and look for an element `<ResourceGroup name="samadapter-rg" class="IBM.ResourceGroup">`.
  - If the new policy contains such an element, you can activate the new policy from the operations console. See [“Activating automation policies” on page 133](#).
- If the new policy does not contain such an element, do not activate it. If you activate the new policy, the resource definitions of samadapter are removed and the adapter is stopped, which causes the domain to become unavailable in the operations console.

To reactivate the end-to-end automation adapter, use the Define function of the end-to-end automation adapter configuration dialog. This is described in the section "Configuring the end-to-end automation adapter" in IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide.

To prevent the adapter being removed when activating a new policy:
1. Use the sampolicy -s <file> command to save the current policy that contains the resource definitions for samadapter.
2. Copy the resource definitions for samadapter to the new policy that you want to activate.
3. Activate the new policy from the operations console.

**Activating automation policies**

Perform this task to activate an automation policy for a System Automation for Multiplatforms automation domain.

To perform the task, the following prerequisites must be met:
- You must have at least EEZConfigurator privileges.
- The policy pool directory is configured for the domain.
- The automation policies are stored in the domain’s policy pool directory.

Before proceeding, you should also perform the checks described in "Checking automation policies" on page 132

Perform these steps:
1. Open the "Select an automation policy" page in one of these ways:
   - In the console navigation tree, click Tivoli System Automation for Multiplatforms > Operational Tasks > Activate an automation policy. On the "Activate an automation policy" page, select the appropriate automation domain and click Next.
   - Open the SA operations console, select the appropriate automation domain in the topology tree, open the domain’s Policy page, and click Activate new policy
2. Select the automation policy you want to activate from the policy table. The policy must be error-free to be activated.
3. Click Activate to activate the policy. If you try to activate a policy that is already active, you receive a warning.

Policy activation from the SA operations console is non-disruptive, the activation process corresponds to that triggered by the command line activation with sampolicy -r.
- All policy elements that already exist in the system and have the same attributes and attribute values, remain unchanged.
- If attributes have changed, the corresponding element is changed in the system.
- If policy elements do not exist in the system, they are created.
- Elements that exist in the currently active automation policy but not in the new one are removed.
Deactivating the active automation policy
Perform this task to deactivate the currently active automation policy of an automation domain. This may be required, for example, if the policy causes severe problems that cannot be resolved in any other way.

To perform the task, the following prerequisites must be met:
- You must have at least EEZConfigurator privileges.
- The policy pool directory is configured for the domain.
- The automation policies are stored in the domain’s policy pool directory.

You can deactivate an active policy in one of these ways:
- Open the SA operations console, select the automation domain from the topology tree, open the domain’s Policy page, and click Deactivate policy.
- In the console navigation tree, click Tivoli System Automation for Multiplatforms > Operational Tasks > Deactivate current policy. On the "Deactivate active policy" page, select the appropriate automation domain and click Deactivate policy.

Results: The policy is deactivated, the domain is no longer automated. All resources, groups and relationships will be removed, the corresponding resources will be stopped. If the adapter was made highly available, the resources that make the adapter highly available, for example, a service IP and the adapter resource group, will remain unchanged. Therefore, the domain is still highly available.

Managing resources using the operations console
Managing resources means:
- Starting or stopping a resource or a group of resources.
- Excluding a node from automation and including it again.

Working with requests
You start and stop resources by changing their desired state. You can achieve this by submitting start or stop requests that bring a resource online or offline. The desired state of a resource will be changed when your request wins. The resource will only be started or stopped after all relationships have been fulfilled. See the description in "Starting and stopping resource groups and group members" on page 190.

For submitting requests, the following rules apply:
- Online requests can only be submitted against resources in desired state Offline.
- Offline requests can only be submitted against resources in desired state Online.
- Requests cannot be submitted if the current desired state of the resource results from an operator requests. In this case, the operator requests must be canceled.

Note: Due to the internal design of System Automation for Multiplatforms, requests against a resource group are not propagated to members that are floating resources.

Submitting start requests: Perform the following steps to submit a start request:
1. In the resource table, select the resource you want to start.
2. On the General page, click Request Online.
The Request Online panel is displayed.

3. Use the entry field on the Request Online panel to provide a short description of why you want to change the automation goal of the resource to Online. The character semicolon ";" is not allowed in comments.

4. Click Submit to submit the request.

Results:
- A confirmation message is displayed on the information bar, indicating that the request has been submitted for processing.
- After the next refresh, the icon of the resource is changed to indicate that a request has been issued against the resource.
- The request is processed. Processing of the request is complete when the resource has been started. The operator icon turns green.

Submitting stop requests: Perform the following steps to submit a stop request:

1. In the resource table, select the resource you want to stop.

2. On the General page, click Request Offline. The button is only enabled if the resource’s desired state is Online and no other operator requests against the resource exist.
   The Request Offline panel will be displayed.

3. Use the entry field to provide a short description of why you want to change the automation goal of the resource to Offline. The character semicolon ";" is not allowed in comments.

4. Click Submit to submit the request.

Results:
- A confirmation message is displayed on the information bar, indicating that the request has been submitted for processing.
- After the next refresh, the icon of the resource is changed to indicate that a request has been issued against the resource.
- The request is processed. Processing of the request is complete when the resource has been stopped. The operator icon turns green.

Displaying information about an operator request: When an operator has submitted a start or stop request against a resource, an operator request icon appears on the General page for the resource. The icon indicates whether the request has been completed successfully or not:
Table 20. Operator request icons in the information area

<table>
<thead>
<tr>
<th>Operator request icon</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(yellow)</td>
<td>A stop request has been submitted. The yellow operator icon indicates that the observed state of the resource is not Offline yet.</td>
</tr>
<tr>
<td>(yellow)</td>
<td>A start request has been submitted. The yellow operator icon indicates that the observed state of the resource is not Online yet.</td>
</tr>
<tr>
<td>(green)</td>
<td>The green operator icon indicates that the stop request has been completed successfully. The observed state of the resource is Offline.</td>
</tr>
<tr>
<td>(green)</td>
<td>The green operator icon indicates that the start request has been completed successfully. The observed state of the resource is Online.</td>
</tr>
</tbody>
</table>

This is how you can display more information about the request:

- Move the mouse over the operator request icon to display the user ID of the operator who submitted the request.
- Click the operator request icon to bring up the request details panel.

**Displaying request lists:** All requests and votes that have been submitted against a resource are added to the resource’s request list. You can display the list to find out which requests have been issued and which of the requests wins. The list is sorted by priority with the winning request listed at the top.

The list contains information about each request or vote, for example:

- its source (for example, the name of the operator who submitted the request)
- its priority
- the creation date and time

From the Request list panel, you can display detailed information about each of the requests or votes, including the comments that were added by operators when they submitted the request. Note that, due to the internal design of System Automation for Multiplatforms, requests against a resource group are not propagated to members which are floating resources.

**Steps for viewing a request list and the request details:**
Perform the following steps:

1. In the resource table, select the resource whose request list or request details you want to view.

2. On the General page, click **View requests**.
   The Request list is displayed. The list is sorted by priority. The first entry is the winning request.

3. To display the details for a request, select the resource in the list and click **More info**.
   The Request details panel is displayed.
**Canceling requests:** You can cancel operator requests that have been submitted against resources. Votes and requests generated by automation managers cannot be canceled.

This is what happens when you cancel a request:

- When you cancel a request that did not win, you prevent it from being completed at a later time.
- When you cancel the request that is responsible for the current desired state of the resource, you change the desired state of the resource to the opposite if there are no other requests or votes in the request list that will win when the canceled request is removed.
- When you cancel a request, votes that were generated against other resources because of StartAfter or StopAfter relationships are canceled as well.

**Steps for canceling requests:**
Perform the following steps to cancel a request:

1. Select the resource in the resource table.
2. On the General page, click **Cancel request**.
   
   The button is enabled only if there is an operator request in the request list of the resource.

   The text to the left of the **Cancel request** button describes the resource’s probable desired state after the request has been canceled. The probable desired state is calculated as follows:
   
   - If there are other requests or votes in the request list, the winning request determines the expected desired state.
   - If there are no other request or votes in the list, the desired state that is defined in the policy becomes the automation goal.
   - If there are other requests or votes in the request list, the one with the highest priority will win.

   The desired state that is actually set after the cancelation can differ from the expected state, for example, when a new request or vote is generated at the same time or immediately after you canceled the request.

---

**Moving resources**

You can move resources from one node to another. This allows you to reduce the workload of a node or to perform maintenance on a node.

One move option is available:

- **Move away from ...**: Moves the resource group members away from one or more nodes that the resource group members are currently running on.

The System Automation operations console always checks whether the move function is available. The move function is available for top-level resource groups that contain only floating resources (move groups). The **Move** button is only displayed if the move function is available.

You have two methods to move resources away from a node in the System Automation operations console:
Moving resources using the General panel

To move a resource away from a node in the General panel, perform the following steps:

1. In the System Automation operations console select the resource group in the resource table.

2. Select the General panel.

The Move button appears only if the related automation domain runs System Automation for Multiplatforms 3.1.0.2 or higher, and a top-level resource group is selected that contains only floating resources (move groups). In the General panel the move function is available and the resource group is online.

3. Click Move. A panel similar to the Location info panel appears:
The panel displays:

- All non-group resources contained in the currently selected group and the
  nodes that they are located on.
- The state of the resources (Online or Offline).
- The node that the resources are currently running on by highlighting it. In the Move resource panel the resources are currently running on the node TMCC09.

**Note:** The number of lines displayed in the Move resource panel is dependent on the settings you made in Operations console > Menu > Preferences > View.

4. In the example shown in step 3 on page 138 you have the following option:
   - Move away from TMCC09

5. Click Move. A panel appears where you can type a comment. This comment appears in the log file of the domain.

6. Click Submit to submit the command. The system performs the move. While the move is being processed, the Move resource panel is updated with the current status and location information of the resources that are involved.

**Result:** The resource has been moved to another node.

### Moving resources using the Location info panel

To move a resource away from a node using the Location info panel, perform the following steps:

1. In the System Automation operations console select the resource group.

2. In the information area select the Location info panel. The Location info panel appears:
The panel displays:

- All non-group resources contained in the currently selected group and the nodes that they are located on
- The state of the resources (Online or Offline)
- The node that the resources are currently running on by highlighting it. In the Location info panel the resources are currently running on the node TMCC09.

**Note:** The number of lines displayed in the Location info panel is dependent on the settings you made in Operations console > Menu > Preferences > View.

3. In the example shown in step 2 on page 139 you have the following option:
   - **Move away from TMCC09**

4. Click **Move**. A panel appears where you can type a comment. This comment appears in the log file of the domain.

5. Click **Submit** to submit the command. The system performs the move. While the move is being processed, the Location info panel is updated with the current status and location information of the resources that are involved.

**Result:** The resource has been moved to another node.
Managing your user credentials

Storing your user credentials in the credential vault

If a domain joins the operations console for the first time, a yellow warning symbol appears in the Status column of the topology tree:

Click the domain to open the Automation domain authentication page.

Enter a user ID that is valid for the domain. The user ID should be authorized to perform operations on resources in the first-level automation domain that the operations console supports, for example, requesting an automated resource to be online or excluding a node from automation.

If you leave the check box on the page selected, your user credentials for the domain are saved to the credential vault and the user ID will not be required on further attempts to access the domain.

Changing and deleting your user credentials

Perform the following steps to manage your user credentials that are stored in the credential vault:

1. Click Tivoli System Automation for Multiplatforms > Settings > Stored domain credentials in the navigation tree.
2. On the Stored domain credentials page, you have the following options:
   - To change your user credentials for a domain, select the domain from the
     Credentials in credential vault table and click **Edit** to bring up the First-level
     automation domain authentication page.
   - To delete the user credentials for a specific domain from the credential vault,
     select the domain and click **Delete**.
   - To delete your user credentials for all first-level automation domains from
     the credential vault, click **Delete all**.
Creating and modifying automation policies using the policy editor

This section provides an introduction to the graphical policy editor, which you can use to maintain your automation policies.

Understanding the layout of the policy editor

This section describes the layout of the policy editor's user interface. For detailed step-by-step instructions on using the policy editor, see the online help, which is accessible through the Help link in the console toolbar.

The main panel of the policy editor is divided into several areas, as shown in Figure 9:

1. **Policy type**
   The policy type indicates whether the currently loaded topology represents a System Automation Application Manager end-to-end automation policy, or a System Automation for Multiplatforms first-level automation policy.

2. **Common menu bar**
   Use the buttons on the common menu bar for frequently used actions. For example, you can zoom in and out of the topology viewer, and you can pan to adjacent areas of the currently loaded policy.

3. **Filter bar**
   Use the filter bar to apply a variety of filters to the entities displayed in the
topology viewer. You can specify a partial text string (for example, the filter "WAS" will only display resources that contain the string "WAS").

4 Hint area
When performing certain editing tasks, additional selection boxes and buttons appear below the filter bar to help you complete the task. For example, when adding a member to a group, a Select member selection box appears showing the names of the elements you can add to the group. You can make a selection and then click Add member to complete the task.

5 Overview panel
Use the overview panel to identify which part of the policy is currently being displayed in the topology viewer. This is useful when you have zoomed in to a small region of a policy. You can also use the overview panel to pan to a different area of the policy. See “About the overview panel” on page 145.

6 Topology viewer
Use the topology viewer to edit the currently loaded policy. You can create new policy elements and relationships, add members to groups, remove elements, and so on. See “About the topology viewer” on page 144.

7 Property tabs
Use the property tabs to assign property values to the currently selected policy element. Click the Apply button to apply the changes you have made on a particular property tab.

8 Error messages panel
Use the error messages panel to troubleshoot errors in the currently loaded policy. A policy is checked for semantic and syntactical errors when you load it into the policy editor, and whenever you change and then apply properties of elements.

About the topology viewer
The topology viewer displays a graphical representation of the currently loaded automation policy. An icon represents each element in the policy. Group memberships are represented by blue connector lines (an arrow head points from the containing group to the group member). Relationships between elements are represented by pink connector lines (an arrow head points from the source resource towards the target of the relationship).

Resizing: Use the Resize Viewport icon (located in the bottom right corner of the topology viewer) to change the size of the topology viewer. Left-click the icon, hold down the mouse button, and drag the pointer to the required position. For example, to reduce the size of the topology viewer, drag the pointer towards the top left corner.

Panning: Click the triangular symbols around the perimeter of the topology viewer to pan in the given direction. For example, to view the adjacent area further to the right, click Pan East. Alternatively, click Pan Mouse Mode in the common menu bar if you want to drag an adjacent area into view. For example, if you left-click and hold down the mouse button near the centre of the view and then drag the pointer towards the top right corner, you bring the bottom left area more into view.

Zooming: Click the Zoom In or Zoom Out icon on the common menu to zoom in or out of the topology viewer. Click Fit Contents to Viewport to resize
the policy so that it fits in the topology viewer. Alternatively, click Zoom Mouse Mode in the common menu bar if you want to zoom into a defined area. To define the area, hold down the left mouse button on one corner of the area, then drag the mouse pointer to the opposite corner.

**Limiting the scope of the topology viewer:** By default, all policy elements are displayed in the topology viewer. To display a subset of elements, you can use the filter bar to apply a filter. You can also show or hide relationships, group memberships, constituent resources, persistent attributes, and referenced resources using commands on the View menu. For fixed resources like IBM.ServiceIP, associated persistent attributes are automatically created with an automatically generated name. Persistent attributes are hidden by default; you can show them using the View -> Show persistent attributes command. The removal of move groups and fixed resources intentionally does not remove associated persistent attributes in order to prevent potential problems caused when the persistent attribute is shared among various resources. Persistent attributes are automatically displayed if you select the command Create New -> Persistent attributes.

**About the overview panel**
The overview panel displays a scaled view of the entire graphical representation of the policy. If you have zoomed in to a small area of the policy in the topology viewer, the overview panel displays this area as a blue rectangle. You can pan to a different area of the policy by dragging this rectangle to the required location in the overview panel.

**Creating a policy**
In the scenario below, you learn how to use the policy editor to define an automation policy. The cluster, nodes, and resources in this example are based on those described in Chapter 2, “Getting started,” on page 11, which explains how to define them using RSCT and System Automation for Multiplatforms commands. The scenario is broken down into the following steps:

- Stage 1: Select the policy type and domain
- Stage 2: Create application resource apache1
- Stage 3: Create IP address resource apache1IP
- Stage 4: Create a resource group
- Stage 5: Create a dynamic equivalency
- Stage 6: Define relationships
- Stage 7: Save the policy

**Stage 1: Select the policy type and domain**
1. From the Integrated Solutions Console menu, select Tivoli System Automation > SA Policy Management > Create a new policy:
Stage 2: Create application resource apache1

"Defining RSCT resources for a Web server" on page 19 describes how to create an application resource named “apache1” from a definition file. To create the same resource using the policy editor, follow these steps:

1. Right-click in the topology viewer, then select Create new > Move group > IBM.Application:
A default move group is displayed in the topology viewer. Default properties are displayed in the Properties panel:
2. On the General property tab, replace the default name of the move group with the correct name ("apache1"), and specify the nodes to which it applies ("node01", "node02", and "node03"). In the case of a disconnected domain, you need to build a list of nodes in the Nodes with Constituents panel. To add a node to the list, type its name in the field provided and click the **Apply node** button. The figure below shows node03 being added to the list:
3. On the Class Attributes property tab, enter class attribute details in the fields provided, then click the **Apply** button:

**Stage 3: Create IP address resource apache1IP**

“Defining RSCT resources for a Web server” on page 19 describes how to create an IP address resource named “apache1IP” using command line parameters. To create the same resource using the policy editor, follow these steps:
1. Right-click in the topology view, then select **Create new > Move group > IBM.Service.IP**:

![Topology view screenshot](image)

2. Enter attribute details on the General property tab, then click the **Apply** button.
3. Enter attribute details on the Class Attributes property tab, then click the **Apply** button:
Stage 4: Create a resource group

"Creating a resource group for apache1 and apache1IP" on page 26 describes how to add resources.apache1 and apache1IP to a resource group named "apacherg". To do the same thing using the policy editor, follow these steps:
1. Right-click in the topology viewer, then select **Create new > Resource group**:

![Create new Resource group](image1)

2. To add apache1 to the group, right-click the group, then select **Add member**:

![Add member](image2)
3. In the hint area, select apache1 from the **Select member** drop-down list, then click the **Add member** button:

A connector line joining apacherg to apache1 can now be seen in the topology viewer:
4. To add apache1IP to the group, right-click the resource group, then select **Add member**. Instead of using the hint area, you can specify the member by clicking it in the topology viewer:

Connector lines joining apacherg to both apache1 and apache1IP can now be seen in the topology viewer:

**Stage 5: Create a dynamic equivalency**

"Creating an equivalency definition for network adapters" on page 25 describes the command line method of creating a dynamic equivalency to specify the network adapters that can be used to carry the apache1IP address. To create the same dynamic equivalency using the policy editor, follow these steps:
1. Right-click in the topology viewer, then select **Create new > Equivalency**:

![Topology Viewer](image1)

2. In the properties tab, enter equivalency attributes, then click the **Apply** button:

![Properties Tab](image2)
Stage 6: Define relationships

“Defining relationships between the Web server resources” on page 26 describes the command line method of defining relationships between apache1 and apache1IP, and between apache1IP and the netequ equivalency. To define the same relationships using the policy editor, follow these steps:

1. To define the relationship between resources apache1 and apache1IP, right-click the apache1 move group in the topology viewer, then select Create new relationship:

2. In the hint area, select apache1IP from the Select target drop-down list:
3. Click the **OK** button. A relationship connector line joining apache1 to apache1IP can now be seen in the topology viewer:

4. In the Properties panel, replace the default name for the relationship with the correct name ("apache1_dependson_ip1"), select the correct condition (**DependsOn**) from the **Type** drop-down list, then click the **Apply** button:
5. To define the relationship between apache1IP and the netequ equivalency, right-click the apache1IP move group in the topology viewer, then select **Create new relationship**:

A relationship connector line joining apache1IP to netequ can now be seen in the topology viewer:
6. In the Properties panel, replace the default name for the relationship with the correct name ("apache1IP_dependson_netequ”), select the correct condition (DependsOn) from the Type drop-down list, then click the Apply button:

Stage 7: Save the policy
You can save the policy to a local file or to a policy pool.

- To save the policy to a local file:
  1. From the File menu, select Save to local file.
     The File Download dialog box is displayed.
  2. In the File Download dialog box, click Save.
  3. Browse to a folder and specify a file name for the policy file, then click Save.
  4. In the Download Complete dialog box, click Close.
- To save the policy to a policy pool:
  1. From the File menu, select Save to policy pool.
  2. Select the required domain from the Domain drop-down list. Only editable domains are available for selection.
  3. In the Policy file name field, enter a file name for the policy, then click Save.
Chapter 9. Protecting your resources – quorum support

This chapter describes how System Automation for Multiplatforms protects your resources by using configuration and operational quorum.

Overview

A cluster (also known as peer domain) may split into two or more subclusters in case no more communication is possible between the elements in the cluster. Since subclusters are not aware of one another, it may occur that System Automation for Multiplatforms starts a new instance of an application that is already running in one of the other subclusters. For example, if the application requires access to a shared disk to perform failure recovery, data corruption may occur due to simultaneous access to the disk.

Such resources are characterized as critical. A critical resource is a resource that may not be running on more than one node at any point in time. If such a resource is active on two or more separated nodes, then the data integrity of the cluster is endangered. In order to protect such critical resources, System Automation for Multiplatforms ensures that only one of the subclusters survives while the others are dissolved. Thus System Automation for Multiplatforms prevents data corruption that is caused by system failures or network partitioning.

![Figure 10. Quorum – majority of nodes](image)

If a cluster splits into two or more subclusters, the configuration resource manager (ConfigRM) determines which of the subclusters has the majority of nodes. The majority is given when the subcluster has more than half of all defined nodes in the cluster. The subcluster with the majority of nodes will have the operational quorum. It will survive and become the active cluster, while the other subcluster(s) will be dissolved.

The protection of critical resources is achieved by...
Quorum support

- Configuration quorum
- Operational quorum

Configuration quorum

Configuration quorum determines when configuration changes in the cluster will be accepted. The integrity of the cluster definition is ensured by following the majority rule. Operations affecting the configuration of the cluster are only allowed when \( n/2+1 \) nodes are active, where \( n \) is the number of nodes defined in the cluster. However, for some operations, the majority rule can be overridden or different configuration quorum rules apply:

- You can remove nodes using the `rmrpnode` command if exactly half of the nodes is online and if the configuration can be successfully removed from at least one of the offline nodes. You can also use the `-f` option of this command to override the majority rule.
- The quorum rule for the `startrpdomain` command is \( n/2 \), but you can override it with the all nodes (`-A`) option or the local node (`-L`) option.

**Note:** In a tie situation you can start and stop resource groups using the command `chrg -o online/offline group_name`. For more details, refer to RSCT documentation. (See “Related information” on page xiii.)

Operational quorum

Operational quorum is used to decide whether resources can be safely activated without creating conflicts with other resources. Operational quorum is determined based on the number of online nodes and a tie breaker to resolve certain tie situations. A subcluster has operational quorum if it has more than half of the nodes active.

If operational quorum exists, System Automation for Multiplatforms can manipulate resources or resource groups and bring them online. If no quorum exists, System Automation for Multiplatforms cannot take any action on a resource.

Critical resource protection methods on AIX, Solaris, and Linux

If critical resources are active in an AIX, Solaris, or Linux subcluster that has lost quorum, the ConfigRM uses the "CritRsrcProtMethod" attribute of each node in the subcluster to determine in which way the system should be terminated. The protection methods are based on immediate system shutdown by means of kernel panic simulation. There are six protection methods:

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard reset and reboot operating system (default)</td>
<td>1</td>
</tr>
<tr>
<td>Halt operating system</td>
<td>2</td>
</tr>
<tr>
<td>Hard reset and reboot operating system with Sync</td>
<td>3</td>
</tr>
<tr>
<td>Halt with Sync</td>
<td>4</td>
</tr>
<tr>
<td>No protection. System continues operating</td>
<td>5</td>
</tr>
<tr>
<td>Exit and restart RSCT subsystems</td>
<td>6</td>
</tr>
</tbody>
</table>
Quorum support

A protection method with sync sends file system buffers to disc before the operating system is halted or reset. Thus the probability of data loss or data inconsistency on the file system is reduced. Note that protection methods with sync might also be an unsafe solution in certain situations. There might be a chance that the data is already being accessed during file system flush from another application constituent that has just started in the subcluster which gained operational quorum. It must be examined if this is likely to occur in a given system and application environment.

Regardless of the protection method being used, it is highly recommended to use a journaling file system. This can prevent the corruption of the file system itself and greatly enhances file system recovery after a system reset.

The protection methods on the nodes may be different. However, typically the same protection method is set for each node in the entire cluster.

Critical resource protection methods on Windows

To specify how the system is to react upon a complete system failure (BSOD) if critical resources are active in a Windows subcluster that has lost quorum, you use the Windows Startup and Recovery Dialog. For more information, see “Operational quorum values” on page 263.

Types of tie breakers

In the case of a tie in which the cluster is partitioned into subclusters with an equal number of nodes, the configuration resource manager uses a tie breaker to determine which subcluster has operational quorum. A subcluster will have operational quorum if it has the tie breaker reserved.

The following tie breaker types exist:

<table>
<thead>
<tr>
<th>Tie breaker</th>
<th>Operating system or platform</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operator</td>
<td>AIX, Solaris, Linux, Windows</td>
<td>This tie breaker asks the operator or administrator for a decision. Until the administrator explicitly breaks the tie, neither subcluster will have an operational quorum. The operational quorum state is set to “PendingQuorum” and stays in this state until either the network is repaired, failing nodes are repaired and brought online, or the operator determines which subcluster is the winning one and which is the losing one. This is done by invoking the “ResolveOpQuorumTie” action on a node of each active subcluster.</td>
</tr>
<tr>
<td>Fail</td>
<td>AIX, Solaris, Linux, Windows</td>
<td>this actually is a pseudo tie breaker which means that it does not resolve the tie. Neither subcluster will have operational quorum.</td>
</tr>
<tr>
<td>SCSI</td>
<td>Solaris, Linux on System i, Linux on System p, and Linux on System x</td>
<td>This tie breaker assumes that an SCSI disk is shared by all nodes of the cluster. Tie breaker reservation is done by the SCSI reserve or persistent reserve command.</td>
</tr>
<tr>
<td>ECKD</td>
<td>Linux on System z</td>
<td>This tie breaker type assumes that an ECKD disk is shared by all nodes of the cluster. Exclusive access to the disk is done by means of the ECKD reserve command.</td>
</tr>
</tbody>
</table>
### Table 21. Types of tie breakers (continued)

<table>
<thead>
<tr>
<th>Tie breaker</th>
<th>Operating system or platform</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISK</td>
<td>AIX</td>
<td>This tie breaker type enables you to specify an SCSI or SCSI-like physical disk using an AIX device name, and assumes that the SCSI disk is shared by one or more nodes of the cluster. Tie breaker reservation is done by the SCSI reserve or persistent reserve command. If creating a tie breaker of this type, you need to set the DeviceInfo persistent resource attribute to identify the physical disk. Only SCSI and SCSI-like physical disks are supported. Physical disks attached via Fiber Channel, iSCSI and Serial Storage Architecture Connections are suitable.</td>
</tr>
<tr>
<td>EXEC</td>
<td>AIX, Solaris, Linux, Windows</td>
<td>This generic tie breaker implementation calls a custom executable for tie breaker operations. The network tie breaker &quot;samtb_net&quot; is shipped with this product and is implemented as an EXEC tie breaker executable. See &quot;Network tie breaker&quot; on page 175 for details on how to configure the EXEC tie breaker to use the &quot;samtb_net&quot; executable.</td>
</tr>
</tbody>
</table>

If a cluster that consists of an odd number of nodes is inadvertently partitioned into subclusters, the subcluster that has more than half of the nodes available has quorum. For example, on a three node cluster, the subcluster in which two nodes available has quorum. The remaining single-node subcluster gets neither operational nor configuration quorum, which means that no resources will be started on that node. If a critical resource is already running on that node, the critical resource protection method defined will be applied to the node (see “Critical resource protection methods on AIX, Solaris, and Linux” on page 162 and "Critical resource protection methods on Windows” on page 163).

If you have an even number of nodes in the cluster and one of the subclusters consists of half of the nodes in case of a cluster split, then a tie breaker decides which subcluster is allowed to run critical resources. Nodes with critical resources that lose the competition for the tie breaker are subject to resource protection, which means that they will be stopped or rebooted immediately.

To perform an automatic tie breaking without operator action, you need a disk tie breaker (SCSI for Solaris, Linux on System x, Power, or ECKD for Linux on System z, or DISK for AIX). A disk tie breaker is a shared disk which is accessible from all cluster nodes.

**SLES 10 x/Linux systems:** During startup, RSCT removes watchdog modules like i8xx_tco and 16300esb, which may be preloaded in order to be able to load the required softdog module. This means that you cannot have a hardware-related watchdog running in parallel to System Automation for Multiplatforms.

#### VMTIMEBOMB function

The vmtimebomb function which is running for Linux on System z under VM ensures data protection for scenarios where z/VM gets held, but guest Linux systems keep on running. It is a new implementation of the protection methods described in the preceding section. Seen from the perspective of Topology Services,
the vmtimebomb function is only indirectly accessed through a special VM
vmwatchdog module. This module is similar to the ubiquitous Linux softdog
watchdog module.

Watchdog modules are intended to prevent a permanent failure by automatically
restarting the operating system during some sort of catastrophic failure. Typically,
an application actively and regularly 'pings' the watchdog (usually by writing to a
special device) to indicate that it is alive. This serves as an indication of the health
of the system as a whole. Should the application fail to indicate it is alive, this
indicates that there is a serious malfunction.

Something must monitor the status of the watchdog to take action when the
application ceases to give the 'all's well' report. With softdog, this is just the Linux
operating system. In particularly dire circumstances, the operating system may fail
and not be able to respond. With the vmwatchdog, external supervision is handled
by the z/VM Control Program, which can restart the operating system on the
virtual machine even in the worst cases. The vmwatchdog is only supported for 2.6
kernel versions running as guest systems under z/VM 5.1.0 (or later).

Setting critical resources

Use the ProtectionMode persistent attribute to specify whether the resource is
critical. If it is critical, then the Configuration RM (IBM.ConfigRM) decides
whether the resource can be started as requested. The attribute may have the
integer values 0 (non–critical) or 1 (critical). Per default IBM.Application resources
are non–critical, and IBM.ServiceIP resources are critical. If the resource is set to
critical, monitoring will immediately start.

Issue the RSCT command lsrsrc to list the value of the ProtectionMode attribute:

```
lsrsrc IBM.Application Name NodeNameList ProtectionMode
```

Issue the RSCT command chrsrc to define the resource as critical by setting the
ProtectionMode to 1:

```
chrsrc -s "Name='apache1'" IBM.Application ProtectionMode=1
```

To define a resource as non–critical set ProtectionMode to 0:

```
chrsrc -s "Name='apache1'" IBM.Application ProtectionMode=0
```

Issue the following to verify whether critical resources are currently active on a
node for resource class IBM.Application:

```
lsrsrc IBM.Application Name NodeNameList OpState ProtectionMode
```

This provides the following output:

```
resource 1:
    Name = "apache1"
    NodeNameList = {"node1","node2"}
    OpState = 1
    ProtectionMode = 1
resource 2:
    Name = "apache1"
    NodeNameList = {"node1"}
    OpState = 2
    ProtectionMode = 1
resource 3:
    Name = "apache1"
    NodeNameList = {"node2"}
    OpState = 1
    ProtectionMode = 1
```
Critical resource apache1 is active on the node2.

Issue the following to verify if critical resources are currently active on the nodes:

```
lsrsrc IBM.PeerNode Name CritRsrcActive
```

The output is as follows:

```
Resource Persistent and Dynamic Attributes for IBM.PeerNode
resource 1:
    Name    = "node1"
    CritRsrcActive = 0
resource 2:
    Name    = "node2"
    CritRsrcActive = 1
```

Critical resources are active on node2.

**Getting quorum information**

Use the `lsrsrc` command for the IBM.RecoveryRM daemon to obtain the current quorum states.

```
node02:~/build # lsrsrc -ls IBM.RecoveryRM
```

You get the following output:

```
Daemon State:
    My Node Name    : node02
    Master Node Name: node01 (node number = 1)
    Our CVN         : 61035379498
    Total Node Count: 2
    Joined Member Count: 2
    Config Quorum Count: 2
    Startup Quorum Count: 1
    Operational Quorum State : HAS_QUORUM
    In Config Quorum : TRUE
    In Config State : TRUE
```

The meaning of the various attributes is as follows:

**Total Node count**

Is the number of nodes defined in the cluster.

**Joined Member Count**

is the number of IBM.RecoveryRM daemons running in the cluster. This is equivalent to the number of active nodes in the (sub)cluster.

**Config Quorum Count**

is the number of IBM.RecoveryRM daemons that must be active in order to make a configuration change by means of the System Automation for Multiplatforms commands.

**Startup Quorum Count**

Is the number of IBM.RecoveryRM daemons that must be active before the System Automation for Multiplatforms automation engine is activated.

**Operational Quorum State**

Indicates (sub)cluster wide whether this subcluster can survive or must immediately dissolve in case critical resources are running on the node(s) in the subcluster. The operational quorum state is provided by the dynamic attribute OpQuorumState of PeerDomain class. OpQuorumState can have the following values:
Quorum support

0 – HAS_QUORUM
System Automation for Multiplatforms may start resources.

1 – PENDING_QUORUM
Indicates that a tie situation occurred that is not yet resolved. System Automation for Multiplatforms does not start resources.

2 – NO_QUORUM
System Automation for Multiplatforms is not allowed to start resources.

In Config Quorum
Indicates whether enough nodes hosting IBM.RecoveryRM daemons are active to accept configuration changes by System Automation for Multiplatforms commands. Shows TRUE if the total number of "joined" IBM.RecoveryRM daemon group members within the cluster is equal or above the Config Quorum count.

In Config State
Indicates whether the master IBM.RecoveryRM daemon has completed the verification of the system registry content at startup time. If the state equals to FALSE, any System Automation for Multiplatforms command will be rejected.

Enter the following to list OpQuorumState:

`lsrsr IBM.PeerDomain Name OpQuorumState`

You get the following output:

```
Resource Persistent and Dynamic Attributes for:IBM.PeerDomain
resource 1:
   Name = "myCluster"
   OpQuorumState = 0
```

Setting up and administering a tie breaker
The IBM.TieBreaker resource class allows you to configure a tie breaker such as ECKD or SCSI. Additionally, two tie breakers are predefined, Operator and Fail. The operator tie breaker provides an undetermined result when a tie occurs and it is left to the administrator to resolve the tie through granting or denying the operational quorum. When a tie occurs and a tie breaker of type "Fail" is active, the attempt to reserve the tie breaker is always denied. Default tie breaker type is set to 'Operator'.

To list the available tie breaker type:

`lsrsr -c IBM.TieBreaker`

You get the following output on a Linux system running on System x, System p, or System i:

```
Resource Class Persistent Attributes for: IBM.TieBreaker
resource 1:
   AvailableTypes ={"SCSI",""},{"EXEC",""},{"Operator",""},{"Fail",""}
```

To list the tie breaker name:

`lsrsr IBM.TieBreaker`

You get the following output:
Quorum support

Although you can define several tie breaker resources in the resource class IBM.TieBreaker, only one of them can be active in the cluster at the same time. Issue the following command to list the tie breaker that is currently active in the cluster:

```
lsrsrc -c IBM.PeerNode OpQuorumTieBreaker
```

You get the following output:

```
Resource Class Persistent Attributes for: IBM.TieBreaker
resource 1:
    Name           = "FAIL"
    Type           = "FAIL"
    DeviceInfo     = ""
    ReprobeData    = ""
    ReleaseRetryPeriod = 0
    HeartbeatPeriod = 0
    PreReserveWaitTime = 0
    PostReserveWaitTime = 0
    NodeInfo       = {}

resource 2:
    Name           = "Operator"
    Type           = "Operator"
    DeviceInfo     = ""
    ReprobeData    = ""
    ReleaseRetryPeriod = 0
    HeartbeatPeriod = 0
    PreReserveWaitTime = 0
    PostReserveWaitTime = 0
    NodeInfo       = {}

resource 3:
    Name           = "myTieBreaker"
    Type           = "SCSI"
    DeviceInfo     = "ID=0 LUN=0 CHAN=0 HOST=2"
    ReprobeData    = ""
    ReleaseRetryPeriod = 0
    HeartbeatPeriod = 5
    PreReserveWaitTime = 0
    PostReserveWaitTime = 0
    NodeInfo       = {}

resource 4:
    Name           = "mytb"
    Type           = "EXEC"
    DeviceInfo     = "PATHNAME=/usr/sbin/rsct/bin/samtb_net
                      Address=192.168.177.2"
    ReprobeData    = ""
    ReleaseRetryPeriod = 0
    HeartbeatPeriod = 30
    PreReserveWaitTime = 0
    PostReserveWaitTime = 30
    NodeInfo       = {}
    ActivePeerDomain = "21"
```

To grant/deny the operational quorum when tie breaker is "Operator":
```
runcmd -c IBM.PeerDomain ResolveOpQuorumTie Ownership=1 (0 to deny)
```
Quorum support

Note: In order to avoid race conditions, the operator tie breaker must be denied first for the losing subcluster(s) before granting it to the subcluster which is supposed to continue.

Using a tie breaker

To create a basic setup for the tie breaker you need a cluster of two (or other even number) of nodes. Also you need a disk that is shared between all nodes of the cluster. The tie breaker disk is shared between all cluster nodes.

Attention: When defining tie breaker resources, be aware that the disk on which IBM.TieBreaker resources are stored should not also be used to store file systems.

The following three examples show how to use a tie breaker with an ECKD, SCSI or DISK device. Note that the tie breaker needs not be formatted or partitioned. Then it will only be marked active without size information (in case of ECKD).

Example 1: ECKD tie breaker setup for a two-node cluster

Note the following when defining the tie breaker disk under VM:

- Full pack minidisk should be defined.
- If minidisk cache is used, its value should be set to off.
- ECKD disk is shared between both nodes.

The ECKD tie breaker type is specific for Linux on System z. If you want to create an ECKD tie breaker object, you need to set the DeviceInfo persistent resource attribute to indicate the ECKD device number. This type of tie breaker uses a reserve/release mechanism and needs to be re-reserved periodically to hold the reservation. For this reason, we strongly recommend that you also specify the HeartbeatPeriod persistent resource attribute when creating a tie breaker of this type. The HeartbeatPeriod persistent resource attribute defines the interval at which the reservation request is re-issued.

Collect the following system information (Linux kernel 2.4):

```
node01:~ # cat /proc/subchannels
Device sch. Dev Type/Model CU in use  PIM PAM POM CHPIDs
---------------------------------------------------------------------
50DE 0A6F 3390/0A 3990/E9 F0 A0 FF 7475E6E7 FFFFFFFF
node01:~ # cat /proc/dasd/devices
50dc(ECKD) at ( 94: 0) is : active at blocksize: 4096, 601020 blocks, 2347 MB
50dd(ECKD) at ( 94: 4) is : active at blocksize: 4096, 601020 blocks, 2347 MB
50de(ECKD) at ( 94: 8) is : active at blocksize: 4096, 601020 blocks, 2347 MB
50df(ECKD) at ( 94: 12) is : active at blocksize: 4096, 601020 blocks, 2347 MB
```

For Linux kernel 2.6 use the `lsccs` command instead of the `cat /proc/subchannels` command.

Perform the following steps to use the tie breaker:

1. Create a tie breaker resource object in IBM.TieBreaker class. `DeviceInfo` shows the ECKD device number. It can be obtained from `/proc/dasd/devices` file.

```
node01:~ # mkrsrc IBM.TieBreaker Name=myTieBreaker Type=ECKD DeviceInfo="ID=50de" HeartbeatPeriod=5
node01:~ # lsrsrc IBM.TieBreaker
Resource Persistent Attributes for: IBM.TieBreaker
resource 1:
    Name = "Operator"
    Type = "Operator"
    DeviceInfo = "ID=50de"
    ReprobeData = ""
    ReleaseRetryPeriod = 0
    HeartbeatPeriod = 0
    PreReserveWaitTime = 0
```
Quorum support

PostReserveWaitTime = 0
NodeInfo = {}

resource 2:
Name = "Fail"
Type = "Fail"
DeviceInfo = ""
ReprobeData = ""
ReleaseRetryPeriod = 0
HeartbeatPeriod = 0
PreReserveWaitTime = 0
PostReserveWaitTime = 0
NodeInfo = {}

resource 3:
Name = "myTieBreaker"
Type = "ECKD"
DeviceInfo = "ID=50de"
ReprobeData = ""
ReleaseRetryPeriod = 0
HeartbeatPeriod = 5
PreReserveWaitTime = 0
PostReserveWaitTime = 0
NodeInfo = {}

2. Change OpQuorumTieBreaker attribute in IBM.PeerNode class to one of the tie breaker resource objects.

```
node01:~ # chrsrc -c IBM.PeerNode OpQuorumTieBreaker="myTieBreaker"
node01:~ # lsrcrc -c IBM.PeerNode
```

Hints: If one node of a two-node cluster is rebooted, the rebooting node may come back extremely quickly. This may disrupt the tie breaking method and cause an undesired reboot of the remaining node. If a node that belongs to a cluster must be manually rebooted, use `halt -nf` instead of `reboot -nf`.

If the node reserving a tie breaker is down and cannot be rebooted, manual access on another node is needed to break the reservation and take it over on the other node. The tie breaker disk can

- be either still attached to the healthy node, provided this node has not been rebooted in the mean time:

```
node01:~ # cat /proc/subchannels
Device sch. Dev Type/Model CU in use PIM PAM POM CHPIDs
--------------------------------------------------------------
50DE 0A6F 3390/0A 3990/E9 F0 A0 FF 7475E6E7 FFFFFFFF
```

```
node01:~ # cat /proc/dasd/devices
50de(ECKD) at ( 94: 8) is dasdc: active at blocksize: 4096,601020 blocks, 2347 MB
```

- be boxed, if this node has been rebooted and cannot recognize the tie breaker disk anymore:

```
node01:~ # cat /proc/subchannels
Device sch. Dev Type/Model CU in use PIM PAM POM CHPIDs
--------------------------------------------------------------
50DE 0A6F  FFFF/00 F0 A0 FF 7475E6E7 FFFFFFFF
```

```
node01:~ # cat /proc/dasd/devices
50de(ECKD) at ( 94: 8) is dasdc : boxed
```
To break the tie breaker disk reservation enter the command
/usr/sbin/rsct/bin/tb_break:
`tb_break -t ECKD /dev/dasdc`

The tie breaker disk should now be reserved by the healthy node.

**Note:** If the `tb_brk` command does not work the first time, issue it again.

**Example 2: SCSI tie breaker setup for a two-node cluster**

The SCSI tie breaker type is specific for Solaris, and Linux on System x, pSeries, and iSeries. If you want to create an SCSI tie breaker object, you need to specify the SCSI device using the `DeviceInfo` persistent resource attribute. If the SCSI configuration is different on different nodes in the cluster, you can also use the `NodeInfo` persistent resource attribute to reflect those differences. This type of tie breaker uses a reserve/release mechanism and needs to be re-reserved periodically to hold the reservation. For this reason, we strongly recommend that you also specify the `HeartbeatPeriod` persistent resource attribute when creating a tie breaker of this type. The `HeartbeatPeriod` persistent resource attribute defines the interval at which the reservation request is re-issued.

**Linux:** SCSI devices on Linux can be identified by four integer values for the attributes HOST, CHAN, ID, and LUN:
`node1:="# dmesg | grep "Attached scsi disk"`

Normally these parameters are identical on each cluster node. For example, for node1 and node2 these are HOST=0 CHAN=0 ID=4 LUN=0. You can then create the tie breaker object:
`mkrsrc IBM.TieBreaker Name=myTieBreaker Type=SCSI DeviceInfo=" HOST=0 CHAN=0 ID=4 LUN=0"`

The four values above may also be different for different nodes (even if the target device is same). In that case the `NodeInfo` field should be used.

Use the four integer values from the command output:
`# dmesg | grep "Attached scsi disk"
Attached scsi disk sdf at scsi2, channel 2, id 4, lun 0`

For disk sdf is HOST=2, CHAN=2, ID=4, LUN=0. For example, a SCSI device is connected to 2 nodes named node1 and node2 and has the following SCSI identifiers:
`node1: HOST=0 CHAN=0 ID=4 LUN=0
node2: HOST=2 CHAN=2 ID=4 LUN=0`

You can then create the tie breaker object as
`# mkrsrc IBM.TieBreaker Name=scsi Type=SCSI DeviceInfo="ID=4 LUN=0"
NodeInfo='["node1", "HOST=0 CHAN=0"], ["node2", "HOST=2 CHAN=2"]'`

System Automation for Multiplatforms handles `DeviceInfo` and `NodeInfo` in such a way that it merges the two strings, `DeviceInfo` first and then `NodeInfo`. For example, for node1 the merged string is
"ID=4 LUN=0 HOST=0 CHAN=0"

which will then be parsed.

Also, any duplicated keywords will be allowed and the last one will be used. Therefore, the same command can be specified as
`# mkrsrc IBM.TieBreaker Name=myTieBreaker Type=SCSI DeviceInfo="ID=4 LUN=0
HOST=0,CHAN=0" NodeInfo='["node2", "HOST=2 CHAN=2"]'"
Quorum support

This simplification may be useful as often and most likely the SCSI id is the same for many nodes.

Hint: If the node reserving a tie breaker is down and cannot be rebooted, manual access on another node is needed to release the SCSI tie breaker disk. To release a disk, run the command:

tb_break [-f] HOST CHAN ID LUN

for example,

/usr/sbin/rsct/bin/tb_break -f HOST=0 CHAN=0 ID=4 LUN=0

Solaris: SCSI devices on Solaris can be identified by the disk device name.

node1:~# format
Searching for disks...done
AVAILABLE DISK SELECTIONS:
  0. c0t0d0 <DEFAULT cyl 8894 alt 2 hd 255 sec 63>
     /pci@0,0/pci1022,745002/pci1000,306003/sd00,0
  1. c5605076801820000000000000064d0 <DEFAULT cyl 96 alt 2 hd 64 sec 32>
     /scsi_vhci/disk@g60050768018200a88000000000000064

You can then create the tie breaker object as

# mkrsrc IBM.TieBreaker Name=myTieBreaker Type=SCSI
DeviceInfo="DEVICE=c56050768018200a88000000000000064d0" HeartbeatPeriod=5

Hint: If the node reserving a tie breaker is down and cannot be rebooted, manual access on another node is needed to release the SCSI tie breaker disk. To release a disk, run the command:

tb_break -b -t SCSI "DEVICE=<device name>"

for example,

# /usr/sbin/rsct/bin/tb_break -b -t SCSI
"DEVICE=c56050768018200a88000000000000064d0"

Example 3: AIX DISK tie breaker setup for a two-node cluster

The DISK tie breaker type is specific to AIX. If you want to create a DISK tie breaker object, you need to set the DevicelInfo persistent resource attribute to indicate the AIX device name. The AIX device name must specify a SCSI or SCSI-like physical disk that is shared by all nodes of the peer domain.

Physical disks attached via Fiber Channel, iSCSI, and Serial Storage Architecture may serve as a DISK tie breaker. However, IDE hard disks do not support the SCSI protocol and cannot serve as a DISK tie breaker. Logical volumes also cannot serve as a DISK tie breaker. This type of tie breaker uses a reserve/release mechanism and needs to be re-reserved periodically to hold the reservation. For this reason, we strongly recommend that you also specify the HeartbeatPeriod persistent resource attribute when creating a tie breaker of this type. The HeartbeatPeriod persistent resource attribute defines the interval at which the reservation request is re-issued.

To print every known physical volume in the system along with its physical disk name, enter the lsvg command:
lsvg

An output similar to the following one is displayed:

hdisk0 000000371e5766b8 rootvg active
hdisk1 000006968340ead54 None
Quorum support

In order to verify that a disk is a SCSI or SCSI-like disk and so a suitable candidate for a DISK tie breaker, use the lsdev command. For example:

lsdev -C -l hdisk1

An output similar to the following one is displayed:
hdisk1 Available 10-60-00-0,0 16 Bit SCSI Disk Drive

In order to serve as a tie breaker disk, the disk must be shared by all nodes of the peer domain. Check the physical volume ID returned by the lspv command to determine if the disk is shared between nodes (in the preceding output for the lspv command, the physical volume ID is listed in the second column; the volume ID for hdisk1 is 000069683404ed54.) Be aware, however, that AIX remembers all disks that have been attached to the system, and the disks listed by the lspv command may no longer be attached. If such a disk was moved to another machine, it might appear as if the disk is shared, when in fact it is no longer attached to the original machine.

The disk on which IBM.TieBreaker resources are stored should not also be used to store file systems. If the nodes of the cluster share more than one disk, it may be difficult to determine which one is the tie breaker disk, and which one is used for application data. The output from the lsdev command shows the SCSI address associated with the disk. (In the preceding output for the lsdev command, the SCSI address is listed in the third column; the SCSI address for hdisk0 is 10-60-00-0,0.) This information will help you to identify the correct disk if you are aware of the address of the disk prior to its installation.

Once you know the device name, you can issue the mkrsrc command:

mkrsrc IBM.TieBreaker Name=myTieBreaker Type=DISK DeviceInfo="DEVICE=/dev/hdisk1" HeartbeatPeriod=5

**Hint:** The tie breaker relies on SCSI-2 reservation, which is not necessarily supported by every combination of storage and driver setup. In order to check that the setup supports SCSI-2 reservation, RSCT comes with the disk_reserve utility that must be invoked with its full path (/usr/sbin/rsct/bin/disk_reserve).

The tie breaker will work correctly if the tie breaker disk can be reserved and unlocked from either node and if the disk cannot be reserved while it is locked by the other node.

**Synopsis:**

**Usage:**


- **-h** - display this help text
- **-v** - verbose
- **-f** - reserve after break (for the -l or -b option)
- **-d sdisk_name** - disk to operate (for example, /dev/sdb)
- **-l** - lock (reserve)
- **-u** - unlock (release)
- **-b** - break
Quorum support

-g sg_device_name (for example, /dev/sg1)

Examples:
/usr/sbin/rsct/bin/disk_reserve -l -f -d /dev/sde
/usr/sbin/rsct/bin/disk_reserve -l -g /dev/sg3

If the node reserving a tie breaker is down and cannot be rebooted, manual access on another node is needed to release the SCSI tie breaker disk. To release the disk, run the command:
/usr/sbin/rsct/bin/tb_break –f –t DISK "DEVICE=/dev/hdisk1"

Execute the lspath command, for example:
lspath -l hdisk2
lspath: 0514-538 Cannot perform the requested function because the specified device does not support multiple paths.

Sample output:
In contrast to the sample above, this sample shows an example that does not support SCSI-2 reservation and hence not the tie breaker.
#lspath -l hdisk2
Enabled hdisk2 fscsi0
Failed hdisk2 fscsi0
Failed hdisk2 fscsi0
Failed hdisk2 fscsi0
Failed hdisk2 fscsi0
Enabled hdisk2 fscsi0
Enabled hdisk2 fscsi0
Enabled hdisk2 fscsi1
Failed hdisk2 fscsi1
Failed hdisk2 fscsi1
Failed hdisk2 fscsi1
Failed hdisk2 fscsi1
Enabled hdisk2 fscsi1
Enabled hdisk2 fscsi1

Testing the suitability of resources as disk tie breakers

Disk tie breakers perform an SCSI-2 reserve against the defined SCSI disk. To be a valid tie breaker, the resource has to meet the following requirements:

- The disk must be accessible from all nodes at the same time (directly attached to each node).
- The SCSI-2 reserve must be able to reserve the disk.
- The SCSI-2 reserve must ‘exclusively’ reserve the disk. A reserve request of any other node will be rejected.

The third requirement may not be applicable to virtual environments like VIO attached disks on System p. Use the following procedure to test whether the exclusive reserve is working for a given disk:

1. Create a 2-node domain and define a disk tie breaker resource. To make sure that no resource is currently online, stop any defined resource and do not create any resources.
2. Log on to the console of both nodes and start a listing of the syslog (for example, tail -f /var/log/messages on a Linux system).
3. Disable the network connection on one of the nodes, for example, unplug the network cable(s) or use the command ifconfig <if> down.
4. Check the output of the syslog listing: after some seconds one of the nodes should show ‘HAS QUORUM’, whereas the other node should show
Quorum support

‘PENDING_QUORUM’. If both nodes show ‘HAS_QUORUM’, then the disk cannot be used, as the reserve is not exclusive. The disk is not able to handle real split brain situations.

If the domain is composed of a larger number of nodes, the test is similar - the network needs to be split in half the number of nodes on both sides, and then the quorum state needs to be determined on one node of either sub cluster.

Storage Area Network (SAN) disks are an example of unsuitable resources to use as disk tie breakers.

Network tie breaker

The network tie breaker provides an alternative to the disk and operator based tie breakers described in the preceding sections of this chapter. It uses an external IP (network instance) to resolve a tie situation.

There may be several reasons to use a network tie breaker, for example:
• There is no possibility to use a disk based tie breaker.
• It is the highest priority of a high availability environment to communicate with instances outside the cluster.

Example: The primary functions of a web server are to deliver web pages to clients outside of the cluster. In order to make this service highly available, the tie breaker should not grant access to a node which is not able to communicate to instances outside of the cluster.

The network tie breaker should only be used for domains where all nodes are in the same IP sub net. Having the nodes in different IP sub nets makes it more likely that both nodes can ping the network tie breaker, while they cannot communicate to each other. Additionally, the default gateway IP address must not be used if it is virtualized by the network infrastructure. Choose an IP address, which can only be reached via a single path from each node in the domain.

Requirements for the network tie breaker

To ensure the network tie breaker functionality the external IP instance must be reachable from all nodes within the highly available cluster. Also the external IP instance must be able to reply to ICMP echo requests (ping). If you install a firewall rule which will blocks ICMP traffic between the cluster nodes and the external IP instance, the network tie breaker will not work. The biggest danger is that the cluster nodes cannot communicate to their peers (cluster split), but both sub clusters are able to reach the external IP instance. Under normal conditions IP ensures the following: If both sub clusters can reach the external gateway, they are also able to communicate to their peers. There might be unusual IP setups which will not ensure this rule (e.g. firewall settings). If you cannot ensure this rule, you cannot use the network tie breaker.

The following table shows the advantages and disadvantages of the two types of tie breakers:
### Quorum support

#### Table 22. Comparison of network-based and disk-based tie breakers

<table>
<thead>
<tr>
<th>Network-based tie breaker</th>
<th>Disk-based tie breaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ No hardware dependency</td>
<td>+ Most secure tie breaker. Hardware ensures that only one instance (node) is able to get the tie breaker.</td>
</tr>
<tr>
<td>+ Evaluates the availability of communication.</td>
<td></td>
</tr>
<tr>
<td>- If the external IP instance is not available in case of a cluster split, no subcluster will get quorum.</td>
<td>- In case of a loss in communication this tie breaker may grant access to a node which is not able to communicate to instances outside of the cluster.</td>
</tr>
<tr>
<td>- There may be error conditions in which a tie situation occurs but more than one node is able to communicate. In this case there is a slight possibility that both subclusters are able to get the tie breaker.</td>
<td></td>
</tr>
</tbody>
</table>

#### Setting up a network tie breaker

The network tie breaker is realized as an RSCT exec tie breaker. See the RSCT documentation if you want to learn more about an exec tie breaker. The network tie breaker executable `samtb_net` is located in the `/usr/sbin/rsct/bin` directory. In the current implementation, the following options must be specified as key value pairs during the creation of the RSCT exec tie breaker:

**Address=**<IP address>

Address of the external IP instance which should be used to resolve the tie situation. Specify an IPv4 address in dotted quote notation, for example 192.168.1.1. Do not use a DNS name, because DNS may not work properly if communication problems occur, which is typically the case during cluster splits. Address is a mandatory option, no default is present.

**Log=**<1/0>

Specify 1 if you want the network tie breaker to write logs to the system log facility (syslog). Default is 1. Allowed values are 1 and 0.

**Count=**<number>

Numbers of ICMP echo requests which are sent to determinate quorum. If the first request gets a response, no further requests are sent. The default value is 2. Allowed value range is 1 through 9.

The following command will create a new network tie breaker:

```
# mkrsrc IBM.TieBreaker Type="EXEC" Name="mynetworktb"
DeviceInfo='PATHNAME=/usr/sbin/rsct/bin/samtb_net Address=192.168.1.1
Log=1' PostReserveWaitTime=30;
```

Activate your network tie breaker as follows:

```
# chrsrc -c IBM.PeerNode OpQuorumTieBreaker="mynetworktb"
```

You can use any regular RSCT command to manipulate the network tie breaker definition. To delete the tie-breaker definition, use the `rmrsrc` command.

**Some background information about the network tie breaker:** The RSCT tie-breaker design is based on the idea that when a node reserves a tie breaker, the tie breaker is no longer available and cannot be reserved by any other node.
As this is not feasible for a network-based approach (network tie breaker), some restrictions were introduced to the base tie-breaker design.

Reserve behavior: After an unsuccessful reservation attempt, no other reservation is possible until the node has joined the cluster again. To ensure this, a file is written to /var/ct/ which indicates that a reservation failed. If this file is present, a reserve command for a tie breaker will always fail. There is an additional process forked which watches quorum and removes the block file if the node has joined the domain again.

The following sample file was created by the network tie breaker as the result of a failing tie breaker reserve operation to the external IP instance 192.168.1.1. It contains the time stamp of the failed reserve operation.

```
# cat /var/ct/samtb_net_blockreserve_192.168.1.1
Mo Jul 4 08:38:40 CEST 2005
```

Configuring a RSCT tie breaker resource for the network tie breaker: This section describes the most important configuration options for an RSCT tie-breaker definition and will give you an idea of how they can be configured for a network tie breaker.

**PostReserveWaitTime=30**

In case of a failing reserve, ConfigRM will periodically call tie breaker reserve operation. Since the network tie-breaker only honors the first reserve attempt and blocks periodical reserve in case there was a failed reserve before, the PostReserveWaitTime should be set to the maximum possible value which is 30 seconds. This will keep the system load low in case a node stays in the pending Quorum state (periodical calls to tie breaker reserve).

**HeartbeatPeriod=30**

After there was a successful reserve, ConfigRM starts calling periodical tie breaker heartbeat operation. To keep the system load low during a cluster split, increase the time between the tie breaker heartbeats or even turn off heart beating by setting HeartbeatPeriod to 0.

Reviewing the system logs of a network tie breaker scenario: Here are the system logs of a two-node cluster (node n1 and node n2). For the error scenario it is assumed there are no critical resources running on both nodes. A network problem will break all available communication paths between the peers, but one peer (n2) is still able to communicate to its gateway (192.168.177.2). After some time communication is established again and both nodes can join the cluster.
Quorum support

Overriding the operational quorum

In order to remove nodes from the cluster, at least one node of the cluster must be online to initiate the `rmrpnodem` command. If there are not enough nodes to ever achieve an operational quorum, there is no chance to adjust the cluster size by administrative means so that the quorum can be reestablished.

If for any reasons the operational quorum function must be deactivated, the persistent attribute `OpQuorumOverride` must be set to 1:

```
chrsrc -c IBM.PeerNode OpQuorumOverride=1
```

In this case operational quorum State is always `HAS_QUORUM` and resource protection is not ensured anymore.

Figure 11. Two-node cluster system logs
Chapter 10. Controlling and administering System Automation for Multiplatforms

This chapter describes various parameters which can be used to control and change the general behavior of System Automation for Multiplatforms. It also gives some insight into the infrastructure of System Automation for Multiplatforms and offers some helpful hints and tips.

Controlling System Automation for Multiplatforms

There are several attributes you can use to change the general behavior of System Automation for Multiplatforms. You can start and stop the automation functionality, define timeout periods, and exclude nodes from automation, for example, for maintenance purposes.

These attributes are:

- **TimeOut**
  Specifies the timeout value in seconds for a start control operation executed by System Automation for Multiplatforms. When the timeout period expires, the operation is repeated if the RetryCount is not exceeded.

- **RetryCount**
  Number of retry attempts that may be performed if a control operation fails or times out.

- **Automation**
  Flag to enable or disable automation by System Automation for Multiplatforms.

- **ExcludedNodes**
  List of nodes on which System Automation for Multiplatforms actively pushes resources away or stops them. Can be used for maintenance purposes, for example.

- **ResourceRestartTimeOut**
  Amount of time in seconds System Automation for Multiplatforms waits to restart resources which were located on a failed node on another node.

- **TraceLevel**
  The trace level can be used to control the amount of trace entries written. The maximum value of 255 results in very detailed tracing, while the value 0 suppresses writing various classes of trace entries. Lowering the trace level is advisable for automation policies with a large number of resources.

The current values of the attributes described above can be listed with the command `lssamctrl`. The attributes are changed with the `samctrl` command. Refer to IBM Tivoli System Automation for Multiplatforms Reference for a listing and description of these commands.

**TimeOut and RetryCount**

The TimeOut attribute is always used in conjunction with the RetryCount attribute:

**TimeOut**

Specifies how long System Automation for Multiplatforms will wait for a resource manager to do something.
RetryCount

Specifies the number of possible control operation attempts System Automation for Multiplatforms will make within the TimeOut period if the control operation is not successful. In general, if the first attempt is not successful, the chances are fairly low that it will work on the second or subsequent attempts.

Start operations

The operation timer is started when System Automation for Multiplatforms sends the first resource start control operation to a resource. When the timer has started, there are three possibilities:

1. The resource changes to the desired state (online or offline) within the timeout period. In this case, no further actions are triggered because the resource is in the state in which System Automation for Multiplatforms wants it to be.

2. The resource rejects the start control operation within the timeout period. What happens next depends on the reject code:
   - If it indicates that the error is recoverable, System Automation for Multiplatforms will continue to issue start control operations against the resource. Every control operation attempt is counted. When the RetryCount value is exceeded, System Automation for Multiplatforms stops issuing further control operations.
   - If the error is not recoverable, the resource will go into a problem state. Whether or not this triggers further automation actions depends on the type of resource against which the start operation was issued:
     - If a fixed resource is affected, no further actions are triggered.
     - If the control operation was issued against a constituent of a floating resource and this constituent is in state Offline or Failed Offline, System Automation for Multiplatforms will attempt to issue the control operations against another constituent of the resource. Note that the constituent that rejected the control operation will remain in an unrecoverable error state until you issue a reset operation against it.

3. The resource does not reach the desired state (online) within the timeout period. In this case, System Automation for Multiplatforms first issues a reset operation against the resource and waits until the reset operation has been accepted and the resource is offline. Then, System Automation for Multiplatforms issues another start control operation against the resource. Every control operation attempt is counted and System Automation for Multiplatforms stops issuing control operations when the RetryCount is exceeded or when the maximum timeout (TimeOut * RetryCount) expires, whichever comes first.

When System Automation for Multiplatforms stops issuing control operations for a fixed resource or for a constituent of a floating resource, the OpState of the resource is set to failed offline. This indicates that the resource is no longer usable and that manual intervention is required to correct the cause of the failure. When the problem has been resolved, the resource must be reset with the RMC command resetsrc.

Note the retry counter is always reset when the resource reaches its desired state because no threshold is implemented. This means, for example, that a resource that is started, stays online for a short period of time, and then stops again, will be restarted by System Automation for Multiplatforms in a loop.

Default values are:
Controlling and administering System Automation for Multiplatforms

- TimeOut = 60
- RetryCount = 3

You use the command `samctrl –t Timeout` to change the TimeOut value and the command `samctrl –r Retry_count` to change the RetryCount value.

The IBM.Application class provides its own timeout value. If you add a resource of class IBM.Application to a group, the general TimeOut value is not used for this resource. As TimeOut value for this group member the larger value of StartCommandTimeout or MonitorCommandPeriod attribute (which are attributes of the IBM.Application resource) is used.

Stop operations
The operation timer is started when System Automation for Multiplatforms first sends a resource stop control operation to a resource. After the timer has started, there are three possibilities:

1. The resource changes to the desired state (offline) within the timeout period. No further actions are triggered.
2. The resource rejects the stop control within the timeout period. What happens next depends on the reject code:
   - If it indicates that the error is recoverable, System Automation for Multiplatforms issues another stop control operation against the resource.
   - If the error is not recoverable, the resource goes into a problem state. Manual intervention is required to get the resource out of the problem state.
3. The resource does not reach the desired state (offline) within the timeout period. In this case, System Automation for Multiplatforms first issues a reset operation against the resource and waits until the resource reaches its desired state (offline).

Automation
This flag indicates whether the System Automation for Multiplatforms automation functionality is enabled or not. If automation is disabled, System Automation for Multiplatforms will stop issuing control operations. The state of resources will remain unchanged.

The default value is AUTO mode, which means that automation is turned on.

You use `samctrl –M F` to enable automation, `samctrl –M T` to disable automation.

ExcludedNodes
This is a list of nodes on which System Automation for Multiplatforms will stop all resources and move them to another node if this is possible. For example, you have floating resource A which can run on four nodes node05, node06, node07 and node08. It is a member of resource group RG_A. After you made the group online it is started on node05. If you add node05 to the list of excluded nodes, System Automation for Multiplatforms will stop the resource on node05 and restart it on one of the other nodes.

Caution: If you exclude a node and one or more mandatory members of a group cannot be restarted on another node, the whole group may be stopped.

By default, the list is empty, which means that all nodes in the peer domain can be used.
You use `samctrl –u a` to add one or more nodes to the list of excluded nodes, `samctrl –u d` to delete nodes from that list and `samctrl –u r` to replace nodes in the list.

**ResourceRestartTimeout**

The ResourceRestartTimeout value specifies the time in seconds that System Automation for Multiplatforms waits before it restarts resources that were located on a failed node on an different node. This gives the resources or the failed node the chance to do some cleaning up before the resources are moved to another system.

The default value is 5 seconds.

You specify the resource restart timeout value with the command `samctrl –o`.

You can specify the trace level with the command `samctrl -l`. The trace level (TraceLevel) determines the amount of trace entries that are written. The default value is 127. The maximum value of 255 results in very detailed tracing. If the value is set to 0, various classes of trace entries are not written. Reducing the trace level is advisable for automation policies with a large number of resources.

**Examples**

To list the current System Automation for Multiplatforms control parameters you use the `lssamctrl` command:

```
lssamctrl
```

Displaying System Automation for Multiplatforms Control Information:

```
SAMControl:
  TimeOut       = 60
  RetryCount    = 3
  Automation     = Auto
  ExcludedNodes = {}
  ResourceRestartTimeOut = 5
  ActiveVersion  = [1.2.0.0,Tue 04 May 2004 12:30:48 PM EDT]
  Enable Publisher  = Disabled
  TraceLevel       = 127
```

To add the node node05 to the list of excluded nodes, this command is entered:

```
samctrl -u a node05
```

To set the RetryCount parameter to 5, this command is entered:

```
samctrl -r 5
```
Managing automation policies for System Automation for Multiplatforms domains

This section gives an overview of how you manage automation policies which determine the automation behavior of System Automation for Multiplatforms.

Automation policies are one of the core elements of System Automation for Multiplatforms. You use them to define the rules that determine the automation behavior (actions and decisions) of System Automation for Multiplatforms. Policies contain definitions of resource groups, relationships between resources and/or groups, and equivalencies. Typically, it is the main task of an administrator to maintain the policies and to ensure that they are correct and can be re-created.

To maintain one or more policies you use the sampolicy command. The command can be used in three different flavors:
1. To save a policy and restore it at a later point in time.
2. To completely replace a policy.
3. To update the active policy.

The sampolicy command handles resource groups, relationships, equivalencies, IBM.Application, IBM.ServiceIP, IBM.AgFileSystem, and IBM.Test resources, control parameters (samctrl), and IBM.TieBreaker resources.

Using the sampolicy command to manage policies

The following sections describe how to manage policies using the sampolicy command. For detailed information, see the description of the command in IBM Tivoli System Automation for Multiplatforms Reference.

Disabling sampolicy confirmation requests:
When you issue the sampolicy command to activate, deactivate, or update a policy, you are prompted to confirm the action after the syntax check is complete. To disable the confirmation request, for example, in batch scripts, you can use the -q parameter of the sampolicy command.

Example: To suppress the confirmation request when activating the policy that is specified in the XML file myPolicy.xml, issue the following command:
sampolicy -q -a myPolicy.xml

Saving a policy
System Automation for Multiplatforms uses an XML file to define the automation policy. The ‘Policy XML reference’ chapter in IBM Tivoli System Automation for Multiplatforms Reference provides information on how to define an XML policy file.

Example: To save the active policy to the file myPolicy.xml, use the following command:
sampolicy -s /usr/xml/myPolicy.xml

Activating, restoring, and replacing a saved configuration
You can use the sampolicy command to activate, restore, or replace a policy.

Example: To activate the policy that is specified in the XML file myPolicy.xml, issue the following command:
sampolicy -a /usr/xml/myPolicy.xml
Updating an active policy
You can use the sampolicy command to update the currently active policy in a
non-disruptive way. Update processing is described in more detail in the following
section.

Example: To update the active policy with the elements that are specified in the
XML file update.xml, issue the following command:
sampolicy -u update.xml

Deactivating an active policy
To deactivate the currently active policy, issue the following command:
sampolicy -d

Querying policy information
To display common policy information (for example, PolicyName,
PolicyDescription, PolicyAuthor), you use the -i flag with the sampolicy command.

Example:
To display the common policy information from the file myPolicy.xml, use this
command:
sampolicy -i /usr/xml/myPolicy.xml

Using XML for automation policies
You can define automation policies in XML files. This approach gives you a
number of options:
• You can define the initial automation policy for a System Automation for
  Multiplatforms domain in an XML file.
• You can update an automation policy

XML policy files can be created and edited using the graphical policy editor in
Integrated Solutions Console, as described in "Creating and modifying
automation policies using the policy editor" on page 143

Using XML to define the initial automation policy for a System
Automation for Multiplatforms domain
This scenario outlines steps you need to perform to set up a new System
Automation for Multiplatforms domain by defining and activating the automation
policy. The scenario assumes that you have already performed the following
prerequisite steps:
1. System Automation for Multiplatforms is installed on all nodes that are to form
the cluster.
2. You have defined the cluster and started the cluster as described in "Step 1:
   Defining and administering a cluster" on page 12.

Perform these steps:
1. Define the required XML elements in an XML file as described in IBM Tivoli
   System Automation for Multiplatforms Reference (in chapter "Policy XML
   reference", which also contains an example XML automation policy).
2. Save the XML file to the policy pool directory. This is mandatory if you want to
   manage your automation policies from the SA operations console.
3. Check the automation policy for warnings and errors. Use one of the following
   approaches:
From the command line, issue the command `sampolicy -c`. The results of the check are displayed:

- If no problems are detected, you receive a message, confirming that the automation policy is ready for activation.
- If problems were found during the check, the list of problems is displayed. Any errors that were detected must be resolved before you can activate the policy. If warnings were issued, you should check if the underlying problems cannot be avoided, although the policy can be activated even if they are not resolved.

On the SA operations console, open the “Select an Automation policy” page. On the page, you see whether the automation policy is ready for activation or problems were found during the check. Any errors that were detected must be resolved before you can activate the policy. If warnings were issued, you should check if the underlying problems cannot be avoided, although the policy can be activated even if they are not resolved. For a detailed description, refer to “Checking automation policies” on page 132.

4. Activate the automation policy. Use one of the following approaches:

- From the command line, issue the following command:
  
  `sampolicy -a <file_name>`

  where `<file_name>` is the fully qualified name of the XML file in which the automation policy to be activated is defined.

- On the SA operations console, open the “Select an automation policy” page, select the automation policy, and click **Activate policy**. In general, this is equivalent to issuing the `sampolicy -r` command. For initial policy activation, however, it is equivalent to the `sampolicy -a` command.

**Using XML to update an automation policy**

Perform the following steps:

1. Save the active automation policy to an XML file. To do so, issue the following command from a command line:

   `sampolicy -s <file_name>`

   where `<file_name>` stands for the fully qualified name of the XML file to which the automation policy is to be saved. You should perform this step regardless of whether the currently active automation policy resulted from the activation of an XML policy file or was defined using command line commands. Among others, this has the benefit that you have a backup of the currently active policy for future activation and reference.

2. Create a copy of the XML file and save it under a different, meaningful name to ensure that the definitions of currently active automation policy are retained.

3. Make the desired changes to the automation policy in the XML policy file. This can mean, for example, deleting or creating new resources and groups, changing group memberships, changing attribute values, adding, removing, or changing relationships.

4. Check the automation policy for warnings and errors. Use one of the following approaches:

   - From the command line, issue the following command:

     `sampolicy -c <file_name>`

     where `<file_name>` stands for the fully qualified name of the XML file to be checked. The results of the check are displayed:
Controlling and administering System Automation for Multiplatforms

- If no problems are detected, you receive a message, confirming that the automation policy is ready for activation.
- If problems were found during the check, the list of problems is displayed. Any errors that were detected must be resolved before you can activate the policy. If warnings were issued, you should check if the underlying problems cannot be avoided, although the policy can be activated even if they are not resolved.

  • On the SA operations console, open the "Select an Automation policy" page. On the page, you see whether the automation policy is ready for activation or problems were found during the check. Any errors that were detected must be resolved before you can activate the policy. If warnings were issued, you should check if the underlying problems cannot be avoided, although the policy can be activated even if they are not resolved. For a detailed description, refer to "Checking automation policies" on page 132.

When you are sure that the automation policy is ready for activation, you have several options for activating the policy:

Disruptive activation

In this type of policy activation, the new automation policy completely replaces the old, which means that all currently active definitions are removed and all resources are stopped before the new policy file is activated. The effect is the same as if you would deactivate the active automation policy and activate a new one.

To activate an automation policy in this way, use this command:

```
sampolicy -a <file_name>
```

where `<file_name>` stands for the fully qualified name of the XML file that contains the automation policy to be activated.

Delta-activation with preservation of currently active elements that are absent in the new policy

This type of policy activation is non-disruptive and leaves elements untouched that are currently defined in the active policy but not mentioned in the new policy.

To activate an automation policy in this way, use this command:

```
sampolicy -u <file_name>
```

where `<file_name>` stands for the fully qualified name of the XML file that contains the automation policy to be activated.

Activating an automation policy in this way has the following effects:

- New elements that are defined in the XML policy file are added to the currently active automation policy.
- Any changes in the XML policy file to elements that exist in the currently active automation are applied.
- Elements that are defined identically in both the currently active and the new automation policy remain untouched.
- Elements that exist in the currently active but not in the new automation policy remain untouched.

This policy activation approach allows you to extend and change the currently active automation policy by using XML policy files that only contain new elements and changes to existing elements. Note, however,
that such XML policy files must be complete in the sense that they must be well-formed, valid, and error-free to be activated.

Delta-activation with deletion of currently active elements that are absent in the new policy

This type of policy activation is non-disruptive but elements that exist in the currently active but not in the automation policy to be activated are removed.

To activate an automation policy in this way, use this command:

```
sampolicy -r <file_name>
```

where `<file_name>` stands for the name of the XML file that contains the automation policy to be activated. This type of policy activation is also performed when you activate an automation policy from the SA operations console (see "Activating automation policies" on page 133). Activating an automation policy in this way has the following effects:

- Elements that are defined identically in both the currently active and the new automation policy remain untouched.
- New elements that are defined in the XML policy file are added to the currently active automation policy.
- Any changes in the XML policy file to elements that exist in the currently active automation are applied.
- Elements that exist in the currently active but not in the new automation policy are removed.

This type of policy activation enables you to extend and change the currently active automation policy by using XML policy files that only contain new elements, changes to existing elements, and those elements that are to be left untouched. Note, however, that such XML policy files must be complete in the sense that they must be well-formed, valid, and error-free to be activated.

Using policy templates

If you use large policies that contain multiple repeatable values, or want to move a policy from one environment to another more quickly, you can use the template processing feature that is provided by the `sampolicy` command.

Instead of actual values, you use parameters in the XML file that look like `%%parname%%`, where `parname` is the name of the parameter. The parameter must be defined in the "var" element of the XML policy template file.

The element has the following attributes:

- name: name of parameter
- value: value of the parameter

You can use one parameter in many places and in multiple files but need to change it only in one place. It helps in migrating the policy from one environment to another.

Additionally, it is possible to include other XML files in the XML policy template file which makes working with an XML policy template very comfortable.

You can attach new XML files using the `include` element. Its node value must contain the path to the attached file; the path can be relative.
The template XML file has a different root element: AutomationPolicyTemplate. If you want to activate or update a policy using templates you must use the -t parameter with the sampolicy command.

Example:
sampolicy -a -t top.xml

Contents of a top.xml file:
<AutomationPolicyTemplate productID="SAM" version="2.3.0"
xmlns="http://.ibm.com/TSA/Policy.xsd"
xmlns:xsi="http://.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://.ibm.com/TSA/Policy.xsd SAMPolicyTemplate.xsd ">
  <PolicyInformation>
    <PolicyName>template</PolicyName>
    <AutomationDomainName>%%domain%%</AutomationDomainName>
    <PolicyToken>1.0</PolicyToken>
    <PolicyDescription>MyDescription</PolicyDescription>
    <PolicyAuthor>admin</PolicyAuthor>
  </PolicyInformation>
  <var name="domain" value="lnx"/>
  <var name="node1" value="lnxcm11x"/>
  <include>internal.xml</include>
</AutomationPolicyTemplate>

Contents of internal.xml:
<?xml version="1.0" encoding="UTF-8"?>
<AutomationPolicy productID="SAM" version="2.3"
xmlns="http://.ibm.com/TSA/Policy.xsd"
xmlns:xsi="http://.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://.ibm.com/TSA/Policy.xsd SAMPolicy.xsd">
  <ControlInformation>
    <Timeout>60</Timeout>
    <RetryCount>3</RetryCount>
    <ResourceRestartTimeout>5</ResourceRestartTimeout>
  </ControlInformation>
  <Resource name="T1" class="IBM.Test" node="%%node1%%">
    <ClassAttributesReference>
      <IBM.TestAttributes name="IBM.Test.T1"/>
    </ClassAttributesReference>
  </Resource>
  <IBM.TestAttributes name="IBM.Test.T1" >
    <TimeToStart>0</TimeToStart>
    <TimeToStop>0</TimeToStop>
    <WriteToSyslog>0</WriteToSyslog>
  </IBM.TestAttributes>
</AutomationPolicy>

The policy shown above will create a simple IBM.Test resource using templates. The resulting XML policy will contain all included XML files and the substituted parameters. It will be saved in the current folder with the name of the top-level file and the and the added suffix .complete.tmp.

The file top.xml.complete.tmp that is generated from the two example files above (top.xml and internal.xml), using the command sampolicy -a -t top.xml looks like this:
<?xml version="1.0" encoding="UTF-8"?>
<AutomationPolicy productID="SAM" version="2.3"
xmlns="http://.ibm.com/TSA/Policy.xsd"
xmlns:xsi="http://.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://.ibm.com/TSA/Policy.xsd SAMPolicy.xsd">
Controlling and administering System Automation for Multiplatforms

<PolicyInformation>
  <PolicyName>template</PolicyName>
  <AutomationDomainName>lnx</AutomationDomainName>
  <PolicyToken>1.0</PolicyToken>
  <PolicyDescription>MyDescription</PolicyDescription>
  <PolicyAuthor>admin</PolicyAuthor>
</PolicyInformation>

<ControlInformation>
  <Timeout>60</Timeout>
  <RetryCount>3</RetryCount>
  <ResourceRestartTimeout>5</ResourceRestartTimeout>
</ControlInformation>

<Resource name="T1" class="IBM.Test" node="lnxcm11x">
  <ClassAttributesReference>
    <IBM.TestAttributes name="IBM.Test.T1"/>
  </ClassAttributesReference>
</Resource>

<IBM.TestAttributes name="IBM.Test.T1">
  <TimeToStart>0</TimeToStart>
  <TimeToStop>0</TimeToStop>
  <WriteToSyslog>0</WriteToSyslog>
</IBM.TestAttributes>

</AutomationPolicy>
Starting and stopping resource groups and group members

You use start and stop requests to bring resource groups or individual resource group members online or offline. Start and stop requests always overrule the NominalState value of a resource group. The nominal state itself is not changed.

You can submit start and stop request from the SA operations console or from the command line.

From the command line, you use the following commands:

• To start or stop a resource group, enter:
  `rgreq -o <start|stop>`

• To start or stop or stop a resource group member, enter:
  `rgmbrreq -o <start|stop>`

Scope of start and stop requests

The scope of a start or stop request depends on whether it is submitted against a resource group or against a particular member of a resource group:

• A request that is submitted against a resource group affects all members of the group. Additional resources may be affected if relationships are defined in the policy for the resource group or any of its members or both.

• A request that is submitted against a resource group member only affects the particular resource. Additional resources may be affected if the resource has relationships.

Note: A request that is issued against a mandatory member of a resource group may also have an impact on the resource group's observed and composite states. This happens when the request brings the member into an observed state that conflicts with the desired state of the resource group. In this case, the observed state of the resource group changes to a “pending” state and its composite state changes to “error”.

Example:

When an operator submits a stop request against a mandatory member of a resource group whose desired state is Online, the observed state of the resource group changes to "pending online" and its composite state changes to "error" when the member resource is stopped.

Note that start and stop requests from the same source replace each other because the request list for a resource group or managed resource can only contain one request from each source.

Changing the priority of start and stop requests

The priority level and the value of its Source attribute determine the priority of a request. Depending on how a request is submitted, the Source attribute and the priority level of the request are set to default values. You can override the default values when you issue the commands `rgreq` and `rgmbrreq` by using the `-S` option to specify the source and the `-p` option to set the priority level of the request. For more information on how requests are prioritized, see "Prioritization of requests" on page 195.
Canceling start and stop requests

Start and stop requests are persistent, they are retained indefinitely in a resource's request list. To remove them from the list, they must be canceled. When you cancel the currently active request and the request list contains further requests, the request ranking next on the list is activated.

- To cancel requests in the request list of a resource group, enter:
  `rgreq –o cancel`
- To cancel requests in the request list of a resource group member, enter:
  `rgmbrreq –o cancel`

Displaying request lists

You use the `lsrgreq` command to display submitted requests.

Example scenario

This example scenario illustrates how an operator can change the operational state of a resource group by issuing and canceling requests and how the `lsrgreq` command can be used to display a resource's request list and to track the success of requests. For a detailed description of the commands that are used in the examples below, see [IBM Tivoli System Automation for Multiplatforms Reference](#).

The NominalState of the group "top-rg" is Offline, but its OpState is Online:

```
lnxcm3x:# lsr -g top-rg | grep State
Displaying Resource Group information:
For Resource Group "top-rg".

Resource Group 1:
  NominalState = Offline
  OpState      = Online
  TopGroupNominalState = Offline
```

The output of the `lsrgreq` command shows that the group's OpState results from an Active start request from end-to-end automation:

```
lnxcm3x:# lsrgreq -L -g top-rg
Displaying Resource Group request information:
All request information
For Resource Group "top-rg".

Resource Group 1:
  ResourceGroup = top-rg
  Priority = High
  Action = start
  Source = Automation
  NodeList = {}
  ActiveStatus = Active
  Token = 8f5697eb5f84c0f044995b3d00040a5b
  UserID =
  MoveStatus = None
```

To bring the group "top-rg" offline, an operator issues a stop request:

```
rgreq –o stop top-rg
```

The output of the `lsrgreq` command shows that the operator's stop request was added to the request list but did not win. The stop request remains inactive because it was issued with the default priority (low) and could not win against the higher ranking start request from end-to-end automation:
The output of the `lsrg` command shows that the group stays online:

```
lnxcm3x:~ # lsrg -g top-rg
Displaying Resource Group information:
For Resource Group "top-rg".

Resource Group 1:
  ResourceGroup = top-rg
  Priority = High
  Action = start
  Source = Automation
  NodeList = {}
  ActiveStatus = Active
  Token = 8f5697eb5f84c0f044995b3d00040a5b
  UserID =
  MoveStatus = None

Resource Group 2:
  ResourceGroup = top-rg
  Priority = low
  Action = stop
  Source = Operator
  NodeList = {}
  ActiveStatus = InActive
  Token = 8f5697eb5f84c0f044995dad00007b338
  UserID =
  MoveStatus = None
```

To bring the group offline, the operator issues a high-priority stop request:

```
rgreq -p high -o stop top-rg
```

The output of the `lsrgreq` command shows that this request has replaced the operator's previous low-priority request, won against the start request from end-to-end automation, and has become the active request:

```
lnxcm3x:~ # lsrgreq -L -g top-rg
Displaying Resource Group request information:
All request information
For Resource Group "top-rg".

Resource Group 1:
  ResourceGroup = top-rg
  Priority = High
  Action = start
  Source = Automation
  NodeList = {}
  ActiveStatus = InActive
  Token = 8f5697eb5f84c0f044995b3d000040a5b
  UserID =
  MoveStatus = None

Resource Group 2:
  ResourceGroup = top-rg
  Priority = High
  Action = stop
  Source = Operator
  NodeList = {}
  ActiveStatus = InActive
  Token = 8f5697eb5f84c0f044995dad00007b338
  UserID =
  MoveStatus = None
```
Controlling and administering System Automation for Multiplatforms

NodeList = {}
ActiveStatus = Active
Token = 8f5697eb5f84c0f044996004000368b1
UserID =
MoveStatus = None

The group is stopped due to the stop request:

```
lnxcm3x:# lsrg -g top-rg | grep State
Displaying Resource Group information:
For Resource Group "top-rg".
```

```
Resource Group 1:
  NominalState = Offline
  OpState = Offline
  TopGroupNominalState = Offline
```

To bring the "top-rg" group online again, the operator must only cancel the last stop request, an additional start request is not required:

```
rgreq -o cancel top-rg
```

The output examples below show that the "top-rg" group is started after the request is canceled because the request from end-to-end automation becomes active again:

```
lnxcm3x: # lsrgreq -L -g top-rg
Displaying Resource Group request information:
All request information
For Resource Group "top-rg".
```

```
Resource Group 1:
  ResourceGroup = top-rg
  Priority = High
  Action = start
  Source = Automation
  NodeList = {}
  ActiveStatus = Active
  Token = 8f5697eb5f84c0f044995b3d0004b5
  UserID =
  MoveStatus = None
```

```
lnxcm3x:# lsrg -g top-rg | grep State
Displaying Resource Group information:
For Resource Group "top-rg".
```

```
Resource Group 1:
Resource Group 1:
  NominalState = Offline
  OpState = Online
  TopGroupNominalState = Offline
```

Using requests to lock and unlock resource groups and resources

You use lock requests to freeze resource groups and individual resource group members in their current state and prevent them from being automated. To resume the automation of locked resource groups and resources, you use unlock requests. You issue lock and unlock requests from the command line:

- To lock a resource group, use this command:
  ```
  rgreq -o lock
  ```
- To lock a particular resource group member, use this command:
  ```
  rgmbrreq -o lock
  ```
Lock requests are persistent, they are retained indefinitely in a resource's request list. To remove them from a request list, you must unlock the resource. If the lock request is active when you unlock the resource, the request ranking next on the list is activated.

- To unlock a locked resource group and to remove the lock request from the request list, use this command:
  `rgreq -o unlock`
- To unlock a locked resource group member and to remove the lock request from the request list, use this command:
  `rgmbrreq -o unlock`

### Scope of lock requests
The scope of a lock request can be determined in the following way:

- The locked resource is within the scope of the lock request.
- All members of a resource group are within the scope of a lock request if the group itself is within its scope.
- If the target of a StartAfter, DependsOn, or DependsOnAny relationship is within the scope of a lock request, the source also is within its scope.
- If the source of a StopAfter relationship is within the scope of a lock request, the target also is within its scope.

### Changing the priority of requests
To prioritize requests, System Automation for Multiplatforms utilizes the priority level and the value of the source attribute of the request. Depending on how a request is submitted, the Source attribute and the priority level of the request are set to default values. You can override the default values when you issue the commands `rgreq` and `rgmbrreq` by using the `-S` option to specify the source and the `-p` option to set the priority level of the request. For more information on how requests are prioritized, see "Prioritization of requests" on page 195.

Lock requests from the same source replace each other because the request list of a resource group or resource group member can only contain one request from each source. You can use the `lsrgreq` command to display a request list.

### Example scenario
This example scenario illustrates how an operator can change the operational state of a resource group by issuing and canceling requests and how the `lsrgreq` command can be used to display a resource's request list and to track the success of requests. For a detailed description of the commands that are used in the examples below, see [IBM Tivoli System Automation for Multiplatforms Reference](#).

The NominalState of the resource group "top-rg" is Online, but its OpState is Offline:

```
lnxcm3x:# lsrg -g top-rg | grep State
Displaying Resource Group information:
For Resource Group "top-rg".

Resource Group 1:
    NominalState   = Online
    OpState        = Offline
    TopGroupNominalState = Online
```

The output of the `lsrgreq` command shows that the group is offline because a lock request prevents from being started:
Controlling and administering System Automation for Multiplatforms

$ lsrgreq -L -g top-rg$
Displaying Resource Group request information:
All request information
For Resource Group "top-rg".

Resource Group 1:
  ResourceGroup = top-rg
  Priority = High
  Action = lock
  Source = Operator
  NodeList = {}
  ActiveStatus = Inactive
  Token = 8f5697eb5f84c0f044995b3d00040a5b
  UserID =
  MoveStatus = None

In an attempt to bring the group "top-rg" online, an operator unlocks the group:
$ rgreq -o unlock top-rg$

Now the request list for "top-rg" looks like this:

$ lsrgreq -L -g top-rg$
Displaying Resource Group request information:
All request information
No requests were found.

Prioritization of requests

Sources and default priorities of requests

The default value to which the Source attribute of a request is set and the request’s default priority level depend on how the request is submitted (Submitter) and on the type of resource against which it is issued (Target resource). Table 23 shows the default values.

Table 23. Sources and priority levels of start and stop requests

<table>
<thead>
<tr>
<th>Submitter</th>
<th>Target resource</th>
<th>Source attribute value</th>
<th>Default priority level</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA operations console in direct-access or first-level automation mode</td>
<td>Resource or resource group</td>
<td>Operator</td>
<td>high</td>
</tr>
<tr>
<td>System Automation for Multiplatforms command-line interface</td>
<td>Resource or resource group</td>
<td>Operator</td>
<td>low Possible values: low, high, force</td>
</tr>
<tr>
<td>External scheduler (for example, Tivoli Workload Scheduler or a cron job)</td>
<td>Resource or resource group</td>
<td>ExtSched</td>
<td>High Possible values: low, high, force</td>
</tr>
<tr>
<td>End-to-end automation engine</td>
<td>End-to-end automation resource reference that references a System Automation for Multiplatforms resource or resource group</td>
<td>Automation</td>
<td>high</td>
</tr>
</tbody>
</table>
How requests are prioritized

Requests that are submitted against a resource group are prioritized in the following way:

- Any request takes precedence over the NominalState of the resource group.
- If requests have different priority levels, the priority level determines a request’s priority rank.
- If requests have the same priority level, the Source attributes are evaluated to determine the priority rank of each request. [Figure 12] illustrates how the Source attribute value influences the priority ranking of requests.

![Figure 12. Priority ranking of requests](image)

[Figure 12] illustrates the following facts:
- Requests from source “Operator” take precedence over requests from sources “Automation” and “ExtSched”.

Controlling and administering System Automation for Multiplatforms

Table 23. Sources and priority levels of start and stop requests (continued)

<table>
<thead>
<tr>
<th>Submitter</th>
<th>Target resource</th>
<th>Source attribute value</th>
<th>Default priority level</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA operations console in end-to-end automation mode</td>
<td>End-to-end automation resource reference that references a System Automation for Multiplatforms resource or resource group</td>
<td>Automation</td>
<td>high</td>
</tr>
<tr>
<td>End-to-end automation manager command line interface</td>
<td>End-to-end automation resource reference that references a System Automation for Multiplatforms resource or resource group</td>
<td>Automation</td>
<td>high</td>
</tr>
</tbody>
</table>
Requests from source "Automation" take precedence over requests from source "ExtSched". This means that you can overrule requests from source "Automation" by issuing command line requests using the option -p high. To overrule requests from source "Operator", you must specify the option -p force.

**Priority of operations console requests**

All online and offline requests that are submitted from the System Automation for Multiplatforms or end-to-end automation operations console result in start and stop requests from source "Automation". Such requests have a higher priority than the NominalState. Operators can overrule such requests from the command line by submitting commands with the priority Force.

**Priority of end-to-end automation requests**

When requests are issued against end-to-end automation resources, the resulting requests against the referenced first-level automation resources are always issued by the end-to-end automation manager, which is why their priority level is always set to "High" and their Source attribute value is set to "Automation". This is true regardless of whether the original request against the end-to-end automation resource was issued by the end-to-end automation manager or an operator.

---

### Moving resource groups with the rgreq command

You can move individual resource groups to a different cluster node without affecting all resources currently running on the same node. This is desirable, for example, in load-balancing scenarios, in which workload and performance objectives may already be achieved when a single or a few resource groups are moved to a different node.

Adjustments to the placement of resources can be done using the `rgreq -o move` command.

**Scope of a move**

The scope of a move is all the members of a top-level resource group. Resources that are dependent upon one or more resources involved in the move might be affected, i.e. stopped and started. A move request cannot be issued against a single managed resource.

If the MemberLocation attribute of the top-level resource group is set to collocated, no list of nodes has to be provided with the `rgreq` command. In this case, all resources are located on the same node and will be moved away from that node. If the resource group is not collocated, a list of one or more nodes has to be specified with the -n option of the `rgreq` command. All resources will be moved away from those nodes.

A move request issued against a resource group that contains only fixed resources will not be accepted. Also, a move request issued against an offline resource group will be rejected. While there is a move request already in progress, another move request on the same resource group will be rejected.

If a resource within the move scope has a Dependency against an equivalency, then all other resources that have a Dependency against the same equivalency will also be added to the move scope, and thus stopped and restarted during the move operation. To avoid this behavior, different equivalencies need to be created for the unrelated resources even if the members of the equivalencies are the same.
Processing of a move request

Starting with release 2.3, the behavior of System Automation for Multiplatforms during the offline phase has changed. In earlier releases, all resources within the move scope were stopped. Now, a node exclusion approach is used in which only resources within the move scope that are currently running on the nodes specified in the move request or which have a dependency on these resources are stopped.

In the online phase of the move, the binder performs a new node placement for all resource group members, and the resource group is restarted.

Note that if it turns out to be impossible to restart the mandatory members of the top-level resource group while honouring the list of nodes to move away from, the list will be ignored and resources may be restarted on it. Likewise, if the only place to restart a resource is the original system it was running on, then it will be restarted there if it is a mandatory resource.

A move request is automatically removed when the move action is carried out. The MoveStatus dynamic attribute of the resource group being moved will show values indicating the progress of the move.

You may, however, need to cancel a move request if the resources are not moved back to the original node although the move operation does not complete. You cancel a move request with the following command:

```
rqreq -o movecancel
```

Move and relationships

In addition to performing a move request on the members of a top level resource group, there may be other resource groups and/or resource group members outside the moved top level group, which placement must be adjusted according to defined relationships constraints. This applies to the following relationships:

- Collocation
- AntiCollocation
- DependsOn
- DependsOnAny
- StopAfter
- ForcedDownBy

In addition, the Affinity and AntiAffinity relationships might not be fulfilled after the move.
Using shadow resources

A shadow resource has the following characteristics:

- It is a fixed resource of class IBM.Application that monitors the OpState of another fixed resource.
- It is member of an equivalency whose SelectFromPolicy attribute value is set to NoControl, which specifies that System Automation for Multiplatforms only evaluates the OpState of the member resources but does not start or stop them.

You must use shadow resources if you want to define a relationship of type DependsOn from a floating resource to one out of two or more fixed resources that are running on different nodes, because the DependsOn relationship triggers both a StartAfter and a StopAfter behavior.

If you set up such a DependsOn relationship without defining shadow resources, an undesired automation behavior occurs:

![Diagram of Scenario 1: DependsOn relationship, without shadow resources]

In Scenario 1, System Automation for Multiplatforms will stop the floating resource even if only one of the fixed resources is not available, and System Automation for Multiplatforms can only restart the floating resource if both of the fixed resources are available.

To achieve the desired automation behavior, you create shadow resources that are members of a NoControl equivalency:
In Scenario 2, the ForceDown behavior of the DependsOn relationship will still cause System Automation for Multiplatforms to stop the floating resource if the operational state of the currently selected member of the equivalency changes to Offline, but now System Automation for Multiplatforms can restart the floating resource again if at least one other member of the equivalency is online.

Defining shadow resources

To create shadow resources, perform the following steps:

1. Create the command scripts for the shadow resources.

   Although System Automation for Multiplatforms will only use the MonitorCommand, both a valid StartCommand and a valid StopCommand must also be supplied. The MonitorCommand must query the OpState of the resource it is a shadow for. This can be achieved in one of the following ways:

   - by querying the OpState using
     ```bash
     'lsrsrc -s 'Name like "<res> "'IBM.Application OpState
     ```
   - by duplicating the MonitorCommand of the original resource
   - by saving the OpState of the original resource to a file and reading it within the MonitorCommand of the shadow resource

   **Example:**

   The following sample script contains the required commands for the shadow resource of the resource “fixed_rs1”:

   ```bash
   #!/bin/ksh
   # shadow_sample.sh
   # init section
   #
   Action=${1:-status}
   ```
ResName=${2:-myresource}

UNKNOWN=0
ONLINE=1
OFFLINE=2
FAILED_OFFLINE=3

export CT_MANAGEMENT_SCOPE=2 # necessary to execute SAMP commands

# main section

case ${Action} in
  start)
    # is not executed .. so irrelevant
    RC=0
    ;;

  stop)
    # is not executed .. so irrelevant
    RC=0
    ;;

  status|*)
    RCval=$(lsrsrc -xt -s 'Name="'$ResName'"" IBM.Application OpState)
    RCx=${RCval:-2}
    case ${RCx} in
      [1]*) RC=${ONLINE}
        ;;
      [0]*) RC=${OFFLINE}
        ;;
    esac
    #logger -i -t "$(basename $0)" "$ResName monitored: $RC"
  esac

exit ${RC}

2. Create the shadow resources, using a command like the following for each of the resources:

```
# mkrsrc IBM.Application \
Name="fixed_rs1_shadow" \
ResourceType=0 \
NodeNameList="{'node1'}" \
UserName="root" \
StartCommand="/samplepath/shadow_sample.sh start fixed_rs1" \
StopCommand="/samplepath/shadow_sample.sh stop fixed_rs1" \
MonitorCommand="/samplepath/shadow_sample.sh status fixed_rs1" \
MonitorCommandPeriod=10 \
RunCommandsSync=1
```

3. Create an equivalency with a SelectFromPolicy attribute value of NoControl.

To create the equivalency containing the fixed resources created in step 2, use the following command:

```
# mkequ <equ-name> -p A,NoControl \nIBM.Application:<fixed-resource1>:<node-name1>,<fixed-resource2>:<node-name2>[
...
```

**Example:**

```
mkequ -p A,NoControl shadow_equ IBM.Application:fixed_rs1_shadow:node1,fixed_rs2_shadow:node2
```

To define an ordered policy, use -p 0,NoControl.

4. Create the DependsOn relationship from the floating resource to the equivalency 'shadow_equ':

```
# mkrel -p DependsOn -S IBM.Application:float1 -G IBM.Equivalency:shadow_equ \n  float1-depon-shadow_equ
```
Setting up security for System Automation for Multiplatforms clusters

This section describes how you set up non-root security for the command line interface of System Automation for Multiplatforms on AIX, Solaris, and Linux systems. For System Automation for Multiplatforms on Windows systems, the Administrator user ID is valid only.

On Linux, Solaris, and AIX systems, by default, only the user root has the required authority to perform operational tasks in System Automation for Multiplatforms and to make changes to the automation policy of System Automation for Multiplatforms, while all other users only have read access.

The security concept of System Automation for Multiplatforms is based on that of the RSCT component RMC, which implements security authorization using an access control list (ACL) file. Specifically, RMC uses the ACL file on a particular node to determine the permissions that a user must have in order to access resource classes and their resource instances. Since the System Automation resource managers are internally implemented as an RMC application, the same set of ACL control rules must be used in order to allow non-root users to manage (define, undefine, or change) the System Automation related resource classes (IBM.ResourceGroup, IBM.ManagedRelationship, IBM.Equivalency, IBM.ManagedResource, IBM.CHARMControl, IBM.Application and IBM.ServiceIP) and to start and stop the corresponding resource groups.

For detailed information on how to set up RMC ACL files, see the following sections in IBM RSCT Administration Guide:

- “Managing user access to resources using RMC ACL files” in Chapter 4 ("Managing and monitoring resources using RMC and resource managers")
- “Configuring the global and local authorization identity mappings” in Chapter 7 ("Understanding and administering cluster security services")

RSCT and RMC security authorization support is capable of managing user access based on individual resource classes and single nodes, for example, user access can be restricted to a specific RMC resource class on a particular node within the cluster. This level of authorization setting, however, is very complex and requires a clear understanding of the nature of each individual RMC resource class. Therefore it is recommended that you create roles for a System Automation for Multiplatforms operator and a System Automation for Multiplatforms administrator with general settings that will allow non-root users to manage all resource classes from any node defined within the cluster. By using the procedure described below you will create these two roles:

- “sa_admin” for an administrator
- “sa_operator” for an operator

To create the roles, perform the following steps (note that root authority is required):

1. Create the user IDs that are to be authorized to manage System Automation for Multiplatforms on all nodes. Use the appropriate command for your operating system, for example, on Linux use:
   
   # /usr/sbin/useradd ernie
   # /usr/sbin/useradd bert

2. Create a group for the user IDs on all nodes. Use the appropriate command for your operating system, for example, on Linux use:
   
   # /usr/sbin/groupadd sagroup
3. Add the group to the user IDs on all nodes. Use the appropriate command for your operating system, for example, on Linux use:

   # /usr/sbin/usermod -G sagroup ernie
   # /usr/sbin/usermod -G sagroup bert

   **Note:** Make sure to set the following environment variable for all users of System Automation for Multiplatforms on all nodes (peer domain scope):

   ```
   CT_MANAGEMENT_SCOPE=2
   ```

   You can set the variable permanently if you set it in the profile.


   The file is used to track the System Automation for Multiplatforms history. All commands that modify the resources of the automation manager (IBM.RecoveryRM) are logged to that file.

   By default, the file is owned by the user group root:

   ```
   -rw-r--r-- 1 root root 204 Oct 4 22:00 /var/ct/IBM.RecoveryRM.log
   ```

   You need to change the group ownership to “sagroup”, using the appropriate command for your operating system, for example, on Linux use:

   ```
   /bin/chgrp sagroup /var/ct/IBM.RecoveryRM.log
   ```

   You also need to change the file permission to 664:

   ```
   # /bin/chmod 664 /var/ct/IBM.RecoveryRM.log
   -rw-rw-r-- 1 root sagroup 204 Oct 4 22:00 /var/ct/IBM.RecoveryRM.log
   ```

   **Note:** If the file /var/ct/IBM.RecoveryRM.log does not exist after the initial installation of System Automation for Multiplatforms, you can create a dummy file using the `user/bin/touch` command:

   ```
   # /usr/bin/touch /var/ct/IBM.RecoveryRM.log
   ```

5. Modify the file /var/ct/cfg/ctsec_map.global on all nodes.

   You need to add the following entries for the user IDs “ernie” and “bert” to the RSCT global authorization identity mapping file (/var/ct/cfg/ctsec_map.global) on every node in the cluster:

   ```
   unix:ernie@<cluster>=sa_operator
   unix:ernie@<any_cluster>=sa_operator
   unix:bert@<cluster>=sa_admin
   unix:bert@<any_cluster>=sa_admin
   ```

   The file is used to map a local user ID on a node to a global user ID within the System Automation for Multiplatforms domain. In the example, the local user ID “ernie” is mapped to the global user ID “sa_operator”, and the local user ID “bert” is mapped to the global user ID “sa_admin”.

   You can easily authorize additional local user IDs for System Automation for Multiplatforms by adding additional lines to this global map file (on all nodes), and mapping these to the desired role (System Automation for Multiplatforms Operator or System Automation for Multiplatforms Administrator).

   **Note:** If the file /var/ct/cfg/ctsec_map.global does not exist on a node, copy the default file /usr/sbin/rsct/cfg/ctsec_map.global to the directory /var/ct/cfg, and add the new entries to the file /var/ct/cfg/ctsec_map.global. Do not remove any entries from the file /var/ct/cfg/ctsec_map.global that exist in the default file you copied. The /var/ct/cfg/ctsec_map.global files on all nodes within the cluster must be identical.

6. Modify the file /var/ct/cfg/ctrmc.acls on all nodes. You need to add the following entries for the global user IDs “sa_operator” and “sa_admin” to the
Controlling and administering System Automation for Multiplatforms

RMC ACL file (/var/ct/cfg/ctrmc.acls) on every node in the cluster and comment the line starting with LOCALHOST, for example:

```plaintext
# The following stanza contains default ACL entries.
# These entries are appended
# to each ACL defined for a resource class and
# are examined after any entries
# explicitly defined for a resource class
# by the stanzas in this file,
# including the OTHER stanza.

DEFAULT
root@LOCALHOST   *   rw
# LOCALHOST   *   r // comment this line out!
none:root   *   rw // give root access to all
none:sa_admin   *   rw // append this row for saadmin
none:sa_operator   *   rso // append this row for saoperator
```

7. Once the above modifications are complete the following command must be executed on every node in the cluster to activate the changes:

```plaintext
# /usr/bin/refresh -s ctrmc
```

8. Additional changes required to use sampolicy and *samadapter commands:

   a. Access to configuration files:
   ```plaintext
   # /bin/chgrp -R sagroup /opt/IBM/tsamp/sam/cfg
   # /bin/chmod g+ws /opt/IBM/tsamp/sam/cfg
   # /bin/chmod g+w /opt/IBM/tsamp/sam/cfg/*
   ```

   b. Access to log files:
   ```plaintext
   # /bin/chgrp -R sagroup /var.ibm/tivoli/common/eez/logs
   # /bin/chmod g+ws /var.ibm/tivoli/common/eez/logs
   # /bin/chmod g+w /var.ibm/tivoli/common/eez/logs/*
   ```

   c. Access to configuration files in the /etc directory.
   If there is no directory /etc/opt/IBM/tsamp/sam/cfg, create it by using
   ```plaintext
   # /bin/mkdir -p /etc/opt/IBM/tsamp/sam/cfg
   ```
   Then set the permissions appropriately:
   ```plaintext
   # /bin/chgrp -R sagroup /etc/opt/IBM/tsamp/sam/cfg
   # /bin/chmod g+ws /etc/opt/IBM/tsamp/sam/cfg
   # /bin/chmod g+w /etc/opt/IBM/tsamp/sam/cfg/*
   ```

9. Optional adjustments required for working with the “sam.policies” package:

Pre-canned policies for various applications are provided in the install package “sam.policies”, which can be downloaded from the following location:

```
http://catalog.lotus.com/wps/portal/topal
```

To allow a user who has the “sa_admin” role to set up these pre-canned policies, the permissions and the ownership of the /usr/sbin/rsct/sapolicies directory must be changed after the “sam.policies” package is installed on all nodes:

```plaintext
# chmod –R 2775 /usr/sbin/rsct/sapolicies
# chgrp -R sagroup /usr/sbin/rsct/sapolicies
```

**Results:** When the above steps have been completed successfully, the local users “ernie” and “bert” can perform operational tasks of System Automation for Multiplatforms, such as issue start and stop requests against resources, and the local user “bert” can also perform administrative tasks of System Automation for Multiplatforms, such as define and modify policies.

**Limitations of the non-root security setup**

The following list summarizes the limitations of the non-root security setup described above:
Controlling and administering System Automation for Multiplatforms

- Regular user cannot view the contents of the RMC resource manager trace file (for example, the trace of the IBM.RecoveryRMd daemon). All RMC Resource Manager daemons use the RMC framework library utility to create trace files and core images under the /var/ct/<cluster> directory. Since these resource managers can only be started by a superuser (user ID root) through the /usr/bin/startsrm command, the files that are created belong to the user ID root. This implies that a non-root user cannot collect debug and trace information by using the /usr/sbin/rsct/bin/ctsnap command.
- The following commands can only be invoked with root authority because they utilize Tivoli logging, which only works properly if the log files are maintained with root rights:
  - The sampolicy command
  - The samadapter command to start the end-to-end automation adapter
- The granularity of the ACL objects is based on resource classes, not on resources. This means that a regular user is either allowed to modify resources of a resource class or not, but it is not possible to grant or deny permissions on a resource basis, for example, a database administrator cannot be authorized only for database resources.
- The “sa_operator” role can modify resources by changing attribute values for the resources. This is a result of the “s” permission, which is needed for issuing System Automation for Multiplatforms requests. Without the “s” permission, users who have this role would not be able to perform any useful task, but with this permission, they are allowed to change and set attributes.

The following table shows which role or authority is required to perform typical System Automation for Multiplatforms tasks.

<table>
<thead>
<tr>
<th>Task</th>
<th>Authority</th>
<th>Roles</th>
<th>Permissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product installation</td>
<td>root</td>
<td>System Administrator</td>
<td>Installing and upgrading System Automation for Multiplatforms</td>
</tr>
<tr>
<td>Cluster management</td>
<td>root / sa_admin</td>
<td>System Administrator / System Automation for Multiplatforms Administrator</td>
<td>Defining, starting, stopping, and monitoring clusters and individual RMC Resource Managers</td>
</tr>
<tr>
<td>Resource definition and System Automation for Multiplatforms policy definition</td>
<td>root / sa_admin</td>
<td>System Administrator / System Automation for Multiplatforms Administrator</td>
<td>Defining, removing, changing resources, and setting up automation policies</td>
</tr>
<tr>
<td>Automation operation</td>
<td>root / sa_admin / sa_operator</td>
<td>System Administrator / System Automation for Multiplatforms Administrator and Operator</td>
<td>Issuing Online and Offline request, and resetting and monitoring resource groups and individual resources</td>
</tr>
</tbody>
</table>
### Table 24. Authorizations and roles for performing System Automation for Multiplatforms tasks (continued)

<table>
<thead>
<tr>
<th>Task</th>
<th>Authority</th>
<th>Roles</th>
<th>Permissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collecting trace and debug data for problem determination</td>
<td>root</td>
<td>System Administrator</td>
<td>Access to all system and application trace (log) files. (See the list of limitations above.)</td>
</tr>
<tr>
<td>Security setup</td>
<td>root</td>
<td>System Administrator</td>
<td>Defining, changing, and removing the security setup described in this section.</td>
</tr>
<tr>
<td>Adapter setup</td>
<td>root / sa_admin</td>
<td>System Administrator / System Automation for Multiplatforms Administrator</td>
<td>Defining, changing, and removing the configuration of the end-to-end automation</td>
</tr>
</tbody>
</table>
Diagnosing System Automation for Multiplatforms Resources

To get more information about resources managed by System Automation for Multiplatforms you can use the `samdiag` command, which is documented in IBM Tivoli System Automation for Multiplatforms Reference. This command is mainly intended to be used in situations where it is not obvious to the user what is happening on the system and why.

**Note:** The `samdiag` command provided for Release 1 of System Automation for Multiplatforms could only be executed on the node on which the master daemon was running. So running `samdiag` on a Release 2 daemon will generate an error if Release 1 and Release 2 daemons coexist in the same cluster, and if the master daemon is on a Release 1 node.

To get information about a resource group called "apacherg", you use the command:

`samdiag -g apacherg`

**Output:**

```
Diagnosis::Resource: apacherg/ResGroup/IBM.ResourceGroup
  type: CHARM Resource Group
  Status -
    Observed: Offline - SoftDown
    Desired: Offline - Requested Offline
    (Nominal: Offline - Nominal State: Offline)
    Automation: Idle - CharmBase trigger linked
    Startable: Yes - Resource is startable
    Binding: Unbound - Binding status initialized
    Compound: Satisfactory - Satisfactory

  Resource Based Quorum: None
  Members and Memberships:
    +---bind/HasMember ---> RA/Float/IBM.Test

  Group Constraint: None
  Binding Constraints:
    Flags: None
    Orders:
      Outstanding Order: None - Resource is Unavailable

  Dependencies:
    Start: Satisfied
      +---InCluster ---> Cluster
    Stop: Satisfied
    Binding exceptions:
      There are unbound members.

  Static Relationships:
    +---InCluster ---> Cluster

  Dynamic Relationships:
    +---bind/HasMember ---> RA/Float/IBM.Test
```

The following provides an interpretation of some of the information given in the example:

- The ObservedState should show the same value as OpState and NominalState. If this is not the case, contact the support center serving your location.
- Different values for DesiredState and NominalState indicate that a request has been issued against the resource group.
- The AutomationState can be 'busy' or 'idle'. 'Busy' means that the System Automation for Multiplatforms daemon (IBM.RecoveryRM) is waiting for another resource manager to start or stop a resource. After this has been
completed the AutomationState changes to ‘idle’. If this is not the case, contact the support center serving your location.

- A Startable state of ‘No’ indicates that some relationships, for example a DependsOn relationship, are not correctly set.
- A BindingState of ‘Unbound’ comes along with a Observedstate ‘Offline’. It indicates that the resource group is offline. Before it can be set to online a binding step must be performed which chooses the correct constituent. Then the BindingState is set to ‘Bound’. A BindingState of ‘Bound’ for resource groups being offline is an error.
- A CompoundState of ‘Satisfactory’ indicates that ObservedState and DesiredState are the same. CompoundStates ‘Inhibited’, ‘Denied’, or ‘Broken’ indicate errors like relationships which have not been fulfilled or ‘broken’ resources.
- Resource Based Quorum equals ‘None’ means that all floating resources of the group do not support a Resource Based Quorum of their own and the Resource Based Quorum flag is not set.
- ‘Bind/HasMember’ describes the relationship between resource group apacherg and its floating member RA/Float/IBM.Test. When performing the binding step mentioned above a constituent is selected for each of the ‘Bind/HasMember’ relationships. This constituent is temporarily bound to the resource group before this resource group is set to online.
- Outstanding Orders refer to the AutomationState. If AutomationState is not ‘idle’, the pending command is shown here.
- Start/Stop Dependencies show when a policy prevents the start of a resource.
- Binding exceptions provide a closer explanation of BindingState.
- Static relationships mean that all constituent resources and resource groups are members of the cluster.
- Dynamic relationships are temporary relationships caused by the binding step.
Using the TEC event interface with System Automation for Multiplatforms

This sections gives an overview of how you can use the Tivoli Enterprise Console (TEC) event interface to notify system administrators of important events in clusters that are automated by System Automation for Multiplatforms.

Whenever the configuration of System Automation for Multiplatforms or the state of an automated resource changes or when problems are encountered, the TEC event interface can be used to notify the system administrator.

There are two ways to notify the administrator:

1. Sending events to the IBM Tivoli Enterprise Console. As a prerequisite the TEC publisher function must be enabled (see "Enabling the TEC publisher function" on page 210).

2. Publishing System Automation for Multiplatforms internal attributes into the RSCT infrastructure. The administrator must subscribe to one of the Event Resource Manager (ERRM) scripts in order to get event information. See IBM Reliable Scalable Cluster Technology for Linux, Technical Reference, SA22–7983, for instructions on how to do this; also see the description in the preceding section "Generating events in case of failures" on page 218.

What is the Tivoli Enterprise Console?

The Tivoli Enterprise Console (TEC) is a rule-based event management application that uses a central server to process incoming events. TEC acts as a central collection point for alarms and events from a variety of sources, including those of Tivoli applications, Tivoli partner applications, customer applications, network management platforms, and relational database systems.

What are events?

Events are units of information which can represent performance data or also can indicate problems, status or changes regarding resources. System Automation for Multiplatforms usually sends events when the help of an administrator is required.

A language called Basic Recorder of Objects in C (BAROC) is used to define the structure of events and their properties. These definitions are stored in files with the extension .baroc.

Sending events to the TEC

System Automation for Multiplatforms uses the TEC event interface ‘Tivoli Event Integration Facility (EIF)’ to send events to the TEC. The events will be sent to the TEC server non-TME event port.

The following reasons cause System Automation for Multiplatforms to send events to the TEC:

- The status of a resource or cluster changed.
- A resource was added or deleted.
- A relationship was added or deleted.
- A request was added or deleted.
Enabling the TEC publisher function

In order to receive and view the events coming from different sources like programs, systems or network devices you have to enable the TEC publisher by performing the following steps:


2. Copy files /usr/sbin/rsct/samples/tec/samPublisher.conf and /usr/sbin/rsct/samples/tec/TECPublisher.conf into /etc/Tivoli/tec on all System Automation for Multiplatforms cluster nodes.

3. Customize the publisher configuration file /etc/Tivoli/tec/samPublisher.conf and the TEC EIF file /etc/Tivoli/tec/TECPublisher.conf on each cluster node.

4. Enable the publisher with the command samctrl –e TEC on a System Automation for Multiplatforms cluster node. By default, the publisher is disabled. Use the command samctrl –e P to enable all Publishers specified in the Publisher configuration file at once.

Publisher configuration file

The publisher configuration file /etc/Tivoli/tec/samPublisher.conf specifies a list of all target consumers and their parameters. This is the syntax format of the publisher configuration file:

```bash
# Publisher configuration file
# file name: /etc/Tivoli/tec/samPublisher.conf
#
# File format:
#    <keyword>=<value>
#
# Publisher - unique name of the publisher
#    name length: 1-8 characters
#    valid characters: '0'-'9', 'A'-'Z', 'a'-'z' and '_'
# LibraryPath - name of the publisher library
# ConfigPath - full path to the TEC EIF configuration file
#
# Multiple entries of the Publisher, LibraryPath and ConfigPath can be specified.
# One triplet for each publisher target consumer.
# Maximum supported publishers: 15
#
# Online Update section ---------------------------------------------
# End Online Update section ---------------------------------------------

Publisher=TEC
LibraryPath=libTECPublisher.so
ConfigPath=/etc/Tivoli/tec/TECPublisher.conf

Publisher=TEC2
LibraryPath=libTECPublisher.so
ConfigPath=/etc/Tivoli/tec/TECPublisher2.conf
```

Figure 15. Syntax format and sample of the publisher configuration file

The following syntax rules apply:

- lines starting with # and blank lines will be ignored
- parameter format: <keyword>=<value>
- keyword "Publisher" starts a new triplet of "Publisher", "LibraryPath" and "ConfigPath" parameters
Controlling and administering System Automation for Multiplatforms

- keyword "Publisher" specifies the unique name of the publisher
- keyword "LibraryPath" specifies the full path to the publisher library
- keyword "ConfigPath" specifies the full path to the TEC EIF configuration file
- keyword "Publisher" value must have the length of 1-8 characters
- keyword "Publisher" value can only have the following characters: '0'-'9', 'A'-'Z', 'a'-'z' and ' '
- keyword "Publisher" values "RMC", "P", "ALL" and "STOPALL" are reserved.
- the maximum of 15 publishers can be specified in this configuration file

TEC EIF configuration file /etc/Tivoli/tec/TECPublisher.conf

The TEC EIF configuration file specifies all parameters needed to connect to a specific TEC server. The file name must match the name specified as "ConfigPath" parameter in the publisher configuration file.

The syntax format of the TEC EIF file for the TEC publisher is the existing TEC EIF configuration file syntax.
To avoid that the TEC is flooded with a huge amount of messages, filters are provided in the # TEC EIF configuration file. Per default all filters are enabled, which results in only critical messages being sent to the TEC. If you want additional messages to be sent to the TEC, disable the corresponding filter by using the comment character #.

## Enabling the publisher

The Publisher function is disabled by default. To query the status of the publisher, issue the following command:

```
nodel:/usr/sbin/rsct/samples/tec # lssamctrl
```
The following System Automation for Multiplatforms control information is displayed:

SAMControl:
- TimeOut = 60
- RetryCount = 3
- Automation = Auto
- ExcludedNodes = {}
- ResourceRestartTimeOut = 5
- ActiveVersion = [1.2.0.0,Fri Apr 16 16:05:50 2004]
- EnablePublisher = Disabled

To enable the TEC publisher, issue this command on any node:
```
node1:/usr/sbin/rsct/samples/tec # samctrl –e TEC
```

To disable the TEC publisher, issue this command on any node:
```
node1:/usr/sbin/rsct/samples/tec # samctrl –d TEC
```

To enable all publishers specified in the Publisher configuration file, issue this command on any node:
```
node1:/usr/sbin/rsct/samples/tec # samctrl –e P
```

To disable all publishers specified in the Publisher configuration file, issue this command on any node:
```
node1:/usr/sbin/rsct/samples/tec # samctrl –d P
```

**Setting a new language locale for the TEC event messages**

TEC event messages are always in the language which is the default system locale on the node where the IBM Tivoli System Automation for Multiplatforms master is running.

**Note:** Resource names in TEC event messages can be corrupted, if the user created the resources (mkrg, mkrsrc) in a shell with a different locale than the default system locale, or the terminal program has a different character set translation defined than the shell locale. To solve this problem, the system and shell locales must have identical settings and the character translation of the terminal program must be set accordingly. If the shell locale changes and resources are already created with the old shell locale setting, all resources must be deleted and have to be recreated with the new shell locale.

If the user chooses to adjust the default system locale to his preferred shell settings, then this change has to be done on all nodes of the cluster. Do the following to perform this:

1. Stop the cluster using the `stoprpdomain` command.
2. Edit the file containing the default system locale, set the appropriate values, and save the file.

**SUSE Linux**

File: `/etc/sysconfig/language`

Keywords: 
- `RC_LANG="<NewLocale>"` (replace `<NewLocale>` with your locale setting)
- `ROOT_USES_LANG="yes"`
- All keywords starting with `RC_LC_` must be set to empty strings `""`
- e.g. `RC_LC_ALL=""`

Chapter 10. Controlling and administering System Automation for Multiplatforms 213
Controlling and administering System Automation for Multiplatforms

Run `/etc/SUSEconfig` to apply the changes to your system. You can also use the `yast2 sysconfig` system configuration tool to apply the changes.

RedHat Linux

File: `/etc/sysconfig/i18n`
Keywords: `LANG="<NewLocale>">`
(replace `<NewLocale>` with your locale setting)

AIX

File: `/etc/environment`
Keywords: `LANG="<NewLocale>">`
(replace `<NewLocale>` with your locale setting)

3. Reboot the system.
4. Repeat the steps on all nodes in the cluster.
5. Start the cluster using the `startrpdomain` command.

Publishing System Automation for Multiplatforms internal attributes into the RSCT infrastructure

This function makes System Automation for Multiplatforms internal attributes known to the RSCT infrastructure. For this purpose, the resource classes `IBM.ResourceGroup`, `IBM.Equivalency`, and `IBM.ManagedResource` are extended with the dynamic data structure attribute `AutomationDetails`. The dynamic data structure `AutomationDetails` has the following attributes:

- **CompoundState** – overall status of the resource including group dependencies. Shows how far the resource has reached the DesiredState. For example, “Satisfactory” means that the resource or resource group has reached the requested user status.
- **DesiredState** – user requested status of the resource. For example, "online" means that the user requested that the resource should be online.
- **ObservedState** – actual status of the resource from an automation perspective. For example, "online" means that the resource is currently online.
- **BindingState** – status indicating if the resource is bound to a specific system. For example, "bound" means that the resource is currently bound to a specific system.
- **AutomationState** – status indicating if the resource is currently being automated. For example, "Idle" means that System Automation for Multiplatforms is currently not trying to start or stop the resource.
- **ControlState** – status indicating if the resource can be controlled by automation. For example, "startable" means that it is currently possible to start this resource.
- **HealthState** – health status of the resource. This is reserved for future releases.

The `lsequ` and `lsrg` commands have been extended to show these attributes.

Each change in the value of one of these attributes indicates a status change of a resource and will be published to RSCT. If the TEC publisher is enabled (see “Enabling the TEC publisher function” on page 210), these status changes are also shown as TEC events.
Enabling GDPS/PPRC Multiplatform Resiliency for zSeries

Today, businesses and companies depend on disaster recovery solutions to recover critical data. This is why System Automation for Multiplatforms supports GDPS/PPRC Multiplatform Resiliency on System z (xDR).

Geographically Dispersed Parallel Sysplex (GDPS) is an application availability and disaster recovery solution which is highly customized to work with your z/OS environment. It provides disaster and failure recovery from a single point of control and ensures data consistency. For more information about GDPS, see the IBM Redbooks publication GDPS Family - An Introduction to Concepts and Capabilities, which can be downloaded at http://ibm.com/redbooks.

System Automation for Multiplatforms allows you to extend GDPS/PPRC for Linux systems running on System z. It provides a coordinated disaster recovery solution for systems running on zSeries, including z/OS, Linux on System z under z/VM, and Linux on System z running native in the LPAR.

Notes:
1. If you want to use the xDR functionality, particular versions of z/VM, Linux on System z, GDPS, and System Automation for Multiplatforms must be installed. For detailed information about the available functionality and the required versions, refer to the GDPS manuals. System Automation for Multiplatforms supports xDR for Linux on System z only.
2. The xDR naming conventions require that the names of clusters and nodes must not exceed 32 characters. For xDR, cluster names are not case-sensitive. To use xDR, System Automation for Multiplatforms must be customized as described in the GDPS manuals.
3. English is the only language supported by xDR and GDPS.

Supported GDPS version

xDR requires GDPS V3.3 with APAR PK30315 or higher

Supported Linux distributions
The following Linux distributions are supported for xDR:
- xDR for Linux on System z running under z/VM requires one of the following operating systems:
  - SUSE SLES9 SP2
  - SLES10 SP1
  - RHEL 4.5
- xDR for Linux on System z running native in LPAR requires the following:
  - SUSE SLES 10 SP1
- For native LPAR support, GDPS 3.4 is required.
Resource verifications are usually performed at configuration time when a resource is defined to System Automation for Multiplatforms. Then the user immediately gets notified when a problem occurs, and the definition of a new resource fails.

This may, however, not be the case after a resource has been defined and then a configuration change occurs. System Automation for Multiplatforms gets notified after the configuration change has been made, and then it has to accept and react to these changes, which eventually may lead to one or more resources becoming invalid.

This may happen when you, for example, define a resource using the `mkrsrc` command, change the values of a resource using the `chrsrc` command, or remove a defined resource using the `rmrsrc` command.

In order to verify such configuration changes and to convey the validity of a resource to the user, the resource classes IBM.ResourceGroup, IBM.ManagedResource, IBM.Equivalency, and IBM.ManagedRelationship contain a dynamic attribute `ConfigValidity`. `ConfigValidity` contains a string which explains why the resource is invalid.

Use the `lsrsrc –Ad` command to display the value of `ConfigValidity` together with the values of the other dynamic attributes of a resource.

The following verifications are performed:

- **A resource group’s AllowedNode attribute is empty**
  When a node is removed, it may cause an equivalency to contain an empty member list. A resource group will become invalid if it uses this empty equivalency as its “AllowedNode” attribute. When this happens, the “ConfigValidity” dynamic attribute of the resource group will contain string “Allowed node list is empty”.

- **Intersection of nested resource group’s AllowedNode is empty**
  In a collocated resource group, all nested inner resource groups and the containing resource group must have at least one node in common. If there is only one node in common, and that node is removed, all the resource groups become invalid. When this happens, the “ConfigValidity” dynamic attribute of the resource group will contain the string “No common node in collocated nested resource group”.

- **No node to run a resource**
  In a resource group, there may be the case when there is only one node in common between the resource group’s AllowedNode and a member resource’s NodeNameList. If this node is removed, the resource group becomes invalid. When this happens, the “ConfigValidity” dynamic attribute of the resource group will contain string “No common node to start a resource”.

- **No node to satisfy a relationship**
  In a DependsOn relationship, with implied collocation, the NodeNameList of the source and NodeNameList of the target resource must have at least one node in common. If this node is removed, the relationship becomes invalid. When this happens, the “ConfigValidity” dynamic attribute of the managed relationship will contain string “No common node between source and target”.

- **Cannot satisfy an AntiCollocated relationship**
If two mandatory floating resources in a resource group have an AntiCollocated relationship among each other, and removal of nodes leaves only one node in the resource group’s AllowedNode, the resource group becomes invalid. When this happens, the "ConfigValidity" dynamic attributes of the resource group and the AntiCollocation managed relationship will contain string "An AntiCollocated relationship cannot be satisfied".

- **Cannot satisfy an AntiCollocated relationship - 2**
  A removal of nodes causes two floating resources to have only one constituent left on the same node. But the two have an AntiCollocated relationship. When this happens, the "ConfigValidity" dynamic attributes of the AntiCollocated managed relationship will contain string "An AntiCollocated relationship cannot be satisfied".

- **Propagated invalidity**
  Any invalid inner resource group will cause all enclosing resource groups to become invalid. When this happens, the "ConfigValidity" dynamic attribute of the affected resource group will contain string "An enclosed resource group is invalid."
Enabling IPv6 Support

System Automation for Multiplatforms supports IPv6 on the following platforms: AIX, Linux, Solaris, and Windows Server 2008. In order to use IPv6 with System Automation, you need to have the operating system set up for IPv4 and IPv6. Normal RSCT cluster operations use IPv4 connections, but IBM.ServiceIP resources can be defined to use IPv6 addresses.

In order to enable the IPv6 support in RSCT and System Automation for Multiplatforms, you need to issue the following command:

```bash
chrsrc -c IBM.NetworkInterface IPv6Support=1
```

This creates additional IBM.NetworkInterface resources for IPv6–enabled interfaces as well. You now have two IBM.NetworkInterface resources per physical interface: one for IPv4 and one for IPv6. Examples of how to create IBM.ServiceIP resources with IPv6 addresses are contained in "What is the IBM.ServiceIP resource class?" on page 236. Note that a new IBM.ServiceIP class attribute named Netprefix has been defined for use with IPv6.

Hints and tips

This section provides various hints and tips which are helpful when operating a cluster with System Automation for Multiplatforms.

Rebooting a node

Do not reboot a node when any resources are running on this node. First stop all running resources, using the following command:

```bash
samctrl -u a <node_name>
```

Now you can safely reboot the node.

This command also prevents resources from being started on this node. Re-enable the node for running resources by entering:

```bash
samctrl -u d <node_name>
```

Stopping a node

Before you stop a node (e.g. with the RSCT command `stoprpnode`), you must exclude this node from automation with the following command:

```bash
samctrl -u a <node_name>
```

The exclude must be done even if there are currently no resources online on the node.

After you have started the node again and it is online, you must reestablish automation on that node by entering:

```bash
samctrl -u d <node_name>
```

Generating events in case of failures

RSCT has the ability to generate events in case a dynamic attribute of a resource changes. This is done by the event response resource manager (ERRM). The event response resource manager provides a set of commands that enable you to monitor events of interest (called conditions) and have the RMC system react in particular ways (called responses) if the event occurs.
Controlling and administering System Automation for Multiplatforms

So, for example, you can subscribe for the OpState of a resource group and get an email if the status changes. You can also monitor different resources in your system which are critical. For further explanation how to generate such events, see the chapter on basic resource monitoring in *IBM Reliable Scalable Cluster Technology for Linux, Technical Reference, SA22-7893.*
Chapter 11. Automating System Automation for Multiplatforms resources

This chapter describes how to automate resources for which System Automation for Multiplatforms provides resource managers.

Using the Global Resource Manager

This section describes the characteristics of the Global Resource RM.

The Global Resource RM (IBM.GblResRM) provides the following two resource classes:

1. IBM.Application:
   This class allows additional types of resources (for example, business applications) to be monitored and controlled through the RMC subsystem. These resources can then be automated or recovered by management applications such as System Automation for Multiplatforms.

2. IBM.ServiceIP:
   This class is used to manage IP addresses that can be started, stopped, and moved between network adapters and nodes within a peer domain under the control of the RMC subsystem. These IP addresses will typically be provided to clients that are connecting to some service that is running within the domain. System Automation for Multiplatforms can be used to keep the service and its associated IP address active, even through failures within the domain.

The resource manager (and access to its classes) is operable in peer domain mode only.

The following subsections describe the external characteristics of the resource classes that are supported by this resource manager. Each of the subsections describes one particular resource class, including its persistent and dynamic attributes, actions, etc.

What is the IBM.Application resource class?

The IBM.Application resource class allows new types of floating and fixed resources to be created, monitored, and controlled through the RMC subsystem. These resources can then be automated or recovered by System Automation for Multiplatforms. In order to create a new resource, the following three scripts (resp. commands) must be provided:

1. A start script (or command) to bring the resource online.
2. A stop script (or command) to take the resource offline.
3. A script (or command) to monitor the resource through polling.

Besides these scripts, there are the following basic parameters to the IBM.Application resource class:

1. The name of the resource.
2. The nodes where the resource can run.
3. A user name used to start/stop/monitor the application.
4. Method to be used to synchronously or asynchronously start/stop the application.
5. Different timeouts values.
6. Characterization of the resource as either critical or non-critical.
7. Determination of the monitoring frequency.
8. Identification of the resource as fixed or floating.

Each generic resource that is instrumented through the IBM.Application resource class is considered to be a global resource meaning that it is not tied to a single node. However, the resource may be defined to exist on only a subset of the nodes of the cluster. For each generic resource, one instance of the IBM.Application resource class must be created. This instance is called an aggregate resource since it represents the floating resource that can move between nodes. In addition, there will be one instance of the IBM.Application resource class for each node where the generic resource exists. These are called constituent resources of the aggregate resource. Constituent resources are fixed resources in the sense that they exist on exactly one node of the cluster. Figure 17 illustrates the difference between aggregate and constituent resources.

![Diagram of aggregate and constituent resources](image)

The constituent resources are automatically created or deleted as the definition of the aggregate resource is changed. Most management operations are done through the aggregate resource, but some applications may choose to monitor or operate on the constituents directly. Changes made to the aggregate resource are automatically applied to all constituents, whereas the change of an attribute of a constituent affects this constituent only and is not delivered to other resources (for example, the constituent on one node may have a different start command or monitoring interval).

When nodes are removed from the NodeNameList of the aggregate resource, the constituents are deleted automatically.

**Attributes used by IBM.Application**

This section describes the attributes that are used by resources of the IBM.Application resource class.

When a resource of this class is created with the RMC command `mkrsrc`, it must have these attributes:

- Name
- StartCommand
- StopCommand
- MonitorCommand
Resources of this class may have these attributes:

- **NodeNameList**
- **ResourceType**
- **StartCommandTimeout**
- **StopCommandTimeout**
- **MonitorCommandTimeout**
- **MonitorCommandPeriod**
- **RunCommandsSync**
- **ProtectionMode**

Resources of this class have the following dynamic attribute:

- **OpState**

**Name attribute**

The Name persistent attribute is a user-defined name for the generic application resource. Both the aggregate and constituent resource will have the same value for this attribute. A value for this attribute must be specified when a new IBM.Application resource is created and it must be unique. The attribute must be of type character string.

**NodeNameList attribute**

The NodeNameList persistent attribute is an array of strings that indicates which nodes the IBM.Application resource is available on. If the resource is floating, the Global Resource RM will ensure that there is one constituent (i.e. fixed) resource for each node name in this list. Constituent resources are implicitly created or deleted as necessary to match the entries in this list. Constituent resources will only contain one entry in this array since they are fixed resources and thus only available on one node. This attribute for a floating resource is implicitly modified if a constituent resource is explicitly removed by the administrator so that the aggregate and constituent resource relationship is always consistent. This list may be empty for a floating resource which means that it is not available anywhere and constituents may be added separately. The list may contain at most one name if the resource is fixed (i.e. ResourceType=0). If no name is given for a fixed resource, then RMC will provide a default because the fixed resource is tied to a node and therefore cannot be created without a node name or node ID. The value of this attribute for aggregate resources may be changed with the **chrsrc** command. An attempt to modify this attribute for a constituent resource will generate an error.

**ResourceType attribute**

With the ResourceType persistent attribute you identify whether the resource is fixed or floating. An integer value of 0 indicates that the resource is fixed, a value of 1 indicates that it is a floating resource. This attribute defaults to floating if not specified when a new IBM.Application resource is created.

**StartCommand attribute**

The value of the StartCommand persistent attribute contains the...
exact command and arguments that will be executed when the resource manager receives a start request for the corresponding resource instance. The command is only executed by constituent resources even if the online request was issued to the aggregate resource. In this case, the resource manager will choose a constituent resource if not specified to execute the online request. The command is executed under the user id specified with the UserName attribute. The command is run with the authority and environment of the specified user. Whether the resource manager waits for the command to complete is controlled by the RunCommandsSync attribute (see below for details). The command name must be a character string and it must be an absolute path (i.e. it must begin with a `/`). It must exist and be executable on each node where the resource is accessible (i.e. where there is a constituent). This attribute must be specified when a new IBM.Application resource is defined.

The StartCommand of an IBM.Application resource must adhere to the POSIX.2 standard with respect to directory paths; it must not contain characters required multi-byte encodings.

The command may return the following values:

0 Command has run successfully.
!= 0 Error occurred during command processing.

See “How the resource manager handles return codes of the StartCommand, StopCommand, and MonitorCommand” on page 229.

StopCommand attribute

The value of the StopCommand persistent attribute contains the exact command and arguments that will be executed when the resource manager receives a stop request for the resource instance. A stop request for the aggregate will cause all constituent to run the StopCommand. All other aspects related to running the command are the same as for StartCommand. This attribute must be specified when a new IBM.Application resource is defined.

The StopCommand of an IBM.Application resource must not contain NLS characters in its path or in any of its options.

The command may return the following values:

0 Command has run successfully.
!= 0 Error occurred during command processing.

See “How the resource manager handles return codes of the StartCommand, StopCommand, and MonitorCommand” on page 229.

MonitorCommand attribute

The value of the MonitorCommand persistent attribute contains the exact command and arguments that will be executed periodically
to determine or update the operational state (OpState attribute) of the resource. The exit value from the command is used as the new OpState of the resource:
  Unknown=0
  Online=1
  Offline=2
  Failed Offline=3
  Stuck Online=4
  Pending Online=5
  Pending Offline=6

At least the Online and Offline status should be set by the MonitorCommand script.

The IBM.GblResRM runs this command every MonitorCommandPeriod seconds when there are any subscribers to the OpState dynamic attribute. All other aspects of running the command are identical to those described under StartCommand. To avoid consuming system resources, this command should be as efficient as possible.

The command should return within a reasonable amount of time, usually within a few seconds, at least within 360 seconds. The name of the MonitorCommand must be an absolute path (i.e. it must begin with a ‘/’). It must exist and be executable on each node where the resource is accessible (i.e. where there is a constituent).

This attribute must be specified when a new IBM.Application resource is defined. To learn more about the return value of the MonitorCommand attribute see “Important issues when defining IBM.Application resources” on page 228 and “How the resource manager handles return codes of the StartCommand, StopCommand, and MonitorCommand” on page 229.

The MonitorCommand of an IBM.Application resource must not contain NLS characters in its path or in any of its options.

**MonitorCommandPeriod attribute**

The value of the MonitorCommandPeriod persistent attribute specifies the amount of time (number of seconds) to wait between invocations of the MonitorCommand. This period is started after the prior invocation completes. This attribute must be of type integer and it must be greater than 0. It defaults to 5 seconds.

Since the period starts after the prior invocation completes, the value of the MonitorCommandTimeout attribute can be greater than the value of the MonitorCommandPeriod attribute.

**MonitorCommandTimeout attribute**

With the MonitorCommandTimeout persistent attribute you specify the amount of time a monitor command is allowed to run before it is killed via killpg(). If the command times out, the operational state (OpState attribute) of the resource is set to Unknown=0. This attribute must be of type integer and it must be greater than or equal to 0. The value must be lower than 360 seconds. The default value for this attribute is 5 seconds.

**StartCommandTimeout attribute**

With the StartCommandTimeout persistent attribute you specify
the amount of time a start command is allowed to run before it is killed via killpg(). This attribute also specifies the amount of time after which System Automation for Multiplatforms expects the resource to be online; that is, System Automation for Multiplatforms uses this value instead of the default timeout given with the control parameters. This attribute must be of type integer and it must be greater than or equal to 0. A value of 0 for this attribute means no timeout. The attribute is not used if the RunCommandsSync attribute is set to 0. The default value for this attribute is 5 seconds.

StopCommandTimeout attribute
With the StopCommandTimeout persistent attribute you specify the amount of time a stop command is allowed to run before it is killed via killpg(). This attribute must be of type integer and it must be greater than or equal to 0. A value of zero means no timeout. The default value for this attribute is 5 seconds.

RunCommandsSync attribute
You use the RunCommandsSync persistent attribute to control whether the start/stop commands are executed synchronously with the online() / offline() method. If the value of this attribute is set to the integer value 1, which is the default, then the response to the online() / offline() methods will not be completed until the command completes or times out. Any stderr / stdout outputs will be returned in the response for this case. If the value of this attribute is 0, then the IBM.GblResRM will “fire and forget” the start/stop commands. As soon as the fork/exec completes successfully, the resource manager forgets about them and they run completely unattached from the resource manager. This attribute defaults to 1. Timeouts will not be applied to commands when this attribute is set to 0.

If your StartCommand is the application executable and the command does not return in a certain amount of time, but is running as long as your application is running, you cannot use the synchronous mode. In case of the synchronous mode the resource manager would kill the command after the StartCommand timeout has expired.

Use RunCommandsSync=1 when you know how long the StartCommand normally takes to complete. Adapt the StartCommandTimeout attribute to this time. During heavy system load this time may be longer even in case of a normal start up without any error conditions.

Use RunCommandsSync=0 when the command does not complete until your application is running (e.g. the executable of your application). If you want to run the application executable directly (asynchronous mode), but use the synchronous command mode and send the command to the shells background, redirect I/O file descriptors. If the resource manager is still connected to the process, I/O descriptors the StartCommand timeout will kill your process.

The RunCommandsSync attribute also controls the environment the start/stop and monitor command is executed in. You can either select a basic environment for the commands or the resource manager can run a full-blown login environment including the
user profile and shell configuration files. You can compare this to the system command 'su' (switch user). Either you keep the current environment or you can do a 'su –' and run the full profile (login shell) for the selected user. Following values are allowed for the RunCommandsSync resource attribute:

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Run commands in asynchronous mode. Run commands in a limited environment.</td>
</tr>
<tr>
<td>1 (default)</td>
<td>Run commands in synchronous mode. Run commands in a limited environment.</td>
</tr>
<tr>
<td>2</td>
<td>Run commands in asynchronous mode. Run profile and login scripts for the specified user and build the users login environment.</td>
</tr>
<tr>
<td>3</td>
<td>Run commands in synchronous mode. Run profile and login scripts for the specified user and build the users login environment.</td>
</tr>
</tbody>
</table>

Based on the operating system (AIX, Solaris, or Linux) the limited environment for the user contains following environment variables (values of the variables represent a sample user):

- SHELL=/bin/bash
- USER=myuser
- PATH=:/usr/ucb:/bin:/usr/bin
- LOGIN=myuser
- PWD=/home/myuser
- LANG=de_DE.UTF-8
- HOME=/home/myuser
- SHLVL=2
- LOGNAME=myuser

If this is not sufficient for the specified start/stop and monitor commands, either set the necessary variables in the start/stop or monitor command itself, or put them into the users profile (or shell configuration) and select the login environment for the commands (RunCommandsSync=2|3).

**UserName attribute**

The UserName persistent attribute defines a user name under which the MonitorCommand, StartCommand and StopCommand are run. The commands are run with the authority and environment of the specified user. A check will be made on each node to ensure that the user name exists whenever the configuration of the resource is modified. This attribute must be specified when a new IBM.Application resource is defined. The attribute must be of type character string.

**ProtectionMode attribute**

You use the ProtectionMode persistent attribute to specify whether the resource is critical. If it is critical, then the IBM.ConfigRM decides if the resource can be started as requested. (For further detail on this behavior, see Chapter 9, “Protecting your resources – quorum support,” on page 161).

The attribute may have the integer values 0 (Non-Critical) or 1 (Critical), it defaults to Non-critical. If the resource is set to Critical, monitoring will immediately start, even if no subscriber to this resource exists.
OpState attribute

The value of this dynamic state attribute contains the operational state of the resource as determined by the exit code from running the MonitorCommand periodically. The possible values for this attribute as defined by the state transition diagram in the RMC architecture are:

- Unknown=0
- Online=1
- Offline=2
- Failed Offline=3
- Stuck Online=4
- Pending Online=5
- Pending Offline=6

This attribute is available from both the aggregate resource and all constituent resources. The value for an aggregate resource is a roll-up of the states from each of the constituents.

HealthCommand

Reserved for future use.

HealthCommandPeriod

Reserved for future use

HealthCommandTimeout

Reserved for future use

InstanceName

Reserved for future use

InstanceLocation

Reserved for future use

Actions used by IBM.Application

This section describes the actions that can be performed on resources of the IBM.Application resource class.

The refreshOpState action

In normal operation the MonitorCommandPeriod attribute determines the interval, the OpState of an IBM.Application resource is evaluated by its monitoring script. In case a resource is able to detect a failure by itself, it is possible to trigger an immediate run of the monitoring script which is monitoring the resource. This will result in an immediate OpState refresh of the application resource.

For example, to refresh the OpState of the resource 'WebServer' running on node02, issue the following command on any node:

```
runact -s "Name='WebServer' && NodeNameList=['node02']" IBM.Application refreshOpState
```

SendEIFevent Reserve for future use.

Important issues when defining IBM.Application resources

Keep the following in mind when defining an IBM.Application resource:

1. In order to satisfy goal driven automation System Automation for Multiplatforms resources are always monitored. No matter if the resource is started or stopped, the monitor command has to determine the actual state of
Global resource RM

the resource. Do not put the monitor command on a file system which is not always present (for example, a NFS mount which is part of a policy and is only mounted if the NFS resource is started). If you see an IBM.Application OpState of “unknown”, check the system log and take care the GblResRM resource manager can access the monitor command at any time.

2. The GblResRM resource manager will kill any command which is running longer than the timeout value allows. If you see IBM.Application resources reporting OpState “unknown” during heavy system load, the command probably takes longer to complete than the timeout value of this resource allows. If there is more CPU time, this will not lead to problems. The monitor command will be able to complete and report a valid OpState again. Check the system log to see if the command was killed cause of the timeout. If commands are killed during normal operation, check and adjust the timeout values in your resource definition.

3. The monitor must clearly identify the process or application it is responsible for. If the monitor, for instance looks into the process table for a particular process, this process must be uniquely identified. If the monitor monitors another process than automation actually controls, the behavior of the whole resource would be somehow unpredictable.

4. If you try to automate base system services like printer spooler or mail transport agent, take care these services are only started or stopped by the automation and not by the system runlevel or init process.

5. If you have a more sophisticated monitor which could not only determine if a process exists but also if the application is able to deliver service, there is no possibility to report this fact to the automation engine and trigger a stop or shutdown of this application. If the monitor finds out that the process or application is available but hang or stuck, the monitor itself should kill this process/application. Automation will recognize the application is not available anymore (next monitoring) and will try to restart the application in place or move it to another node.

6. If you try to remove an online resource of class IBM.Application (e.g. with rmgrc), the command will be rejected. You can force removal by setting Force=1. For example, to remove a resource called WebServer you would enter:

   rmgrc -s "Name=='WebServer' && ResourceType==1" IBM.Application Force=1

7. The soft limit for the number of open files is 1024. The hard limit is at the maximum, so the soft limit can be adjusted if needed within the scripts executed as StartCommand, StopCommand, or MonitorCommand for a resource.

   To increase the limit, add the following statement below the header of the *Command scripts for the resource:

   # ulimit -S -n <new-limit-for-number-of-open-files>

   This only has an impact on applications requiring a huge number of open files, like, for example, WebSphere Application Server.

How the resource manager handles return codes of the StartCommand, StopCommand, and MonitorCommand

The resource manager handles the return codes for the StartCommand, StopCommand, and MonitorCommand as follows:

StartCommand:

1. If the StartCommand was able to start the resource it should return a value of 0 to indicate that the resource was properly started and should go online within the next few seconds.
Global resource RM

2. If the StartCommand was not able to start the resource it should return a value other than 0. This signals the automation not to start the resource again and set the resource operational state to Failed Offline. This indicates that you have to manually intervene to fix the resource. When then the resource is able to start, reset the Failed Offline operational state with the `resetsrc` command. Note that whenever a StartCommand failed, automation will issue a StopCommand to ensure application leftovers of the failing start are removed.

3. If the StartCommand completes successfully and returns a value of 0, but the MonitorCommand of the resource does not report an online state after a certain amount of time (depending on the settings in the automation control configuration), the automation will try to restart the resource. The total number of start attempts is defined in the RetryCount attribute (see the description of the `lssamctrl` command in [IBM Tivoli System Automation for Multiplatforms Reference](#)), the default value is 3. If the operational state of the resource does not change to Online after the last attempt, a StopCommand is issued and the resource is set to Failed Offline.

4. If the StartCommand of a resource is not able to complete within the time specified in the StartCommandTimeout attribute, the resource manager will kill the StartCommand and treat the start like a failing start command as described under item 2.

5. In case the StartCommand was valid when the resource was defined, but is later removed or not present (for example, because of a missing NFS mount), the start procedure is treated like a failing StartCommand as described under item 2.

StopCommand:

1. If the StopCommand was able to stop the resource it should return a value of 0 to indicate that the resource was properly stopped and should go offline within the next few seconds.

2. There is no mechanism to handle a failed StopCommand in IBM Tivoli System Automation for Multiplatforms. The StopCommand may indicate a failing stop of the application by returning a value other than 0, but this will not trigger an automation action.

   If the resource does not reach the OpState Offline after a certain amount of time, IBM Tivoli System Automation for Multiplatforms issues a reset operation against the resource, which results in a second invocation of the StopCommand. This second invocation of the StopCommand can be determined within the StopCommand script by checking the SA_RESET environment variable, which is set to 1 for the second invocation. A StopCommand can honor this, like in the following example:

   ```bash
   #/bin/sh
   # A sample stop/reset automation script for the lpd application
   if [ $SA_RESET == 1 ]; then
     killall -9 lpd
     exit $?
   else
     /etc/init.d/ldpd stop
     exit $? 
   fi
   ```

   If a second execution of the StopCommand is not desired, the script could just exit, for example:

   ```bash
   #/bin/sh
   # A sample stop/reset automation script for the lpd application
   if [ $SA_RESET == 1 ]; then
     exit 0
   ```
Global resource RM

```
else
    /etc/init.d/lpd stop
    exit $?;
fi
```

This reset feature was introduced in IBM Tivoli System Automation for
Multiplatforms, release 2.2. However, rewriting existing policy scripts is not
required as they will work as before. But by evaluating the SA_RESET
environment variable, an enhanced stop behavior can now be easily achieved.
If the second StopCommand also fails to bring the operational state of the
resource to Offline, the OpState of the resource is set to Stuck Online. At this
time, operator invention is required to stop the resource. After the resource has
finally stopped and the OpState has changed to Offline, IBM Tivoli System
Automation for Multiplatforms will again control this resource.

3. If the StopCommand of a resource is not able to complete within the
   StopCommandTimeout, the resource manager will kill the command and treat
   the stop like a failing StopCommand as described under item 2 on page 230.

4. A non-zero return code of the StopCommand or reset operation will result in
   the resource being set to automation state 'Problem'. This situation can only be
   recovered by a successful execution of resetsrc against this resource. Since the
   execution of resetsrc against a resource triggers the StopCommand of the
   resource, the StopCommand needs to run successfully as well (that is, it must
deliver a return code of zero). In general, it is recommended that the
   StopCommand of a resource be configured to return with a non-zero return
   code only if there is a real problem to stop the resource. For all other situations,
   the StopCommand should be configured to return with a zero return code in
   order to prevent the resource from ending up in the 'Problem' state.

5. In case the stop command was valid when the resource was defined, but is
   later removed or not present (for example, because of a missing NFS mount),
   the stop procedure is treated like a failing StopCommand as described under
   item 2 on page 230.

MonitorCommand:

1. If the MonitorCommand was able to determine the operational state of a
   resource it should return one of the valid RMC operational states (see page
   228). Keep in mind that in this case 0 is not the return value for RMC
   operational state Online, but the return value for operational state Unknown,
   which is the most critical state for the automation. A resource with an unknown
   operational state will not be automated anymore, and this may also affect other
   resources which have dependencies to this resource.

2. If the monitor command of a resource is not able to complete within the
   MonitorCommandTimeout, the resource manager will kill the
   MonitorCommand and set the RMC operational state to Unknown, which
   indicates a major problem with the resource. There will be no automation with
   this resource, until the MonitorCommand returns an operational state other
   than Unknown.

3. In case the monitor command was valid when the resource was defined but is
   later removed or not present (for example, because of a missing NFS mount),
   the operational state is set to Unknown, which indicates a major problem with
   the resource.

4. In both cases the MonitorCommand may continue to report valid RMC
   operational states after system load has decreased or NFS is present again, and
   now automation will continue automating the resource.

5. During the execution of the StartCommand of a resource, the OpState of the
   resource will always be set to Pending Online, regardless of the actual OpState
reported by a MonitorCommand, to reflect that the resource is actually in the process of starting. There are two exceptions - if the MonitorCommand returns Online or Failed Offline – because these OpState values indicate that the resource is already online or broken and cannot be started.

6. During the execution of the StopCommand of a resource, the OpState of the resource will always be set to Pending Offline, regardless of the actual OpState reported by a MonitorCommand, to reflect that the resource is actually in the process of stopping. There are three exceptions - if the MonitorCommand returns Offline, Failed Offline, or Stuck Online – because these OpState values indicate that the resource is already offline or stuck and cannot be stopped.

How the Global Resource Manager creates processes for the StartCommand, StopCommand, and MonitorCommand of IBM.Application resources

The first process in the UNIX/Linux kernel is the init process which creates and starts (spawns) the system resource controller (src). As shown in the following figure, the system resource controller is responsible for the System Automation for Multiplatforms resource managers.

```
Init+-atd
   |+-srcmstr
```

The Global Resource Manager creates additional processes to run the start/stop and monitor commands of an IBM.Application resource. All commands created by the Global Resource Manager run in the shell of the specified user. Here is some pseudo code of this functionality:

```
{ fork;
  if child
    switch to specified user ID;
    run the users default shell and execute the command e.g.
    bash -c /usr/bin/mycommand;
  endif
}
```

The command itself, such as mycommand in the previous shown pseudo code, can be any executable including shell scripts which may create additional processes or use job control.

The following describes various scenarios:

1. IBM.Application is defined in asynchronous command mode:

   StartCommand="/usr/bin/mycommand"
   RunCommandsSync=0

   In this case the Global Resource Manager creates (forks) a process for the command, then it completely detaches from the new process and closes all file descriptors to the new process.
   The Global Resource Manager does not take care of the new process anymore, and in theory the command could run forever.
   Since the new process does not have a parent anymore, it becomes an orphan and is adopted by the init process. When the process eventually ends, init will collect the return code of the process.

2. IBM.Application is defined in synchronous command mode:
Global resource RM

StartCommand="/usr/bin/mycommand"
RunCommandsSync=1

Here the Global Resource Manager does not detach itself from the newly created (forked) process. File descriptors are open, and the resource manager waits for completion of the command:

- If the command returns a bad return code, messages from stderr are captured and written to the Global Resource Manager trace and error block of the start/stop command.
- If the command does not return within the time specified in the StartCommandTimeout attribute, the Global Resource Manager sends SIGKILL to the forked process (the users default shell). SIGKILL is propagated to all child processes of the user shell and therefore all child processes should end.

3. IBM.Application is defined in synchronous command mode, but uses job control of the user shell:

StartCommand="/usr/bin/mycommand &"
RunCommandsSync=1

In this setup, the user default shell will not end until all child processes which have open file descriptors to the shell have ended. Here the time specified in the StartCommandTimeout attribute also applies to mycommand, even it is run in background of the user shell. So if mycommand runs longer as the time specified in the StartCommandTimeout attribute, the Global Resource Manager will send SIGKILL to the user shell and it will propagate SIGKILL to all background process.

If you want to make this setup working, you must make sure that all file descriptors of the shell child process are detached from the shell. This may look like:

StartCommand="/usr/bin/mycommand > /dev/null 2>&1 &"
RunCommandsSync=1

Now the user shell can end right after it has created (forked) the process for the command and has detached from the file descriptors of the new process. The command (mycommand) itself behaves as described in the first scenario, it becomes an orphan and is adopted by the init process.

Example: Implement the lpd printer spooler as an IBM.Application resource

The following example shows how to prepare the lpd printer spooler on a SUSe based Linux system to be managed by System Automation for Multiplatforms.

1. Remove the lpd from the default runlevel of the system. If you want to run this resource as a floating resource on more than one node, you have to check the runlevel on each node.

2. For the start and stop command of the IBM.Application use the default init scripts shipped with your lp daemon:

StartCommand: /etc/init.d/lp start
StopCommand: /etc/init.d/lp stop

3. For the monitor command we use a simple shell script which checks for the lpd process in the process table:

File: /root/lpmon

```
#!/bin/bash

OPSTATE_ONLINE=1
OPSTATE_OFFLINE=2
```
Alternatively you can use the **pidmon** command shipped with System Automation for Multiplatforms. It basically searches the process table for a given command string. If the command string was found, the RMC OpState is returned. See [IBM Tivoli System Automation for Multiplatforms Reference](https://www.ibm.com) for a detailed description of this command.

MonitorCommand: `/root/lpmon`

or

MonitorCommand: `/usr/sbin/rsct/bin/pidmon '/usr/sbin/lpd'`

4. In case of a floating resource make sure that all nodes can access the start/stop and monitor command under the same path. Since the lpd is a small and simple application, the default Start-/Stop- and MonitorCommandTimeout values (default is 5 seconds) can be used. In order to start lpd via the init scripts provide root as the user name for the IBM.Application.

Now the IBM.Application resource can be defined using the **mkrsrc** command:

```
# mkrsrc IBM.Application \
Name = "line_printer_daemon" \
ResourceType = 1 \
StartCommand = "/etc/init.d/lpd start" \
StopCommand = "/etc/init.d/lpd stop" \
MonitorCommand = "/usr/sbin/rsct/bin/pidmon '/usr/sbin/lpd' " \
MonitorCommandPeriod = 15 \
MonitorCommandTimeout = 5 \
StartCommandTimeout = 5 \
StopCommandTimeout = 5 \
UserName = "root" \
RunCommandsSync = 1 \
ProtectionMode = 0 \
NodeNameList = "{\'node01\','node02\'}"
```

This command results in three resources being created: An aggregate resource named "line_printer_daemon" which can potentially be brought online on nodes "node01" and "node02" and two constituent resources also named "line_printer_daemon", one on node "node01" and the other on node "node02". If a start request is issued against the aggregate resource, then the Global Resource RM chooses one of the constituents and starts it with the script (or command) specified with the StartCommand attribute.

**Configuring a supporting resource for an IBM.Application resource**

If you use an IBM.Application in combination with an IBM.Equivalency and DependsOn relationship, automation will choose a resource from the equivalency and provide this resource as a supporting resource to the IBM.Application.
In this sample configuration automation picks an application Ax from the equivalency and takes care application B will be depend on this application (see the chapters Chapter 5, “Using equivalencies,” on page 51 and Chapter 6, “Using managed relationships,” on page 57). Since more resources from the equivalency can fulfill the location constrains of the DependsOn relationship (in this sample A1 and A2 can run on Inxcm1), automation will provide the information which resource was selected to the start command of resource B.

If necessary the start script of resource B can use this information to pass special parameters to its application, or activate dedicated code which has to be carried out in combination with the selected resource from the equivalency.

To pass this information to the start command, the resource manager will set following environment variables in the start command environment of the resource:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA_SUPPORTING_RESOURCE_RH</td>
<td>The resource handle of the supporting resource, such as &quot;0x601d 0xffff 0xcac15160 0xbad91087 0x0f933128 0x5888898&quot;</td>
</tr>
<tr>
<td>SA_SUPPORTING_RESOURCE_NAME</td>
<td>The name of the supporting resource, such as &quot;A1&quot;</td>
</tr>
</tbody>
</table>

Here is some sample code which shows how the start script for resource B can honor the supporting resource environment variables:

```bash
# start command for resource B with logic for
# supporting resource
...
if [ $SA_SUPPORTING_RESOURCE_NAME = "A1" ]
then
  # start resource B and connect it to supporting
  # resource A1
  ...
fi

if [ $SA_SUPPORTING_RESOURCE_NAME = "A2" ]
# start resource B and connect it to supporting
```
What is the IBM.ServiceIP resource class?

The IBM.ServiceIP resource class is used to manage IP addresses that can be started, stopped and moved between adapters and nodes within a peer domain. Each resource of this class identifies one IP address. These IP addresses will typically be provided to clients that are connecting to some service that is running within the domain. A recovery management application can be used to keep the service and its associated IP address active, even through failures within the domain.

A Service IP address is an aggregate resource with one constituent resource per node where the administrator wants to allow that resource to be potentially brought on line. It is a floating aggregate resource since it can only have one constituent active at a time.

The IBM.ServiceIP resource class uses the following basic parameters:
1. The name of the resource.
2. The nodes where the resource can run.
3. The IP address which can be moved.
4. The netmask of the IP address.

The broadcast address and flag for the network interface will be taken from the parent interface where the ServiceIP resource is aliased on during the start.

Keep in mind that IBM.ServiceIP will generate a static routing entry for the network the IBM.ServiceIP is in. Take care that this network/route will not destroy the network configuration of the device the ServiceIP is aliased on.

Also automation does not take care of dynamic routing. If you specify a ServiceIP which is not in the subnet of the parent network interface, this physical device will host two different networks. Be sure to properly set up routing outside the automation cluster in order to support this network.

Characteristics of IBM.ServiceIP

The IBM.Service IP class has two different flavours:

1. **IBM.ServiceIP automatically chooses a suitable network interface**
   - If the IBM.ServiceIP receives a start request, the resource manager tries to choose a suitable network interface. If it can find one, the service IP is aliased on the that interface. This interface is called the supporting resource or supporting network interface.

   In order to determine a suitable network interface, the resource manager compares the IPAddress attribute of the IBM.ServiceIP resource to the IP address of all existing network interfaces. If there is a suitable network (subnet) match found, the resource manager assigns the alias address to that network interface. If no matching subnet is found, the online request fails.

   In case of the automatic interface choosing algorithm automation has no chance to evaluate the actual state of the network interface. As long as the network interface is configured and running, the ServiceIP can be assigned to the network interface. Keep in mind that UNIX/Linux do not change the status of
a network interface, for example if the network cable is unplugged. Even if the
device driver is able to detect the missing cable, the interface will stay
configured and running.

If you want to exploit RSCT heartbeat mechanism to detect a network interface
failure like a missing cable, use the supporting resource setup of the
IBM.ServiceIP described in the next section.

Example
An IBM.ServiceIP is defined as:

IPAddress=192.168.1.5
NetMask="255.255.255.0"
NodeNameList="{'node01','node02'}"

The following network interfaces are available in the cluster:

1. IPAddress=192.168.1.1
   Netmask=255.255.255.0
   NodeNameList="{'node01'}"
2. IPAddress=9.152.172.91
   Netmask=255.255.255.0
   NodeNameList="{'node02'}"
3. IPAddress=192.168.2.1
   Netmask=255.255.255.0
   NodeNameList="{'node03'}"

Only interface Number 1 is able to hold the service IP. All other interfaces do
not match the address (subnet) of the service IP. Calls to start the IBM.ServiceIP
on other nodes than node01 will fail, because there is no suitable supporting
network interface.

2. IBM.ServiceIP receives a supporting resource to alias the service IP address
   on

   In this case there are no limitations for the IPAddress and the NetMask
   attribute of the IBM.ServiceIP resource. The service IP must have a DependsOn
   relationship (see "DependsOn relationship" on page 68) to an equivalency of
   network interfaces (see "Chapter 5, "Using equivalencies," on page 51 for
   supporting resources). If System Automation for Multiplatforms decides to
   bring a service IP resource on a particular node online, it picks a suitable
   network interface from the equivalency of network interfaces and provides it to
   the IBM.ServiceIP as supporting resource. This is the interface hosting the alias.

   In this configuration automation is able to monitor the operational state of the
   network interface the service IP address is aliased on. In case of an interface
   failure detected by RSCT heartbeat, automation will force down the
   dependency chain and stop the ServiceIP. If there is another online network
   interface in the equivalency, automation will choose this device to assign the
   ServiceIP on.

   Example:
   An IBM.ServiceIP is defined as:
   IPAddress=9.152.192.1
   NetMask="255.255.255.0"
   NodeNameList="{'node01','node02','node03'}"

   The following network interfaces are available in the cluster:
   1. IPAddress=192.168.1.1
      Netmask=255.255.255.0
      NodeNameList="{'node01'}"
Global resource RM

2. IPAddress=192.168.1.2
   Netmask=255.255.255.0
   NodeNameList="{'node02'}"
3. IPAddress=192.168.1.3
   Netmask=255.255.255.0
   NodeNameList="{'node03'}"

All three network interfaces form an equivalency.
The following picture illustrates the setup:

If System Automation for Multiplatforms decides to start the service IP, it will pick
a network interface from the equivalency. In this example the service IP can float
between node01, node02 and node03, because all three nodes have a network
interface in the equivalency which can act as supporting resource.

Attributes used by IBM.ServiceIP
This section describes the attributes that are used by resources of the IBM.ServiceIP
resource class.

When a resource of this class is created with the RMC command mkrsrc, it must
have these attributes:
- Name
- IPAddress

Resources of this class may have these attributes:
- NodeNameList
- ResourceType
- NetMask (or NetPrefix)
- ProtectionMode

Resources of this class have the following dynamic attribute:
- OpState

Name attribute
The Name persistent attribute is a user defined name for this
service IP address (e.g. mail-server-ip). Both the aggregate and
constituent resource will have the same value for this attribute.
A value for this attribute must be specified when a new
IBM.ServiceIP resource is created and it must be unique. The attribute must be of type character string.

**NodeNameList attribute**

The NodeNameList persistent attribute is an array of strings that indicates which nodes the IBM.ServiceIP resource is available on. If the resource is floating, the Global Resource RM will ensure that there is one constituent (fixed) resource for each node name in this list. Constituent resources are implicitly created or deleted as necessary to match the entries in this list. Constituent resources will only contain one entry in this array since they are fixed resources and thus only available on one node. This attribute for a floating resource is implicitly modified if a constituent resource is explicitly removed by the administrator so that the aggregate and constituent resource relationship is always consistent. This list may be empty for a floating resource which means that it is not available anywhere and constituents may be added separately. The list may contain at most one name if the resource is fixed (i.e. ResourceType=0). If no name is given for a fixed resource, then RMC will provide a default because the fixed resource is tied to a node and therefore cannot be created without a node name or node id.

The value of this attribute for aggregate resources may be changed with the `chrsrc` command. An attempt to modify this attribute for a constituent resource will generate an error.

**ResourceType attribute**

You use the ResourceType persistent attribute to identify whether the resource is fixed or floating. An integer value of 0 indicates that the resource is fixed, a value of 1 indicates that it is a floating resource. This attribute defaults to floating if not specified when a new IBM.ServiceIP resource is created.

**IPAddress attribute**

With the IPAddress persistent attribute you specify the IP address that will be aliased onto a network interface where the resource is brought online. This attribute is required when a new IBM.ServiceIP resource is created. For IPv4, the IP address must be given in ‘dotted decimal’ notation as a character string; for example, 9.152.80.251. For IPv6, you can use the standard formats for IPv6 addresses; for example, 2001:db8::1428:57ab. Note that for IPv6, link-local addresses are not allowed for this attribute.

**NetMask attribute**

You use the NetMask persistent attribute to specify the netmask that will be assigned to the IP address defined in the IPAddress attribute. The attribute must be given as a character string, for example 255.255.255.0. Setting the NetMask attribute is not allowed if the IPAddress attribute contains an IPv6 address or if the NetPrefix is specified.

**NetPrefix attribute**

This attribute specifies the netprefix value of the IPv6 address given in the IPAddress attribute. It is valid for both IP address ranges: IPv6 and IPv4. It serves as mandatory replacement for the NetMask attribute for IPv6 as well as optional replacement for...
IPv4. This attribute takes integer values only; for example, 80. Do not specify a forward slash or other characters in front of the number value.

**ProtectionMode attribute**

The ProtectionMode persistent attribute specifies whether the resource is critical. If it is critical, then the IBM.ConfigRM decides if the resource can be started as requested. (For further detail on this behavior, see Chapter 9, "Protecting your resources – quorum support," on page 161. The attribute may have the integer values 0 (Non-Critical) or 1 (Critical), it defaults to Critical.

**OpState attribute**

The value of this dynamic state attribute contains the operational state of the resource as determined by the resource manager. Typical values for this state are Online (value is 1) and Offline (value is 2) meaning that the IP address is either operational or not operational.

**What happens when an IBM.ServiceIP is started?**

If the resource manager is able to assign the IP address on the selected network interface, the following functions will be carried out:

1. On Linux on System z with OSA network hardware an IP address takeover will be initiated for the specified IP address.
2. The ServiceIP is created as an IP alias on the selected network device.
3. To invalidate ARP (Address Resolution Protocol) cache entries of other IP host which may have the ServiceIP with a wrong hardware address (MAC address) in their cache, an unsolicited/gratuitous ARP packet is broadcast into the network. (Not applicable for IPv6 addresses.)
4. There is a system log message created to record the fact a ServiceIP is started on the specified network interface.

Also note that

- An IP address alias may generate a new entry in the system routing table of the host (in case the ServiceIP network differ form the network of the device).
- The IBM.ServiceIP will not modify your default gateway settings or other network/routing configurations.

**Example 1: Define an IP address as an IBM.ServiceIP resource**

In order to define an IP address which has the address IP 9.152.172.11, netmask 255.255.255.0 and potentially runs on the nodes node05 and node06, issue the following RMC command:

```
mkrsrc IBM.ServiceIP
    Name="WebServerIP"
    NodeNameList="{'node05','node06'}"
    IPAddress=9.152.172.11
    NetMask=255.255.255.0
```

**Example 2: Define an IP address as an IBM.ServiceIP resource and use an IBM.Equivalency of network interfaces**

As shown in the preceding example, define an IP address which has the address IP 9.152.172.11, netmask 255.255.255.0 and potentially runs on the nodes node05 and node06:
Global resource RM

```plaintext
mkrsrc IBM.ServiceIP
   Name="WebServerIP"
   NodeNameList="{'node05','node06'}"
   IPAddress=9.152.172.11
   NetMask=255.255.255.0
```

The nodes node05 and node06 have each more than one network interface. To form an equivalency containing eth1 device of node node05 and node node06, type:

```plaintext
mkequ MyInterfaces IBM.NetworkInterface:eth1:node05,eth1:node06
```

Now you can connect the ServiceIP with the equivalency:

```plaintext
mkrel -p dependson -S IBM.ServiceIP:WebServerIP -G IBM.Equivalency:MyInterfaces
```

**Example 3: Define an IPv6 address as an IBM.ServiceIP**

You need to enable IPv6 support before you can use the following example. This is described in Chapter 10, "Controlling and administering System Automation for Multiplatforms," on page 179, section "Enabling IPv6 Support" on page 218. In order to define an IPv6 address which has the address IP 2001:db8::1428:57ab, netprefix 64 and potentially runs on the nodes node05 and node06, issue the following RMC command:

```plaintext
mkrsrc IBM.ServiceIP
   Name="WebServerIP"
   NodeNameList="{'node05','node06'}"
   IPAddress=2001:db8::1428:57ab
   NetPrefix=64
```

You can also use an IBM.Equivalency for IPv6 resources as described in the previous example. Since you now have two IBM.NetworkInterface resources defined per physical interface, one for IPv4 and one for IPv6, you need to adjust the command slightly to create the equivalency:

```plaintext
mkequ -S "Name='eth1' & IPVersion=6" MyInterfaces IBM.NetworkInterface
```

The extra -S argument selects only those IBM.NetworkInterface resources which have the name eth1 and the IPVersion attribute set to 6. Now you can connect your resource to the equivalency:

```plaintext
mkrel -p dependson -S IBM.ServiceIP:WebServerIP
   -G IBM.Equivalency:MyInterfaces WebIP_depon_MyInterfaces
```

**Solaris IP Multipath support**

Solaris systems with IP Multipath (IPMP) configurations are supported with the following setup:

The RSCT infrastructure is harvesting Solaris IPMP groups as resources of class IBM.NetworkInterface.

The Name attribute of the harvested resources consists of the an @ character as a prefix followed by the original name of the IPMP group. For example, if the IPMP group has the name haip, the name attribute of the corresponding IBM.NetworkInterface is @haip.

To control the placement of resources of class IBM.ServiceIP to interfaces which are member of an IPMP group, you have to form an equivalency of IBM.NetworkInterfaces. The members of that equivalency have to be the harvested IBM.NetworkInterface resources which correspond to the IPMP group.
Global resource RM

Note: The usage of physical interfaces in this context will lead to unpredictable results.

Sample: Solaris IPMP groupname haip with RSCT IBM.NetworkInterface Name = @haip

# ifconfig -a
bge0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
   inet 9.152.21.226 netmask fffffff0 broadcast 9.152.23.255
     groupname haip
     ether 0:3:ba:be:ff:fd
bge1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
   inet 0.0.0.0 netmask ff000000 broadcast 0.255.255.255
     groupname haip
     ether 0:3:ba:be:ff:fe

# lsrrsrc -s "NodeNameList = {'node1'} " -x IBM.NetworkInterface Name IPAddress SubnetMask CommGroup HeartbeatActive HardwareAddress NodeNameList
resource 1:
   Name = "bge0"
   IPAddress = "9.152.21.226"
   SubnetMask = "255.255.252.0"
   CommGroup = ""
   HeartbeatActive = 0
   HardwareAddress = "00:03:ba:be:ff:fd"
   NodeNameList = {'node1'}
resource 2:
   Name = "bge1"
   IPAddress = "0.0.0.0"
   SubnetMask = "255.0.0.0"
   CommGroup = ""
   HeartbeatActive = 0
   HardwareAddress = "00:03:ba:be:ff:fe"
   NodeNameList = {'node1'}
resource 3:
   Name = "@haip"
   IPAddress = "9.152.21.226"
   SubnetMask = "255.255.252.0"
   CommGroup = "CG1"
   HeartbeatActive = 1
   HardwareAddress = ""
   NodeNameList = {'node1'}

mkrsrc IBM.ServiceIP Name="WebServerIP" NodeNameList='"node1","node2"'
   IPAddress=9.152.172.11 NetMask=255.255.255.0

mkequ IPMPInterfaces IBM.NetworkInterface:@haip:node1,@haip:node2
mkrel -p dependson -S IBM.ServiceIP:WebServerIP
   -G IBM.Equivalency:IPMPInterfaces WebIp_depon_IPMPInterfaces
Using the Test Resource Manager

This section describes the characteristics of the Test resource manager.

The IBM Test resource manager (IBM.TestRM) manages test resources and provides functions to manipulate the operational state of these resources. The resource manager is operational in a peer domain mode only and provides the resource class IBM.Test. IBM.TestRM does not control real resources.

What is the IBM.Test resource class?

The IBM.Test resource class allows new types of fixed and floating resources to be created, monitored and controlled through the RMC subsystem. These resources are not real resources but just the containers to define, monitor and control them. These resources can then be automated or recovered by System Automation for Multiplatforms. The purpose of the IBM.Test class is to provide a lightweight and easy-to-handle resource to simulate automation scenarios without the overhead of real resources. Each resource controlled by the IBM.Test class is considered to be a globalized resource which divides in one aggregate and one constituent on every node the resource is defined. See Figure 17 on page 222 for details.

The IBM.Test resource class provides a set of persistent resource attributes to simulate the behavior of real resources.

Attributes used by IBM.Test

This section describes the persistent attributes that are used by the IBM.Test resource class.

When a resource of this class is created with the RMC command `mkrsrc`, it must have the following persistent attribute

• Name

Resources of this class may have these attributes:

• NodeNameList
• ResourceType
• ForceOpState
• TimeToStart
• TimeToStop
• WriteToSyslog
• MoveTime
• MoveFail

Resources of this class have the following dynamic attributes:

• OpState
• MoveState
• OpQuorumState

Name attribute

The Name persistent attribute is a user-defined name for the test resource. Both the aggregate and constituent resource will have the same value for this attribute.

A value for this attribute must be specified when a new IBM.Test resource is created and it must be unique.
Test resource manager

NodeNameList attribute
The NodeNameList persistent attribute is an array of strings that indicates which nodes the IBM.Test resource is available on. If the resource is floating, the TestRM will ensure that there is one constituent (fixed) resource for each node name in this list. Constituent resources are implicitly created or deleted as necessary to match the entries in this list. Constituent resources will only contain one entry in this array since they are fixed resources and thus only available on one node. This attribute for a floating resource is implicitly modified if a constituent resource is explicitly removed by the administrator so that the aggregate and constituent resource relationship is always consistent. This list may be empty for a floating resource which means that it is not available anywhere and constituents may be added separately.

ResourceType attribute
You use the ResourceType persistent attribute to identify whether the resource is fixed or floating. An integer value of 0 indicates that the resource is fixed, a value of 1 indicates that it is a floating resource. This attribute defaults to fixed if not specified when a new IBM.Test resource is created.

ForceOpState attribute
You use this attribute to initiate an OpState change of the test resource via the RMC chrsrc command. This could be used to simulate a failure in the resource. The last state change is saved in this persistent resource attribute. Specifying this attribute during the creation of the resource has no effect. Normally ForceOpState changes should be done on constituent resources, as the aggregate resource collects the OpState of the whole resource. Allowed values of this attribute are:
- Unknown=0
- Online=1
- Offline=2
- Failed Offline=3
- Stuck Online=4
- Pending Online=5
- Pending Offline=6

TimeToStart
After a test resource receives the start command, the TimeToStart attribute specifies the amount of time (in seconds) it takes a resource to change its OpState from pending online to online. The default value is 0 seconds, then the resource immediately goes online.

TimeToStop
After a test resource receives the stop command, the TimeToStop attribute specifies the amount of time (in seconds) it takes a resource to change its OpState from pending offline to offline. The default value is 0 seconds, then the resource immediately goes offline.

WriteToSyslog
A resource of the class IBM.Test is capable to log online, offline and ForceOpState events in the syslog facility. You use the WriteToSyslog attribute to turn on/off the writing to the syslog daemon. Allowed values of this attribute are:
OpState attribute

The value of this dynamic state attribute contains the operational state of the resource. IBM.Test resource OpState follows the RMC start/stop commands or the ForceOpState event from an operator or test script (automated testcase). The possible values for this attribute as defined by the state transition diagram in the RMC architecture are:

- Unknown=0
- Online=1
- Offline=2
- Failed Offline=3
- Stuck Online=4
- Pending Online=5
- Pending Offline=6

MoveTime Reserved for internal use.
MoveFail Reserved for internal use.
MoveState Reserved for internal use.
OpQuorumState Reserved for internal use.

Example: Create a test resource and manipulate its OpState

In order to create an IBM.Test resource on 2 nodes, issue the following RMC command:

```
mkrsrsc IBM.Test \
    Name="mytest" \
    NodeNameList="{'node01','node02'}" \
    ResourceType=1 \
    TimeToStart=5 \
    TimeToStop=2 \
    WriteToSyslog=1
```

The following command causes a constituent on node02 to change its OpState to Failed Offline. If the resource is automated by System Automation for Multiplatforms, the automation manager starts the resource on another node.

```
chrsrc -s "Name='myTest' && NodeNameList=['node02']" IBM.Test ForceOpState=3
```
Test resource manager
Chapter 12. Automating RSCT Storage resource manager resources

IBM Tivoli System Automation for Multiplatforms exploits the Storage resource manager (StorageRM) functionality provided by RSCT, which allows you to automate file system resources of class IBM.AgFileSystem. The extent of support provided by IBM Tivoli System Automation for Multiplatforms for automating these resources varies depending on a number of factors, such as the type of storage device on which the file system resides, the operating system of the node to which the storage device is attached, and whether storage entities are harvested or user-defined:

• The platform-specific scope of support is outlined in “Supported storage devices.”

• To automate IBM.AgFileSystem resources that are automatically harvested, no further additional manual configuration is required. To automate file systems that are not automatically harvested, you use user-defined IBM.AgFileSystem resources. The following sections describe the steps you need to perform to define and automate these resources:
  - “Creating user-defined IBM.AgFileSystem resources” on page 251
  - “Automating user-defined IBM.AgFileSystem resources in an LVM environment” on page 251

For additional information about the storage resource manager, refer to the following sections in Chapter 6 - “Understanding and administering the storage resource manager” of RSCT Administration Guide:

• Configuring file system mounting on IBM.AgFileSystem resources
• Storage resource manager prerequisites

Supported storage devices

System Automation for Multiplatforms on Solaris supports storage devices as follows:

• Only user-defined IBM.AgFileSystem resources are supported. These resources can be mounted and unmounted by taking the resources online and offline respectively.

• All Solaris file system types (for example, UFS, ZFS) are supported.

• Storage entities of class IBM.VolumeGroup, IBM.Partition, IBM.Disk and IBM.LogicalVolume are not supported.

• The automatic harvesting of IBM.AgFileSystem resources is supported in version 3.1.0.1. of System Automation for Multiplatforms.

• For information about network file system support on Solaris, see “Support for network file systems” on page 251.

For information about storage device support on Windows, see “Components not available for System Automation for Multiplatforms on Windows” on page 264.

The following sections give a platform-specific overview of the scope of support that is available for automating IBM.AgFileSystem resources on AIX and Linux.
Automation support for single-path and multipath storage devices of the IBM TotalStorage DS4000 family

AIX
Full support is available for SPIO and MPIO storage devices:

- Harvested IBM.AgFileSystem resources can be automated.
  - IBM.AgFileSystem resources are harvested if they are of type jfs or jfs2 and reside on storage entities that are harvested themselves (storage entities of class IBM.LogicalVolume, IBM.VolumeGroup, IBM.Disk).
- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).
- SCSI-2 reservation is supported for SPIO and MPIO storage devices using the Redundant Disk Array Controller (RDAC) driver.

Limitations:
- No striping
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.

Linux on POWER and Linux on System x
Full support is available for single-path I/O (SPIO) storage devices, and for multipath storage I/O (MPIO) devices with Redundant Disk Array Controller (RDAC) device drivers:

- Harvested IBM.AgFileSystem resources can be automated.
  - IBM.AgFileSystem resources are harvested if they are of type ext2, ext3, or reiserfs and reside on storage entities that are themselves harvested (storage entities of class IBM.LogicalVolume, IBM.Partition, IBM.VolumeGroup, IBM.Disk).
- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).
- SCSI-2 reservation is supported for disks harvested from the RDAC driver.
- Linux RAID (/dev/md devices) is supported.

Limitations:
- File systems created on md devices without using LVM are not harvested, they can only be automated using user-defined IBM.AgFileSystem resources.
- md devices themselves are only harvested as IBM.Disk resources if a physical volume has been created on the md device using the pvcreate command.
- SCSI-2 reservation is not supported for non-RDAC drivers or for md devices themselves.
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.
- EVMS is not supported, which includes any Volume Groups/Logical Volumes created or managed by EVMS.
- For SLES 10 and RHEL 5, harvesting of storage entities of class IBM.Disk, IBM.VolumeGroup, IBM.LogicalVolume, and IBM.AgFileSystem and IBM.Partition is supported. File systems may be automated if the limitations for md devices listed above are met.

Linux on System z
md devices are harvested as IBM.Disk resources if a physical volume has been created on the md device using the pvcreate command.
Limitations:

- Only user-defined IBM.AgFileSystems resources or IBM.AgFileSystem resources residing on harvested md devices can be automated. Resource harvesting for other disks is not supported. Even if harvesting of other disk resources is successful, the harvested resources cannot be automated.
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.
- SCSI reservation is not supported.

Support for single-path and multipath storage devices on other IBM TotalStorage disk system families

**AIX**

Full support is available for SPIO and MPIO storage devices:

- Harvested IBM.AgFileSystem resources can be automated.
  
  IBM.AgFileSystem resources are harvested if they are of type jfs or jfs2 and reside on storage entities that are harvested themselves (storage entities of class IBM.LogicalVolume, IBM.VolumeGroup, IBM.Disk).

- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).

- SCSI-2 reservation is supported for SPIO and MPIO storage devices using the Redundant Disk Array Controller (RDAC) driver.

**Limitations:**

- No mirroring
- No striping
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.

**Linux on POWER and Linux on System x**

Full support is available for single-path I/O (SPIO) storage devices and limited support for multipath storage I/O (MPIO) devices for the IBM TotalStorage Disk systems.

- Harvested IBM.AgFileSystem resources can be automated.
  
  IBM.AgFileSystem resources are harvested if they are of type ext2, ext3, or reiserfs and reside on storage entities that are harvested themselves (storage entities of class IBM.LogicalVolume, IBM.VolumeGroup, IBM.Disk).

- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).

- Linux RAID (/dev/md devices) is supported.

**Limitations:**

- User-defined IBM.AgFileSystem resources can be automated only if the disk hosting the file system has the same device name on all nodes of the cluster, and if the below limitations for md devices are met.
- SCSI-2 reservation is supported for the RDAC driver only.
- File systems created on md devices without using LVM are not harvested, they can only be automated using user-defined IBM.AgFileSystem resources.
- md devices themselves are only harvested as IBM.Disk resources if a physical volume has been created on the md device using the pvcreate command.
- EVMS is not supported, which includes any Volume Groups/logical Volumes created or managed by EVMS.
Linux on System z
md devices are harvested as IBM.Disk resources if a physical volume has been created on the md device using the `pvcreate` command.

Limitations:
- Only user-defined IBM.AgFileSystems resources or IBM.AgFileSystem resources residing on harvested md devices can be automated. Resource harvesting for other disks is not supported. Even if harvesting of other disk resources is successful, the harvested resources cannot be automated.
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.
- SCSI reservation is not supported.

Support for non-DS4000 single-path storage devices

AIX
Full support is provided for single-path storage devices:
- Harvested IBM.AgFileSystem resources can be automated.
  IBM.AgFileSystem resources are harvested if they are of type `jfs` or `jfs2` and reside on storage entities that are harvested themselves (storage entities of class IBM.LogicalVolume, IBM.VolumeGroup, IBM.Disk).
- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).
- SCSI-2 reservation is supported.

Limitations:
- No mirroring
- No striping
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.

Linux on POWER and Linux on System x
Limited support is provided:
- Harvested IBM.AgFileSystem resources can be automated.
  IBM.AgFileSystem resources are harvested if they are of type `ext2`, `ext3`, or `reiserfs` and reside on storage entities that are harvested themselves (storage entities of class IBM.LogicalVolume, IBM.Partition, IBM.VolumeGroup, IBM.Disk).
- User-defined IBM.AgFileSystem resources can be automated (for example, network file systems).

Limitations:
- Support for SCSI reservation is limited. Perform a disk reserve operation to check whether SCSI reservation is available.
- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.

Linux on System z
md devices are harvested as IBM.Disk resources if a physical volume has been created on the md device using the `pvcreate` command.

Limitations:
- Only user-defined IBM.AgFileSystems resources or IBM.AgFileSystem resources residing on harvested md devices can be automated. Resource harvesting for
other disks is not supported. Even if harvesting of other disk resources is successful, the harvested resources cannot be automated.

- User-defined IBM.AgFileSystem resources can only be automated if the disk hosting the file system has the same device name on all nodes of the cluster.
- SCSI reservation is not supported.

Support for network file systems
The following applies to Linux on POWER, Linux on System x, Linux on System z, Solaris, and AIX:

Network file systems are not harvested. To automate such file systems, you use user-defined IBM.AgFileSystem resources.

Note that network file systems of class IBM.AgFileSystem can only be automated and monitored successfully if the root user of the importing system has write access to the file system.

Cascaded usage of file systems is not possible:

With IBM Tivoli System Automation for Multiplatforms you can define a highly available NFS server, where the exported file systems are automated as resources of class IBM.AgFileSystem which reside on a shared disk medium. The NFS server itself is automated as a resource of class IBM.Application which may float on systems that have access to the shared disk medium. When an additional system imports the network file systems, however, the imported file systems must not already exist as user-defined IBM.AgFileSystem resources on the importing system, otherwise, monitoring of the file systems will fail and the resources will go into OpState 3 (FAILED OFFLINE).

Creating user-defined IBM.AgFileSystem resources

To automate file systems which are not automatically harvested (for example, NFS devices), you use the mkrsrc to create the corresponding IBM.AgFileSystem resources.

Example:

```
mkrsrc IBM.AgFileSystem Name=<name> DeviceName=<device-name>
Vfs=<file-system-type> MountPoint=<mount-point>
NodeNameList={<list-of-node-names>}
```

Automating user-defined IBM.AgFileSystem resources in an LVM environment

To automate file systems that reside on logical volumes and are not automatically harvested, you must define the dependencies between the file systems, logical volumes, and volume groups using IBM Tivoli System Automation for Multiplatforms means. The following actions must be taken:

1. Define a resource of class IBM.Application to represent the volume group.
   Create scripts for the StartCommand, StopCommand, and MonitorCommand attributes of the IBM.Application resource which contain the proper commands for activating, deactivating, and status reporting for the volume group.
2. Define a resource of class IBM.AgFileSystem to represent the file system.
3. Define a DependsOn relationship in which the file system resource is the source and the application resource representing the volume group is the target.
Attributes used by IBM.AgFileSystem

This section describes the persistent attributes that are used by the IBM.AgFileSystem resource class. When a resource of this class is created with the RMC command mkrsrc, it must have the following persistent attributes:

- Name
- Vfs
- DeviceName
- MountPoint

Resources of this class may have these attributes:

- NodeNameList
- ResourceType
- ProtectionMode
- PreOnlineMethod
- Force

Resources of this class have the following dynamic attributes:

- OpState

Attribute details

Name attribute
The Name persistent attribute is a user-defined name for this mount point (for example, mail-server-data). Both the aggregate and constituent resource will have the same value for this attribute. A value for this attribute must be specified when a new IBM.AgFileSystem resource is created and it must be unique. The attribute must be of type character string.

Vfs attribute
Use the Vfs persistent attribute to specify the type of file system the resource will represent.

DeviceName attribute
Use the DeviceName persistent attribute to specify the name of an existing file system (for example, /dev/fs1v03).

MountPoint attribute
Use the MountPoint persistent attribute to specify an existing mount point.

NodeNameList attribute
The NodeNameList persistent attribute is an array of strings that indicates which nodes the IBM.AgFileSystem resource is available on. If the resource is floating, the Storage Resource RM will ensure that there is one constituent (fixed) resource for each node name in this list. Constituent resources are implicitly created or deleted as necessary to match the entries in this list. Constituent resources will only contain one entry in this array because they are fixed resources and are thus only available on one node. This attribute for a floating resource is implicitly modified if a constituent resource is explicitly removed by the administrator so that the aggregate and constituent resource relationship is always consistent. This list may be empty for a floating resource, which means that it is not available anywhere and constituents may be added separately. The list may contain at most one name if the resource is fixed (that is, ResourceType=0). If no name is given for a fixed resource, RMC will provide a default name
because the fixed resource is tied to a node and therefore cannot be created without a node name or node ID. The value of this attribute for aggregate resources may be changed with the `chrsrc` command. An attempt to modify this attribute for a constituent resource will generate an error.

**ResourceType attribute**

Use the ResourceType persistent attribute to identify whether the resource is fixed or floating. An integer value of 0 indicates that the resource is fixed; a value of 1 indicates that it is a floating resource. This attribute defaults to fixed if not specified when a new IBM.AgFileSystem resource is created.

**ProtectionMode attribute**

The ProtectionMode persistent attribute specifies whether the resource is critical. If it is critical, the IBM.ConfigRM decides if the resource can be started as requested. (For further detail on this behavior, see Chapter 9, "Protecting your resources – quorum support," on page 161. The attribute may have the integer values 0 (Non-Critical) or 1 (Critical), it defaults to Non-Critical.

**PreOnlineMethod attribute**

This setting determines whether the storage resource manager will, by default, use the `fsck` command to check and repair file systems and, if so, the exact format of the `fsck` command. For details, refer to *RSCT Administration Guide*, Chapter 6 - Understanding and administering the storage resource manager.

**Force attribute**

Specify Force=1 if an IBM.AgFileSystem resource that specifies the same DeviceName attribute value already exists. This will be the case if the file system was harvested by the storage resource manager. The Force attribute is not needed if the device is not already known to the storage resource manager (as in the case of NFS).

**OpState attribute**

The value of this dynamic state attribute contains the operational state of the resource as determined by the resource manager. Typical values for this state are Online (value is 1) and Offline (value is 2) meaning that the mount point is either mounted or not.
Chapter 13. System Automation for Multiplatforms on Windows: Particularities and restrictions

This chapter summarizes the specific behavior of System Automation for Multiplatforms on Windows. It concentrates on describing differences to System Automation for Multiplatforms on AIX, Solaris, and Linux systems but does not make the previous chapters obsolete which provide general information about this product that is valid for all platforms.

Using System Automation for Multiplatforms on Windows systems

As noted in previous chapters, basic UNIX Administrator skills are required when you are using System Automation for Multiplatforms on Windows systems.

Contents of the SA for Multiplatforms

The installation program creates the folder Tivoli SA MP Base in your start menu. It contains the following items:

- Folder Documentation
  This folder contains a softcopy of the user documentation for this product. To display the PDF files, a "PDF reader" must be available on the system. Note that Documentation folder is not available if you opted against installing it.

- Configure Automation Adapter - cfgsamadapter
  Run the graphical configuration program for configuring the System Automation for Multiplatforms automation adapter.

- IBM Tivoli System Automation – Shell
  This is the main shell application to be used for interacting with RSCT and System Automation for Multiplatforms on Windows.

- List Defined Resources – lssam
  When a domain has been set up and resources have been created, you can use this utility to display all defined resources and resource groups and their states. It starts an SA MP shell with the command lssam -top.

- SA MP Version
  A link to a text file containing information about the currently installed System Automation for Multiplatforms version.

- Uninstall SA for Multiplatforms
  A link to uninstall the product.

Working with the IBM Tivoli System Automation Shell

All RSCT and System Automation for Multiplatforms related commands must be issued from within the IBM Tivoli System Automation –Shell which is located by default in the folder Start > All Programs > SA for Multiplatforms. It is not possible to issue RSCT or System Automation for Multiplatforms commands from the standard Windows command prompt window nor can they be issued by using the Run Windows command.

This IBM Tivoli System Automation Shell is based on a Korn Shell process which is installed together with the Windows component "Subsystem for UNIX-based
Applications. Therefore, it can be further customized (for example, by setting the command-line-editor of choice) by changing the files /etc/profile.lcl or <user_home>/.kshrc.

After you have set up your first automation domain, it does not matter on which system of this domain you log in. You can issue any RSCT or System Automation for Multiplatforms command from any node on the domain.

To log in, you can use either one of the following methods:

• Directly log in on the Windows system and open a IBM Tivoli System Automation Shell.
• Open a connection from a PC by using "Open a Remote Connection", and open a IBM Tivoli System Automation Shell.
• If you have configured the SUA telnet or rsh daemon process (see the description in IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide) you can access any node of the automation domain by using telnet or rsh. This access is only supported for System Automation for Multiplatforms installed on Windows Server 2003.

Resource instrumentation with resource classes

A resource is a physical (hardware) or logical (software) entity that provides some service to other components. Each resource belongs to a specific resource class. All resources support a common set of operations and, within their classes, have a unique set of attributes and methods. All RSCT controlled resources are globally accessible within the cluster. Resource classes are implemented by Resource Managers as described in previous chapters of this guide. System Automation for Multiplatforms on Windows allows you to control and automate the following resource types, using the resource class IBM.Application:

• SUA Applications (Binary)
• SUA shell scripts
• Windows Applications (EXE)
• Windows Batch Files
• Visual Basic scripts
• Windows Services
• Volumes and shared resources

IBM.Application

Automating native SUA applications and shell scripts: Binary applications that are written for or ported to the SUA subsystem and shell scripts (.sh and .ksh) can be automated using the IBM.Application resource class that is provided by the Global Resource Manager. The behavior is identical to that of such resources on AIX, Solaris, and Linux systems and will therefore not be described in more detail here.

Automating Windows applications (Binary): On SUA subsystems, the command runwin32 makes finding and running Win32 binaries easier. In the directory /usr/contrib/win32/bin, the subsystem even provides a collection of shell scripts that are named after many of the cmd.exe built-in commands and many of the Windows utilities normally found in the Windows system directory; that is, %SystemRoot%\system32 on Windows Server 2003 or %SystemRoot%\System32 on Windows Server 2008.
In addition to these shell scripts, the SUA shells also provide a case-insensitive command searching mechanism that makes it easier to invoke Win32 programs directly, without having to use the indirection of the scripts in /usr/contrib/win32/bin.

Once a Win32-based application is started, it interacts directly with the Windows subsystem and, typically, there is no further connection with the SUA subsystem. The IBM.Application resource class can also control native Windows applications (.exe). These applications can either be launched directly or by using a shell script or Windows batch file, which can, for example, be used for mapping return codes.

Note that Win32 GUI applications (for example, Notepad) can be automated but the graphical user interfaces will not be displayed on the desktop.

**Automating Windows batch files:** The IBM.Application resource class can be used to automate Windows applications that are started, stopped, and monitored using Windows batch files (.bat). These batch files can be specified as Start/Stop and Monitor commands of the resource by prepending the correct Windows batch-file processor cmd.exe. In this case, the batch file must return the required return codes indicating the current status of the resource to Tivoli System Automation for Multiplatforms. The batch file must end with a call of exit <RC> in order to achieve this. It must not use the /b flag of the exit command because this would only set the environment variable ERRORLEVEL with the return code, which cannot be observed within the shell environment of the Subsystem for UNIX-based Applications.

**Example:**
StartCommand "cmd.exe /C XYZStart.bat"
StopCommand = "cmd.exe /C XYZStop.bat"
MonitorCommand = "cmd.exe /C XYZMonitor.bat"

**Automating Visual Basic Scripts:** Many utility programs on Windows are implemented using Visual Basic Scripts (VBScripts). These scripts can be launched from both the Windows command shell and the Kornshell provided by Interix using the cscript utility. This is an example of a very simple example script that opens a Windows dialog box and prompts for an integer value to be returned as exit value to the caller:
```
sExitCode = InputBox ( "Enter Exit Code (0 - 255)", "Script", "0")
If Not IsNumeric(sExitCode) Then sExitCode = 0
WScript.Quit(sExitCode Mod 255)
```

After execution of this script with cscript, the exit code which was specified by the user can be prompted by observing the %ERRORLEVEL% environment variable or the $? value in the Kornshell, respectively.

**Example:**
StartCommand "cscript XYZStart.vbs"
StopCommand = "cscript XYZStop.vbs"
MonitorCommand = "cscript XYZMonitor.vbs"

For more information refer to http://.microsoft.com/technet/interopmigration/unix/sfu/runwin32.mspx

**Automating Windows Services:** Many applications on the Windows operating system are registered as service, which are controlled by the Windows Service Manager. Depending on the settings, this controller will start them during system startup and even react on failures, for example, by restarting them. In order to be
controllable by the Windows Service Manager, each of these application must
register a method for startup, shutdown, and for delivering the current status back
to the Windows Service Manager.

To make these services controllable by Tivoli System Automation for
Multiplatforms, the utility samservice is used. The utility can perform these tasks:
• Contact the Windows Service Manager to:
  – Start or Stop a service with given name
  – Query the current operational state of the service
• Map the return codes from the Windows Service Manager to the defined return
codes of Tivoli System Automation for Multiplatforms
• Disconnect from the Windows Service Manager

The samservice utility can be used to create resources of class IBM.Application. It
is possible to specify this utility for the Start/Stop and Monitor commands as
shown in the example below. The samservice utility is shipped as component of
System Automation for Multiplatforms on Windows only.

**Note:** When using samservice to automate Windows services, the particular
Windows service should no longer be automatically started by the Windows
Service Manager.

**Example:**
StartCommand "samservice –start XYZ"
StopCommand = "samservice –stop XYZ"
MonitorCommand = "samservice –monitor XYZ"

**Automating Windows Shared Resources:** Windows systems allow for sharing
disk and other resources with other systems by the “shared resource” mechanism.
Whole disks (that is, volumes) or Windows folders (that is, a particular directory
tree) can be shared. These shared resources are published using the Universal
Naming Convention (UNC).

Example of a shared resource name: \computername\resourcename

A Windows system can access a shared resource by simply accessing the UNC
name or by mapping the UNC name to a drive letter. The “net.exe use” command
can be used to map a shared resource to a driver letter.

Example: “net.exe use X: \server\res” maps the shared resource “res” from
system “server” to drive letter “X”.

**Automating NFS shares:** Windows systems with Subsystem for UNIX-based
Applications installed allow for using NFS shares. NFS shares can be mounted
using the “mount” command and unmounted using the “umount” command.

**IBM.ServiceIP**
The creation of a service IP address on Windows is quite similar to the process
described for AIX, Solaris, and Linux in earlier chapters of this document. The only
difference can be seen in the “Name” field of harvested network interfaces. On
Windows, the unique ID of each network connection is used to represent the
harvested network interface. So instead of the common identifications “eth0” or
“en0”, the network interface name is similar what you see in the example below.
> lsrc IBM.NetworkInterface

Resource Persistent Attributes for IBM.NetworkInterface

resource 1:

Name = "F8A5F7C5-68C9-4F39-8779-4AFA80385145"
DeviceName = ""
IPAddress = "9.152.23.169"
SubnetMask = "255.255.252.0"
Subnet = "9.152.20.0"
CommGroup = "CG1"
HeartbeatActive = 1
Aliases = {}
DeviceSubType = 0
LogicalID = 0
NetworkID = 0
NetworkID64 = 0
PortID = 0
HardwareAddress = ""
DevicePathName = "Local Area Connection"
ActivePeerDomain = "Domain1"
NodeNameList = {"SAXBOPT10C"}

The “DevicePathName” value of class IBM.NetworkInterface reflects the name of interfaces explored under Windows “Network Connections”. This helps to recognize the harvested network interfaces to be used for the creation of a IBM.ServiceIP resource.

**IP Address restrictions:** In contrast to other platforms supported by System Automation for Multiplatforms, Windows Server 2003 and Windows Server 2008 take a different approach to associating IP addresses with network interfaces. Other platforms have the notion of a base IP address assigned to a network interface whereas all other IP addresses are designated as alias IP addresses. In particular, an IP address assigned using an IBM.ServiceIP resource is an alias IP address on these platforms. On Windows Server 2003 and Windows Server 2008, all IP addresses are co-equal and there is no notion of base and alias IP addresses. In particular, an IP address assigned using an IBM.ServiceIP resource has the same precedence as any other IP addresses assigned to a network interface on the Windows Server platforms.

This different approach of associating IP addresses with network interfaces on Windows Server 2003 and Windows Server 2008 has the following consequence: when System Automation for Multiplatforms queries the system for all IP addresses associated with a network interface, it will classify the first IP address reported for the network interface as base IP address. The base IP address is used by System Automation for Multiplatforms to construct an IBM.NetworkInterface resource for a network interface. Since the Windows Server platform always reports the “minimum” IP address as first IP address, the “minimum” IP address will always be the one used for the IBM.NetworkInterface resource representing the network interface. The terms “minimum” and “less than” applied to IP addresses are defined with respect to comparing the four dotted values which the IP address is comprised of.

Example: The IP addresses 192.168.0.11 and 192.168.0.100 are currently assigned to a network interface. The IP address 192.168.0.11 is less than 192.168.0.100 because the first three values are equal and the fourth value of the first IP address is less than the fourth value of the second IP address. For this reason, the IP address 192.168.0.11 is the minimum IP address of the network interface and thus, the base IP address. An IBM.NetworkInterface resource is harvested for the network interface with the IP address 192.168.0.11.
Defining a new IP address for a network interface which is "less than" the current "minimum" IP address or removing the current "minimum" IP address will thus change the base IP address for a network interface. In particular, if an IP address assigned using an IBM.ServiceIP resource is "less than" the current "minimum" IP address, this IP address will become the new base IP address. When a new IP address becomes the base IP address of a network interface, this will be recognized by System Automation for Multiplatforms and the associated IBM.NetworkInterface resource will be changed. If the new base IP address is in a different subnet than the former IP address, the IBM.NetworkInterface is moved to a different communication group which can cause a currently running automation domain to fail.

Example: The IP addresses 192.168.0.11 and 192.168.0.100 are currently assigned to a network interface. The IP address 192.168.0.11 is the minimum IP address of the network interface and thus, the base IP address. The new IP address 10.0.0.1 is associated with the network interface by bringing an IBM.ServiceIP resource online. IP address 10.0.0.1 is the new minimum IP address and becomes the base IP address. System Automation for Multiplatforms detects this change and modifies the IBM.NetworkInterface resource for the network interface accordingly.

Attention: On Windows Server 2003 and Windows Server 2008, never change the "minimum" IP address of a network interface by either adding an IP address to, removing an IP address from, or modifying an existing IP address on a network interface. This includes the usage of IBM.ServiceIP resource for doing so. If you change the "minimum" IP address of a network interface by any means, this operation can severely disturb the operation of a running System Automation for Multiplatforms automation domain.

Duplicate IP Addresses: Windows Server 2003 and Windows Server 2008 perform active detection of duplicate IP addresses. If a new IP address is added (or an existing IP address is modified), a check is made to see if the new IP address is already in use on the network link (that is, subnet) to which the network interface is connected. If a new IP address is classified as duplicate, it is made inactive by the operating system. When displaying the IP addresses of the system, Windows Server reports all duplicate IP addresses as such. When an IP address is assigned to a network interface using an IBM.ServiceIP resource, Windows Server performs duplicate IP address detection as well. If the IP address is classified as duplicate, the IBM.ServiceIP resource loses control over the IP address on this system. System Automation for Multiplatforms reports that the state of the IBM.ServiceIP resource is undefined.

Attention: Make sure that IP addresses assigned to a network interface using an IBM.ServiceIP resource are always unique on the network link to which the network interface is connected, otherwise System Automation for Multiplatforms is not able to start, stop, and monitor the IBM.ServiceIP resource.

If you create a duplicate IP address using an IBM.ServiceIP resource, you need to remove this IP address manually. This is done by using the Windows netsh.exe command. Perform the following steps:

1. Select Start > All Programs > Accessories > Command Prompt
2. Run the following command: netsh.exe interface ipv4 show ipaddresses
3. Identify the name of the interface to which the duplicate IP address is assigned.
4. Run the following command: netsh.exe interface ipv4 delete address="<Interface name>" address="<Duplicate IP address>"
5. Run the following command: `netsh.exe interface ipv4 show ipaddresses`
6. Check that the duplicate IP address was removed.

---

**Startup procedure**

**Checking that System Resource Controller (SRC) is running**

The System Resource Controller (SRC) is the bootstrap process that is started first and then brings up all other daemon processes required by RSCT and Tivoli System Automation for Multiplatforms to run. It is also responsible for restarting these daemon processes on failure.

First introduced on AIX, the System Resource Controller provides a set of commands and subroutines to make it easier for system managers to create and control subsystems. The SRC was designed to minimize the need for operator intervention. It provides a mechanism to control subsystem processes using a common command line interface.

The SRC supports the following set of administrative commands:

- **srcmstr**
  - Starts the System Resource Controller

- **startsrc**
  - Starts a subsystem or a group of subsystems

- **stopsrc**
  - Stops a subsystem or a group of subsystems

- **refresh**
  - Refreshes a subsystem

- **traceon**/**traceoff**
  - Turns tracing of a subsystem on or off

- **lssrc**
  - Retrieves the status of a subsystem

The System Resource Controller on Windows is started automatically as Windows service at system startup. It can be found in the Windows services list under the Name “System Resource Controller.”
All other components of RSCT and of System Automation for Multiplatforms are started using the init.d daemon process which is configured in /etc/r2.d.

**Restarting SRC and RMC**

The following procedure describes how to bounce all components of RSCT/RMC. It should only be used if all other options for interacting with Tivoli System Automation for Multiplatforms (for example, stoprpnode, stoprpdomain) fail.

To bounce all components, complete the following steps:

1. Stop all resource managers, hats and hags subsystems by issuing `stopsrc -a`.
2. Check if all processes are stopped by issuing `lssrc -a`.
3. Optionally, issue a `kill -9` command against the remainder of subsystems still shown in the output of the `lssrc -a` command.
4. Stop the System Resource Controller from the Windows Service Manager.
5. Optionally, issue a `kill -9` command against the System Resource Controller process if the process named `srcmstr` is still shown as output of the `ps -e -o pid,cmdnam -X unix | grep srcmstr` command.
6. Restart the System Resource Controller from the Windows Service Manager.
7. Start the RMC subsystem by issuing the command `rmcctrl -s`.

**Dead-Man-Switch (DMS) to protect critical resources**

In a high availability environment it is crucial that at most one instance of a critical resource is running. A typical example of a critical resource is the read/write access to a shared disk, which, when given to more than one node at one time, would result in utter corruption of the file system structure unless a cluster-aware file system takes care of proper locking.

The quorum algorithms of the ConfigRM prevent this scenario from happening, but that is only guaranteed if ConfigRM, HATS and HAGS are given enough system resources to perform their calculations. The dead man switch kicks in if this lower level RSCT infrastructure can no longer be relied on to handle critical resources; process starvation or dead locks are examples. The DMS requires
periodic access within a time period; if the access is missed, the kernel itself
triggers an immediate reboot, the ultimate method to prevent a critical resource
from being started twice.

On Linux systems, this functionality is implemented by using the “reboot”
respectively the “halt” system call and by using a so called “softdog” module. On
Solaris and AIX, the device driver hDMS is used for this purpose. On Windows,
the device driver ibmhadms.sys is used. It can be used to:
• Immediately stop a system on request
• Stop the system if it has not been rearmed for a while (Dead-Man-Switch)

The device driver uses a low-level kernel routine to initiate a “Blue Screen of
Death” (BSOD), the ultimate method to immediately stop a Windows system,
which is comparable to the “halt” method provided on UNIX and Linux systems.

**Operational quorum values**

If critical resources are active on a subcluster that has lost quorum, the ConfigRM
decides in which way the system should be terminated. On AIX, Solaris, and
Linux, there are six different protection methods which can be configured by the
“CritRscProtMethod” attribute on each node. For System Automation for
Multiplatforms on Windows, the ibmhadms.sys device driver is not configurable in
this way.

Whether and how a system recovers from a BSOD can be configured within the
“Startup and Recovery” dialog of the advanced section of the system properties.
This can be defined in the Windows System Properties > “Advanced” Tab >
“Startup and Recovery” section > Settings. Desired configuration changes must be
specified separately on each system.

For System Automation for Multiplatforms on Windows, the protection method
table looks as follows:

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Value</th>
<th>System Automation for Multiplatforms on Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard reset and reboot operating system (default)</td>
<td>1</td>
<td>Whether or not the system is rebooted after BSOD is specified in “System Properties”</td>
</tr>
<tr>
<td>Halt operating system</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hard reset and reboot operating system with Sync</td>
<td>3</td>
<td>Not supported – will be handled in the same way as values 1 and 2 above</td>
</tr>
<tr>
<td>Halt with Sync</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>No protection. System continues operating</td>
<td>5</td>
<td>Supported</td>
</tr>
<tr>
<td>Exit and restart RSCT subsystems</td>
<td>6</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Depending on the Windows setup, the operating system will restart automatically
after a BSOD occurs.
Components only available for System Automation for Multiplatforms on Windows

**SAMSERVICE utility**

This utility can be used to easily start, stop and monitor Windows services. It contacts the Windows service manager process, queries the current states, and returns the state after it has been translated to the RSCT state values. Each start or stop command to samservice will result in a start or stop request against the Windows service manager as well. If a service cannot be started or stopped by the samservice utility, try to run the start or stop command manually from the Windows service manager and debug its results.

Components not available for System Automation for Multiplatforms on Windows

Storage Resource Manager is available on AIX, Solaris, and Linux but not on Windows. This means that control scripts specified for IBM.Application resources must be used to control storage resources (disks).
IPv6 support is not available on Windows Server 2003 R2.
Appendix A. Setting up a highly available network

When talking about setting up a highly available network, we should distinguish between two situations:

- A more reliable cluster communication in which the cluster infrastructure (RSCT) takes care that all available communication paths are used to ensure cluster integrity and configuration data replication.
- A representation of a highly available IT service in which automation takes care of an IP address (further called ServiceIP) which represents an IT service to clients outside the cluster.

The same communication interface is often used for both tasks, but this is not necessarily and often not the best way to do. The following section describes this.

Considerations when planning a highly available network setup

This section is intended to help you to understand the complexity of a highly available network and provides some questions which can help you to plan the setup of a highly available network.

What makes high availability of a network infrastructure difficult?

Ensuring high availability of the network infrastructure is not as easy as plugging in another network interface in an existing node. Of course you can plug in the new device and configure it with an address from the existing network, but this will not work in case the "wrong" interface dies. Here is an example from a Linux setup:

```
eth0 Link encap:Ethernet  HWaddr 00:00:00:00:00:00
  inet addr:192.168.1.1  Bcast:255.255.255.0  Mask:255.255.255.0
  inet6 addr:fe80::0/108 Scope:Link
 UP BROADCAST RUNNING NOARP MULTICAST MTU:1492 Metric:1
 RX packets:147264 errors:0 dropped:0 overruns:0 frame:0
 TX packets:1557216 errors:0 dropped:0 overruns:0 carrier:0
 collisions:0 txqueuelen:100
 RX bytes:871548285 (831.3 Mb)  TX bytes:674939866 (643.6 Mb)
```

```
eth1 Link encap:Ethernet  HWaddr 00:00:00:00:00:00
  inet addr:192.168.1.1  Bcast:255.255.255.0  Mask:255.255.255.0
  inet6 addr:fe80::0/108 Scope:Link
 UP BROADCAST RUNNING MULTICAST MTU:1492 Metric:1
 RX packets:297557 errors:0 dropped:0 overruns:0 frame:0
 TX packets:289815 errors:0 dropped:0 overruns:0 carrier:0
 collisions:0 txqueuelen:100
 RX bytes:30153627 (28.7 Mb)  TX bytes:38726921 (36.9 Mb)
```

```
ServiceIP 192.168.1.45
eth0 192.168.1.254
eth1 192.168.1.1
```

```
KernIP routing table
Destination Gateway Genmask Flags Metric Ref Use Face
192.168.1.0   *   255.255.255.0  U   0   0   0  eth1
192.168.1.0   *   255.255.255.0  U   0   0   0  eth0
default 192.168.1.254  0.0.0.0  UG  0   0   0  eth1
```

Figure 21. Problems planning a highly available network

Every static configured network device causes an entry in the routing table. The routing algorithm chooses the first matching route out of this table. In this
example, device eth1 on node lnxcm1 fails. As eth1 is the first entry in the routing table, the node cannot send packages out to the network although there is another working network interface (eth0).

This chapter gives you a couple of suggestions how you can make a more sophisticated approach to ensure the high availability of your network infrastructure. Like in many other cases the most complicate approach (dynamic routing) will be the best, but you may want to consider the second or third best approach to keep complexity and implementation effort low.

**Things to clarify before planning a high availability network**

Clarify the following questions before starting to plan your high availability network:

1. What kind of a high availability network do you need? Is it necessary to move a ServiceIP from one interface to another on the same node, or is it also appropriate to switch to another node which has a working interface in the required subnet?
2. Can you implement additional IP subnets or do you have to use an existing network infrastructure?
3. Do you only work in the scope of our cluster nodes or are you able to implement/deploy network services on other nodes outside of the automation cluster?
4. What kind of network hardware do you have?
5. How much effort do you want to invest?
6. How much complexity do you want to introduce?
7. What skills do we have?

Depending on the answers to the questions above you may want to choose one of the setups described in this chapter to develop your own high availability network strategy.

### Detecting network interface failures in single-node and two-node clusters

If you are running a single-node or two-node cluster, additional configuration is required to detect network interface failures: The cluster software periodically tries to reach each network interface in the cluster. If one interface fails on one node of a two node cluster, the interface on the other node will also be flagged as offline because it does not receive a response from its peer.

To avoid this behavior, the cluster software must be configured to contact a network instance outside of the cluster. Best practise is to use the default gateway of the subnet the interface is in.

On each node, create the following file:

```
/var/ct/cfg/netmon.cf
```

Each line of this file should contain the machine name or IP address of the external network instance. IP addresses should be specified in dotted decimal format.

This is an example of a netmon.cf file:
Network setup

# This is default gateway for all interfaces in the subnet 192.168.1.0
192.168.1.1

# This is default gateway for all interfaces in the subnet 192.168.2.0
gw.de.ibm.com

Additional recommendation when running on Linux on zSeries under zVM

In addition to creating the netmon.cf file as described above, you are recommended to turn off broadcast for all communication groups when running IBM Tivoli System Automation for Multiplatforms on Linux on zSeries in a zVM environment. The RSCT heartbeat mechanism performs a broadcast ping from time to time, especially when a network interface adapter is not available. The reason for this feature is to find out whether the network interface adapter that sends this broadcast ping is still operational (this can be determined according to whether other machines reply to this broadcast ping or not). This feature is not needed if the netmon.cf file is set up correctly as described above, as in that case, there are other well-known network interface adapters to be checked for availability. While a broadcast ping on a standalone machine does not present a performance issue, it will have a negative impact on performance if the machines are running in a zVM environment. This is because all other systems running under this zVM and within the same network segment (same IP network and net mask) will reply to this broadcast ping request. As a result, even VM guest systems that are idle and currently paged out will be loaded into the zVM just to reply to this ping. Depending on the number of guest systems running under this zVM, this may decrease the performance of the whole z/VM system.

In order to prevent a negative impact on performance, the following setup changes are highly recommended:

- Get all the communication groups of the cluster:
  
  # lscmg

- Turn off broadcast for all communication groups:
  
  # chcomg -x b <communication group> ...

  For example:
  
  chcomg -x b CG1

- Use the lscmg command again to verify that broadcast is turned off.
**Network setup**

**Two-node cluster, each node has one ethernet interface**

The following network setup is given:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>Device</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster node</td>
<td>lnxcm1</td>
<td>eth0</td>
<td>9.152.172.1/24</td>
</tr>
<tr>
<td>Cluster node</td>
<td>lnxcm2</td>
<td>eth0</td>
<td>9.152.172.2/24</td>
</tr>
<tr>
<td>Router</td>
<td>gw</td>
<td>eth0</td>
<td>9.152.172.254/24</td>
</tr>
<tr>
<td>ServiceIP</td>
<td>–</td>
<td>–</td>
<td>9.152.172.3/24</td>
</tr>
</tbody>
</table>

In this setup, the cluster communication and the presentation of the highly available IT service use the same communication path, the 9.152.172.0 network.

Automation can assign the ServiceIP either on the lnxcm1 interface eth0 or on the lnxcm2 interface eth0. If one interface fails, automation moves the ServiceIP to the other node. Thus it satisfies the policy which requires assigning the ServiceIP on a running network interface.

In this setup the failure of one network interface will lead to a break in the cluster communication with all the problems described in Chapter 9, “Protecting your resources – quorum support,” on page 161. If the communication breaks as shown in Figure 23 on page 271, the tie breaker decides which node is able to go on with automation. If this is node lnxcm1, automation will find on lnxcm1 no online network interface to assign the ServiceIP on.
In this example the network 9.152.172.0 served two purposes:
1. Representing the network for the highly available IT service.
2. Used for internal cluster communication.

Sample System Automation for Multiplatforms policy:

```sh
lnxcm1# mkequ NetInt IBM.NetworkInterface:eth0:lnxcm1,eth0:lnxcm2
lnxcm1# mrsrc IBM.ServiceIP Name="SIP" IPEndPoint="9.152.172.3" NetMask="255.255.255.0" NodeNameList="{'lnxcm1','lnxcm2'}"
lnxcm1# mkrsrc IBM.Equivalency:NetInt:eth0
lnxcm1# addrgmbr -g rg IBM.ServiceIP:SIP
lnxcm1# mkrel -p dependson -S IBM.ServiceIP:SIP -G IBM.Equivalency:NetInt
```

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very easy setup.</td>
<td>Each communication problem leads to cluster split.</td>
</tr>
<tr>
<td>Less network hardware</td>
<td>ServiceIP moves only between nodes.</td>
</tr>
<tr>
<td>required.</td>
<td></td>
</tr>
</tbody>
</table>

Figure 23. Two nodes, one interface – interface failure
Two-node cluster, each node has two network interfaces

Before starting with this setup keep in mind that it is not possible to have more than one static configured network interface in the same IP subnet. Each IP address will cause an entry in the kernel routing table. In case of two interfaces in the same subnet there will be 2 routes for the same subnet. If the interface, which created the first entry, fails the communication for this subnet will break down even if there is another interface which still is able to communicate.

Two physically separated networks, move ServiceIP between nodes

The following network setup applies:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>Device</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster node</td>
<td>lnxcm1</td>
<td>eth0</td>
<td>9.152.172.1/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>192.168.1.1/24</td>
</tr>
<tr>
<td>Cluster node</td>
<td>lnxcm2</td>
<td>eth0</td>
<td>9.152.172.2/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>192.168.1.2/24</td>
</tr>
<tr>
<td>Router</td>
<td>gw</td>
<td>eth0</td>
<td>9.152.172.254/24</td>
</tr>
<tr>
<td>ServiceIP</td>
<td>-</td>
<td>-</td>
<td>9.152.172.3/24</td>
</tr>
</tbody>
</table>

There are now two networks 192.168.1.0 and 9.152.172.0 for the cluster communication. If there is a failure in one network interface the cluster will not break.

- Network 9.152.172.0 represents the network for the highly available IT service.
- Network 192.168.1.0 makes cluster internal communication more reliable.

Since only the network of the ServiceIP is connected to the gateway, a failure of interface eth0 on lnxcm1 will cause the automation to move the ServiceIP to the
interface eth0 on the other node lnxcm2. Because of the physical separation of the two networks it is not possible to move the ServiceIP from eth0 to eth1 within the same node.

The sample System Automation for Multiplatforms policy is the same as shown on page on page 271.

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy setup.</td>
<td>ServiceIP moves only between nodes.</td>
</tr>
<tr>
<td>Redundancy in cluster communication.</td>
<td></td>
</tr>
</tbody>
</table>

Three logical networks in one physical network, move ServiceIP between network interfaces

Another network setup is required to not only move the ServiceIP between nodes in the cluster but also between interfaces within one node. Needed is a separate logical network for each interface of a node, and an additional network for the ServiceIP. Choosing an existing network (one of eth0 or eth1) would cause routing problems. Make sure to connect all interfaces to the same physical network. This allows each interface to hold addresses of all the logical networks.

The following network setup applies:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>Device</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster node</td>
<td>lnxcm1</td>
<td>eth0</td>
<td>192.168.1.1/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>192.168.2.1/24</td>
</tr>
<tr>
<td>Cluster node</td>
<td>lnxcm2</td>
<td>eth0</td>
<td>192.168.1.2/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>192.168.2.2/24</td>
</tr>
<tr>
<td>Router</td>
<td>gw</td>
<td>eth0</td>
<td>9.152.172.254/24</td>
</tr>
<tr>
<td>ServiceIP</td>
<td>-</td>
<td>–</td>
<td>9.152.172.3/24</td>
</tr>
</tbody>
</table>

Network 9.152.172.0 represents the network for the highly available IT service. Network 192.168.1.0 represents the first cluster internal communication network.

Figure 25. Two nodes, two interfaces, one physical network

- Network 9.152.172.0 represents the network for the highly available IT service.
- Network 192.168.1.0 represents the first cluster internal communication network.
Network setup

- Network 192.168.2.0 represents the second cluster internal communication network.

Sample System Automation for Multiplatforms policy:

```
lnxcm1# mkequ NetInt
IBM.NetworkInterface:eth0:lnxcm1,eth1:lnxcm1,eth0:lnxcm2,eth1:lnxcm2

lnxcm1# mkrsrc IBM.ServiceIP Name="SIP" IPAddress="9.152.172.3"
NetMask="255.255.255.0" NodeNameList="{'lnxcm1', 'lnxcm2'}"

lnxcm1# mkrg rg

lnxcm1# addrgmbr -g rg IBM.ServiceIP:SIP

lnxcm1# mkrel -p dependson -S IBM.ServiceIP:SIP -G IBM.Equivalency:NetInt
```

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy setup.</td>
<td>3 logical networks in 1 physical network.</td>
</tr>
<tr>
<td>Redundancy in cluster communication.</td>
<td>Traffic of 3 networks on 1 physical medium.</td>
</tr>
<tr>
<td>ServiceIP can move between interfaces and nodes</td>
<td></td>
</tr>
</tbody>
</table>

Two physically separated networks, dynamic routing and VIPA

A detailed description of this setup is beyond the scope of this manual. Basically the ServiceIP is assigned to a virtual network within the kernel of a cluster node. Dynamic routing on all cluster nodes and the gateway makes sure that a route to the ServiceIP is established.

The following network setup applies:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>Device</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster node</td>
<td>lnxcm1</td>
<td>eth0</td>
<td>9.152.170.1/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>9.152.171.1/24</td>
</tr>
<tr>
<td>Cluster node</td>
<td>lnxcm2</td>
<td>eth0</td>
<td>9.152.170.2/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>9.152.171.2/24</td>
</tr>
<tr>
<td>Router</td>
<td>gw</td>
<td>eth0</td>
<td>9.152.170.254/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>9.152.171.254/24</td>
</tr>
<tr>
<td>ServiceIP</td>
<td>-</td>
<td>--</td>
<td>9.152.172.3/24</td>
</tr>
</tbody>
</table>
Advantage | Disadvantage
---|---
There is no dependency to the physical network device. | Complicate setup.
Concept of finding dynamically the best way to a host (IP address). | Dynamic routing required.
No need to move ServiceIP between interfaces | Setup is not restricted to the cluster nodes; gateway also has to support dynamic routing.

**Interface bonding**

Several physical network interfaces are bonded together to one logical network device. The operating system has to support this feature with a special bonding device driver. Consult your operating system documentation how to configure interface bonding on your system. Make sure that you configure HA (high availability) bonding and ensure your network interface cards support the interface failure detection mechanism your bonding driver requires.

The following network setup applies:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>Device</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster node</td>
<td>lnxcm1</td>
<td>eth0</td>
<td>9.152.172.1/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>9.152.172.1/24</td>
</tr>
<tr>
<td>Cluster node</td>
<td>lnxcm2</td>
<td>eth0</td>
<td>9.152.172.2/24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eth1</td>
<td>9.152.172.2/24</td>
</tr>
<tr>
<td>Router</td>
<td>gw</td>
<td>eth0</td>
<td>9.152.172.254/24</td>
</tr>
<tr>
<td>ServiceIP</td>
<td>-</td>
<td>-</td>
<td>9.152.172.3/24</td>
</tr>
</tbody>
</table>
Network setup

![Network setup diagram]

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy setup.</td>
<td>Operating system has to support interface bonding.</td>
</tr>
<tr>
<td>Redundancy in cluster communication.</td>
<td>Network interface hardware may has to support interface failure detection (for example, MII link monitoring).</td>
</tr>
<tr>
<td>There is no need to move ServiceIP between devices on the same node.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B. Troubleshooting

Use this information to troubleshoot problems you might experience while using IBM Tivoli System Automation for Multiplatforms.

Troubleshooting general System Automation for Multiplatforms errors and problems

This section covers the following topics:

- "How automation works" on page 277
  This topic summarizes important concepts of System Automation.
- "How to obtain troubleshooting information" on page 279
  Use this topic to learn how you can gather information about automated resources and resource groups.
- "Error analysis" on page 288
  Use the error scenarios described in this topic to learn how to understand, isolate, and resolve errors that are reported by System Automation. The following errors are discussed:
  - "A resource has an OpState of Failed Offline" on page 288
  - "A resource group has an OpState of Failed Offline" on page 289
  - "StartCommand timed out" on page 290
  - "StopCommand timed out" on page 291
  - "MonitorCommand timed out" on page 291
- "Problem analysis" on page 293
  Use the problem scenarios described in this topic to effectively troubleshoot problems related to System Automation that are typically not indicated by error messages. The following problems are discussed:
  - "A resource does not start" on page 293
  - "A resource group does not start" on page 294
  - "A resource does not stop" on page 294
  - "A resource group does not stop" on page 295
  - "No failover occurs after a node is excluded" on page 295
  - "No failover occurs after a node crash or reboot" on page 296
  - "Cluster cannot be set up" on page 297

How automation works

This topic summarizes important concepts of System Automation. For detailed information, refer to Chapter 7, “How System Automation for Multiplatforms processes the system information,” on page 91.

Automation manager

The automation manager consists of the binder and the logic deck.

Binder: The binder is responsible for finding a placement for the members of a resource group when the group is started, or when the resources must be started on a different node because the node on which they were previously running has crashed or is rebooted. The corresponding task is the so-called binding step, and the result of the task is reflected in the BindingState of a resource.
Troubleshooting

The following BindingStates for a resource exist:

**Unbound**
The resource is not bound and therefore Offline. System Automation has not yet tried to find a placement for the resource.

**Bound**
The resource has been bound to a node, and the resource is either running on that node or System Automation will start the resource on the node after all dependencies on other resources have been fulfilled.

**Sacrificed**
System Automation could not find a placement for the resource. There is no node on which this resource could be started, which is why the resource will not be started by System Automation.

If the placement of more than one resource group results in a conflict, the priority value controls which group loses the conflict, which means that it is not placed and its binding state is set to Sacrificed. This is illustrated in the following example:

**Scenario:** In a two-node cluster, resource group “RG1” contains resource “R1”, and resource group “RG2” contains resource “R2”. “R1” depends on “R2”. “R1” depends on “R2”. Both are started. Subsequently, “R1” fails. The resulting behavior depends on the priority of the resource groups:

- If “RG1” and “RG2” have the same priority, “R1” is not restarted.
- If “RG1” has at least a priority of 21 (and “RG2” has a priority of 0), “R2” is stopped and “R2” and “R1” are started on a different node.

**Logic deck:** The logic deck is responsible for sending out the start and stop orders for the individual resources to bring them Online or Offline. When sending out the orders, the logic deck ensures that all start and stop dependencies defined in the automation policy are fulfilled.

**Important internal resource states**
System Automation maintains information about many internal states for each resource. The most important are:

**DesiredState**
The state that System Automation anticipates for a resource is called DesiredState; it is the state the resource should be in when requests and votes from other resources are taken into account. The DesiredState is either Online or Offline.

The DesiredState of a resource is not necessarily identical to the value of the NominalState attribute of the resource group that contains the resource, because requests and votes that are submitted against a resource have a higher priority than the NominalState (for detailed information about request priorities, see “Starting and stopping resource groups and group members” on page 190).

**ObservedState**
The ObservedState is the actual state of the resource. It is monitored by the resource’s resource manager, for example, with the MonitorCommand for a resource of class IBM.Application, and any state change is reported to the automation manager.

The goal of System Automation is to ensure that all resources’ ObservedState values match their DesiredState.
How to obtain troubleshooting information

Before you begin: Collect debug and trace data
At any time of operation, the subsystems of System Automation and RSCT write debug and trace data to trace files on the local disk. The trace files are created as ring buffers to limit the amount of space taken up by the files. When the available space is exceeded, the trace files are overwritten. Depending on the number of resources and the activity on the nodes, large amounts of data may be logged to these files and they may be overwritten at any time.

To ensure that no debug and trace data is lost and to increase the probability that all required diagnostic information is available should you need to contact IBM support, it is recommended that you collect all trace data before starting your troubleshooting activities, which may generate trace output.

To collect all debug and trace data, use the `ctsnap` command:

```
/usr/sbin/rsct/bin/ctsnap
```

The command generates a compressed file in the `/tmp` directory.

**Tip:** If you are sure that the problem is automation-related, for example, when a resource does not start, it is usually sufficient to collect the RecoveryRM trace. To save the RecoveryRM trace to a file, complete the following steps:

1. Find out on which node the RecoveryRM master daemon is running. To do this, issue the following command:
   ```
   lssrc -ls IBM.RecoveryRM | grep Master
   ```
2. Save the RecoveryRM trace to an output file. To do this, issue the following command on the master node:
   ```
   rpttr /var/ct/<domain-name>/log/mc/IBM.RecoveryRM/trace > <output-file-name>
   ```

Using the system log as information source
The messages that are generated by all subsystems of System Automation and by RSCT are the first source of information in troubleshooting and problem determination:

- **Linux:** Messages are written to the system log (`/var/log/messages`).
- **AIX:** The system logger is not configured by default. Messages are written to the error log. To be able to obtain the debug data, it is recommended that you configure the system logger in the file `/etc/syslog.conf`. When you have made the necessary changes, you must recycle the `syslogd` using the command `refresh -s syslogd`. The location of the log file is defined in `/etc/syslog.conf`.
- **Solaris:** Messages are written to the system log (`/var/adm/messages`).
- **Windows:** Messages are logged to `/var/adm/log/messages`. The log file can be formatted using `/usr/sbin/rsct/bin/fcslogrpt`.

Messages are logged to the system log for the following events:

- **Start of a subsystem, for example:**
  ```
  Mar 10 13:04:15 node1 RecoveryRM[5482]: (Recorded using libc_fdf.c.a cv 2):::Error ID: 824....zyK22/WWN/8cU0B....................:::Reference ID: :::::Template ID: 0:::Details File: :::::Location:
  RSCT,IBM.RecoveryRM.d.C,1.21.1.2,135 ::::RECOVERYRM_INFO_0_ST IBM.RecoveryRM daemon has started.
  ```
- **Stop of a subsystem, for example:**
Troubleshooting

Error of a subsystem, for example:

Mar 10 13:04:14 node1 srcmstr: src_error=-9035, errno=0,
module='srchevn.c'@line:'251', 0513-035 The IBM.RecoveryRM Subsystem ended
abnormally. SRC will try and restart it.

Messages related to the Quorum state of the subcluster, for example:

Mar 9 16:13:07 node1 ConfigRM[31411]: (Recorded using libct_ffdc.a cv 2):::Error ID: :::Reference ID: :::Template ID: 0:::Details File: :::Location:
RSCT,PeerDomain.C,1.99.11.1,15510 :::CONFIGRM_HASQUORUM_ST The
operational quorum state of the active peer domain has changed to
HAS_QUORUM. In this state, cluster resources may be recovered and
controlled as needed by management applications.

Start and stop of an IBM.ServiceIP resource, for example:

Mar 8 09:41:08 node1 GblResRM[1886]: (Recorded using libct_ffdc.a cv 2):::Error
ID: :::Reference ID: :::Template ID: 0:::Details File: :::Location:
RSCT,ServiceIP.C,1.2.5,1360 :::GBLRESRM_IPONLINE IBM.ServiceIP
assigned address on device. IBM.ServiceIP 10.67.78.89 eth1:1

Mar 8 09:42:44 node1 GblResRM[1886]: (Recorded using libct_ffdc.a cv 2):::Error
ID: :::Reference ID: :::Template ID: 0:::Details File: :::Location:
RSCT,ServiceIP.C,1.2.5,1434 :::GBLRESRM_IPOFFLINE IBM.ServiceIP
removed address. IBM.ServiceIP 10.67.78.89

A StartCommand, StopCommand, or MonitorCommand for a resource of class
IBM.Application runs into a timeout, for example:

Mar 13 10:25:55 node1 GblResRM[24275]: (Recorded using libct_ffdc.a cv 2):::Error
ID: :::Reference ID: :::Template ID: 0:::Details File: :::Location:
RSCT,Application.C,1.2.1,2434 :::GBLRESRM_MONITOR_TIMEOUT
IBM.Application monitor command timed out. Resource name resource1

Tip: In addition to the data that is logged by default, it is recommended that you
log the execution of the StartCommand and StopCommand of
IBM.Application resources to a specific location.

The scripts that are provided with the pre-canned policies for System
Automation log all executions of the StartCommand and StopCommand for a
resource to the system log by default. The following sample output shows the
data that is written to the system log when a resource is started out of the
StartCommand of the pre-canned policy for the NFS server:

Mar 13 10:34:31 node1 /usr/sbin/rsct/sapolicies/nfsserver/nfsserverctrl-
server:[27230]: NFS server started

Using audit logs as information source

IBM.RecoveryRM master daemon audit log

The IBM.RecoveryRM master daemon maintains an audit log in which it
records all requests, error responses to the requests, important information
about the current policy, and about binding issues. To view the log, issue the
rpttr command on the node on which the master daemon of the
IBM.RecoveryRM is running:

rpttr /var/ct/<domain_name>/log/mc/IBM.RecoveryRM/trace_summary

Example:
The following example shows the four records that appear in the audit log for the following four events:

- An operator issues a start request against resource group "A".
- This effects a start request against its member resource "RA".
- A stop request is issued against resource group "A".
- This effects a stop request against its member resource "RA".

```
12:16:20.181285 T(1096711088) _RCD Online Request against RA on node saxb02
12:16:35.727970 T(1096711088) _RCD Offline Request against RA on node saxb02
```

**GblResRM daemon audit log**

On every node, the GblResRM daemon maintains an audit log in which it records any execution of a start command, stop command, or reset operation for a resource, every start or stop of a ServiceIP resource, and every OpState change of a resource. To view such an audit log, issue the `rpttr` command locally on a node:

```
rpttr /var/ct/<domain_name>/log/mc/IBM.GblResRM/trace_summary
```

**Example:**

The following example shows the records that appear in the audit log when a resource ("appfloata") and an IBM.ServiceIP resource ("ip") are stopped.

```
12:51:08.864796 T(4152898784) _GBD Taking application resource offline: Name=appfloata
handle=0x6028 0xffff 0xff2f99d1 0x13fbb275 0x1046ca5c 0x98691b60
12:51:10.877355 T(4152603872) _GBD Stop command for application resource "appfloata" (handle 0x6028 0xffff 0xff2f99d1 0x13fbb275 0x1046ca5c 0x98691b60) succeeded with exit code 0
handle=0x6028 0xffff 0xff2f99d1 0x13fbb275 0x1046ca5c 0x98691b60
12:51:12.961970 T(4152898784) _GBD Resource "ip" (handle 0x6029 0xffff 0xff2f99d1 0x13fbb275 0x1046ca5c 0x98691b60):
IP address 10.47.77.97 has been successfully taken offline on network interface "eth0:0".
12:51:12.962272 T(4152898784) _GBD Monitor reports: No network device flagged UP with IP address 10.47.77.97.
Taking resource "ip" (handle 0x6029 0xffff 0xff2f99d1 0x13fbb275 0x1046ca5c 0x98691b60) offline.
```

**Using the command history**

The file `/var/ct/IBM.RecoveryRM.log` is used to record the IBM Tivoli System Automation command history. It contains entries for all System Automation commands that have been executed locally on the node.

**Example:**

User root invoked *chrg -o Online samadapter-rg on Thu Mar 1 16:51:53 CET 2007
User root invoked *chrg -o Offline samadapter-rg on Thu Mar 1 17:00:14 CET 2007
User root invoked *chrg -o Online samadapter-rg on Thu Mar 1 17:00:20 CET 2007

**Using commands to gather information**

Use the commands described in this section to obtain detailed information about resources and resource groups if the information provided in the system log was not sufficient to resolve the problem. For best results, you should invoke the commands in the sequence in which they are listed.

For detailed information about the commands, refer to [IBM Tivoli System Automation for Multiplatforms Reference](#).

The following commands can be used to gather information about resources and resource groups:

1. `lsrg -Ab -V -g <resource-group-name>`
Troubleshooting

The command `lsrg -Ab` shows all information about the resource groups that are defined in the automation policy. Add the `-V` option to also display the most important automation details, including the DesiredState, ObservedState and BindingState. When analyzing the output, make sure to check if the value of the ConfigValidity attribute indicates a configuration problem (for more information about the ConfigValidity attribute, refer to “Dynamically verifying resources” on page 216).

This example shows the information that is displayed for a resource group that is Offline:

```
node1:~ # lsrg -Ab -V -g rg1
Starting to list resource group information.

Displaying Resource Group information:
All Attributes
For Resource Group "rg1".

Resource Group 1:

Name = rg1
MemberLocation = Collocated
Priority = 0
AllowedNode = ALL
NominalState = Offline
ExcludedList = {}
Subscription = {}
Owner =
Description =
Instruction =
ActivePeerDomain = domain1
OpState = Offline
TopGroup = rg1
MoveStatus = [None]
ConfigValidity =
AutomationDetails[CompoundState] = Satisfactory
[DesiredState] = Offline
[ObservedState] = Offline
[BindingState] = Unbound
[AutomationState] = Internal
[StartableState] = Yes
[HealthState] = Not Applicable
```

This example shows the information that is displayed for a resource group that is Online:

```
node1:~ # lsrg -Ab -V -g rg1
Starting to list resource group information.

Displaying Resource Group information:
All Attributes
For Resource Group "rg1".

Resource Group 1:

Name = rg1
MemberLocation = Collocated
Priority = 0
AllowedNode = ALL
NominalState = Online
ExcludedList = {}
Subscription = {}
Owner =
Description =
Instruction =
ActivePeerDomain = domain1
OpState = Online
TopGroup = rg1
MoveStatus = [None]
ConfigValidity =
```
2. **lsrg -m**

The command displays the operational state of all managed resources.

**Example:**

```bash
node1:~ # lsrg -m
```

Displaying Member Resource information:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM.ServiceIP:ip1</td>
<td>True</td>
<td>rg1</td>
<td>Online</td>
<td>Nominal</td>
<td>node1</td>
</tr>
<tr>
<td>IBM.Application:app1</td>
<td>True</td>
<td>rg1</td>
<td>Online</td>
<td>Nominal</td>
<td>node1</td>
</tr>
<tr>
<td>IBM.Application:app2</td>
<td>True</td>
<td>rg2</td>
<td>Offline</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. **lssamctrl**

The command `lssamctrl` displays the global automation parameters, for example, the list of excluded nodes and the RetryCount, which specifies the maximum number of retries for the StartCommand when the resource does not start at the first attempt.

**Example:**

```bash
node1:~ # lssamctrl
```

Displaying SAM Control information:

<table>
<thead>
<tr>
<th>SAMControl:</th>
</tr>
</thead>
<tbody>
<tr>
<td>TimeOut = 60</td>
</tr>
<tr>
<td>RetryCount = 3</td>
</tr>
<tr>
<td>Automation = Auto</td>
</tr>
<tr>
<td>ExcludedNodes = {}</td>
</tr>
<tr>
<td>ResourceRestartTimeOut = 5</td>
</tr>
<tr>
<td>ActiveVersion = [2.2.0.1,Thu Jun 14 08:00:09 2007]</td>
</tr>
<tr>
<td>EnablePublisher = Disabled</td>
</tr>
<tr>
<td>TraceLevel = 31</td>
</tr>
<tr>
<td>ActivePolicy = []</td>
</tr>
</tbody>
</table>

4. **lsrgreq -L {-m}**

The command `lsrgreq -L` lists all requests that were issued against resource groups. Such requests are either issued directly from the command line using the command `rgreq`, or implicitly from the operations console, or by an end-to-end automation manager. In the latter case, the Source of the request is ‘Automation’.

**Example:**

```bash
node1:~ # lsrgreq -L
```

Displaying Resource Group request information:

| ResourceGroup Priority Action Source NodeList Active UserID MoveStatus |
|--------------------------|----------------|----------------|-----------------|-----------------|----------------|
| rg1                      | low            | Start          | Operator ()     | Active          | None            |

If the command is executed with the `-m` option, all requests against resource group members are displayed.

**Example:**

```bash
node1:~ # lsrgreq -L -m
```

Displaying Member Resource request information:

<table>
<thead>
<tr>
<th>Member Resource 1:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class:Resource:Node[ManagedResource] = IBM.Application:app1</td>
</tr>
</tbody>
</table>
5. **Issam**

The command **Issam** displays summary information about the operational states of the managed resources on a per node basis. It provides additional information about excluded nodes and requests that were issued against resources or resource groups.

**Example:**

In this example, the command **Issam** was used to find out why the resource 'app1' is offline although the nominal state of the resource group 'rg1', of which it is a member, is Online.

```bash
node1:~ # lssam
Online IBM.ResourceGroup:rg1 Nominal=Online
  |- Online IBM.ServiceIP:ip1
    |- Online IBM.ServiceIP:ip1:node1
    `- Offline IBM.ServiceIP:ip1:node2 Node=Excluded
  `- Offline IBM.Application:app1 Request=Offline
    |- Offline IBM.Application:app1:node1
    `- Offline IBM.Application:app1:node2 Node=Excluded

Offline IBM.ResourceGroup:rg2 Nominal=Offline
  `- Offline IBM.Application:app2
    `- Offline IBM.Application:app2:node2 Node=Excluded
```

The output shows why 'app1' is offline although the nominal state of resource group 'rg1' is Online (*Nominal=Online*):

- Node 'node2' is in the list of excluded nodes (*Node=Excluded*), which is why all resources on the node, including 'app1', are offline.
- An Offline request has been issued against 'app1' (*Request=Offline*), which explains why it is also offline on node 'node1'.

Note that the most important information is highlighted in color in the output.

6. **Isequ –Ab**

The command **Isequ –Ab** is used to display all resources of type "equivalency" that are defined in the automation policy. Typically, equivalencies are used to define the network interfaces that can be used by resources of type IBM.ServiceIP.

When analyzing the output of the command, make sure to check if the value of the ConfigValidity attribute indicates a configuration problem (for more information about the ConfigValidity attribute, refer to "Dynamically verifying resources" on page 216).

**Example:**

```bash
node1:~ # lsequ –Ab
Displaying Equivalency information:
All Attributes

Equivalency 1:
  Name = eq1
  MemberClass = IBM.NetworkInterface
  Resource:Node[Membership] = {eth0:node1,eth0:node2}
  SelectString = ""
  SelectFromPolicy = ANY
  MinimumNecessary = 1
  Subscription[Consumer,...] = [{EEZ,All,None}]
  ActivePeerDomain = domain1
  Resource:Node[ValidSelectResources] = {eth0:node1,eth0:node2}
```
Appendix B. Troubleshooting  285

Troubleshooting

```
Resource:Node[InvalidResources] = {}
ConfigValidity =
AutomationDetails[CompoundState] = Undefined
```

The attribute Resource:Node[ValidSelectResources] must contain resources, especially if a dynamic SelectString is used. Subsequently, the OpState of the valid resources must be checked:

```
# lsrsrc IBM.<MemberClass-attribute-value> Name NodeNameList OpState
```

7. **Isrel –Ab**

The command `Isrel –Ab` is used to display all relationships that are defined in the automation policy.

**Example:**

```
nodel:~ # Isrel –Ab
```

Displaying Managed Relationship Information:

All Attributes

Managed Relationship 1:

- **Class:** Resource:Node[Source] = IBM.Application:app1
- **Class:** Resource:Node[Target] = {IBM.Application:app2}
- **Relationship** = StartAfter
- **Conditional** = NoCondition
- **Name** = app1_StartAfter_app2
- **ActivePeerDomain** = domain1
- **ConfigValidity** =

When analyzing the output, check the relationships for completeness (for example, both the Source and the Target of a relationship must be defined), and check if the value of the ConfigValidity attribute indicates a configuration problem (for more information about the ConfigValidity attribute, refer to “Dynamically verifying resources” on page 216).

8. **lsrsrc -A d -c IBM.CHARMControl**

This command lists the public attribute of the CHARMControl class, which represents the automation engine itself. The command currently only returns the value of the **Automating** attribute, which is usually 1. In rare cases, the configuration of resource groups and relationships may become damaged, preventing automation from working properly. When this happens, the value of the **Automating** attribute changes to 0, indicating that automation has been stopped.

Stopping automation instead of recycling the RecoveryRM (as was the case in System Automation for Multiplatforms releases prior to 2.3) allows you to collect diagnostic information and to even fix the configuration. A relationship with a non-existing target, for example, is considered a damaged configuration. Once such relationship is corrected or deleted, automation can be resumed by stopping and restarting the current master RecoveryRM with the commands `stpsrc` and `startsrc`.

9. **samdiag** (for deep analysis)

The command `samdiag` is used to display detailed state information for an individual resource. The command can also be used to externalize all internal variables of a resource from the automation manager. The command is very helpful in problem analysis, but it is not intended for daily use because it generates a large amount of information.

**Example:**

```
nodel:~ # samdiag -g rg1
```

Displaying information for the following:

Resource Group "rg1."

Diagnosis::Resource: rg1/ResGroup/IBM.ResourceGroup
Troubleshooting

type: CHARM Resource Group
Status -
  Observed: Online - Available
  Desired: Online - Requested Online
  (Nominal: Online - Nominal State: Online)
  Automation: Idle - CharmBase trigger linked
  Startable: Yes - Resource is startable
  Binding: Bound - Bound
  Compound: Satisfactory - Satisfactory

Resource Based Quorum: Not Supported - CharmBase trigger linked
Members and Memberships:
  +++HasMember  ---> app1/Fixed/IBM.Application/node1
  +++HasMember  ---> ip1/Fixed/IBM.ServiceIP/node1
  +++bind/HasMember  ---> app1/Float/IBM.Application
  +++bind/HasMember  ---> ip1/Float/IBM.ServiceIP

Group Constraint: Collocated
Binding Constraints:
  Flags:
    None
  Orders:
    Outstanding Order: None - Resource is Available

Dependencies:
  Start: Satisfied
    +++InCluster  ---> Cluster
  Stop: Satisfied

Binding exceptions:
  None

Static Relationships:
  +++InCluster  ---> Cluster

Dynamic Relationships:
  +++bind/HasMember  ---> app1/Float/IBM.Application
  +++bind/HasMember  ---> ip1/Float/IBM.ServiceIP

To query the details for a specific resource, the command has to be executed as follows:

samdiag IBM.<resource-class-name>::<resource-name>::<node-name>

For example, for the resource 'app1' on node 'node1':
node1:~ # samdiag IBM.Application:app1:node1
Displaying information for the following:
Resource "IBM.Application:app1:node1":

Diagnosis::Resource: app1/Fixed/IBM.Application/node1
  type: Fixed Resource
Status -
  Reported: Online - Online
  Observed: Online - Online
  Desired: Online - Requested Online
  (Nominal: Offline - Defaulted: offline)
  Automation: Idle - Idle - Online completed
  Startable: Yes - Resource is startable
  Binding: Bound - Bound
  Compound: Satisfactory - Satisfactory

Resource Based Quorum: Not Supported - CharmBase trigger linked
Groups and Aggregates:
  +++HasMember  ---> rg1/ResGroup/IBM.ResourceGroup
  +++bind/HasMember  ---> rg1/ResGroup/IBM.ResourceGroup

Binding Constraints:
  Flags:
    None
  Orders:
    Outstanding Order: None - Idle - Online completed

Dependencies:
  Start: Satisfied
    +++RunsOn  ---> node1/Node/IBM.PeerNode
Stop: Satisfied
<---HasMember
Static Relationships:
+---RunsOn
Dynamic Relationships:
<---bind/HasMember

--- rg1/ResGroup/IBM.ResourceGroup
--- node1/Node/IBM.PeerNode
--- rg1/ResGroup/IBM.ResourceGroup
Troubleshooting

Error analysis

Use the error scenarios described in this topic to learn how to effectively troubleshoot errors that are reported by System Automation.

The following error scenarios are discussed in this section:

- "A resource has an OpState of Failed Offline"
- "A resource group has an OpState of Failed Offline" on page 289
- "A resource has an OpState of Stuck Online" on page 290
- "StartCommand timed out” on page 290
- “StopCommand timed out” on page 291
- “MonitorCommand timed out” on page 291

A resource has an OpState of Failed Offline

This error has three possible causes:

The cluster node is not Online

If a cluster node is not Online, all resources that are defined on the node have an OpState of Failed Offline. In such a case, the problem is not resource- but node-related.

The MonitorCommand of the resource returns with return code 3 (== Failed Offline)

To find out if this is the case, you execute the MonitorCommand manually and check the return code of the command. Perform these steps:

1. Get the value of the MonitorCommand attribute for the resource:
   
   ```bash
   # lsrsrc -s 'Name="<resource_name>"' IBM.Application Name MonitorCommand
   ```
2. Execute the MonitorCommand
3. Get the return code of the MonitorCommand:
   
   ```bash
   # echo $?
   ```

If the return code is 3 (Failed Offline), investigate why the MonitorCommand itself returns this value and resolve the problem. After the problem has been resolved, the resource should have an OpState of Offline.

System Automation has set the resource to ‘Failed Offline’ because previous attempts to start the resource failed

If the MonitorCommand returns 2 (Offline) but the resource has an OpState of ‘Failed Offline’, this indicates either that the execution of the StartCommand for this resource returned with an error (not 0 or timeout) or that System Automation could not start the resource within the number of attempts defined in the RetryCount attribute (see the description of the Issamctrl command above).

To investigate the problem, do this:

1. Check the system log for messages indicating a timeout for the StartCommand for this resource.
2. If there is no such message, check the appropriate log files for the application that is behind the resource. Identify and correct all problems.
3. Check the audit trail.
   The following audit trail entries indicate problems in the start script:
4. Finally, use the following command to reset the resource from the ‘Failed Offline’ state:

```
# resetrsrc -s 'Name="<resource_name>" && NodeNameList="<node_name>"' \
IBM.Application
```

Now the resource should have an OpState of Offline and System Automation will start the resource again if the desired state of the resource is Online.

**A resource group has an OpState of Failed Offline**

If the resources of a resource group do not start and the resource group has an OpState of ‘Failed Offline’ this indicates that the binder was unable to find a placement for the resources and furthermore the resource group’s BindingState is Sacrificed. Check this using the following command:

```
# lsrsg -Ab -V -g <resource_group_name>
```

If the BindingState is Sacrificed, do this:

- Check the audit trail for entries like in the following example:

```
9:22:46.520729 T(229390) _RCD Online request injected:
A/ResGroup/IBM.ResourceGroup
09:22:46.522817 T(229390) _RCD RIBME-Hist for <NULL>:
 BINDER: Bind A/ResGroup/IBM.ResourceGroup
09:22:46.532464 T(229390) _RCD RIBME-Hist for <NULL>:
 BINDER: Resource RA/Fixed/IBM.Test/saxb02 hsa no usable options
09:22:46.532467 T(229390) _RCD RIBME-Hist for <NULL>:
 BINDER: Resource RB/Fixed/IBM.Test/saxb03 hsa no usable options
```

The example shows a resource group that is collocated but has two fixed members on different nodes preventing the binder from placing the resources; this is expressed as ‘having no usable option’.

- Check that there are no resources with OpState ‘Failed Offline’ within that group
- Check the ExcludedList attribute (lssamctrl) for excluded nodes
- Check that all relationships for starting the group are fulfilled
- Check that the equivalencies in the automation policy all have ‘valid’ and Online members:

```
# lsequ -Ab
– Check the attributes Resource:Node[ValidSelectResources] and MemberClass
  There should be ‘valid’ resources listed
– Check the ‘OpState’ of the resources in the equivalency:

```
# lsrsrc <resource_class> Name NodeNameList OpState
```

where `<resource_class>` represents the MemberClass above

An OpState of 1 indicates that the resource is Online, an OpState of 2 indicates that it is Offline.
Troubleshooting

A resource has an OpState of Stuck Online

There are two possible reasons, why a resource has an OpState of Stuck Online:

- In rare cases, the MonitorCommand of a resource returns with return code 4 (== Stuck Online). This can be checked by executing the MonitorCommand manually and checking the command’s return code:
  1. Retrieve the value of the MonitorCommand attribute for this resource:
     ```bash
     lsrc -s "Name='<resource_name>'" 'IBM.Application Name MonitorCommand'
     ```
  2. Execute the MonitorCommand.
  3. Retrieve the return code of the MonitorCommand:
     ```bash
     echo $?
     ```

     If the return code is 4 (Stuck Online), investigate why the MonitorCommand itself returns this value. After this problem has been resolved, the resource should have an OpState of Offline.

- The second and more likely reason for a resource to have an OpState of Stuck Online (if the MonitorCommand returns 1 (Online) or 6 (Pending Offline), but the resource has an OpState of ‘Stuck Online’) is that a the resource could not be stopped by System Automation for Multiplatforms previously, and System Automation for Multiplatforms has finally set the resource to Stuck Online. This is the case if the execution of the StopCommand for this resource and a subsequent reset against that resource failed to bring the resource offline.

  This error cannot be recovered by System Automation for Multiplatforms and manual intervention is required. After investigating why the resource did not stop, an operator must stop the resource. When the OpState of the resource is evaluated as Offline at the next execution of the MonitorCommand, System Automation for Multiplatforms will again take control of this resource, and no further manual steps are required.

Timeout messages are found in the system log

StartCommand timed out: A message is logged in the system log if the StartCommand for a resource did not finish within the time period defined in the StartCommandTimeout attribute for this resource. There are two possible causes for this problem:

- The value defined in the StartCommandTimeout attribute is too low

  To check the value of the attribute, perform these tasks:
  1. Determine the actual setting of this attribute for this resource:
     ```bash
     # lsrc -s 'Name='<resource_name>' ' IBM.Application Name \ StartCommandTimeout
     ```
  2. Determine how long it takes to execute the StartCommand for this resource

     Important: This should never be done on a running production system, but only during maintenance, or on another test system!
     Keep in mind that the time needed to execute the command may increase if the system is under load.
  3. Compare the actual setting of the timeout value with the time needed to execute the command
  4. If necessary, adjust the StartCommandTimeout value:
     ```bash
     # chrsrc -s 'Name='<resource_name>' ' IBM.Application StartCommandTimeout=<new_value_in_seconds>
     ```

     This change can be done dynamically.
Troubleshooting

- The StartCommand execution results in a hang situation, because one of the statements within the executed script hangs.
  - Investigation requires manual execution of the StartCommand
    - **Important:** This should never be done on a running production system, but only during maintenance!
  - If it turns out that the script is not finishing (hanging), additional debugging can be activated by adding `set -x` as the second line of the StartCommand script.
  - Identify the statement that results in the hang and correct the problem. Note that this is out of the scope of System Automation.

**StopCommand timed out:** A message is logged in the system log if the StopCommand for a resource did not finish within the time period defined in the StopCommandTimeout attribute for this resource. There are two possible causes for this problem:

- The value defined within the StopCommandTimeout attribute is too low
  To check if this is the case, do this:
  1. Determine the actual setting of this attribute for this resource:
     
     ```
     # lsrc -s 'Name="<resource_name>"' IBM.Application Name \ StopCommandTimeout
     ```
  2. Determine how long it takes to execute the StopCommand for this resource
     - **Important:** This should never be done on a running production system, but only during maintenance, or on another test system!
     Keep in mind that the time needed to execute the command may increase if the system is under load.
  3. Compare the actual setting of the timeout value with the time needed to execute the command.
  4. Adjust the StopCommandTimeout value if needed:
     
     ```
     # chrsrc -c 'Name="<resource_name>"' IBM.Application \ StopCommandTimeout=<new_value_in_seconds>
     ```
     
     This change can be done dynamically.

- The StopCommand execution results in a hang situation, because one of the statements within the executed script hangs.
  - Investigation requires manual execution of the StopCommand
    - **Important:** This should never be done on a running production system, but only during maintenance!
  - If it turns out that the script is not finishing (hanging), additional debugging can be activated by adding `set -x` as the second line of the StopCommand script.
  - Identify the statement that results in the hang and correct the problem. Note that this is out of the scope of System Automation.

**MonitorCommand timed out:** A message is logged in the system log if the MonitorCommand for a resource did not finish within the time period defined in the MonitorCommand attribute for this resource. There are two possible causes for this problem:

- The value defined within the MonitorCommandTimeout attribute is too low
  To check if this is the case, do this:
  1. Determine the actual settings of the MonitorCommand attributes for this resource:
Troubleshooting

# lsrc -s 'Name=<resource_name>' | IBM.Application Name \ MonitorCommand MonitorCommandTimeout MonitorCommandPeriod

2. Determine how long it takes to execute the MonitorCommand for this resource by issuing the MonitorCommand directly on the command line.
   Keep in mind that the time needed to execute the command may increase if the system is under load.

3. Compare the actual setting of the timeout value with the time needed to execute the command.

4. Adjust the MonitorCommandTimeout value if needed:
   ```bash
   # chrsrc -c 'Name=<resource_name>' | IBM.Application \ MonitorCommandTimeout=<new_value_in_seconds>
   ```
   This change can be done dynamically, however, the value of the MonitorCommandTimeout attribute must be lower than or equal to the value of the MonitorCommandPeriod attribute.

• The MonitorCommand execution results in a hang situation, because one of the statements within that script executed is hanging.
  − Investigation requires manual execution of the MonitorCommand
  − If it turns out that the script is not finishing (hanging), then additional debugging can be activated by adding `set -x` as the second line of the MonitorCommand script.
  − Determine the statement that results in the hang and correct the problem.
    Note that this is out of the scope of System Automation
Problem analysis

Use this section to learn how to analyze and resolve the following problems:

- "A resource does not start"
- "A resource group does not start" on page 294
- "A resource does not stop" on page 294
- "A resource group does not stop" on page 295
- "No failover occurs after a node is excluded" on page 295
- "No failover occurs after a node crash or reboot" on page 296

A resource does not start

If a resource does not start, perform the following steps:

1. Check for messages related to the execution of the StartCommand for that resource in the system log, the appropriate application log, and the process table (ps -ef). If the StartCommand has not been executed at all, proceed with Step 2, otherwise investigate why the application does not come Online.

2. Check for Operational Quorum:
   
   ```
   # lssrc -ls IBM.RecoveryRM | grep Quorum
   ```
   
   If Operational Quorum == HAS_QUORUM, proceed with Step 3. If not, find out how many nodes are Online, using:
   
   ```
   # lsrpn ode
   ```
   
   Operational quorum requires that either more than half of the nodes in the cluster are online or exactly half of the nodes are online and the tie breaker has been reserved:
   
   - If less than half of the nodes are online, start additional nodes.
   - If exactly half of the nodes are online, check the attribute of the active tie breaker:
     
     ```
     # lsrsrc -c IBM.PeerNode OpQuorumTieBreaker
     ```
     
     If the value of this attribute is Operator, the tie breaker must set manually:
     
     a. Deny ownership of the tie breaker to the node that should not get it (if the other node is still Online):
        
        ```
        # runact -c IBM.PeerDomain ResolveOpQuorumTie Ownership=0
        ```
        
     b. Grant ownership of the tie breaker to the node that should get it:
        
        ```
        # runact -c IBM.PeerDomain ResolveOpQuorumTie Ownership=1
        ```
        
   The best way to ensure that the problem does not reoccur is to define an automatic disk or network tie breaker, which ensures that the tie breaker is reserved automatically.

   Now, check the setting for the active tie breaker:
   
   ```
   # lsrsrc -s 'Name=<name-of-active-tie-breaker>' IBM.TieBreaker
   ```
   
   Check that the disk is correctly allocated for a disk tie breaker or that the IP address is available for a network tie breaker. Note that these settings can only be changed if **configuration quorum** is established, which means that more than half of the nodes must be online.

3. Check for requests against the resource or the containing resource group:

   ```
   # lsrgreq -L
   # lsrgreq -L -m
   # lssam
   ```
Troubleshooting

If a stop request was issued against the resource or resource group, this explains why the resource stays offline. To resolve the problem, cancel the request.

4. Check that the cluster is in automation mode and not in manual mode, which would prevent System Automation from starting resources, and that the node or nodes are not in the list of excluded nodes, because System Automation cannot start resources on excluded nodes. Use the following command:
   ```
   # lssamctrl
   ```
   If the value of the Automation attribute is Manual, the cluster is in manual mode. The mode can be set to Auto using:
   ```
   # samctrl -M F
   ```
   If there are nodes in the list of excluded nodes, they can be deleted from the list using:
   ```
   # samctrl -u d <node_name>
   ```

5. Check the DesiredState, ObservedState, and BindingState of the resource using this command for all nodes:
   ```
   # samdiag IBM.<resource-class>:<resource-name>[;:<node-name>]
   ```
   If the BindingState for the resource is Sacrificed on all nodes, this indicates that the binder was not able to find a placement for this resource that satisfies all relationships to other resources. Typically, this problem occurs at the time when an automation policy is created or changed.

A resource group does not start
A resource group is comprised of a number of resources. If none of the resources of the group is starting, perform the following steps:
1. Identify which of the resources should start first by evaluating the relationships.
2. Find out why this resource does not start by proceeding as described in section “A resource does not start” on page 293. Make sure to check for requests against the resource group, and evaluate all relationships in which the resource group is defined as source resource. To determine the BindingState of the resource group, use one of these commands:
   ```
   # lsrg -Ab -V -g <resource-group-name>
   # samdiag -g <resource-group-name>
   ```

A resource does not stop
If a resource does not stop, perform the following steps:
1. If a resource does not stop after the StopCommand was executed, System Automation for Multiplatforms will issue a reset operation against the resource, which triggers the execution of the StopCommand. If the resource still does not reach the OpState Offline, the OpState of the resource will finally be set to Stuck Online. To resolve the problem, proceed as described in “A resource has an OpState of Stuck Online” on page 290.
2. If the StopCommand for the resource has not been executed, check for requests against the resource or the containing resource group, using one of these commands:
   ```
   # lsrgrq -L
   # lsrgrq -L -m
   # lssam
Troubleshooting

If there is a start request against the resource or resource group, check if the request can be canceled.

3. Check that the cluster is in automation mode, not in manual mode:
   
   # lssamctrl

   If the value for the attribute Automation is Manual, the cluster is in manual mode, and System Automation will not stop any resources. This can be set to Auto using:
   
   # samctrl -M F

4. Check if there are relationships from other resources that prevent this resource from stopping, especially, check for the following relationships:

   - StartAfter (a StartAfter relationship keeps the dependent resource online)
   - DependsOn and DependsOnAny (both relationships implicitly include a StartAfter relationship, which will keep the dependent resource online)

5. Check if there are StopAfter relationships to other resources which prevent this resource from stopping (if the target resource needs to stay online, the source resource will also stay online).

A resource group does not stop

A resource group is comprised of a number of resources. If none of the resources of the group is stopping, perform the following steps:

1. Identify which of the resources should stop first by evaluating the relationships.

2. Find out why this resource does not stop by proceeding as described in section “A resource does not stop” on page 294. Make sure to check for requests against the resource group, and evaluate all relationships in which the resource group is defined as target resource.

No failover occurs after a node is excluded

If no failover occurs after a node is excluded, perform the following steps:

1. Check that the cluster is in automation mode and not in manual mode, and a node is available on which the resources can be started using:
   
   # lssamctrl

   If the value for attribute Automation is Manual, the cluster is in manual mode, and System Automation will not start any resources. This can be set to Auto using:
   
   # samctrl -M F

   Display the list of nodes that are online and compare it with the list of excluded nodes using:
   
   # lsrpnode

   If there are too many or all nodes in the list of excluded nodes, you can delete nodes from the list using the following command:
   
   # samctrl -u d <node_name>

2. Check if there are resources with OpState ‘Failed Offline’. If there are, proceed as described in section “A resource does not start” on page 293.

3. Check if there are resources with OpState ‘Stuck Online’. If there are, proceed as described in “A resource has an OpState of Stuck Online” on page 290.

4. Check the BindingState of the resource group to be started first:

   - If the Binding State is ‘Sacrificed’, System Automation could not find a placement for the resources.
Troubleshooting

Check the equivalencies for valid member resources and check that the OpState of these resources is Online, using:

```
# lsequ -Ab
```

Check the attribute ValidSelectResources.
Check that the resources are Online:

```
# lsrsr IBM.<equivalency_class> Name NodeNameList OpState
```

- If the BindingState is 'Bound', System Automation could not start the resources. Proceed as described in section "A resource does not start" on page 293.

5. Check for relationships that cannot be fulfilled without the excluded node.

No failover occurs after a node crash or reboot

To analyze and resolve the problem, perform the following steps:

1. Check for operational quorum:

```
# lssrc -ls IBM.RecoveryRM | grep Quorum
```

If Operational Quorum == HAS_QUORUM, proceed with Step 2. If not, find out how many nodes are Online, using:

```
# lsrpnode
```

Operational quorum requires that either more than half of the nodes in the cluster are online or exactly half of the nodes are online and the tie breaker has been reserved:

- If less than half of the nodes are online, start additional nodes.
- If exactly half of the nodes are online, check the attribute of the active tie breaker:

```
# lsrsr -c IBM.PeerNode OpQuorumTieBreaker
```

If the value of this attribute is Operator, the tie breaker must be set manually:

a. Deny ownership of the tie breaker to the node that should not get it (if the other node is still online):

```
# runact -c IBM.PeerDomain ResolveOpQuorumTie Ownership=0
```

b. Grant ownership of the tie breaker to the node that should get it:

```
# runact -c IBM.PeerDomain ResolveOpQuorumTie Ownership=1
```

The best way to ensure that the problem does not reoccur is to define an automatic disk or network tie breaker, which ensures that the tie breaker is reserved automatically.

Now, check the setting for the active tie breaker:

```
# lsrsr -s 'Name="<name-of-active-tie-breaker>"' IBM.TieBreaker
```

Check that the disk is correctly allocated for a disk tie breaker or that the IP address is available for a network tie breaker. Note that these settings can only be changed if configuration quorum is established, which means that more than half of the nodes must be online.

2. Check that the cluster is in automation mode and not in manual mode, which would prevent System Automation from starting resources, and check that the node or nodes are not in the list of excluded nodes, because System Automation cannot start resources on excluded nodes. Use the following command:

```
# lssamctrl
```
If the value of the Automation attribute is Manual, the cluster is in manual mode. The mode can be set to Auto using:

```
# samctrl –M F
```

If there are nodes in the list of excluded nodes, you can delete nodes from the list using the following command:

```
# samctrl –u d <node_name>
```

3. Depending on whether the entire resource group or only one of its members does not start, proceed as described in the appropriate section above.

**No reset occurs after a start control operation timeout**

The operation timer is started when IBM Tivoli System Automation first sends a resource start control operation to a resource. If the resource does not reach the desired state (online) within the timeout period and IBM Tivoli System Automation fails to issue a reset operation against the resource, perform the following steps:

1. Stop the cluster using the `stoprpdomain` command.
2. Restart the cluster using the `startrpdomain` command.

**Other problems**

If the problem is most likely related to the automation manager, you should try recycling the automation manager (IBM.RecoveryRM) before contacting IBM support. This can be done using the following commands:

1. Find out on which node the RecoveryRM master daemon is running using the following command:
   ```
   lssrc –ls IBM.RecoveryRM | grep Master
   ```

2. On the node running the master, retrieve the PID and kill the automation manager:
   ```
   lssrc –ls IBM.RecoveryRM | grep PID
   kill -9 <PID>
   ```

As a result, an automation manager on another node in the domain will take over the master role, and proceeds with making automation decisions. The `src` subsystem will restart the killed automation manager immediately.

**Cluster cannot be set up**

See "Step 1: Defining and administering a cluster" on page 12 for information on how to avoid problems with setting up a cluster. A typical obstacle, for example, not having set the environment variable CT_MANAGEMENT_SCOPE, is also mentioned there.

**Remote node unable to access shared disks after node crash**

If your AIX node crashes and the remote node is unable to access the shared disks (that is, if they are locked), this may be because your shared volume groups are not enhanced concurrent capable. For instructions on making shared volume groups enhanced concurrent capable on AIX, see *IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide*.

**Reporting problems**

Problems for which no troubleshooting information is available should be reported as PMRs against the product IBM Tivoli System Automation for Multiplatforms. When you report the problem, supply the following information:

- The debug and trace data you collected before troubleshooting (see "Before you begin: Collect debug and trace data" on page 279)
Troubleshooting

- A short description of the automation policy that was active when the problem occurred. To retrieve the required information you can use the command `sampolicy -s`.
- A short description of the tasks performed before the error occurred

Known problems and limitations

**Tabular output is incorrectly aligned for multi-byte languages**
Tabular output of line commands may not be aligned correctly when displayed in a shell that has been set to a locale for a multi-byte language. This is caused by problem in the Perl scripting language due to which the width of multi-byte characters is not calculated correctly.
An automation domain is not displayed in the topology tree

If a first-level automation domain is not visible in the topology tree in the operations console, perform the following steps:

1. Check if the adapter is running. To do this, issue the following command on one of the nodes of the domain:
   ```shell
   samadapter status
   ```
   If the adapter is running, a message like in the following example comes up:
   ```shell
   samadapter is running on sapb13
   ```
   If the adapter is automated, a message like in the following example comes up:
   ```shell
   Automated ResourceGroup 'samadapter-rg' runs on sapb13
   ```
   Make a note of the name of the node on which the adapter runs (in the example this is sapb13) and proceed with step 4.

2. If the adapter is not running, issue the following command to check if the domain is online:
   ```shell
   lsrpdomain
   ```
   A message like in the following example comes up:
   ```shell
   Name OpState RSCTActiveVersion MixedVersions TSPort GSPort
domain1 Online 2.5.5.1 No 12347 12348
   ```
   If OpState is not Online, start the domain.

3. If the domain is online, start the adapter with the following command:
   ```shell
   samadapter start
   ```
   After the start message has appeared, reissue the following command:
   ```shell
   samadapter status
   ```

4. If the adapter is running, check again in the operations console if the domain now appears in the topology tree.

5. If the domain is still not visible, you need the connection information that you specified in the adapter configuration dialog to resolve the problem. Perform the following steps:
Troubleshooting

a. Launch the adapter configuration dialog of System Automation for Multiplatforms. To do this, issue the following command on a node in the domain:
   ```
cfgsamadapter
   ```

b. On the entry panel of the configuration dialog, click **Configure**.

c. Open the Adapter page on the Configure panel and record the values that appear in the following fields for future reference:
   - **Host name or IP Address**
   - **Request port number**

   This is the connection information the host running the operations console uses to reach the adapter on any of the nodes in the domain.

d. Open the page host using adapter and record the values that appear in the following fields for future reference:
   - **Host name or IP Address**
   - **Event port number**

   This is the connection information the adapter on any of the nodes in the domain uses to reach the host running the operations console.

6. Check if the operations console can be reached from each node in the domain. A simple test is `ping <operations console host>`.

   If there is a firewall between the nodes of the domain and the host running the operations console, check with the network administrator if the firewall permits a connection between the node (page Adapter: **Host name or IP Address**) and the host running the operations console (page Host using adapter: **Host name or IP Address** and **Event port number**).

7. The adapter determines whether SSL must be used for the communication with the operations console. To check the SSL settings of the adapter, launch the adapter configuration dialog using the command `cfgsamadapter`. On the Security page of the configuration dialog, verify that the SSL settings are correct. If the **Enable SSL** check box is selected, the operations console must also be configured to support SSL in direct access mode.

   Check this by running `/cfgdirect.sh` in the directory `<isc_runtime_root>/AppServer/profiles/default/Tivoli/EEZ`. Verify that the settings for the `eez-ssl-truststore`, `eez-ssl-keystore`, `eez-ssl-keystore-password`, and `eez-ssl-keystore-alias` keys used are the same as those as shown on the Security page of the configuration dialog. (For more information about the adapter configuration dialog, see [IBM Tivoli System Automation for Multiplatforms Installation and Configuration Guide](#).

   You may not be able to stop the adapter using the **samadapter stop** command if the key chosen in adapter is incorrect or does not exist. In this case, use the following command to determine the process-ID:
   ```
   ps ax | grep sam.adapter
   ```

   Then use the following command to terminate the samadapter process:
   ```
   kill <process-ID>
   ```
Troubleshooting

8. On the host running the operations console, use `netstat` to find out if it is listening for events on the event port defined in `Event port number`.

   When the event port number is set to 2002 on a Windows host, `netstat` brings up a message like in the following example:

   ```
   C:\>netstat
   Active Connections
   Proto Local Address Foreign Address State
   ... TCP E2EHOST:2002 sapb13.boeblingen.de.ibm.com:45688 ESTABLISHED ...
   ```

   If `netstat` does not display any information about the event port defined in `Event port number`, open the file `/etc/hosts` (on Windows the file is located in `C:\WINDOWS\system32\drivers\etc\hosts`) and verify that the loopback address (127.0.0.1) is not related to the actual host name. The loopback address should be related to localhost only.

   For example, the entry in `/etc/hosts` may look like the following:

   ```
   127.0.0.1 localhost.localdomain localhost
   ```

9. Check if each node in the domain can be reached from the operations console.
   A simple test is `ping <hostname or IP Address>`.

   If there is a firewall between the host running the operations console and the nodes of the domain, check with the network administrator if the firewall permits a connection between the host running the operations console (page host using adapter: `Host name or IP Address` and `Request port number`) and the node (page Adapter: `Host name or IP Address`).

10. On the on node on which the adapter is running, use `netstat` to find out if it is listening on the event port defined in `Request port number`.

    For example, when the Request port number is set to 2001, `netstat` brings up a message like this on AIX, Solaris, and Linux hosts:

    ```
    saph13:~ # netstat -atn |grep 2001
tcp 0 0 9.152.20.113:2001 :::* LISTEN
    ```

11. When the communication between all ports has been established correctly (see the descriptions above), check whether the EEZ Publisher is running. The EEZ Publisher must be running on the master node of System Automation for Multiplatforms.

    To check if the Publisher is running, perform the following steps:

    a. Issue the following command on one of the nodes of the automation domain:

       ```
       - issue lssamctrl
       ```

       If the Publisher is enabled, you will receive output like in the following example:

       ```
       safl03:~ # lssamctrl | grep Publisher
       EnablePublisher = EEZ
       ```

    b. Issue the following command on the master node of System Automation for Multiplatforms (see “Administering the recovery resource manager” on page 17 how to find out the master node):

       ```
       ps ax
       ```

       You should receive output like in the following example:

       ```
       safl04:~ # ps ax | grep Publisher
       25756 ? S 0:00 TECPublisher
       /etc/opt/IBM/tsamp/sam/cfg/EEZPublisher.conf EEZ
       ```
c. Issue the following command on the IBM Tivoli System Automation for Multiplatforms node on which the adapter is running:
```
netstat
```
You should receive output like in the following example:
```
Safli03:~ # netstat -atn | grep 5539
tcp 0 0 :::5539 :::* LISTEN
tcp 0 0 9.152.21.82:5539 9.152.20.92:32793 ESTABLISHED
```
If the Publisher is not running or communication on port 5539 cannot be established, perform the following steps:

a. Check that the file `/etc/Tivoli/tec/samPublisher.conf` contains the following entry:
   ```
   #--SAMP-EEZ:
   Publisher=EEZ
   LibraryPath=libTECPublisher.so
   ConfigPath=/etc/opt/IBM/tsamp/sam/cfg/EEZPublisher.conf
   ```

b. Check that the file `/etc/opt/IBM/tsamp/sam/cfg/EEZPublisher.conf` contains the following entries:
   ```
   ServerLocation=adapter_ip_address
   ServerPort=5539
   ```
   The value specified for `adapter_ip_address` in the file must match the value provided on the Adapter tab of the adapter configuration dialog.

12. If the domain still does not appear in the operations console, contact IBM support and provide diagnostic information:

a. On each node in the domain, find out where the trace files are located. The trace files can be found in the `/eez/logs` subdirectory of the Tivoli Common Directory. To find the path to the Tivoli Common Directory, issue the following command:
   ```
cat /etc/ibm/tivoli/common/cfg/log.properties
```
   The command returns the path to the Tivoli Common Directory, for example:
   ```
Tivoli_common_dir=/var/ibm/tivoli/common
```
   This means that the trace files can be found in the following directory:
   ```
/var/ibm/tivoli/common/eez/logs
```

b. Use `tar` to package all files in the directory and provide the archive to IBM support.

---

Time stamps displayed in the operations console are not in local time

The values that appear in the time stamps in the operations console are derived from the time zone settings of the operating system where the Integrated Solutions Console server is installed. If the time stamps are not in local time, check the time zone settings on your Integrated Solutions Console server.
Typically, you use the configuration tools that are provided with the operating system to change the time settings:

- On AIX, you can configure the time settings with the smit or smitty system configuration tool. Use the menu entries **System environments —> Change/Show Date and Time** to adjust the time settings.

- On SuSE Linux, you can use the yast2 or yast system configuration tools. Use the menu entries **System —> Date and Time (SLES-9)** or **System —> Set Time Zone (SLES-8)**.

- On Red Hat Linux distributions, you can use the configuration tools redhat-config-time or system-config-time.

- On Windows, you can adjust the time settings with the Date and Time entry on the Control Panel.

You may have to restart your operating system for the changes to take effect.

**Tip:**

**AIX, Solaris, Linux:** If you have modified the time zone settings as described above but the values displayed in the time stamps in the operations console are still inappropriate, you can set the environment variable **TZ** to resolve the problem.

**Examples:**

- To set the time zone for Berlin, Germany, use the following command:
  
  ```
  export TZ="Europe/Berlin"
  ```

- To set the time zone to US Eastern Standard Time, use the following command:
  
  ```
  export TZ="US/Eastern"
  ```

**After a node reboot, fixed resources show incorrect error states in the operations console**

Fixed resources that are displayed in the operations console show incorrect error states after a node reboot. The correct states will be displayed once the resources are started on this node.

**Viewing log files in a multilingual environment**

You can use the argument `client.encoding.override=UTF-8 JVM` to configure an application server for UCS Transformation Format. This format enables an application server to handle most character encodings. You should specify this argument if you require multiple-language encoding support in the administrative console. This may be the case if you have some automation domains running in locales with encodings other than what is specified in the client browser.

For example, if you want to use the SA operations console to view the log for a first-level automation domain running in a German locale, but you have set the default language of your browser to Japanese, some German special characters that may appear in the log will not be displayed correctly in the Japanese browser.

To configure an application server for UCS Transformation Format, perform these steps:

1. In the administrative console, click **Servers > Application servers** and select the server you want to enable for UCS Transformation Format.
Troubleshooting

2. Then, under Server Infrastructure, click **Java and Process Management > Process Definition > Java Virtual Machine**.

3. Specify `-Dclient.encoding.override=UTF-8` for **Generic JVM Arguments** and click **OK**. When this argument is specified, UCS Transformation Format is used instead of the character encoding that would be used if the `autoRequestEncoding` option was in effect.

4. Click Save to save your changes.

5. Restart the application server.

**Recovering automatically saved files of the policy editor**

The Policy Editor automatically saves your current work on automation policies on the system where the operations console is installed. It creates temporary files in the directory used by the operating system for this purpose. For example:

**Windows**: `C:\Documents and Settings\Administrator\Local settings\Temp` (as defined in environment variable `%TEMP%`).

**AIX, Linux, z/OS**: `/tmp`

The file name of the autosave files is constructed based on the following pattern:

```
.eezautosave_<Timestamp>_<ISC_user_ID>_<ISC_session_ID>_<internal ID>.xml
```

If you experience a session timeout or some other kind of problem with the policy editor, you can retrieve your work by continuing with the contents of the automatically saved files.

1. Transfer the newest autosaved file to your local machine, for example, by using `ftp`.

2. Open the policy editor and select **Edit an existing policy**. Load the locally saved policy.

Notes:

1. On Unix platforms, these files are hidden using a “.”. To see them, use the `ls -a` command.

2. All automatically saved files are deleted when the WebSphere Application Server is terminated.
Appendix C. Using IBM Support Assistant

IBM Support Assistant is a free, stand-alone application that you can install on any workstation. IBM Support Assistant saves you time searching product, support, and educational resources and helps you gather support information when you need to open a problem management record (PMR) or Electronic Tracking Record (ETR), which you can then use to track the problem.

You can then enhance the application by installing product-specific plug-in modules for the IBM products you use. The product-specific plug-in for Tivoli System Automation for Multiplatforms provides you with the following resources:

- Support links
- Education links
- Ability to submit problem management reports

Installing IBM Support Assistant and the Tivoli System Automation for Multiplatforms plug-in

To install the IBM Support Assistant V3.0, complete these steps:

- Go to the IBM Support Assistant Web Site:
  

- Download the installation package for your platform. Note that you will need to sign in with an IBM user ID and password (for example, a MySupport or developerWorks® user ID). If you do not already have an IBM user ID, you may complete the free registration process to obtain one.

- Uncompress the installation package to a temporary directory.

- Follow the instructions in the *Installation and Troubleshooting Guide*, included in the installation package, to install the IBM Support Assistant.

To install the plug-in for Tivoli System Automation for Multiplatforms, complete these steps:

1. Start the IBM Support Assistant application. IBM Support Assistant is a Web application that is displayed in the default, system configured Web-browser.
2. Click the **Updater** tab within IBM Support Assistant.
3. Click the **New Products and Tools** tab. The plug-in modules are listed by product family.
4. Select **Tivoli > Tivoli System Automation for Multiplatforms**.
5. Select the features you want to install and click **Install**. Be sure to read the license information and the usage instructions.
6. Restart IBM Support Assistant.
Appendix D. Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user’s responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Mail Station P300
2455 South Road
Poughkeepsie New York 12601-5400
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS"
WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

**Trademarks**

- IBM, the IBM logo, ibm.com, AIX, DB2, developerWorks, HACMP, NetView, Tivoli, Tivoli Enterprise, Tivoli Enterprise Console, WebSphere, and z/OS are trademarks of International Business Machines Corporation in the United States, other countries, or both. IBM Redbooks and the IBM Redbooks logo are registered trademarks of IBM.
- Adobe, Acrobat, Portable Document Format (PDF), and PostScript are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, other countries, or both.
- Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
- Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.
- UNIX is a registered trademark of The Open Group in the United States and other countries.
- Other company, product, and service names may be trademarks or service marks of others.
Index

Special characters
(EIF) Tivoli Event Integration Facility 209

A
about this book xi
access roles
   for IBM Tivoli System Automation for Multiplatforms 115
additional info page
   overview 132
adding a member resource to resource group 47
addrrgmbr command 47
addrpnode command 15
affinity relationship 84
aggregate resource 222
AllowedNode resource group attribute 39
AntiAffinity relationship 85
AntiCollocated relationship 82
attributes
   dynamic 3, 30
   NodeNameList 31
   persistent 3, 30
   resource attributes, description of 3, 29
SelectFromPolicy 31
used by equivalencies 53
used by IBM.Application 222
used by IBM.ServiceIP 238
used by IBM.Test 243
used by managed relationships 58
used by resource groups 38
used by resources 30
audience of this book xi
audit log
   GblResRM 281
   RecoveryRM master daemon 280
automation
   events affecting automation 94
   parameters 181
automation manager
   binder 277
   logic deck 278
   overview 277
automation policy management 183
AutomationDetails attribute 44

B
baroc file 210
binder
   definition 91
   binding states 277

C
changing
   a resource group 49
   an equivalency 55
   attributes of resource group members 49
   chequ command 55
   chrg command 49
   chrgmbr command 49
   clusters
      adding nodes to 15
      defining and administering 12
      removing 16
      taking offline 16
   collocated relationship 79
   commands
      addrrgmbr 47
      addrpnode 15
      chequ 55
      chrg 49
      chrgmbr 49
      lssequ 54
      lsrel 88
      lsrg 48
      lsrpnode 16
      mkequ 54
      mkrel 88
      mkrg 47
      mkrpdomain 12
      preprpnode 12
      rmequ 55
      rmrel 89
      rmrg 50
      rmrgmbr 49
      rmrpdomain 16
      rmrpnode 16
      smactrl 218
      samdiag 207
      startrpdomain 12, 16
      startrpnode 15
      stoprpdomain 16
      stoprpnode 16
   Condition attribute 59
   conditions for location relationships 77
   configuration quorum 5, 162
   ConfigValidity
      dynamic attribute 216
      troubleshooting 281, 284, 285
      verifying resources 216
   constituent resources 222
   creating
      an equivalency 54
      relationships 88
      resource groups 26, 47
   credential vault
      Integrated Solutions Console 141
   critical resource 161
   DependsOn
      behavior schemes 68
      relationship attribute 68
      start behavior 68
      stop behavior 70
   DependsOnAny relationship attribute 74
diagnosing resources 207
dynamic attribute 30
   AutomationDetails 38
   ConfigValidity 216
definition 3
   MoveStatus 44, 198
   OpQuorumState 166
   OpState 32, 38, 223, 238
   TopGroup 38

E
e-mail address xiii
   equivalencies
      attributes used by 53
      changing 55
      description of 4, 51
      listing 54
      making (creating) 54
      removing 55
      rules for using 52
   establishing an equivalency 54
   establishing the node list for resource group members 47
   event response resource manager (ERRRM) 218
   events affecting automation 94
   ExcludedList attribute 42
   ExcludedNodes parameter 181

F
fail tiebreaker 167
fixed resources
   definition 31
floating resources
   definition 32
force down behavior of DependsOn 68
ForcedDownBy relationship 75

G
gateways, default
   netmon.cf 268
   GblResRM
      audit log 281
highlighting xii
highly available network 267
how to use this book xi

I
IBM.Application 222
IBM.Equivalency resource class 134
IBM.RecoveryRM 17
IBM.ServiceIP 268
attributes 238
definition 236
IBM.TestRM 243
IFNotOffline condition 78
IFNotOnline condition 78
IOffline condition 78
IFOnline condition 77
information area overview 130
Integrated Solutions Console credential vault 141
logging on 117
interface bonding 275
ISO 9000 xii
IsStartable relationship 86

L
listing an equivalency 54
relationships 88
resource groups 48
location relationships 77
conditions IFNotOffline 78
IFNotOnline 78
IOffline 78
IFOnline 77
logging in
Integrated Solutions Console 117
lsed command 54
lsrel command 88
lsrg command 48
lsrpnode command 16

M
making a resource group 47
managed relationships attributes used by 58
description of 57
managed resource description of 4, 30
Mandatory attribute of ManagedResource class 46
MemberClass equivalency attribute 53
MemberLocation resource group attribute 40
MemberOf attribute 46
Membership equivalency attribute 53
mkequ command 54
mkrcon command 47
mkrpdomain command 12
MoveStatus, dynamic attribute 44, 198
moving resources 137
persistent attribute (continued)
MonitorCommand 224
MonitorCommandPeriod 225
MonitorCommandTimeout 225
Name 41, 58, 223, 238, 243
NetMask 239
NetPrefix 239
NodeNameList 31, 223, 239, 244
NominalState 41
OpQuorumOverride 178
Priority 41
ProtectionMode 165, 227, 240
Relationship 59
ResourceType 31, 223, 239, 244
RunCommandsSync 226
SelectFromPolicy 53
SelectString 53
Source 58
StartCommand 223
StartCommandTimeout 225
StopCommand 224
StopCommandTimeout 226
Target 58
TimeToStart 244
TimeToStop 244
UserName 227
WriteToSyslog 244
policy automation policy management 183
modifying a policy 183
policy based automation 1
replacing a policy 183
sample policies 1
saving and restoring a policy 183
policy editor
layout 143
policy page
overview 131
preprpnode command 12
prerequisite knowledge for this book xi
Priority resource group attribute 41
ProtectionMode attribute 165
publisher configuration file 210
publishing System Automation for Multiplatforms internal attributes 214

Q
quorum configuration quorum 5, 162
getting information 166
majority of nodes 161
operational quorum 5, 162
tie breaker 163

R
recovery resource manager (IBM.RecoveryRM) 17
RecoveryRM master daemon audit log 280
relationship DependsOn 68
DependsOnAny 74
Relationship attribute 59
relationships
Affinity 84
AntiAffinity 85
AntiCollocated 82
Collocated 79
creating 88
ForcedDownBy 75
IsStartable 86
listing 88
location relationship 77
removing 89
StartAfter 60
StopAfter 66
relationships page
overview 132
removing a member resource from a resource
group 49
an equivalency 55
relationships 89
resource groups 50
removing a member resource
resource group 49
request
cancel 137
Request
Online 134
stop 135
request details
displaying 136
request list
displaying 136
requests
processing of a move request 198
starting and stopping resource groups
and resources 190
resource
critical 161, 165
description of 3, 29
description of resource attributes 3, 29
description of resource class 4, 29
verifying 216
resource class
definition 4
resource classes
definition 29
IBM.Application 221
IBM.ServiceIP 236
IBM.Test 243
IBM.TieBreaker 167
resource group
adding a member resource to 47
allowing it to become Online 94
attributes used by 38
changing attributes of 49
description of 35
establishing the node list for 47
listing (incl. its resource members) 48
making (creating) 47
moving 197
removing 50
starting 49
states 41
stopping 49
resource group members
changing the attributes of 49
resource manager
Global Resource RM 221
resource class IBM.Application 221
resource class IBM.ServiceIP 236
resource class IBM.Test 243
test resource manager 243
resource protection methods
on Windows 163
resource states
overview 278
state cycle 112
resource type 31
ResourceId attribute 182
resources
attributes used by 30
moving 137
resources section
overview 125
ResourceType attribute for resource 31
rmeq command 55
rml command 89
rmgr command 50
rmrgmbr command 49
rmrpdomain command 16
rmrpnode command 16

S
samctrl command 218
samdiag command 207
SelectFromPolicy attribute 31
SelectFromPolicy equivalency
attribute 53
SelectString equivalency attribute 53
setting up a highly available
network 267
shadow resources
using 199
Source attribute 58
start behavior of DependsOn 68
start request 134
StartAfter relationship 60
StartAfter/IfPossible relationship 63
Starting
resources 134
starting a resource group 49
startrpdomain command 12, 16
startrpnode command 15
state transitions for a resource group 43
stop behavior of DependsOn 70
stop request 135
StopAfter relationship 66
StopCommand
Offline timeout 98
SA_RESET 98
StopCommandTimeout 98
Stopping
resources 135
stopping a resource group 49
stoprpdomain command 16
stoprpnode command 16
System Automation for Multiplatforms
going started 11
setting up a highly available
network 267
system resource controller 232

T
Target attribute 58
TEC (Tivoli Enterprise Console) 209
TEC (Tivoli Enterprise Console) event
interface 209
TEC EIF configuration file 211
TEC publisher function 210
test resource manager 243
tie breaker 5, 167, 176
tie breakers
types 163
TimeOut and RetryCount parameter 179
Tivoli Enterprise Console (TEC) 209
Tivoli Enterprise Console (TEC) event
interface 209
Tivoli Event Integration Facility
(EIF) 209
TopGroup resource group attribute 43
topology tree
hiding domains 124
Located here column 125
overview 123
Status column 124
trademarks 308
troubleshooting
audit log 280
collecting information for 279
commands 281
ConfigValidity 281, 284, 285
ctsnap command 279
system log 279
using audit logs 281

U
user credentials
managing 141
user groups
for IBM Tivoli System Automation for
Multiplatforms 115

V
verifying resources
ConfigValidity 216
VMTIMEBOMB 164

W
Web browsers
configuring 117
JavaScript 117
multiple browser windows 118
security level 117
security settings 117
supported 117

X
xDR
overview 215
supported Linux distributions 215
Readers’ Comments — We’d Like to Hear from You

System Automation for Multiplatforms
Administrator's and User's Guide
Version 3.1

Publication No. SC33-8415-02

We appreciate your comments about this publication. Please comment on specific errors or omissions, accuracy, organization, subject matter, or completeness of this book. The comments you send should pertain to only the information in this manual or product and the way in which the information is presented.

For technical questions and information about products and prices, please contact your IBM branch office, your IBM business partner, or your authorized remarketer.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you. IBM or any other organizations will only use the personal information that you supply to contact you about the issues that you state on this form.

Comments:

Thank you for your support.
Submit your comments using one of these channels:
• Send your comments to the address on the reverse side of this form.
• Send a fax to the following number: FAX (Germany): 07031 16-3456
  FAX (Other Countries): +49 7031 16-3456
• Send your comments via e-mail to: eservdoc@de.ibm.com

If you would like a response from IBM, please fill in the following information:

Name

Address

Company or Organization

Phone No.

E-mail address