



EXPRESSCLUSTER X 4.2 for Linux Hardware Feature Guide

Release 1

NEC Corporation

Apr 10, 2020

TABLE OF CONTENTS:

- 1 Preface** **1**
 - 1.1 Who Should Use This Guide 1
 - 1.2 How This Guide is Organized 1
 - 1.3 EXPRESSCLUSTER X Documentation Set 1
 - 1.4 Conventions 2
 - 1.5 Contacting NEC 3

- 2 Linkage with specific hardware** **5**
 - 2.1 NX7700x series linkage 5
 - 2.2 Express5800/A1080a or Express5800/A1040a series linkage 5

- 3 Linkage with Server Management Infrastructure** **15**
 - 3.1 Overview of the server management infrastructure 15
 - 3.2 Overview of linkage between the server management infrastructure and EXPRESSCLUSTER 15
 - 3.3 Setup of the function to link with the server management infrastructure 16
 - 3.4 Message receive monitor resources 17

- 4 Legal Notice** **21**
 - 4.1 Disclaimer 21
 - 4.2 Trademark Information 21

- 5 Revision History** **23**

1.1 Who Should Use This Guide

The *EXPRESSCLUSTER X Hardware Feature Guide* is intended for system administrators. Detailed information of the functions for linking with specific hardware is covered in this guide.

The guide provides supplemental information to the *EXPRESSCLUSTER X Installation and Configuration Guide*.

For information on the construction and the operation of clusters, refer to the guide.

1.2 How This Guide is Organized

- *2. Linkage with specific hardware*: Provides information on linkage with specific hardware.
- *3. Linkage with Server Management Infrastructure*: Provides information on linkage with Server Management Infrastructure.

1.3 EXPRESSCLUSTER X Documentation Set

The EXPRESSCLUSTER manuals consist of the following six guides. The title and purpose of each guide is described below.

EXPRESSCLUSTER X Getting Started Guide

This guide is intended for all users. The guide covers topics such as product overview, system requirements, and known problems.

EXPRESSCLUSTER X Installation and Configuration Guide

This guide is intended for system engineers and administrators who want to build, operate, and maintain a cluster system. Instructions for designing, installing, and configuring a cluster system with EXPRESSCLUSTER are covered in this guide.

EXPRESSCLUSTER X Reference Guide

This guide is intended for system administrators. The guide covers topics such as how to operate EXPRESSCLUSTER, function of each module, and troubleshooting. The guide is complement to the Installation and Configuration Guide.

EXPRESSCLUSTER X Maintenance Guide

This guide is intended for administrators and for system administrators who want to build, operate, and maintain EXPRESSCLUSTER-based cluster systems. The guide describes maintenance-related topics for EXPRESSCLUSTER.

EXPRESSCLUSTER X Hardware Feature Guide

This guide is intended for administrators and for system engineers who want to build EXPRESSCLUSTER-based cluster systems. The guide describes features to work with specific hardware, serving as a supplement to the Installation and Configuration Guide.

EXPRESSCLUSTER X Legacy Feature Guide

This guide is intended for administrators and for system engineers who want to build EXPRESSCLUSTER-based cluster systems. The guide describes EXPRESSCLUSTER X 4.0 WebManager, Builder, and EXPRESSCLUSTER Ver 8.0 compatible commands.

1.4 Conventions

In this guide, **Note**, **Important**, **See also** are used as follows:

Note: Used when the information given is important, but not related to the data loss and damage to the system and machine.

Important: Used when the information given is necessary to avoid the data loss and damage to the system and machine.

See also:

Used to describe the location of the information given at the reference destination.

The following conventions are used in this guide.

Convention	Usage	Example
Bold	Indicates graphical objects, such as fields, list boxes, menu selections, buttons, labels, icons, etc.	In User Name, type your name. On the File menu, click Open Database.
Angled bracket within the command line	Indicates that the value specified inside of the angled bracket can be omitted.	clpstat -s[-h host_name]
#	Prompt to indicate that a Linux user has logged on as root user.	# clpcl -s -a
Monospace (courier)	Indicates path names, commands, system output (message, prompt, etc.), directory, file names, functions and parameters.	/Linux/4.2/en/server/

Continued on next page

Table 1.1 – continued from previous page

Convention	Usage	Example
Monospace bold (courier)	Indicates the value that a user actually enters from a command line.	Enter the following: # clpcl -s -a
<i>Monospace italic</i> (courier)	Indicates that users should replace italicized part with values that they are actually working with.	clpstat -s [-h host_name]

1.5 Contacting NEC

For the latest product information, visit our website below:

<https://www.nec.com/global/prod/expresscluster/>

LINKAGE WITH SPECIFIC HARDWARE

This chapter explains the settings for the linkage between specific hardware and EXPRESSCLUSTER.

This chapter covers:

- 2.1. *NX7700x series linkage*
- 2.2. *Express5800/A1080a or Express5800/A1040a series linkage*

2.1 NX7700x series linkage

2.1.1 NX7700x series linkage

It can't be used.

2.2 Express5800/A1080a or Express5800/A1040a series linkage

2.2.1 Express5800/A1080a or Express5800/A1040a series linkage

Express5800/A1080a or Express5800/A1040a series linkage takes over operations by linking to EXPRESSCLUSTER if the BMC installed on a server detects an error.

This function uses BMC heartbeat resources and message receive monitor resources.

BMC heartbeat resources monitor whether the target servers are active by using the BMC network.

Setting a message receive monitor resource for Express5800/A1080a or Express5800/A1040a series linkage makes the BMC installed on the server report an error to execute a recovery operation.

2.2.2 Notes on Express5800/A1080a or Express5800/A1040a series linkage

To use this function, the following conditions must be satisfied. Hardware - The hardware and firmware of the BMC must support this function. For the available models, refer to "Servers supporting Express5800/A1080a or Express5800/A1040a series linkage" in "Installation requirements for EXPRESSCLUSTER" in the Getting Started Guide.

- The ipmi service must be started in advance.
- To use this function in the same environment as that for ESMPRO Agent, use the ESMPRO Agent control panel (ESMamsadm) to change the agent event operation after report from **Shutdown** to **No Operation**.

Source name: ESMCOMMONSERVICE, ID: C00001FD, C0000203, C0000454

[Starting the control panel (ESMamsadm)]

- 1) Log in as a root user.
- 2) Move to the directory containing ESMamsadm.

```
# cd /opt/nec/esmpro_sa/bin/
```
- 3) Start the control panel (ESMamsadm).

```
# ./ESMamsadm
```

[Specifying an operation after report for each monitor event]

- 1) Start the control panel (ESMamsadm) and select **Agent Event Setting**.
- 2) Select a source for **Source Name**. (Use the ↑ or ↓ key.)
- 3) Select an event ID for **Event ID**. (Use the ↑ or ↓ key.)
- 4) Click the **Setting...** button. The **Monitor Event Setting** window appears.
- 5) Change **Operation after Report** from **Shutdown** to **No Operation**.

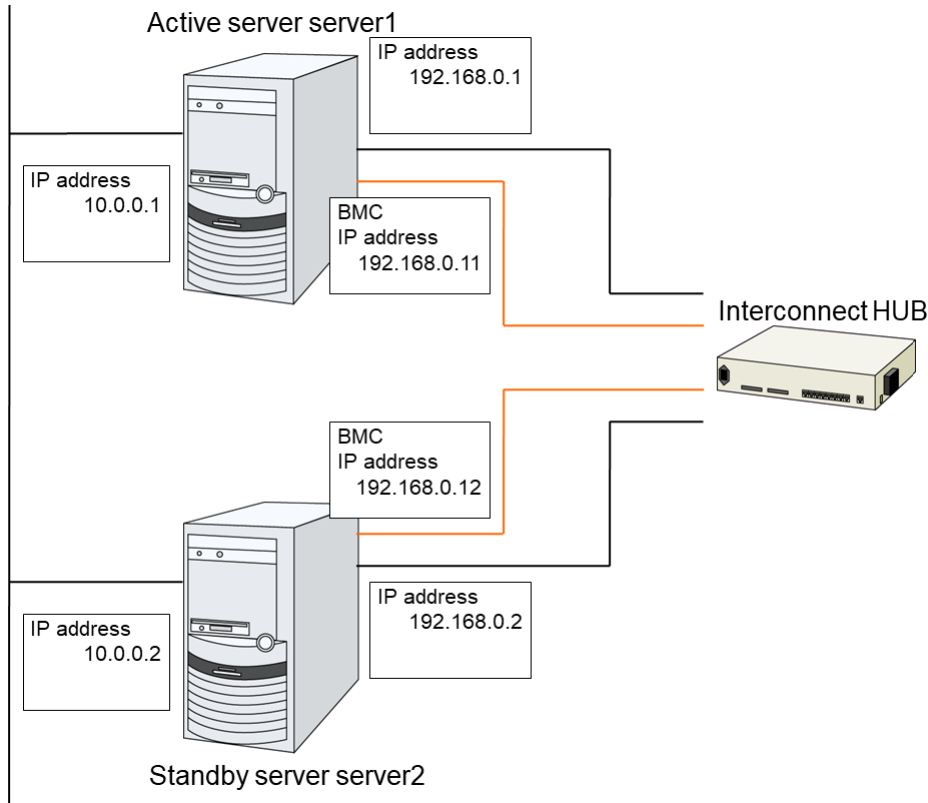
For details, refer to the *ESMPRO Agent User's Guide (Linux)*.

2.2.3 Creating configuration information for using Express5800/A1080a or Express5800/A1040a series linkage

The following describes, using an example, how to create the configuration information needed to use the Express5800/A1080a or Express5800/A1040a series linkage function.

Sample cluster environment

The creation of configuration information is explained below using an example for configuring a cluster environment with the following network configuration. In the figure below, disk configuration and other data are omitted because they are not directly related to the Express5800/A1080a or Express5800/A1040a series linkage function.



The following table lists sample values for the cluster configuration information that is used to configure the cluster system shown in the above figure. The following pages provide a step-by-step procedure for creating cluster configuration information for these conditions. When actually setting values, use the configuration information for the cluster to be configured.

	Parameter to be set	Value	
Cluster configuration	Cluster name	cluster	
	Number of servers	2	
	Number of failover groups	1	
	Number of monitor resources	2	
	• Heartbeat resources	Number of LAN heartbeat resources	2
		Number of COM heartbeat resources	0
		Number of disk heartbeat resources	0
		Number of BMC heartbeat resources	1
First server information (master server)	Server name	server1	

Continued on next page

Table 2.1 – continued from previous page

	Parameter to be set	Value
	Interconnect IP address (dedicated)	192.168.0.1
	Interconnect IP address (backup)	10.0.0.1
	BMC server Web console IP address	192.168.0.11
Second server information	Server name	server2
	Interconnect IP address (dedicated)	192.168.0.2
	Interconnect IP address (backup)	10.0.0.2
	BMC server Web console IP address	192.168.0.12
First group	Type	Failover
	Group name	failover1
	Starting server	All servers
	Number of group resources	-
• Group resource (omit- ted)	-	-
	-	-
	-	-
First monitor resource (default)	Type	User-mode monitor
	Monitor resource name	userw
Second monitor resource	Type	Message receive monitor
	Category	BMCNOTICE
	Keyword (common)	192.168.0.1
	Keyword (Individual server setting: server1)	192.168.0.1
	Keyword (Individual server setting: server2)	192.168.0.2

Continued on next page

Table 2.1 – continued from previous page

	Parameter to be set	Value
	Recovery operation	Executes the failover for the recovery target.
	Recovery target	failover1

Procedure for creating cluster configuration information

The creation of cluster configuration information involves the following three basic steps: creating a cluster, creating groups, and creating monitor resources. The following shows the flow of the procedure for setting items specific to Express5800/A1080a or Express5800/A1040a series linkage. For other items, refer to "Creating cluster configuration information" in the Installation and Configuration Guide.

Note: An operation on cluster configuration information can be performed any number of times. Most settings can be modified later by using the rename function or the properties view function.

1 Creating a cluster

Create a cluster and add servers.

1-1 Adding a cluster

Add a cluster you want to construct and enter its name. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-2 Adding servers

Add servers and set their server names, IP addresses, and other items. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-3 Setting the network configuration

Set the network configuration between servers that constitute the cluster. Set BMC heartbeat resources as an item specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-4 Setting network partition resolution resources

Set the network partition resolution resources. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

2 Creating a failover group

Create a failover group that operates as a unit when failover occurs. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

2-1 Adding a failover group

Add a group that operates as a unit when failover occurs. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

2-2 Adding a group resource

Add a resource that constitutes a group. This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

3 Creating a monitor resource

Create a monitor resource that monitors the specified target in the cluster.

1. Adding a monitor resource (message receive monitor)

Add a monitor resource to use.

1 Procedure for creating a cluster

First, create a cluster. To the created cluster, add a server that is part of the cluster and determine the priorities of the server and heartbeat. Only steps 1 to 3 contain the settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-1 Adding a cluster

This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-2 Adding servers

This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

1-3 Setting the network configuration

Set the network configuration between the servers that constitute the cluster.

Priority	Type	MDC	server1	server2
1	Kernel Mode	Do Not Use	192.168.0.1	192.168.0.2
2	Kernel Mode	Do Not Use	10.0.0.1	10.0.0.2
3	BMC	Do Not Use	192.168.0.11	192.168.0.12

Configure the interconnect among the servers constructing the cluster. Click "Add" to add interconnect and select the type. For "Kernel mode", "User mode", "BMC", "DISK", "Witness HB" and "COM" settings, configure the route which is used for heartbeat. For "Mirror Communication Only" setting, configure the route which is used only for data mirroring communication. Configuring more than one routes is recommended. For "Kernel mode", "User mode", "DISK" and "COM" settings, click each server column cell and set an IP address or device. For "Witness HB" setting, click each server column cell to set "Use" or "Do not use", and then click "Properties" to set detailed settings. Click "Up" or "Down" to configure the priority to preferentially use the LAN only for the communication among the cluster servers. For "Mirror Communication Only" settings, click each server column cell to configure IP addresses. For the communication route which is used for data mirroring communication, select the mirror disk connect name to be allocated to the communication route in MDC column.

1. For the communication route used for BMC heartbeat transmission (interconnect), click a cell in the **Type** column, and then select **BMC**. Click a cell in the column for each server, and then enter the BMC server Web console IP address.

1-4 Setting network partition resolution resources

This item contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

2 Creating a failover group

This step contains no settings that are specific to Express5800/A1080a or Express5800/A1040a series linkage.

3 Creating a monitor resource

Create a monitor resource that monitors the specified target in the cluster. For the Express5800/A1080a or Express5800/A1040a series linkage function, set a message receive monitor. For details on the monitor,

see "2.2.4. Express5800/A1080a or Express5800/A1040a series linkage with message receive monitor resources".

3-1 Adding a monitor resource (Message receive monitor)

Add a monitor resource that monitors the reception of an error report from the server.

1. Click **Next** in **Group**.
2. The **Monitor Resource List** appears. Click **Add**.
3. The **Monitor Resource Definition** dialog box appears. Select the monitor resource type (Message receive monitor) from the **Type** box, and then enter the monitor resource name (mrw1) in the **Name** box. Click **Next**.
4. Enter the monitoring settings. Click **Next** without changing the default values.
5. Enter BMCNOTICE in the **Category** box and the IP address of the OS side of each server that can communicate with the BMC as the IP address of the destination of the error report in the **Keyword** box. Use the **Monitor(special)** tab because the IP addresses differ depending on the server. In this example, enter 192.168.0.1 for **Common** and **server1** and 192.168.0.2 for **server2**.
6. Set **Executing failover the recovery target** for **Recovery Action**. For Express5800/A1080a or Express5800/A1040a series linkage, select this setting for **Recovery Action**.
7. Set a recovery target. Click **Browse**. In the displayed tree view, select **failover1** and then click **OK**. **failover1** is set for **Recovery Target**.
8. Click **Complete**.

This completes the creation of cluster configuration information that is specific to Express5800/A1080a or Express5800/A1040a series linkage.

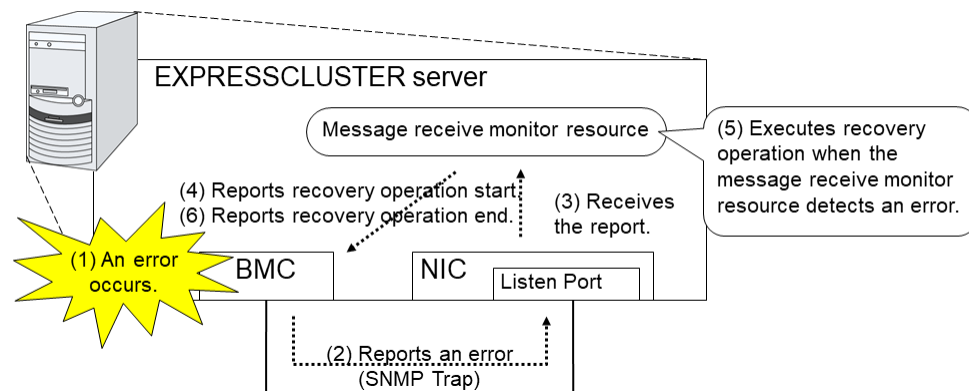
2.2.4 Express5800/A1080a or Express5800/A1040a series linkage with message receive monitor resources

This function enables EXPRESSCLUSTER to immediately execute a recovery operation if the BMC installed on the Express5800/A1080a or Express5800/A1040a series detects an error.

When this function is not used, and the BMC detects an error, server reset and other operations are executed immediately. For this reason, required operations such as application termination are not executed.

When this function is used, even if the BMC detects an error, server reset and other operations are performed after EXPRESSCLUSTER executes recovery.

The following figure is an overview of the operation that is performed when Express5800/A1080a or Express5800/A1040a series linkage with message receive monitor resources is used.



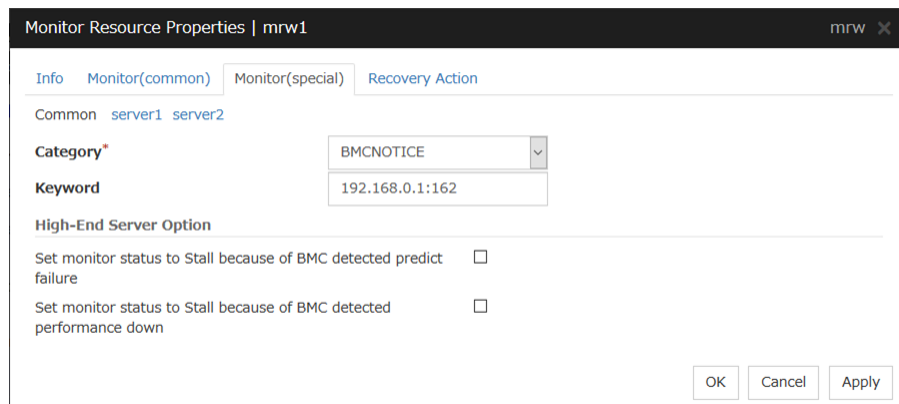
2.2.5 Notes on Express5800/A1080a or Express5800/A1040a series linkage with message receive monitor resources

In addition to the conditions listed in "2.2.2. Notes on Express5800/A1080a or Express5800/A1040a series linkage", the following condition must be satisfied.

- Since SNMP Traps are used for reporting an error from the BMC to a message receive monitor resource, communication using TCP/IP must be enabled between the BMC and OS network interfaces.
Also note the following points when setting BMC linkage:
- Set BMCNOTICE for **Category**.
- Unique values must be specified for **Keyword** on different servers. Specify the IP address and port number on the OS side connected to the BMC on each server as the destination of error reports.

2.2.6 Monitor(special) tab

For more information of other tabs, see "Monitor resource details" in the Reference Guide.



Category (within 32 bytes)

Specify a category.
Specify BMCNOTICE.

Keyword (within 1023 bytes)

Specify an IP address that can communicate with the BMC on each server. In this case, specify the IP address and port number for receiving error reports.

Use individual server settings to specify the values for each server.

The port number can be omitted (default: 162). To set the port number, use the same value for all the message receive monitor resources for the same server.

The format is as follows:

<IP address> [:<Port number>]

LINKAGE WITH SERVER MANAGEMENT INFRASTRUCTURE

This chapter provides an overview of the server management infrastructure included in Enterprise Linux with Dependable Support.

This chapter covers:

- 3.1. *Overview of the server management infrastructure*
- 3.2. *Overview of linkage between the server management infrastructure and EXPRESSCLUSTER*
- 3.3. *Setup of the function to link with the server management infrastructure*
- 3.4. *Message receive monitor resources*

3.1 Overview of the server management infrastructure

This server management infrastructure is included in Enterprise Linux with Dependable Support. This software provides the following functions:

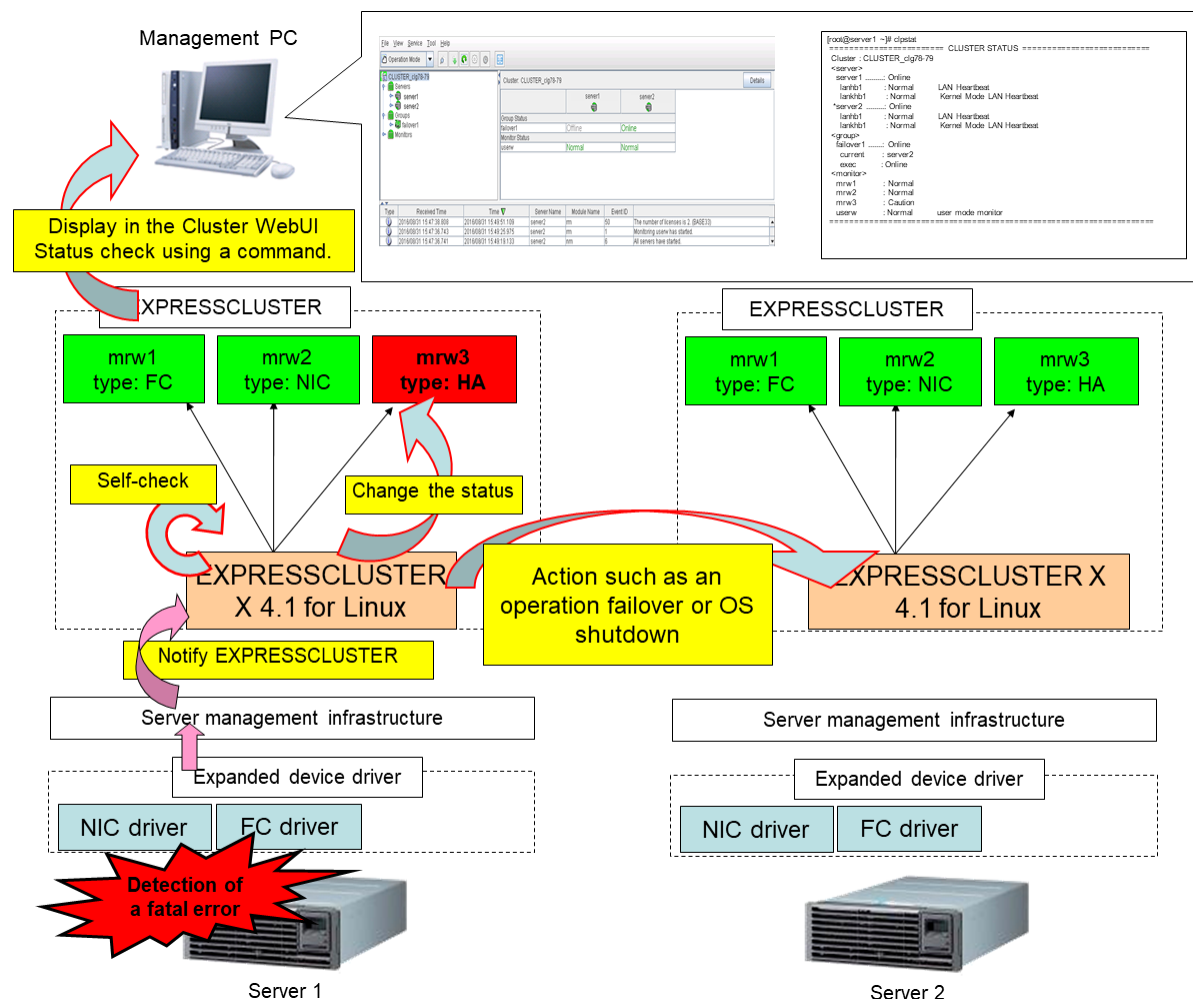
- Recording information about failures detected by the expanded device driver
- Linking with EXPRESSCLUSTER X to perform a failover when the expanded device driver detects a fatal system failure

For details, see the manual for Enterprise Linux with Dependable Support.

3.2 Overview of linkage between the server management infrastructure and EXPRESSCLUSTER

EXPRESSCLUSTER's function for linking with the server management infrastructure is not used for EXPRESSCLUSTER to perform monitoring itself. This linkage function is used for EXPRESSCLUSTER to receive messages spontaneously sent by the driver module and passively perform a failover or other processing.

The following shows an overview:



If a fatal system error occurs, the expanded device driver included in Enterprise Linux with Dependable Support (hereafter referred to as the expanded driver) sends a message to EXPRESSCLUSTER through the server management infrastructure. After receiving such a message, EXPRESSCLUSTER performs the following operations.

- EXPRESSCLUSTER makes the status of the corresponding message receive monitor (mrw) abnormal. The administrator can visually determine that an error was detected by checking the status using the Cluster WebUI or an EXPRESSCLUSTER command.
- When a failure occurs, EXPRESSCLUSTER performs an operation failover or shuts down the OS according to the specified action.

3.3 Setup of the function to link with the server management infrastructure

For details about resources other than the message receive monitor resource, see the EXPRESSCLUSTER manuals below.

- Installing EXPRESSCLUSTER
 - "Installing EXPRESSCLUSTER" in the Installation and Configuration Guide
 - "Registering the license" in the Installation and Configuration Guide

- Creating EXPRESSCLUSTER configuration information
 - "Creating the cluster configuration data" in the Installation and Configuration Guide
 - "Verifying a cluster system" in the Installation and Configuration Guide

To use the function for linking with the server management infrastructure, the message receive monitor resources must be registered with the cluster . To create configuration information, register the necessary message receive monitor resources as described in the manual. For the message receive monitor resources, see "Message receive monitor resources".

- Uploading EXPRESSCLUSTER configuration information
 - "Modifying the cluster configuration data" in the Installation and Configuration Guide

3.4 Message receive monitor resources

The message receive monitor resources monitor error messages reported from outside. This section only covers the part associated with linkage with the server management infrastructure. For other cases, see "Monitor resource details" in the Reference Guide.

3.4.1 Notes on message receive monitor resources

A message receive monitor resource cannot execute any scrip before the final action if it is linked with the server management infrastructure.

Do not use the `clprexec` command, because EXPRESSCLUSTER manages the status of a message receive monitor resource if it is linked with the server management infrastructure.

When the keyword by a message receive monitor resource is specified, if an error is detected in the device specified as the monitor target, an error occurs and the error correction action is performed .

If no device is specified as the keyword by a message receive monitor resource and an error is detected in any device that matches the Category, an error occurs and the error correction action is performed.

3.4.2 Category by a message receive monitor resource

A message receive monitor resource receives the following message types when it is linked with the server management infrastructure.

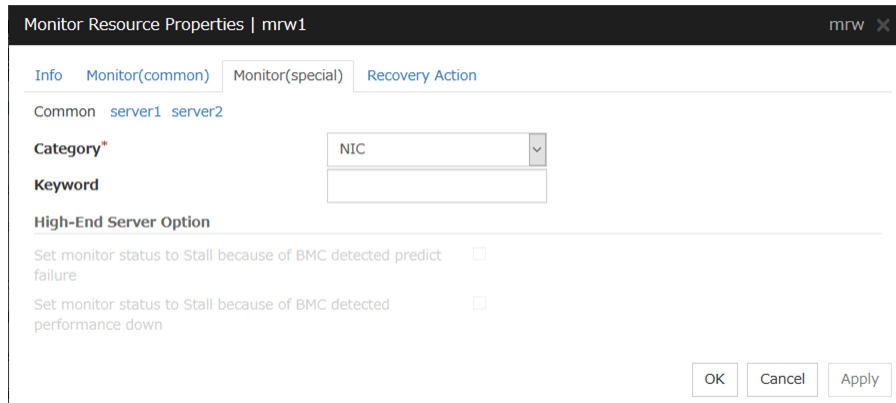
1. NIC
Monitors the error messages of network interface cards.
2. FC
Monitors the error messages of Fibre Channel.
3. HA/SS
Monitors the error messages of the EXPRESSCLUSTER X HA/StorageSaver.
4. HA/AM
Monitors the error messages of the EXPRESSCLUSTER X HA/ApplicationMonitor.
5. HA/RS
Monitors the error messages of the EXPRESSCLUSTER X HA/ResourceSaver.

6. SPS

Monitors the error messages of the SPS.

3.4.3 Monitor(special) tab

For more information on the **Info** tab and the **Monitor (common)** tab, see "Monitor resource details" in the Reference Guide.



The screenshot shows a window titled "Monitor Resource Properties | mrw1" with a close button. It has four tabs: "Info", "Monitor(common)", "Monitor(special)", and "Recovery Action". The "Monitor(special)" tab is active. Below the tabs, there are two server names: "server1" and "server2". The "Category*" field is a dropdown menu with "NIC" selected. The "Keyword" field is an empty text box. Under the "High-End Server Option" section, there are two checkboxes: "Set monitor status to Stall because of BMC detected predict failure" and "Set monitor status to Stall because of BMC detected performance down", both of which are currently unchecked. At the bottom right, there are three buttons: "OK", "Cancel", and "Apply".

Category (within 32 bytes)

Specify a category.

Be sure to select a default character string from the list box.

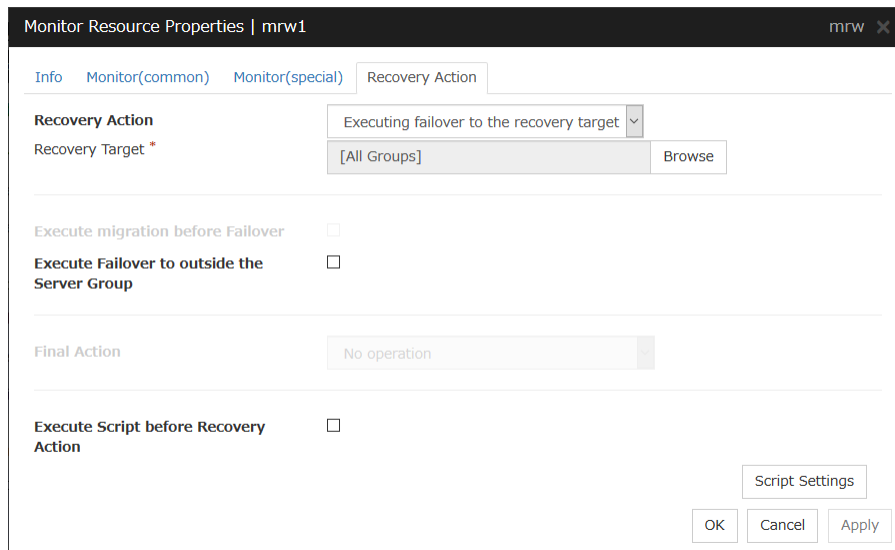
Keyword (within 1023 bytes)

Specify a monitor target.

3.4.4 Recovery Action tab

For more information on the **Info** tab and the **Monitor (common)** tab, see "Monitor resource details" in the Reference Guide.

Specify the recovery target and the action upon detecting an error. For message receive monitor resources, select "Restart the recovery target", "Executing failover to the recovery target", or "Execute the final action" as the action to take when an error is detected. However, if the recovery target is inactive, the recovery action is not performed.



Recovery Action

Select the action to take when a monitor error is detected.

- Executing recovery script
Let the recovery script run upon the detection of a monitor error.
- Restart the recovery target
Restart the group or group resource selected as the recovery target when a monitor error is detected.
- Executing failover to the recovery target
Perform a failover for the group selected as the recovery target or the group to which the group resource selected as the recovery target belongs when a monitor error is detected.
- Execute the final action
Execute the selected final action when a monitor error is detected.

Execute Failover to outside the Server Group

Configurable only for message receive monitor resources. Specify whether to fail over to a server group other than the active server group upon the reception of an error message.

Execute Script before Recovery Action

This setting is disabled when linking with the server management infrastructure. Execute the script before the operation to be performed upon error detection, as selected for the recovery action.

For details about the settings other than the above, see "Recovery Action tab" in "Monitor resource properties" in "Monitor resource details" in the Reference Guide.

LEGAL NOTICE

4.1 Disclaimer

- Information in this document is subject to change without notice. No part of this document may be reproduced or transmitted in any form by any means, electronic or mechanical, for any purpose, without the express written permission of NEC Corporation.

4.2 Trademark Information

- EXPRESSCLUSTER® is a registered trademark of NEC Corporation.
- Linux is a registered trademark of Linus Torvalds in the United States and other countries.
- RPM is a registered trademark of Red Hat, Inc. or its subsidiaries in the United States and other countries.
- Other product names and slogans written in this manual are trademarks or registered trademarks of their respective companies.

REVISION HISTORY

Edition	Revised Date	Description
1st	Apr 10, 2020	New manual

© Copyright NEC Corporation 2020. All rights reserved.