EXPRESSCLUSTER® X 4.0

for Linux

Getting Started Guide

September 14, 2018
2nd Edition

Revision History

Edition	Revised Date	Description
1st	Apr 17, 2018	New manual.
2nd	Sep 14, 2018	Corresponds to the internal version 4.0.1-1.

Disclaimer

Trademark Information

Table of Contents

Preface

Who Should Use This Guide

EXPRESSCLUSTER Getting Started Guide is intended for first-time users of the EXPRESSCLUSTER. The guide covers topics such as product overview of the EXPRESSCLUSTER, how the cluster system is installed, and the summary of other available guides. In addition, latest system requirements and restrictions are described.

How This Guide is Organized

Section I **Introducing EXPRESSCLUSTER**

Chapter 1 **What is a cluster system?**
Helps you to understand the overview of the cluster system and EXPRESSCLUSTER.

Chapter 2 **Using EXPRESSCLUSTER**
Provides instructions on how to use a cluster system and other related-information.

Section II **Installing EXPRESSCLUSTER**

Chapter 3 **Installation requirements for EXPRESSCLUSTER**
Provides the latest information that needs to be verified before starting to use EXPRESSCLUSTER.

Chapter 4 **Latest version information**
Provides information on latest version of the EXPRESSCLUSTER.

Chapter 5 **Notes and Restrictions**
Provides information on known problems and restrictions.

Chapter 6 **Upgrading EXPRESSCLUSTER**
Provides instructions on how to update the EXPRESSCLUSTER.

Appendix

Appendix A **Glossary**

Appendix B **Index**

EXPRESSCLUSTER X Documentation Set

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

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

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.

Reference Guide
This guide is intended for system administrators. The guide covers topics such as how to operate EXPRESSCLUSTER, function of each module, maintenance-related information, and troubleshooting. The guide is supplement to the *Installation and Configuration Guide*.

EXPRESSCLUSTER X Integrated WebManager Administrator's Guide
This guide is intended for system administrators who manage cluster systems using EXPRESSCLUSTER with Integrated WebManager, and also intended for system engineers who introduce Integrated WebManager. This guide describes detailed issues necessary for introducing Integrated WebManager in the actual procedures.

Conventions

In this guide, **Note**, **Important**, **Related Information** 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.

Related Information:
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 in as root user.	`# clpcl -s -a`
Monospace (courier)	Indicates path names, commands, system output (message, prompt, etc.), directory, file names, functions and parameters.	`/Linux/4.0/en/server/`
Monospace bold (courier)	Indicates the value that a user actually enters from a command line.	Enter the following: **`# clpcl -s -a`**
Monospace italic (courier)	Indicates that users should replace italicized part with values that they are actually working with.	`rpm -i expressclsbuilder-<version_number>-<release_number>.x86_64.rpm`

Contacting NEC

For the latest product information, visit our website below:

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

Section I Introducing EXPRESSCLUSTER

This section helps you to understand the overview of EXPRESSCLUSTER and its system requirements. This section covers:

- Chapter 1 What is a cluster system?
- Chapter 2 Using EXPRESSCLUSTER

Chapter 1 What is a cluster system?

This chapter describes overview of the cluster system.
This chapter covers:

Overview of the cluster system

A key to success in today's computerized world is to provide services without them stopping. A single machine down due to a failure or overload can stop entire services you provide with customers. This will not only result in enormous damage but also in loss of credibility you once enjoyed.

A cluster system is a solution to tackle such a disaster. Introducing a cluster system allows you to minimize the period during which operation of your system stops (down time) or to avoid system-down by load distribution.

As the word "cluster" represents, a cluster system is a system aiming to increase reliability and performance by clustering a group (or groups) of multiple computers. There are various types of cluster systems, which can be classified into the following three listed below. EXPRESSCLUSTER is categorized as a high availability cluster.

High Availability (HA) Cluster

In this cluster configuration, one server operates as an active server. When the active server fails, a standby server takes over the operation. This cluster configuration aims for high-availability and allows data to be inherited as well. The high availability cluster is available in the shared disk type, data mirror type or remote cluster type.

Load Distribution Cluster

This is a cluster configuration where requests from clients are allocated to load-distribution hosts according to appropriate load distribution rules. This cluster configuration aims for high scalability. Generally, data cannot be taken over. The load distribution cluster is available in a load balance type or parallel database type.

High Performance Computing (HPC) Cluster

This is a cluster configuration where CPUs of all nodes are used to perform a single operation. This cluster configuration aims for high performance but does not provide general versatility. Grid computing, which is one of the types of high performance computing that clusters a wider range of nodes and computing clusters, is a hot topic these days.

High Availability (HA) cluster

To enhance the availability of a system, it is generally considered that having redundancy for components of the system and eliminating a single point of failure is important. "Single point of failure" is a weakness of having a single computer component (hardware component) in the system. If the component fails, it will cause interruption of services. The high availability (HA) cluster is a cluster system that minimizes the time during which the system is stopped and increases operational availability by establishing redundancy with multiple servers.

The HA cluster is called for in mission-critical systems where downtime is fatal. The HA cluster can be divided into two types: shared disk type and data mirror type. The explanation for each type is provided below.

Shared disk type

Data must be inherited from one server to another in cluster systems. A cluster topology where data is stored in a shared disk with two or more servers using the data is called shared disk type.

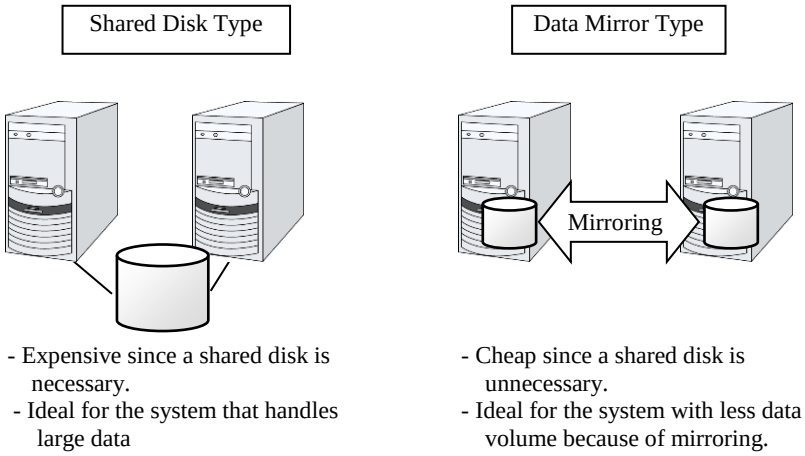


Figure 1-1: HA cluster configuration

If a failure occurs on a server where applications are running (active server), the cluster system detects the failure and applications are automatically started in a standby server to take over operations. This mechanism is called failover. Operations to be inherited in the cluster system consist of resources including disk, IP address and application.

In a non-clustered system, a client needs to access a different IP address if an application is restarted on a server other than the server where the application was originally running. In contrast, many cluster systems allocate a virtual IP address on an operational basis. A server where the operation is running, be it an active or a standby server, remains transparent to a client. The operation is continued as if it has been running on the same server.

File system consistency must be checked to inherit data. A check command (for example, fsck or chkdsk in Linux) is generally run to check file system consistency. However, the larger the file system is, the more time spent for checking. While checking is in process, operations are stopped. For this problem, journaling file system is introduced to reduce the time required for failover.

Logic of the data to be inherited must be checked for applications. For example, roll-back or roll-forward is necessary for databases. With these actions, a client can continue operation only by re-executing the SQL statement that has not been committed yet.

A server with the failure can return to the cluster system as a standby server if it is physically separated from the system, fixed, and then succeeds to connect the system. Such returning is acceptable in production environments where continuity of operations is important.

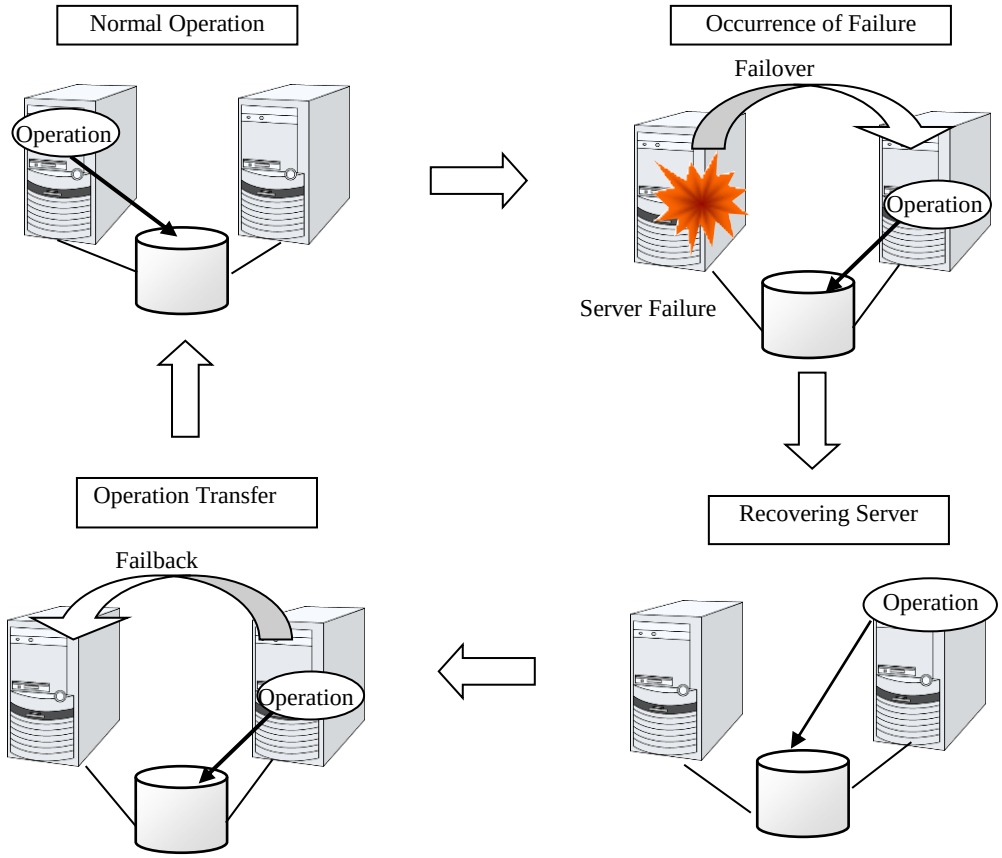


Figure 1-2: From occurrence of a failure to recovery

When the specification of the failover destination server does not meet the system requirements or overload occurs due to multi-directional standby, operations on the original server are preferred. In such a case, a failback takes place to resume operations on the original server.

A standby mode where there is one operation and no operation is active on the standby server, as shown in Figure 1-3, is referred to as uni-directional standby. A standby mode where there are two or more operations with each server of the cluster serving as both active and standby servers is referred to as multi-directional standby.

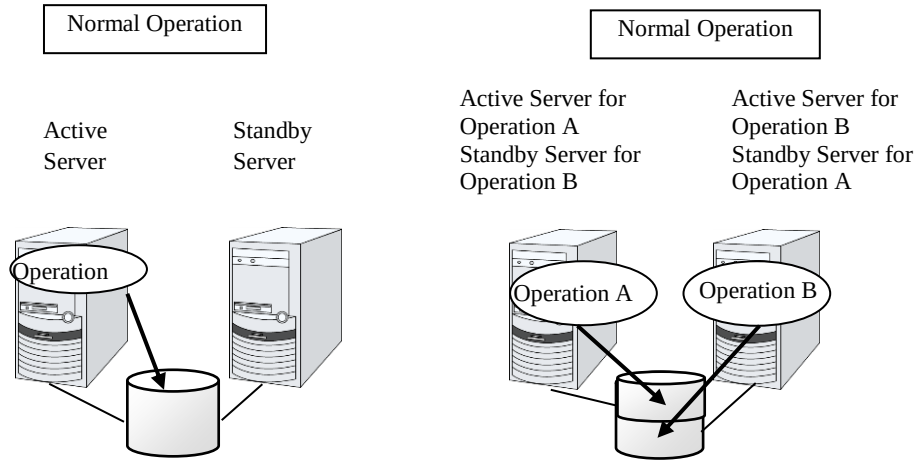


Figure 1-3: HA cluster topology

Data mirror type

The shared disk type cluster system is good for large-scale systems. However, creating a system with this type can be costly because shared disks are generally expensive. The data mirror type cluster system provides the same functions as the shared disk type with smaller cost through mirroring of server disks.

The data mirror type is not recommended for large-scale systems that handle a large volume of data since data needs to be mirrored between servers.

When a write request is made by an application, the data mirror engine not only writes data in the local disk but sends the write request to the standby server via the interconnect. Interconnect is a network connecting servers. It is used to monitor whether or not the server is activated in the cluster system. In addition to this purpose, interconnect is sometimes used to transfer data in the data mirror type cluster system. The data mirror engine on the standby server achieves data synchronization between standby and active servers by writing the data into the local disk of the standby server.

For read requests from an application, data is simply read from the disk on the active server.

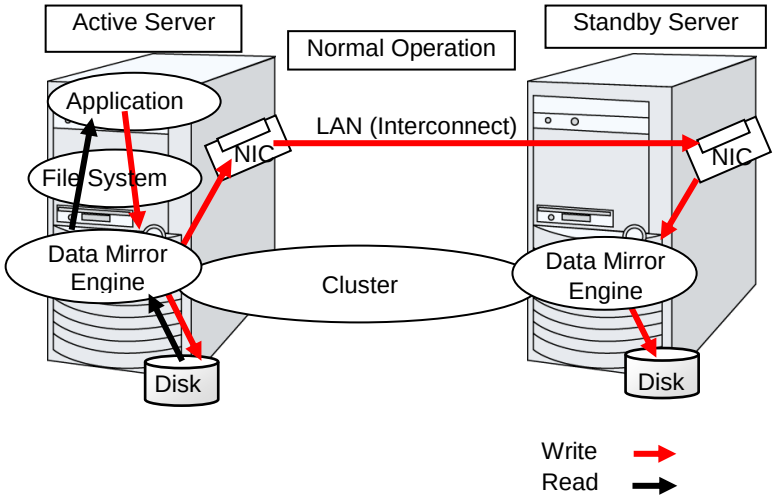


Figure 1-4: Data mirror mechanism

Snapshot backup is applied usage of data mirroring. Because the data mirror type cluster system has shared data in two locations, you can keep the disk of the standby server as snapshot backup without spending time for backup by simply separating the server from the cluster.

Failover mechanism and its problems

There are various cluster systems such as failover clusters, load distribution clusters, and high performance computing (HPC) clusters. The failover cluster is one of the high availability (HA) cluster systems that aim to increase operational availability through establishing server redundancy and passing operations being executed to another server when a failure occurs.

Error detection mechanism

Cluster software executes failover (for example, passing operations) when a failure that can impact continued operation is detected. The following section gives you a quick view of how the cluster software detects a failure.

Heartbeat and detection of server failures

Failures that must be detected in a cluster system are failures that can cause all servers in the cluster to stop. Server failures include hardware failures such as power supply and memory failures, and OS panic. To detect such failures, heartbeat is employed to monitor whether or not the server is active.

Some cluster software programs use heartbeat not only for checking whether or not the target is active through ping response, but for sending status information on the local server. Such cluster software programs begin failover if no heartbeat response is received in heartbeat transmission, determining no response as server failure. However, grace time should be given before determining failure, since a highly loaded server can cause delay of response. Allowing grace period results in a time lag between the moment when a failure occurred and the moment when the failure is detected by the cluster software.

Detection of resource failures

Factors causing stop of operations are not limited to stop of all servers in the cluster. Failure in disks used by applications, NIC failure, and failure in applications themselves are also factors that can cause the stop of operations. These resource failures need to be detected as well to execute failover for improved availability.

Accessing a target resource is a way employed to detect resource failures if the target is a physical device. For monitoring applications, trying to service ports within the range not impacting operation is a way of detecting an error in addition to monitoring whether or not application processes are activated.

Problems with shared disk type

In a failover cluster system of the shared disk type, multiple servers physically share the disk device. Typically, a file system enjoys I/O performance greater than the physical disk I/O performance by keeping data caches in a server.

What if a file system is accessed by multiple servers simultaneously?

Because a general file system assumes no server other than the local updates data on the disk, inconsistency between caches and the data on the disk arises. Ultimately the data will be corrupted. The failover cluster system locks the disk device to prevent multiple servers from mounting a file system, simultaneously caused by a network partition.

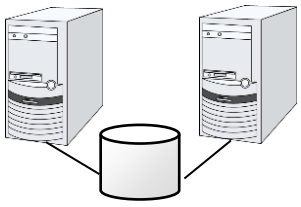


Figure 1-5: Cluster configuration with a shared disk

Network partition (split-brain-syndrome)

When all interconnects between servers are disconnected, failover takes place because the servers assume other server(s) are down. To monitor whether the server is activated, a heartbeat communication is used. As a result, multiple servers mount a file system simultaneously causing data corruption. This explains the importance of appropriate failover behavior in a cluster system at the time of failure occurrence.

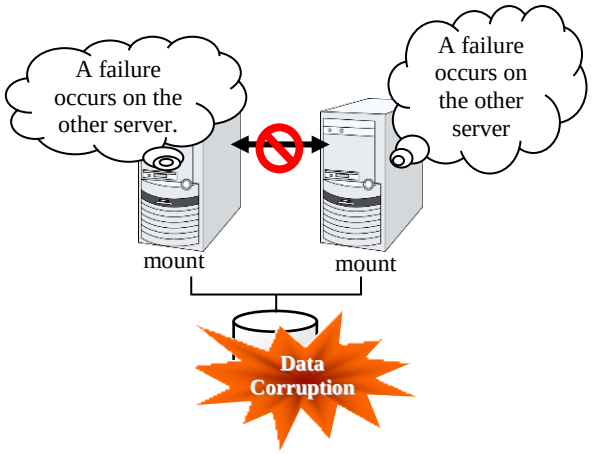


Figure 1-6: Network partition problem

The problem explained in the section above is referred to as “network partition” or “split-brain syndrome.” The failover cluster system is equipped with various mechanisms to ensure shared disk lock at the time when all interconnects are disconnected.

Taking over cluster resources

As mentioned earlier, resources to be managed by a cluster include disks, IP addresses, and applications. The functions used in the failover cluster system to inherit these resources are described below.

Taking over the data

Data to be passed from a server to another in a cluster system is stored in a partition on the shared disk. This means data is re-mounting the file system of files that the application uses on a healthy server. What the cluster software should do is simply mount the file system because the shared disk is physically connected to a server that inherits data.

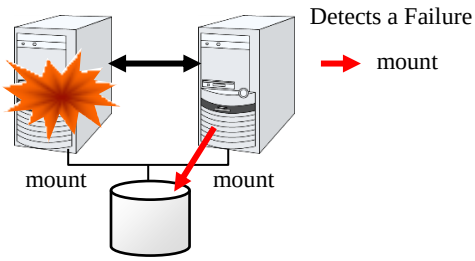


Figure 1-7: Taking over data

Figure 1-7 may look simple, but consider the following issues in designing and creating a cluster system.

One issue to consider is recovery time for a file system. A file system to be inherited may have been used by another server or being updated just before the failure occurred and requires a file system consistency check. When the file system is large, the time spent for checking consistency will be enormous. It may take a few hours to complete the check and the time is wholly added to the time for failover (time to take over operation), and this will reduce system availability.

Another issue you should consider is writing assurance. When an application writes important data into a file, it tries to ensure the data to be written into a disk by using a function such as synchronized writing. The data that the application assumes to have been written is expected to be inherited after failover. For example, a mail server reports the completion of mail receiving to other mail servers or clients after it has securely written mails it received in a spool. This will allow the spooled mail to be distributed again after the server is restarted. Likewise, a cluster system should ensure mails written into spool by a server to become readable by another server.

Taking over the applications

The last to come in inheritance of operation by cluster software is inheritance of applications. Unlike fault tolerant computers (FTC), no process status such as contents of memory is inherited in typical failover cluster systems. The applications running on a failed server are inherited by rerunning them on a healthy server.

For example, when instances of a database management system (DBMS) are inherited, the database is automatically recovered (roll-forward/roll-back) by startup of the instances. The time needed for this database recovery is typically a few minutes though it can be controlled by configuring the interval of DBMS checkpoint to a certain extent.

Many applications can restart operations by re-execution. Some applications, however, require going through procedures for recovery if a failure occurs. For these applications, cluster software allows to start up scripts instead of applications so that recovery process can be written. In a script, the recovery process, including cleanup of files half updated, is written as necessary according to factors for executing the script and information on the execution server.

Summary of failover

To summarize the behavior of cluster software:

- Detects a failure (heartbeat/resource monitoring)
- Resolves a network partition (NP resolution)
- Switches cluster resources
  - Pass data
  - Pass IP address
  - Application Taking over

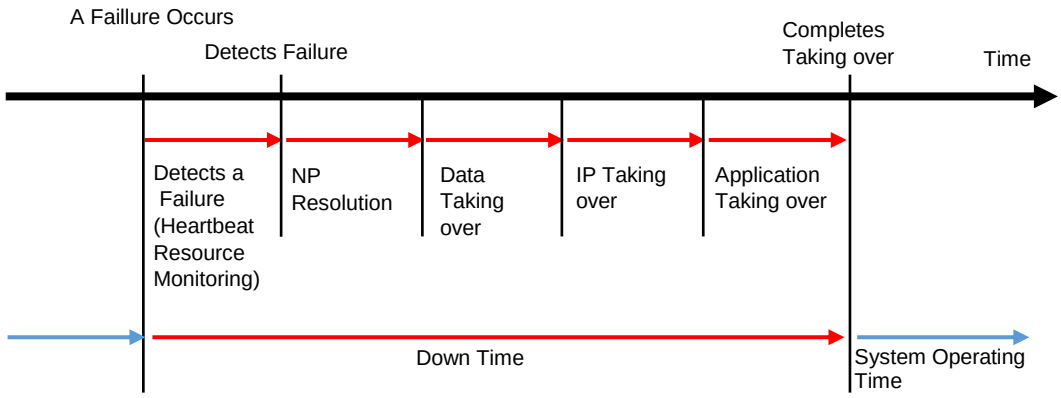


Figure 1-8: Failover time chart

Cluster software is required to complete each task quickly and reliably (see Figure 1-8). Cluster software achieves high availability with due consideration on what has been described so far.

Eliminating single point of failure

Having a clear picture of the availability level required or aimed is important in building a high availability system. This means when you design a system, you need to study cost effectiveness of countermeasures, such as establishing a redundant configuration to continue operations and recovering operations within a short period of time, against various failures that can disturb system operations.

Single point of failure (SPOF), as described previously, is a component where failure can lead to stop of the system. In a cluster system, you can eliminate the system's SPOF by establishing server redundancy. However, components shared among servers, such as shared disk may become a SPOF. The key in designing a high availability system is to duplicate or eliminate this shared component.

A cluster system can improve availability but failover will take a few minutes for switching systems. That means time for failover is a factor that reduces availability. Solutions for the following three, which are likely to become SPOF, will be discussed hereafter although technical issues that improve availability of a single server such as ECC memory and redundant power supply are important.

- Shared disk
- Access path to the shared disk
- LAN

Shared disk

Typically a shared disk uses a disk array for RAID. Because of this, the bare drive of the disk does not become SPOF. The problem is the RAID controller is incorporated. Shared disks commonly used in many cluster systems allow controller redundancy.

In general, access paths to the shared disk must be duplicated to benefit from redundant RAID controller. There are still things to be done to use redundant access paths in Linux (described later in this chapter). If the shared disk has configuration to access the same logical disk unit (LUN) from duplicated multiple controllers simultaneously, and each controller is connected to one server, you can achieve high availability by failover between nodes when an error occurs in one of the controllers.

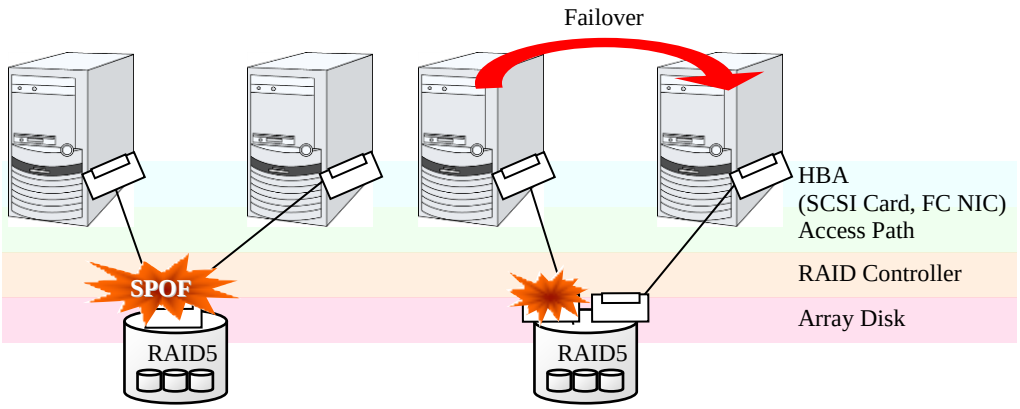


Figure 1-9: Example of the shared disk RAID controller and access paths being SPOF (left) and an access path connected to a RAID controller

With a failover cluster system of data mirror type, where no shared disk is used, you can create an ideal system having no SPOF because all data is mirrored to the disk in the other server. However you should consider the following issues:

- Disk I/O performance in mirroring data over the network (especially writing performance)
- System performance during mirror resynchronization in recovery from server failure (mirror copy is done in the background)
- Time for mirror resynchronization (clustering cannot be done until mirror resynchronization is completed)

In a system with frequent data viewing and a relatively small volume of data, choosing the data mirror type for clustering is a key to increase availability.

Access path to the shared disk

In a typical configuration of the shared disk type cluster system, the access path to the shared disk is shared among servers in the cluster. To take SCSI as an example, two servers and a shared disk are connected to a single SCSI bus. A failure in the access path to the shared disk can stop the entire system.

What you can do for this is to have a redundant configuration by providing multiple access paths to the shared disk and make them look as one path for applications. The device driver allowing such is called a path failover driver. Path failover drivers are often developed and released by shared disk vendors. Path failover drivers in Linux are still under development. For the time being, as discussed earlier, offering access paths to the shared disk by connecting a server on an array controller on the shared disk basis is the way to ensure availability in Linux cluster systems.

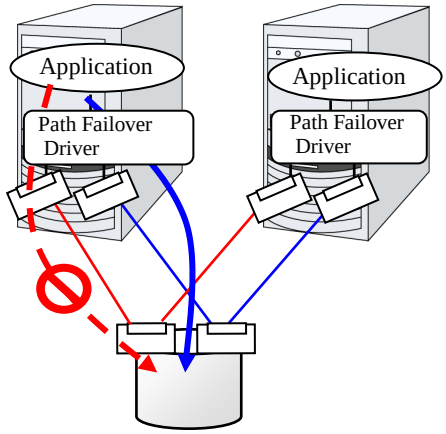


Figure 1-10: Path failover driver

LAN

In any systems that run services on a network, a LAN failure is a major factor that disturbs operations of the system. If appropriate settings are made, availability of cluster system can be increased through failover between nodes at NIC failures. However, a failure in a network device that resides outside the cluster system disturbs operation of the system.

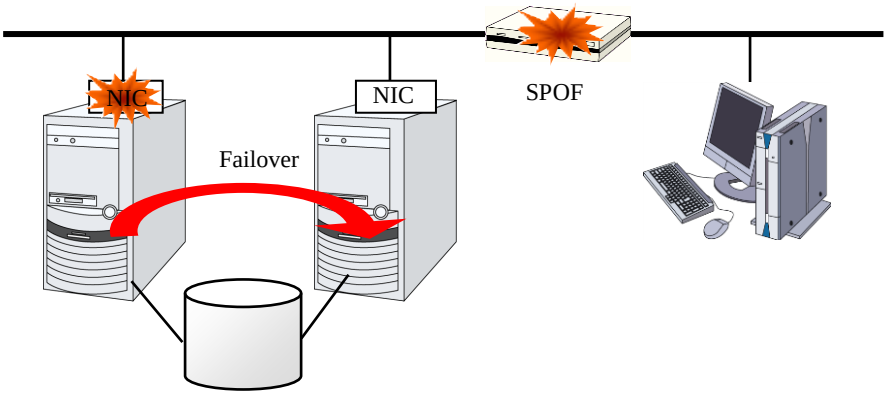


Figure 1-11: Example of router becoming SPOF

LAN redundancy is a solution to tackle device failure outside the cluster system and to improve availability. You can apply ways used for a single server to increase LAN availability. For example, choose a primitive way to have a spare network device with its power off, and manually replace a failed device with this spare device. Choose to have a multiplex network path through a redundant configuration of high-performance network devices, and switch paths automatically. Another option is to use a driver that supports NIC redundant configuration such as Intel's ANS driver.

Load balancing appliances and firewall appliances are also network devices that are likely to become SPOF. Typically they allow failover configurations through standard or optional software. Having redundant configuration for these devices should be regarded as requisite since they play important roles in the entire system.

Operation for availability

Evaluation before staring operation

Given many of factors causing system troubles are said to be the product of incorrect settings or poor maintenance, evaluation before actual operation is important to realize a high availability system and its stabilized operation. Exercising the following for actual operation of the system is a key in improving availability:

- Clarify and list failures, study actions to be taken against them, and verify effectiveness of the actions by creating dummy failures.
- Conduct an evaluation according to the cluster life cycle and verify performance (such as at degenerated mode)
- Arrange a guide for system operation and troubleshooting based on the evaluation mentioned above.

Having a simple design for a cluster system contributes to simplifying verification and improvement of system availability.

Failure monitoring

Despite the above efforts, failures still occur. If you use the system for long time, you cannot escape from failures: hardware suffers from aging deterioration and software produces failures and errors through memory leaks or operation beyond the originally intended capacity. Improving availability of hardware and software is important yet monitoring for failure and troubleshooting problems is more important. For example, in a cluster system, you can continue running the system by spending a few minutes for switching even if a server fails. However, if you leave the failed server as it is, the system no longer has redundancy and the cluster system becomes meaningless should the next failure occur.

If a failure occurs, the system administrator must immediately take actions such as removing a newly emerged SPOF to prevent another failure. Functions for remote maintenance and reporting failures are very important in supporting services for system administration. Linux is known for providing good remote maintenance functions. Mechanism for reporting failures are coming in place. To achieve high availability with a cluster system, you should:

- Remove or have complete control on single point of failure.
- Have a simple design that has tolerance and resistance for failures, and be equipped with a guide for operation and troubleshooting.
- Detect a failure quickly and take appropriate action against it.

Chapter 2 Using EXPRESSCLUSTER

This chapter explains the components of EXPRESSCLUSTER, how to design a cluster system, and how to use EXPRESSCLUSTER.

This chapter covers:

What is EXPRESSCLUSTER?

EXPRESSCLUSTER is software that enhances availability and expandability of systems by a redundant (clustered) system configuration. The application services running on the active server are automatically inherited to a standby server when an error occurs in the active server.

EXPRESSCLUSTER modules

EXPRESSCLUSTER consists of following three modules:

EXPRESSCLUSTER Server

A core component of EXPRESSCLUSTER. This includes all high availability functions of the server. The server functions of the Cluster WebUI, WebManager, and Builder are also included.

Cluster WebUI / WebManager

A tool to manage EXPRESSCLUSTER operations. Uses a Web browser as a user interface. The WebManager is installed in EXPRESSCLUSTER Server, but it is distinguished from the EXPRESSCLUSTER Server because the WebManager is operated from the Web browser on the management PC.

Builder

A tool for editing the cluster configuration data. The Builder uses a web browser as a user interface like the Cluster WebUI and WebManager. The following two versions of Builder are provided: the offline version, which is installed on your terminal as software independent of EXPRESSCLUSTER Server, and the online version, which is opened by clicking the setup mode icon on the WebManager screen toolbar or Setup Mode on the View menu. The Builder needs to be installed separately from the EXPRESSCLUSTER Server on the machine where you use the Builder.

Software configuration of EXPRESSCLUSTER

The software configuration of EXPRESSCLUSTER should look similar to the figure below. Install the EXPRESSCLUSTER Server (software) on a Linux server, and the Builder on a management PC or a server. Because the main functions of Cluster WebUI, WebManager, and Builder are included in EXPRESSCLUSTER Server, it is not necessary to separately install them. However, to use the Builder in an environment where EXPRESSCLUSTER Server is not accessible, the offline version of Builder must be installed on the PC. The Cluster WebUI, WebManager, or Builder can be used through the web browser on the management PC or on each server in the cluster.

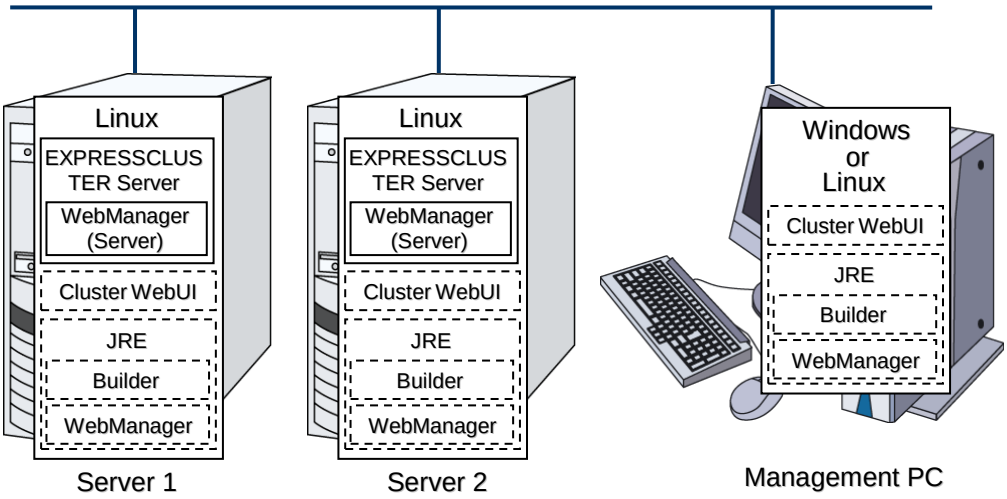


Figure 2-1 Software configuration of EXPRESSCLUSTER

How an error is detected in EXPRESSCLUSTER

There are three kinds of monitoring in EXPRESSCLUSTER: (1) server monitoring, (2) application monitoring, and (3) internal monitoring. These monitoring functions let you detect an error quickly and reliably. The details of the monitoring functions are described below.

What is server monitoring?

Server monitoring is the most basic function of the failover-type cluster system. It monitors if a server that constitutes a cluster is properly working.

EXPRESSCLUSTER regularly checks whether other servers are properly working in the cluster system. This way of verification is called “heartbeat communication.” The heartbeat communication uses the following communication paths:

Primary Interconnect

Uses an Ethernet NIC in communication path dedicated to the failover-type cluster system. This is used to exchange information between the servers as well as to perform heartbeat communication.

Secondary Interconnect

Uses a communication path used for communication with client machine as an alternative interconnect. Any Ethernet NIC can be used as long as TCP/IP can be used. This is also used to exchange information between the servers and to perform heartbeat communication.

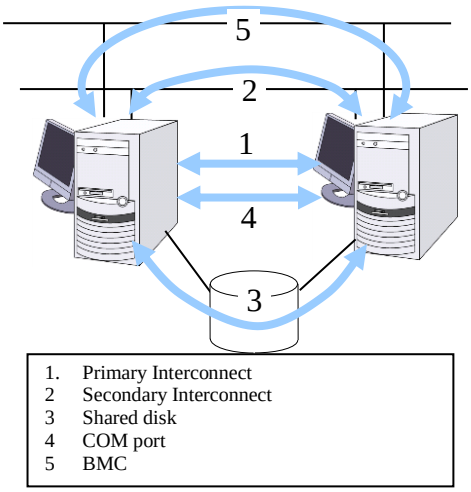


Figure 2-2 Server monitoring

Shared disk

Creates an EXPRESSCLUSTER-dedicated partition (EXPRESSCLUSTER partition) on the disk that is connected to all servers that constitute the failover-type cluster system, and performs heartbeat communication on the EXPRESSCLUSTER partition.

COM port

Performs heartbeat communication between the servers that constitute the failover-type cluster system through a COM port, and checks whether other servers are working properly.

BMC

Performs heartbeat communication between the servers that constitute the failover-type cluster system through the BMC, and checks whether other servers are working properly.

Having these communication paths dramatically improves the reliability of the communication between the servers, and prevents the occurrence of network partition.

Note:
Network partition refers to a condition when a network gets split by having a problem in all communication paths of the servers in a cluster. In a cluster system that is not capable of handling a network partition, a problem occurred in a communication path and a server cannot be distinguished. As a result, multiple servers may access the same resource and cause the data in a cluster system to be corrupted.

What is application monitoring?

Application monitoring is a function that monitors applications and factors that cause a situation where an application cannot run.

Activation status of application monitoring

An error can be detected by starting up an application from an exec resource in EXPRESSCLUSTER and regularly checking whether a process is active or not by using the pid monitor resource. It is effective when the factor for application to stop is due to error termination of an application.

Note:

An error in resident process cannot be detected in an application started up by EXPRESSCLUSTER. When the monitoring target application starts and stops a resident process, an internal application error (such as application stalling, result error) cannot be detected.

Resource monitoring

An error can be detected by monitoring the cluster resources (such as disk partition and IP address) and public LAN using the monitor resources of the EXPRESSCLUSTER. It is effective when the factor for application to stop is due to an error of a resource which is necessary for an application to operate.

What is internal monitoring?

Internal monitoring refers to an inter-monitoring of modules within EXPRESSCLUSTER. It monitors whether each monitoring function of EXPRESSCLUSTER is properly working. Activation status of EXPRESSCLUSTER process monitoring is performed within EXPRESSCLUSTER.

- Critical monitoring of EXPRESSCLUSTER process

Monitorable and non-monitorable errors

There are monitorable and non-monitorable errors in EXPRESSCLUSTER. It is important to know what can or cannot be monitored when building and operating a cluster system.

Detectable and non-detectable errors by server monitoring

Monitoring condition: A heartbeat from a server with an error is stopped

Example of errors that can be monitored:

- Hardware failure (of which OS cannot continue operating)
- System panic

Example of error that cannot be monitored:

- Partial failure on OS (for example, only a mouse or keyboard does not function)

Detectable and non-detectable errors by application monitoring

Monitoring conditions: Termination of applications with errors, continuous resource errors, and disconnection of a path to the network devices.

Example of errors that can be monitored:

- Abnormal termination of an application
- Failure to access the shared disk (such as HBA[1] failure)
- Public LAN NIC problem

Example of errors that cannot be monitored:

- Application stalling and resulting in error. EXPRESSCLUSTER cannot monitor application stalling and error results. However, it is possible to perform failover by creating a program that monitors applications and terminates itself when an error is detected, starting the program using the exec resource, and monitoring application using the PID monitor resource.

[1] HBA is an abbreviation for host bus adapter. This adapter is not for the shared disk, but for the server.

Network partition resolution

Upon detecting that a heartbeat from a server is interrupted, EXPRESSCLUSTER determines whether the cause of this interruption is an error in a server or a network partition. If it is judged as a server failure, failover (activate resources and start applications on a healthy server) is performed. If it is judged as a network partition, protecting data is given priority over operations and a processing such as emergency shutdown is performed.
The following is the network partition resolution method:

- ping method

Related Information:

For the details on the network partition resolution method, see Chapter 7, “Details on network partition resolution resources” of the Reference Guide.

Failover mechanism

Upon detecting that a heartbeat from a server is interrupted, EXPRESSCLUSTER determines whether the cause of this interruption is an error in a server or a network partition before starting a failover. Then a failover is performed by activating various resources and starting up applications on a properly working server.

The group of resources which fail over at the same time is called a "failover group." From a user's point of view, a failover group appears as a virtual computer.

Note:

In a cluster system, a failover is performed by restarting the application from a properly working node. Therefore, what is saved in an application memory cannot be failed over.

From occurrence of error to completion of failover takes a few minutes. See the figure 2-3 below:

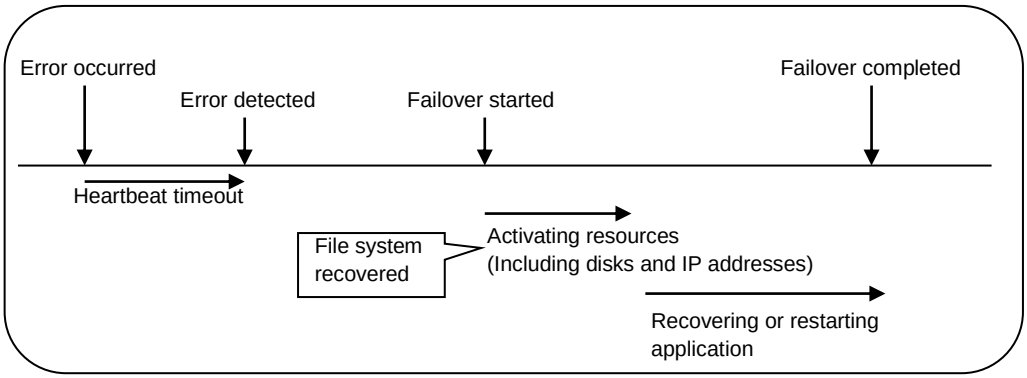


Figure 2-3 Failover time chart

Heartbeat timeout

- The time for a standby server to detect an error after that error occurred on the active server.
- The setting values of the cluster properties should be adjusted depending on the application load. (The default value is 90 seconds.)

Activating various resources

- The time to activate the resources necessary for operating an application.
- The resources can be activated in a few seconds in ordinary settings, but the required time changes depending on the type and the number of resources registered to the failover group. For more information, refer to the *Installation and Configuration Guide*.

Start script execution time

- The data recovery time for a roll-back or roll-forward of the database and the startup time of the application to be used in operation.
- The time for roll-back or roll-forward can be predicted by adjusting the check point interval. For more information, refer to the document that comes with each software product.

Failover resources

EXPRESSCLUSTER can fail over the following resources:

Switchable partition

- Resources such as disk resource, mirror disk resource and hybrid disk resource.
- A disk partition to store the data that the application takes over.

Floating IP Address

- By connecting an application using the floating IP address, a client does not have to be conscious about switching the servers due to failover processing.
- It is achieved by dynamic IP address allocation to the public LAN adapter and sending ARP packet. Connection by floating IP address is possible from most of the network devices.

Script (exec resource)

- In EXPRESSCLUSTER, applications are started up from the scripts.
- The file failed over on the shared disk may not be complete as data even if it is properly working as a file system. Write the recovery processing specific to an application at the time of failover in addition to the startup of an application in the scripts.

Note:
In a cluster system, failover is performed by restarting the application from a properly working node. Therefore, what is saved in an application memory cannot be failed over.

System configuration of the failover type cluster

In a failover-type cluster, a disk array device is shared between the servers in a cluster. When an error occurs on a server, the standby server takes over the applications using the data on the shared disk.

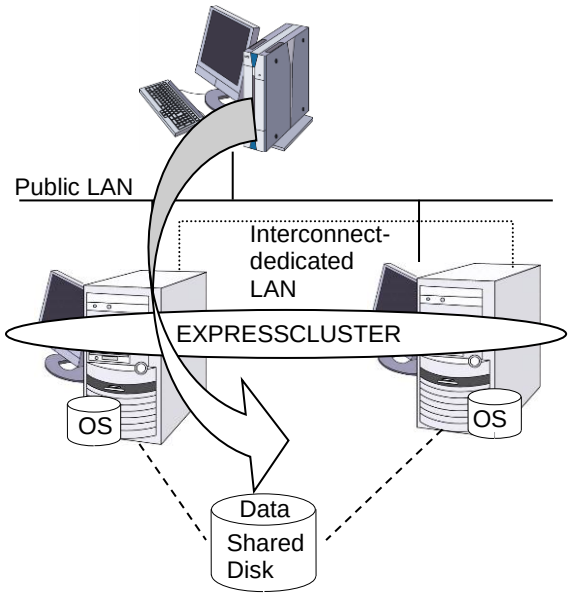


Figure 2-4 System configuration

A failover-type cluster can be divided into the following categories depending on the cluster topologies:

Uni-Directional Standby Cluster System

In the uni-directional standby cluster system, the active server runs applications while the other server, the standby server, does not. This is the simplest cluster topology and you can build a high-availability system without performance degradation after failing over.

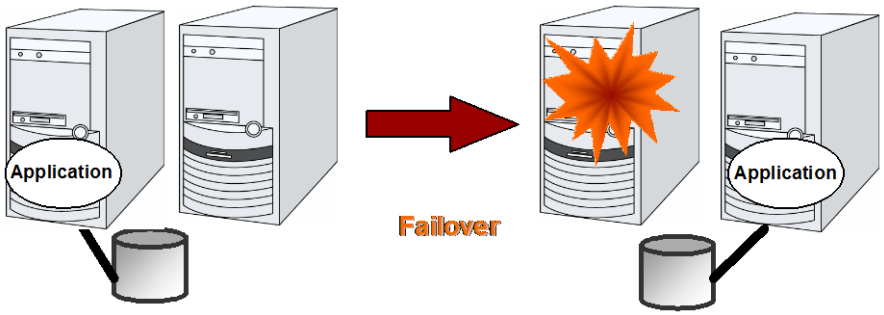


Figure 2-5 Uni-directional standby cluster system

Same Application Multi Directional Standby Cluster System

In the same application multi-directional standby cluster system, the same applications are activated on multiple servers. These servers also operate as standby servers. The applications must support multi-directional standby operation. When the application data can be split into multiple data, depending on the data to be accessed, you can build a load distribution system per data partitioning basis by changing the client's connecting server.

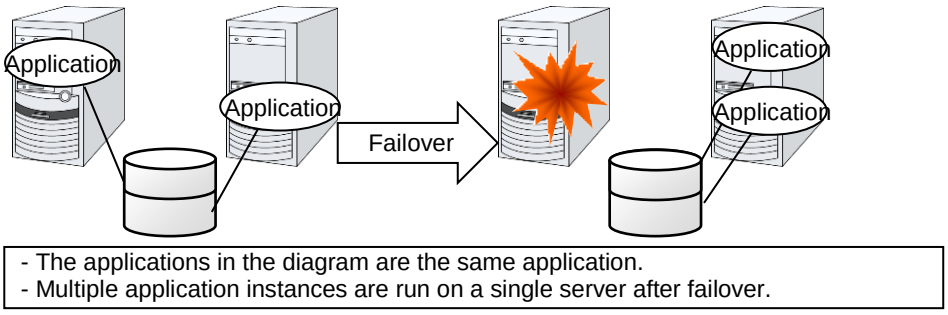


Figure 2-6 Same application multi directional standby cluster system

Different Application Multi Directional Standby Cluster System

In the different application multi-directional standby cluster system, different applications are activated on multiple servers and these servers also operate as standby servers. The applications do not have to support multi-directional standby operation. A load distribution system can be built per application unit basis.

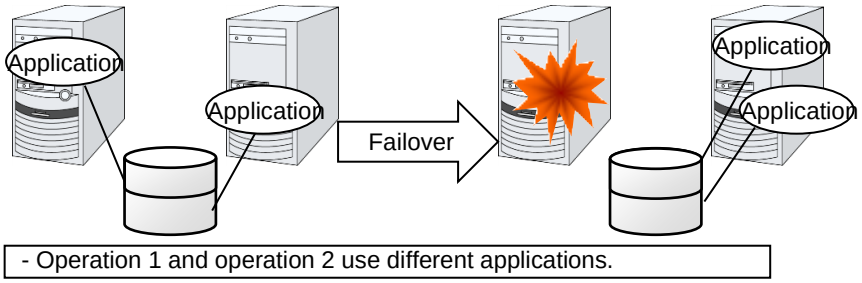


Figure 2-7 Different application multi directional standby cluster system

Node to Node Configuration

The configuration can be expanded with more nodes by applying the configurations introduced thus far. In a node to node configuration described below, three different applications are run on three servers and one standby server takes over the application if any problem occurs. In a uni-directional standby cluster system, one of the two servers functions as a standby server. However, in a node to node configuration, only one of the four server functions as a standby server and performance deterioration is not anticipated if an error occurs only on one server.

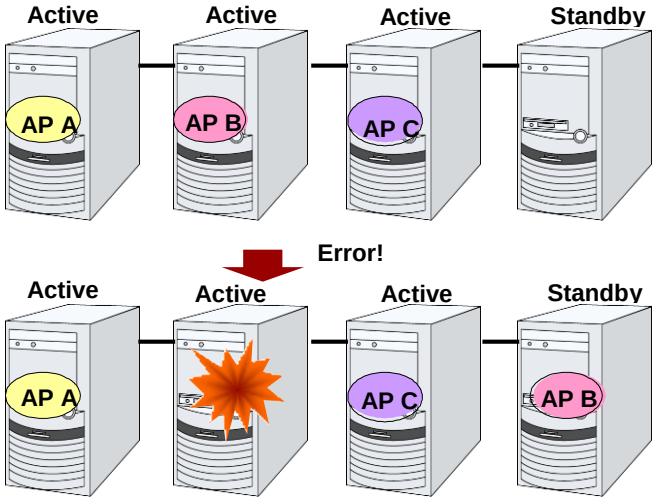


Figure 2-8 Node to Node configuration

Hardware configuration of the shared disk type cluster

The hardware configuration of the shared disk in EXPRESSCLUSTER is described below. In general, the following is used for communication between the servers in a cluster system:

- Two NIC cards (one for external communication, one for EXPRESSCLUSTER)
- COM port connected by RS232C cross cable
- Specific space of a shared disk

SCSI or FibreChannel can be used for communication interface to a shared disk; however, recently FibreChannel is more commonly used.

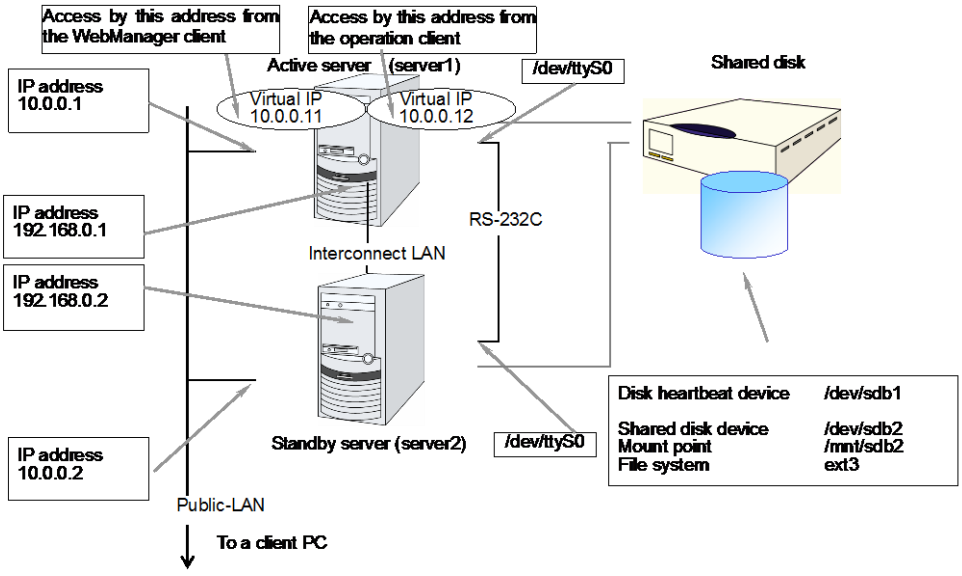


Figure 2-9 Sample of cluster environment when a shared disk is used

Hardware configuration of the mirror disk type cluster

The hardware configuration of the mirror disk in EXPRESSCLUSTER is described below.

Unlike the shared disk type, a network to copy the mirror disk data is necessary. In general, a network is used with NIC for internal communication in EXPRESSCLUSTER.

Mirror disks need to be separated from the operating system; however, they do not depend on a connection interface (IDE or SCSI.)

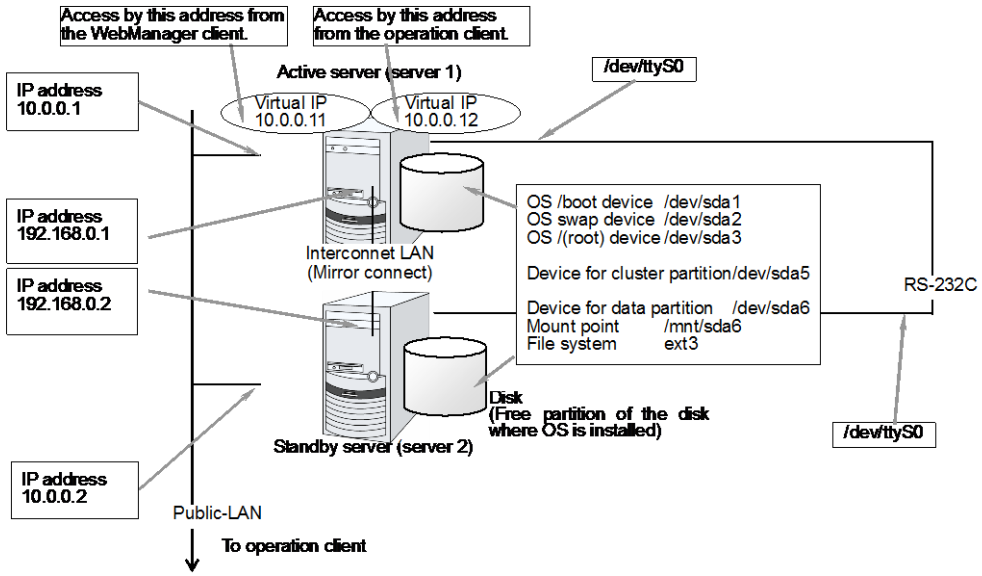


Figure 2-10 Sample of cluster environment when mirror disks are used (when allocating cluster partition and data partition to the disk where OS is installed):

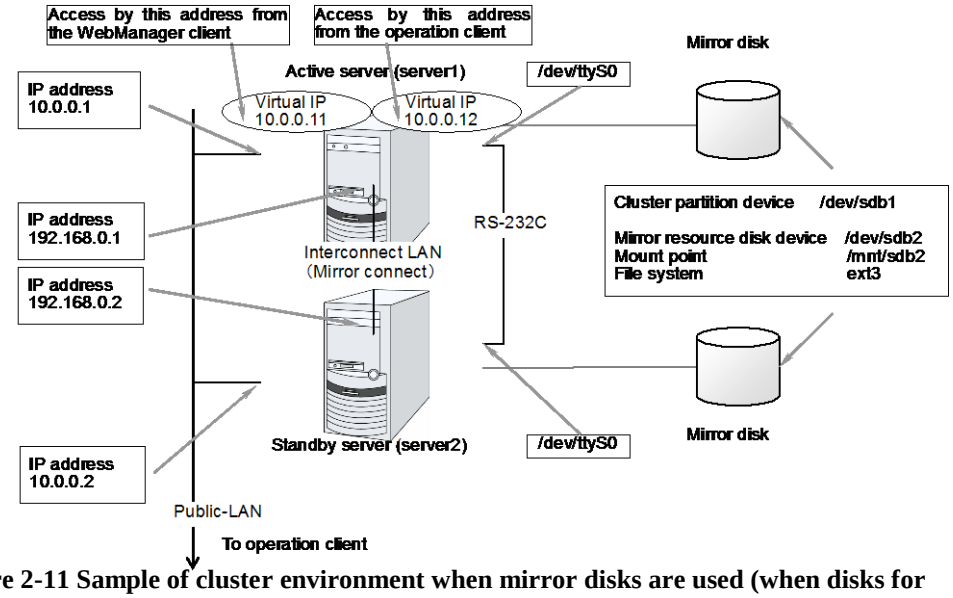


Figure 2-11 Sample of cluster environment when mirror disks are used (when disks for cluster partition and data partition are prepared):

Hardware configuration of the hybrid disk type cluster

The hardware configuration of the hybrid disk in EXPRESSCLUSTER is described below.

Unlike the shared disk type, a network to copy the data is necessary. In general, NIC for internal communication in EXPRESSCLUSTER is used to meet this purpose.

Disks do not depend on a connection interface (IDE or SCSI).

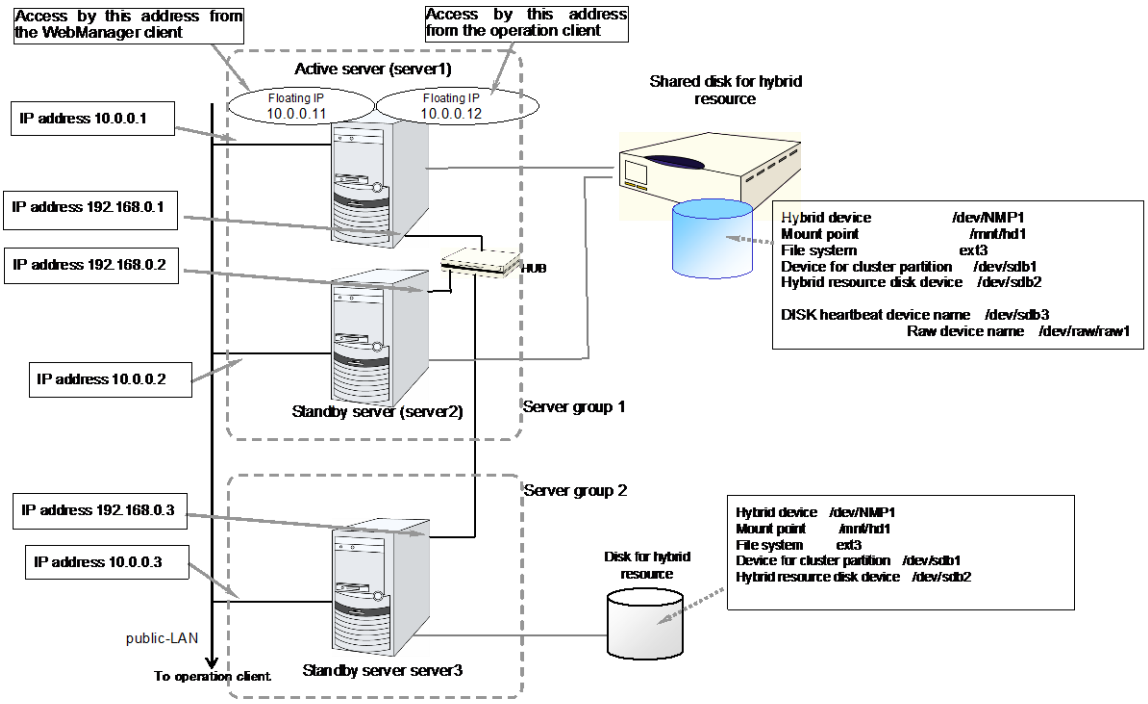


Figure 2-12: Sample of cluster environment where hybrid disks are used (two servers use a shared disk and the third server's general disk are used for mirroring)

What is cluster object?

In EXPRESSCLUSTER, the various resources are managed as the following groups:

Cluster object
Configuration unit of a cluster.

Server object
Indicates the physical server and belongs to the cluster object.

Server group object
Groups the servers and belongs to the cluster object.

Heartbeat resource object
Indicates the network part of the physical server and belongs to the server object.

Network partition resolution resource object
Indicates the network partition resolution mechanism and belongs to the server object.

Group object
Indicates a virtual server and belongs to the cluster object.

Group resource object
Indicates resources (network, disk) of the virtual server and belongs to the group object.

Monitor resource object
Indicates monitoring mechanism and belongs to the cluster object.

What is a resource?

In EXPRESSCLUSTER, a group used for monitoring the target is called “resources.” There are four types of resources and are managed separately. Having resources allows distinguishing what is monitoring and what is being monitored more clearly. It also makes building a cluster and handling an error easy. The resources can be divided into heartbeat resources, network partition resolution resources, group resources, and monitor resources.

Heartbeat resources

Heartbeat resources are used for verifying whether the other server is working properly between servers. The following heartbeat resources are currently supported:

LAN heartbeat resource
Uses Ethernet for communication.

Kernel mode LAN heartbeat resource
Uses Ethernet for communication.

COM heartbeat resource
Uses RS232C (COM) for communication.

Disk heartbeat resource
Uses a specific partition (cluster partition for disk heartbeat) on the shared disk for communication. It can be used only on a shared disk configuration.

BMC heartbeat resource

Uses Ethernet for communication via the BMC. This resource can be used only when the BMC hardware and firmware support the communication.

Network partition resolution resources

The following resource is used to resolve a network partition.

PING network partition resolution resource
This is a network partition resolution resource by the PING method.

Group resources

A group resource constitutes a unit when a failover occurs. The following group resources are currently supported:

Floating IP resource (fip)
Provides a virtual IP address. A client can access virtual IP address the same way as the regular IP address.

EXEC resource (exec)
Provides a mechanism for starting and stopping the applications such as DB and httpd.

Disk resource (disk)
Provides a specified partition on the shared disk. It can be used only on a shared disk configuration.

Mirror disk resource (md)
Provides a specified partition on the mirror disk. It can be used only on a mirror disk configuration.

Hybrid disk resource (hd)
Provides a specified partition on a shared disk or a disk. It can be used only for hybrid configuration.

Volume manager resource (volmgr)
Handles multiple storage devices and disks as a single logical disk.

NAS resource (nas)
Connect to the shared resources on NAS server. Note that it is not a resource that the cluster server behaves as NAS server.

Virtual IP resource (vip)
Provides a virtual IP address. This can be accessed from a client in the same way as a general IP address. This can be used in the remote cluster configuration among different network addresses.

VM resource (vm)
Starts, stops, or migrates the virtual machine.

Dynamic DNS resource (ddns)
Registers the virtual host name and the IP address of the active server to the dynamic DNS server.

AWS elastic ip resource (awseip)
Provides a system for giving an elastic IP (referred to as EIP) when EXPRESSCLUSTER is used on AWS.

AWS virtual ip resource (awsvip)
Provides a system for giving a virtual IP (referred to as VIP) when EXPRESSCLUSTER is used on AWS.

AWS DNS resource (awsdns)
Registers the virtual host name and the IP address of the active server to Amazon Route 53 when EXPRESSCLUSTER is used on AWS.

Azure probe port resource (azurepp)
Provides a system for opening a specific port on a node on which the operation is performed when EXPRESSCLUSTER is used on Microsoft Azure.

Azure DNS resource (azuredns)
Registers the virtual host name and the IP address of the active server to Azure DNS when EXPRESSCLUSTER is used on Microsoft Azure.

Monitor resources

A monitor resource monitors a cluster system. The following monitor resources are currently supported:

Floating IP monitor resource (fipw)
Provides a monitoring mechanism of an IP address started up by a floating IP resource.

IP monitor resource (ipw)
Provides a monitoring mechanism of an external IP address.

Disk monitor resource (diskw)
Provides a monitoring mechanism of the disk. It also monitors the shared disk.

Mirror disk monitor resource (mdw)
Provides a monitoring mechanism of the mirroring disks.

Mirror disk connect monitor resource (mdnw)
Provides a monitoring mechanism of the mirror disk connect.

Hybrid disk monitor resource (hdw)
Provides a monitoring mechanism of the hybrid disk.

Hybrid disk connect monitor resource (hdnw)
Provides a monitoring mechanism of the hybrid disk connect.

PID monitor resource (pidw)
Provides a monitoring mechanism to check whether a process started up by exec resource is active or not.

User-mode monitor resource (userw)
Provides a monitoring mechanism for a stalling problem in the user space.

NIC Link Up/Down monitor resource (miiw)

Provides a monitoring mechanism for link status of LAN cable.

Volume manager monitor resource (volmgrw)
Provides a monitoring mechanism for multiple storage devices and disks.

Multi target monitor resource (mtw)

Provides a status with multiple monitor resources.

Virtual IP monitor resource (vipw)

Provides a mechanism for sending RIP packets of a virtual IP resource.

ARP monitor resource (arpw)

Provides a mechanism for sending ARP packets of a floating IP resource or a virtual IP resource.

Custom monitor resource (genw)

Provides a monitoring mechanism to monitor the system by the operation result of commands or scripts which perform monitoring, if any.

VM monitor resource (vmw)
Checks whether the virtual machine is alive.

Message receive monitor resource (mrw)
Specifies the action to take when an error message is received and how the message is displayed on the WebManager.

Dynamic DNS monitor resource (ddnsw)

Periodically registers the virtual host name and the IP address of the active server to the dynamic DNS server.

Process name monitor resource (psw)

Provides a monitoring mechanism for checking whether a process specified by a process name is active.

BMC monitor resource (bmcw)

Provides a monitoring mechanism for checking whether a BMC is active.

DB2 monitor resource (db2w)

Provides a monitoring mechanism for IBM DB2 database.

ftp monitor resource (ftpw)

Provides a monitoring mechanism for FTP server.

http monitor resource (httpw)

Provides a monitoring mechanism for HTTP server.

imap4 monitor resource (imap4w)

Provides a monitoring mechanism for IMAP4 server.

MySQL monitor resource (mysqlw)

Provides a monitoring mechanism for MySQL database.

nfs monitor resource (nfsw)

Provides a monitoring mechanism for nfs file server.

Oracle monitor resource (oraclew)

Provides a monitoring mechanism for Oracle database.

Oracle Clusterware Synchronization Management monitor resource (osmw)

Provides a monitoring mechanism for Oracle Clusterware process linked EXPRESSCLUSTER.

pop3 monitor resource (pop3w)

Provides a monitoring mechanism for POP3 server.

PostgreSQL monitor resource (psqlw)

Provides a monitoring mechanism for PostgreSQL database.

samba monitor resource (sambaw)

Provides a monitoring mechanism for samba file server.

smtp monitor resource (smtpw)

Provides a monitoring mechanism for SMTP server.

Sybase monitor resource (sybasew)

Provides a monitoring mechanism for Sybase database.

Tuxedo monitor resource (tuxw)

Provides a monitoring mechanism for Tuxedo application server.

Websphere monitor resource (wasw)

Provides a monitoring mechanism for Websphere application server.

Weblogic monitor resource (wlsw)

Provides a monitoring mechanism for Weblogic application server.

WebOTX monitor resource (otxsw)

Provides a monitoring mechanism for WebOTX application server.

JVM monitor resource (jraw)

Provides a monitoring mechanism for Java VM.

System monitor resource (sraw)

Provides a monitoring mechanism for the resources specific to individual processes or those of the whole system.

AWS elastic ip monitor resource (awseipw)
Provides a monitoring mechanism for the elastic ip given by the AWS elastic ip (referred to as EIP) resource.

AWS virtual ip monitor resource (awsvipw)
Provides a monitoring mechanism for the virtual ip given by the AWS virtual ip (referred to as VIP) resource.

AWS AZ monitor resource (awsazw)
Provides a monitoring mechanism for an Availability Zone (referred to as AZ).

AWS DNS monitor resource (awsdnsw)
Provides a monitoring mechanism for the virtual host name and IP address provided by the AWS DNS resource.

Azure probe port monitor resource (azureppw)
Provides a monitoring mechanism for probe port for the node where an Azure probe port resource has been activated.

Azure load balance monitor resource (azurelbw)
Provides a mechanism for monitoring whether the port number that is same as the probe port is open for the node where an Azure probe port resource has not been activated.

Azure DNS monitor resource (azurednsw)
Provides a monitoring mechanism for the virtual host name and IP address provided by the Azure DNS resource.

Getting started with EXPRESSCLUSTER

Refer to the following guides when building a cluster system with EXPRESSCLUSTER:

Latest information

Refer to Section II, “Installing EXPRESSCLUSTER” in this guide.

Designing a cluster system

Refer to Section I, “Configuring a cluster system” in the *Installation and Configuration Guide* and Section II, “Resource details” in the *Reference Guide*.

Configuring a cluster system

Refer to the *Installation and Configuration Guide*.

Troubleshooting the problem

Refer to Section III, “Maintenance information” in the *Reference Guide*.

Section II Installing EXPRESSCLUSTER

This section provides the latest information on the EXPRESSCLUSTER. The latest information on the supported hardware and software is described in detail. Topics such as restrictions, known problems, and how to troubleshoot the problem are covered.

- Chapter 3 Installation requirements for EXPRESSCLUSTER
- Chapter 4 Latest version information
- Chapter 5 Notes and Restrictions
- Chapter 6 Upgrading EXPRESSCLUSTER

Chapter 3 Installation requirements for EXPRESSCLUSTER

This chapter provides information on system requirements for EXPRESSCLUSTER.
This chapter covers:

Hardware

EXPRESSCLUSTER operates on the following server architectures:

- x86_64
- IBM POWER (Replicator, Replicator DR, Agents except Database Agent are not supported)
- IBM POWER LE (Replicator, Replicator DR and Agents are not supported)

General server requirements

Required specifications for EXPRESSCLUSTER Server are the following:

- RS-232C port 1 port (not necessary when configuring a cluster with 3 or more nodes)
- Ethernet port 2 or more ports
- Shared disk
- Mirror disk or empty partition for mirror
- CD-ROM drive

When using the off-line Builder upon constructing and changing the existing configuration, the following is required for communication between the off-line Builder and servers:

- A machine to operate the off-line Builder and a way to share files

Servers supporting NX7700x series linkage

The table below lists the supported servers that can use the NX7700x series linkage function of the BMC heartbeat resources and message receive monitor resources. This function cannot be used by servers other than the following.

Server	Remarks
NX7700x/A2010M	Update to the latest firmware.
NX7700x/A2010L	Update to the latest firmware.
NX7700x/A3012M	Update to the latest firmware.
NX7700x/A3012L	Update to the latest firmware.
NX7700x/A3010M	Update to the latest firmware.

Servers supporting Express5800/A1080a and Express5800/A1040a series linkage

The table below lists the supported servers that can use the Express5800/A1080a and Express5800/A1040a series linkage function of the BMC heartbeat resources and message receive monitor resources. This function cannot be used by servers other than the following.

Serve	Remarks
Express5800/A1080a-E	Update to the latest firmware.
Express5800/A1080a-D	Update to the latest firmware.
Express5800/A1080a-S	Update to the latest firmware.
Express5800/A1040a	Update to the latest firmware.

Software

System requirements for EXPRESSCLUSTER Server

Supported distributions and kernel versions

The environment where EXPRESSCLUSTER Server can operate depends on kernel module versions because there are kernel modules unique to EXPRESSCLUSTER.

There are the following driver modules unique to EXPRESSCLUSTER.

Driver module unique to EXPRESSCLUSTER	Description
Kernel mode LAN heartbeat driver	Used with kernel mode LAN heartbeat resources.
Keepalive driver	Used if keepalive is selected as the monitoring method for user-mode monitor resources. Used if keepalive is selected as the monitoring method for shutdown monitoring.
Mirror driver	Used with mirror disk resources.

Regarding supported distributions and kernel versions, please refer to the following web site.

EXPRESSCLUSTER website
→System Requirements
→EXPRESSCLUSTER X for Linux

Note: For the kernel version of Cent OS supported by EXPRESSCLUSTER, see the supported kernel version of Red Hat Enterprise Linux.

Applications supported by monitoring options

Version information of the applications to be monitored by monitor resources is described below.

x86_64

Monitor resource	Monitored application	EXPRESSCLUSTER version	Remarks
Oracle monitor	Oracle Database 12c Release 1 (12.1)	4.0.0-1 or later	
	Oracle Database 12c Release 2 (12.2)	4.0.0-1 or later	
DB2 monitor	DB2 V10.5	4.0.0-1 or later	
	DB2 V11.1	4.0.0-1 or later	
PostgreSQL monitor	PostgreSQL 9.3	4.0.0-1 or later	
	PostgreSQL 9.4	4.0.0-1 or later	
	PostgreSQL 9.5	4.0.0-1 or later	
	PostgreSQL 9.6	4.0.0-1 or later	
	PostgreSQL 10	4.0.0-1 or later	
	PowerGres on Linux 9.1	4.0.0-1 or later	
	PowerGres on Linux 9.4	4.0.0-1 or later	
	PowerGres on Linux 9.6	4.0.0-1 or later	
MySQL monitor	MySQL 5.5	4.0.0-1 or later	
	MySQL 5.6	4.0.0-1 or later	
	MySQL 5.7	4.0.0-1 or later	
	MariaDB 5.5	4.0.0-1 or later	
	MariaDB 10.0	4.0.0-1 or later	
	MariaDB 10.1	4.0.0-1 or later	
	MariaDB 10.2	4.0.0-1 or later	
Sybase monitor	Sybase ASE 15.5	4.0.0-1 or later	
	Sybase ASE 15.7	4.0.0-1 or later	
	Sybase ASE 16.0	4.0.0-1 or later	
SQL Server monitor	SQL Server 2017	4.0.0-1 or later	
Samba monitor	Samba 3.3	4.0.0-1 or later	
	Samba 3.6	4.0.0-1 or later	
	Samba 4.0	4.0.0-1 or later	
	Samba 4.1	4.0.0-1 or later	
	Samba 4.2	4.0.0-1 or later	
	Samba 4.4	4.0.0-1 or later	
	Samba 4.6	4.0.0-1 or later	
NFS monitor	nfsd 2 (udp)	4.0.0-1 or later	
	nfsd 3 (udp)	4.0.0-1 or later	
	nfsd 4 (tcp)	4.0.0-1 or later	

Monitor resource	Monitored application	EXPRESSCLUSTER version	Remarks
	mountd 1 (tcp)	4.0.0-1 or later	
	mountd 2 (tcp)	4.0.0-1 or later	
	mountd 3 (tcp)	4.0.0-1 or later	
HTTP monitor	No specified version	4.0.0-1 or later	
SMTP monitor	No specified version	4.0.0-1 or later	
POP3 monitor	No specified version	4.0.0-1 or later	
imap4 monitor	No specified version	4.0.0-1 or later	
ftp monitor	No specified version	4.0.0-1 or later	
Tuxedo monitor	Tuxedo 12c Release 2 (12.1.3)	4.0.0-1 or later	
Weblogic monitor	WebLogic Server 11g R1	4.0.0-1 or later	
	WebLogic Server 11g R2	4.0.0-1 or later	
	WebLogic Server 12c R2 (12.2.1)	4.0.0-1 or later	
Websphere monitor	WebSphere Application Server 8.5	4.0.0-1 or later	
	WebSphere Application Server 8.5.5	4.0.0-1 or later	
	WebSphere Application Server 9.0	4.0.0-1 or later	
WebOTX monitor	WebOTX Application Server V9.1	4.0.0-1 or later	
	WebOTX Application Server V9.2	4.0.0-1 or later	
	WebOTX Application Server V9.3	4.0.0-1 or later	
	WebOTX Application Server V9.4	4.0.0-1 or later	
	WebOTX Application Server V10.1	4.0.0-1 or later	
JVM monitor	WebLogic Server 11g R1	4.0.0-1 or later	
	WebLogic Server 11g R2	4.0.0-1 or later	
	WebLogic Server 12c	4.0.0-1 or later	
	WebLogic Server 12c R2 (12.2.1)	4.0.0-1 or later	
	WebOTX Application Server V9.1	4.0.0-1 or later	
	WebOTX Application Server V9.2	4.0.0-1 or later	WebOTX update is required to monitor process groups
	WebOTX Application Server V9.3	4.0.0-1 or later	

Monitor resource	Monitored application	EXPRESSCLUSTER version	Remarks
	WebOTX Application Server V9.4	4.0.0-1 or later	
	WebOTX Application Server V10.1	4.0.0-1 or later	
	WebOTX Enterprise Service Bus V8.4	4.0.0-1 or later	
	WebOTX Enterprise Service Bus V8.5	4.0.0-1 or later	
	JBoss Enterprise Application Platform 7.0	4.0.0-1 or later	
	Apache Tomcat 8.0	4.0.0-1 or later	
	Apache Tomcat 8.5	4.0.0-1 or later	
	Apache Tomcat 9.0	4.0.0-1 or later	
	WebSAM SVF for PDF 9.0	4.0.0-1 or later	
	WebSAM SVF for PDF 9.1	4.0.0-1 or later	
	WebSAM SVF for PDF 9.2	4.0.0-1 or later	
	WebSAM Report Director Enterprise 9.0	4.0.0-1 or later	
	WebSAM Report Director Enterprise 9.1	4.0.0-1 or later	
	WebSAM Report Director Enterprise 9.2	4.0.0-1 or later	
	WebSAM Universal Connect/X 9.0	4.0.0-1 or later	
	WebSAM Universal Connect/X 9.1	4.0.0-1 or later	
	WebSAM Universal Connect/X 9.2	4.0.0-1 or later	
System monitor	No specified version	4.0.0-1 or later	

Note: To use monitoring options in x86_64 environments, applications to be monitored must be x86_64 version.

IBM POWER

Monitor resource	Monitored application	EXPRESSCLUSTER version	Remarks
DB2 monitor	DB2 V10.5	4.0.0-1 or later	
PostgreSQL monitor	PostgreSQL 9.3	4.0.0-1 or later	
	PostgreSQL 9.4	4.0.0-1 or later	
	PostgreSQL 9.5	4.0.0-1 or later	
	PostgreSQL 9.6	4.0.0-1 or later	
	PostgreSQL 10	4.0.0-1 or later	

Note: To use monitoring options in IBM POWER environments, applications to be monitored must be IBM POWER version.

Operation environment of VM resources

The followings are the version information of the virtual machines on which VM resources operation are verified.

Virtual Machine	Version	EXPRESSCLUSTER version	Remarks
vSphere	5.5	4.0.0-1 or later	Need management VM
	6.5	4.0.0-1 or later	Need management VM
XenServer	6.5 (x86_64)	4.0.0-1 or later	
KVM	Red Hat Enterprise Linux 6.9 (x86_64)	4.0.0-1 or later	
	Red Hat Enterprise Linux 7.4 (x86_64)	4.0.0-1 or later	

Note: The following functions do not work when ExpressCluster is installed in XenServer.

- Kernel mode LAN heartbeat resources
- Mirror disk resources/Hybrid disk resources
- User-mode monitor resources (keepalive/softdog method)
- Shutdown monitoring (keepalive/softdog method)

Operation environment for JVM monitor

The use of the JVM monitor requires a Java runtime environment. Also, monitoring a domain mode of JBoss Enterprise Application Platform requires Java® SE Development Kit.

Java®Runtime Environment
Version7.0 Update 6(1.7.0_6) or later

Java®Runtime Environment
Version 8.0 Update 11(1.8.0_11) or later

Java®Runtime Environment
Version 9.0 (9.0.1) or later

The tables below list the load balancers that were verified for the linkage with the JVM monitor.

x86_64

Load balancer	EXPRESSCLUSTER version	Remarks
Express5800/LB400h or later	4.0.0-1 or later	
InterSec/LB400i or later	4.0.0-1 or later	
BIG-IP v11	4.0.0-1 or later	
MIRACLE LoadBalancer	4.0.0-1 or later	
CoyotePoint Equalizer	4.0.0-1 or later	

Operation environment for AWS elastic ip resource, AWS virtual ip resource, AWS Elastic IP monitor resource, AWS virtual IP monitor resource, AWS AZ monitor resource

The use of the AWS elastic ip resource, AWS virtual ip resource, AWS Elastic IP monitor resource, AWS virtual IP monitor resource, AWS AZ monitor resource requires the following software.

Software	Version	Remarks
AWS CLI	1.6.0 or later	
Python	2.6.5 or later	Versions starting with 3. are not allowed.

The following are the version information for the OSs on AWS on which the operation of the AWS elastic ip resource, AWS virtual ip resource, AWS Elastic IP monitor resource, AWS virtual IP monitor resource, AWS AZ monitor resource is verified.

The environment where EXPRESSCLUSTER Server can operate depends on kernel module versions because there are kernel modules unique to EXPRESSCLUSTER.

Since the OS is frequently version up that AWS is to provide, when it is not possible to behavior will occur.

Kernel versions which has been verified, please refer to the *Supported distributions and kernel versions*.

x86_64

Distribution	EXPRESSCLUSTER version	Remarks
Red Hat Enterprise Linux 6.8	4.0.0-1 or later	
Red Hat Enterprise Linux 6.9	4.0.0-1 or later	
Red Hat Enterprise Linux 7.3	4.0.0-1 or later	
Red Hat Enterprise Linux 7.4	4.0.0-1 or later	
Cent OS 6.8	4.0.0-1 or later	
Cent OS 6.9	4.0.0-1 or later	
Cent OS 7.3	4.0.0-1 or later	
Cent OS 7.4	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP3	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP4	4.0.0-1 or later	
SUSE Linux Enterprise Server 12 SP1	4.0.0-1 or later	
Oracle Linux 6.6	4.0.0-1 or later	
Oracle Linux 7.3	4.0.0-1 or later	
Ubuntu 14.04.LTS	4.0.0-1 or later	
Ubuntu 16.04.3 LTS	4.0.0-1 or later	

Operation environment for AWS DNS resource, AWS DNS monitor resource

The use of the AWS DNS resource, AWS DNS monitor resource requires the following software.

Software	Version	Remarks
AWS CLI	1.11.0 or later	
Python (When OS is Red Hat Enterprise Linux 6, Cent OS 6, SUSE Linux Enterprise Server 11, Oracle Linux 6)	2.6.6 or later	Versions starting with 3. are not allowed.
Python (When OS is besides Red Hat Enterprise Linux 6, Cent OS 6, SUSE Linux Enterprise Server 11, Oracle Linux 6)	2.7.5 or later	Versions starting with 3. are not allowed.

The following are the version information for the OSs on AWS on which the operation of the AWS DNS resource, AWS DNS monitor resource is verified.

The environment where EXPRESSCLUSTER Server can operate depends on kernel module versions because there are kernel modules unique to EXPRESSCLUSTER.

Since the OS is frequently version up that AWS is to provide, when it is not possible to behavior will occur.

Kernel versions which has been verified, please refer to the *Supported distributions and kernel versions*.

x86_64

Distribution	EXPRESSCLUSTER version	Remarks
Red Hat Enterprise Linux 6.8	4.0.0-1 or later	
Red Hat Enterprise Linux 6.9	4.0.0-1 or later	
Red Hat Enterprise Linux 7.3	4.0.0-1 or later	
Red Hat Enterprise Linux 7.4	4.0.0-1 or later	
Cent OS 6.8	4.0.0-1 or later	
Cent OS 6.9	4.0.0-1 or later	
Cent OS 7.3	4.0.0-1 or later	
Cent OS 7.4	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP3	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP4	4.0.0-1 or later	
SUSE Linux Enterprise Server 12 SP1	4.0.0-1 or later	
Oracle Linux 6.6	4.0.0-1 or later	
Oracle Linux 7.3	4.0.0-1 or later	
Ubuntu 14.04.LTS	4.0.0-1 or later	
Ubuntu 16.04.3 LTS	4.0.0-1 or later	

Operation environment for Azure probe port resource, Azure probe port monitor resource, Azure load balance monitor resource

The following are the version information for the OSs on Microsoft Azure on which the operation of the Azure probe port resource, Azure probe monitor resource, Azure load balance monitor resource is verified.

The environment where EXPRESSCLUSTER Server can operate depends on kernel module versions because there are kernel modules unique to EXPRESSCLUSTER.

Since the OS is frequently version up that Microsoft Azure is to provide, when it is not possible to behavior will occur.

Kernel versions which has been verified, please refer to the *Supported distributions and kernel versions*.

x86_64

Distribution	EXPRESSCLUSTER version	Remarks
Red Hat Enterprise Linux 6.8	4.0.0-1 or later	
Red Hat Enterprise Linux 6.9	4.0.0-1 or later	
Red Hat Enterprise Linux 7.3	4.0.0-1 or later	
Red Hat Enterprise Linux 7.4	4.0.0-1 or later	
CentOS 6.8	4.0.0-1 or later	
CentOS 6.9	4.0.0-1 or later	
CentOS 7.3	4.0.0-1 or later	
CentOS 7.4	4.0.0-1 or later	
Asianux Server 4 SP6	4.0.0-1 or later	
Asianux Server 4 SP7	4.0.0-1 or later	
Asianux Server 7 SP1	4.0.0-1 or later	
Asianux Server 7 SP2	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP3	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP4	4.0.0-1 or later	
SUSE Linux Enterprise Server 12 SP1	4.0.0-1 or later	
Oracle Linux 6.6	4.0.0-1 or later	
Oracle Linux 7.3	4.0.0-1 or later	
Ubuntu 14.04.LTS	4.0.0-1 or later	
Ubuntu 16.04.3 LTS	4.0.0-1 or later	

The following are the Microsoft Azure deployment models with which the operation of the Azure probe port resource is verified. For details on how to set up a Load Balancer, refer to the documents from Microsoft (https://azure.microsoft.com/en-us/documentation/articles/load-balancer-arm/).

x86_64

Deployment model	EXPRESSCLUSTER Version	Remark
Resource Manager	4.0.0-1 or later	Load balancer is required

Operation environment for Azure DNS resource, Azure DNS monitor resource

The use of the Azure DNS resource, Azure DNS monitor resource requires the following software.

Software	Version	Remarks
Azure CLI (When OS is Red Hat Enterprise Linux 6, Cent OS 6, Asianux Server 4, SUSE Linux Enterprise Server 11, Oracle Linux 6)	1.0 or later	Python is not required.
Azure CLI ((When OS is besides Red Hat Enterprise Linux 6, Cent OS 6, Asianux Server 4, SUSE Linux Enterprise Server 11, Oracle Linux 6)	2.0 or later	
Python ((When OS is besides Red Hat Enterprise Linux 6, Cent OS 6, Asianux Server 4, SUSE Linux Enterprise Server 11, Oracle Linux 6)	2.7.5 or later	Versions starting with 3. are not allowed.

The following are the version information for the OSs on Microsoft Azure on which the operation of the Azure DNS resource, Azure DNS monitor resource is verified.

The environment where EXPRESSCLUSTER Server can operate depends on kernel module versions because there are kernel modules unique to EXPRESSCLUSTER.

Since the OS is frequently version up that Microsoft Azure is to provide, when it is not possible to behavior will occur.

Kernel versions which has been verified, please refer to the *Supported distributions and kernel versions*.

x86_64

Distribution	EXPRESSCLUSTER version	Remarks
Red Hat Enterprise Linux 6.8	4.0.0-1 or later	
Red Hat Enterprise Linux 6.9	4.0.0-1 or later	
Red Hat Enterprise Linux 7.3	4.0.0-1 or later	
Red Hat Enterprise Linux 7.4	4.0.0-1 or later	
CentOS 6.8	4.0.0-1 or later	
CentOS 6.9	4.0.0-1 or later	
CentOS 7.3	4.0.0-1 or later	
CentOS 7.4	4.0.0-1 or later	
Asianux Server 4 SP6	4.0.0-1 or later	
Asianux Server 4 SP7	4.0.0-1 or later	
Asianux Server 7 SP1	4.0.0-1 or later	
Asianux Server 7 SP2	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP3	4.0.0-1 or later	
SUSE Linux Enterprise Server 11 SP4	4.0.0-1 or later	

SUSE Linux Enterprise Server 12 SP1	4.0.0-1 or later	
Oracle Linux 6.6	4.0.0-1 or later	
Oracle Linux 7.3	4.0.0-1 or later	
Ubuntu 14.04.LTS	4.0.0-1 or later	
Ubuntu 16.04.3 LTS	4.0.0-1 or later	

The following are the Microsoft Azure deployment models with which the operation of the Azure DNS resource, the Azure DNS monitor resource is verified. For setting about Azure DNS, please refer to the *EXPRESSCLUSTER X 4.0 HA Cluster Configuration Guide for Microsoft Azure (Linux)*

x86_64

Deployment model	EXPRESSCLUSTER Version	Remark
Resource Manager	4.0.0-1 or later	Azure DNS is required

Operation environment for the Connector for SAP

OS and SAP NetWeaver(or later, SAP NW), which confirms the operation of the Connector for SAP presents the version information of the following.

x86_64

SAP NW Version	EXPRESSCLUSTER Version	OS	Cluster configuration	Remark
7.5	4.0.0-1 or later	Red Hat Enterprise Linux 7.3	NAS connection, Shared Disk Type	
		SUSE Linux Enterprise Server 12 SP1		
		Red Hat Enterprise Linux 7.4		

IBM POWER

SAP NW Version	EXPRESSCLUSTER Version	OS	Cluster configuration	Remark
7.5	4.0.0-1 or later	SUSE Linux Enterprise Server 11 SP4	NAS connection, Shared Disk Type	

Hardware and software requirements of the SAP NW, please refer to the documentation of the SAP NW.

Required memory and disk size

Required memory size		Required disk size		Remark
User mode	**Kernel mode**	**Right after installation**	**During operation**	
200MB(*1)	When the synchronization mode is used: 1MB + (number of request queues x I/O size) + (2MB + Difference Bitmap Size x number of mirror disk resources and hybrid disk resources When the asynchronous mode is used: 1MB + (number of request queues x I/O size) + (2MB + (number of asynchronous queues x I/O size) + Difference Bitmap Size) x number of mirror disk resources and hybrid disk resources When the kernel mode LAN heartbeat driver is used: 8MB When the keepalive driver is used: 8MB	300MB	2.0GB	

(*1) excepting for optional products.

Note:

The I/O size is 128 KB for the vxfs file system and 4 KB for file systems other than it.

For the setting value of the number of request queues and asynchronization queues, see "Understanding mirror disk resources" in the *Reference Guide*.

System requirements for the Cluster WebUI

Supported operating systems and browsers

Refer to the website, *http://www.nec.com/global/prod/expresscluster/*, for the latest information. Currently the following operating systems and browsers are supported:

Operating system	Browser	Language
Windows 7 Service Pack 1	Internet Explorer 11	English/Japanese/Chinese
Windows 8	Internet Explorer 10	English/Japanese/Chinese
Windows 8.1	Internet Explorer 11	English/Japanese/Chinese
Windows 10	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2012	Internet Explorer 10	English/Japanese/Chinese
Windows Server 2012 R2	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2016	Internet Explorer 11	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update8	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update9	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update4	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP6	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP7	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP1	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP2	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 6 update6	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Ubuntu 14.04 LTS	Firefox 51.0	English/Japanese/Chinese
Ubuntu 16.04.3 LTS	Firefox 51.0	English/Japanese/Chinese

Note:

When using an IP address to connect to Cluster WebUI, the IP address must be registered to **Site** of **Local Intranet** in advance.

Note:

When accessing Cluster WebUI with Internet Explorer 11, the Internet Explorer may stop with an error. In order to avoid it, please upgrade the Internet Explorer into KB4052978 or later.

Required memory and disk size

Required memory size: 200 MB or more

Required disk size: 50 MB or more

System requirements for the Builder

Supported operating systems and browsers

Refer to the website, *http://www.nec.com/global/prod/expresscluster/*, for the latest information. Currently supported operating systems and browsers are the following:

Operating system	Browser	Language
Microsoft Windows® 7 Service Pack 1	Internet Explorer 11	English/Japanese/Chinese
Microsoft Windows® 8	Internet Explorer 10	English/Japanese/Chinese
Microsoft Windows® 8.1	Internet Explorer 11	English/Japanese/Chinese
Microsoft Windows® 10	Internet Explorer 11	English/Japanese/Chinese
Microsoft Windows Server® 2012	Internet Explorer 10	English/Japanese/Chinese
Microsoft Windows Server® 2012 R2	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2016	Internet Explorer 11	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update8	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update9	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update4	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP1	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 6 update6	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Ubuntu 14.04 LTS	Firefox 51.0	English/Japanese/Chinese
Ubuntu 16.04.3 LTS	Firefox 51.0	English/Japanese/Chinese

Note:

When using an IP address to connect to WebManager, the IP address must be registered to **Site** of **Local Intranet** in advance.

Java runtime environment

Required:

Java™ Runtime Environment, Version 8.0 Update 152 (1.8.0_152) or later

Java™ Runtime Environment, Version 9.0 (9.0.4) or later

Note:

The Java applet version of Builder does not run in Java™ Runtime Environment version 9.0.

Required memory and disk size

Required memory size: 50MB or more

Required disk size: 10MB or more (excluding the size required for Java runtime environment)

Supported EXPRESSCLUSTER versions

Offline Builder version	EXPRESSCLUSTER X version
4.0.0-1	4.0.0-1
	4.0.1-1

Note:

When you use the Offline Builder and the EXPRESSCLUSTER `rpm`, a combination of their versions should be the one shown above. The Builder may not operate properly if they are used in a different combination.

System requirements for the WebManager

Supported operating systems and browsers

Refer to the website, *http://www.nec.com/global/prod/expresscluster/*, for the latest information. Currently the following operating systems and browsers are supported:

Operating system	Browser	Language
Windows 7 Service Pack 1	Internet Explorer 11	English/Japanese/Chinese
Windows 8	Internet Explorer 10	English/Japanese/Chinese
Windows 8.1	Internet Explorer 11	English/Japanese/Chinese
Windows 10	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2012	Internet Explorer 10	English/Japanese/Chinese
Windows Server 2012 R2	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2016	Internet Explorer 11	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update8	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update9	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update4	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP6	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP7	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP1	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP2	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 6 update6	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Ubuntu 14.04 LTS	Firefox 51.0	English/Japanese/Chinese
Ubuntu 16.04.3 LTS	Firefox 51.0	English/Japanese/Chinese

Note:

When using an IP address to connect to WebManager, the IP address must be registered to **Site** of **Local Intranet** in advance.

Java runtime environment

Required:

Java™ Runtime Environment, Version 8.0 Update 162 (1.8.0_162) or later

Java™ Runtime Environment, Version 9.0 (9.0.4) or later

Note:

The Java applet version of WebManager does not run in Java™ Runtime Environment version 9.0.

Required memory and disk size

Required memory size: 50MB or more

Required disk size: 10MB or more (excluding the size required for Java runtime environment)

System requirements for the Integrated WebManager

This section explains system requirements to operate the Integrated WebManager. Refer to the *Integrated WebManager Administrator's Guide* for the Java application version Integrated WebManager.

Supported operating systems and browsers

Currently the following operating systems and browsers are supported:

Operating system	Browser	Language
Windows 7 Service Pack 1	Internet Explorer 11	English/Japanese/Chinese
Windows 8	Internet Explorer 10	English/Japanese/Chinese
Windows 8.1	Internet Explorer 11	English/Japanese/Chinese
Windows 10	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2012	Internet Explorer 10	English/Japanese/Chinese
Windows Server 2012 R2	Internet Explorer 11	English/Japanese/Chinese
Windows Server 2016	Internet Explorer 11	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update8	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 6 update9	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Red Hat Enterprise Linux 7 update4	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP6	Firefox 51.0	English/Japanese/Chinese
Asianux Server 4 SP7	Firefox 51.0	English/Japanese/Chinese
Asianux Server 7 SP1	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 6 update6	Firefox 51.0	English/Japanese/Chinese
Oracle Linux 7 update3	Firefox 51.0	English/Japanese/Chinese
Ubuntu 14.04 LTS	Firefox 51.0	English/Japanese/Chinese
Ubuntu 16.04.3 LTS	Firefox 51.0	English/Japanese/Chinese

Java runtime environment

Required:

Java™ Runtime Environment, Version 8.0 Update 162 (1.8.0_162) or later

Java™ Runtime Environment, Version 9.0 (9.0.4) or later

Note:

The Java applet version of Integrated WebManager does not run in Java™ Runtime Environment version 9.0.

Required memory size and disk size

Required memory size: 50MB or more

Required disk size: 10MB or more (excluding the size required for Java runtime environment)

Chapter 4 Latest version information

This chapter provides the latest information on EXPRESSCLUSTER.
This chapter covers:

Correspondence list of EXPRESSCLUSTER and a manual

Description in this manual assumes the following version of EXPRESSCLUSTER. Make sure to note and check how EXPRESSCLUSTER versions and the editions of the manuals are corresponding.

EXPRESSCLUSTER Internal Version	Manual	Edition	Remarks
4.0.1-1	Installation and Configuration Guide	2nd Edition	
	Getting Started Guide	2nd Edition	
	Reference Guide	2nd Edition	
	Integrated WebManager Administrator's Guide	13th Edition	

New features and improvements

The following features and improvements have been released.

No.	Internal Version	Contents
1	4.0.0-1	Management GUI has been upgraded to Cluster WebUI.
2	4.0.0-1	HTTPS is supported for Cluster WebUI and WebManager.
3	4.0.0-1	The fixed term license is released.
4	4.0.0-1	The maximum number of mirror disk and/or hybrid disk resources has been expanded.
5	4.0.0-1	Volume manager resource and volume manager monitor resource support ZFS storage pool.
6	4.0.0-1	The supported operating systems have been expanded.
7	4.0.0-1	“systemd" is supported.
8	4.0.0-1	Oracle monitor resource supports Oracle Database 12c R2.
9	4.0.0-1	MySQL monitor resource supports MariaDB 10.2.
10	4.0.0-1	PostgreSQL monitor resource supports PowerGres on Linux 9.6.
11	4.0.0-1	SQL Server monitor resource has been added.
12	4.0.0-1	ODBC monitor resource has been added.
13	4.0.0-1	WebOTX monitor resource now supports WebOTX V10.1.
14	4.0.0-1	JVM monitor resource now supports Apache Tomcat 9.0.
15	4.0.0-1	JVM monitor resource now supports WebOTX V10.1.
16	4.0.0-1	The following monitor targets have been added to JVM monitor resource. - CodeHeap non-nmethods - CodeHeap profiled nmethods - CodeHeap non-profiled nmethods - Compressed Class Space
17	4.0.0-1	AWS DNS resource and AWS DNS monitor resource have been added.
18	4.0.0-1	Azure DNS resource and Azure DNS monitor resource have been added.
19	4.0.0-1	Monitoring behavior to detect error or timeout has been improved.
20	4.0.0-1	The function to execute a script before or after group resource activation or deactivation has been added.
21	4.0.0-1	The function to disable emergency shutdown for servers included in the same server group has been added.

No.	Internal Version	Contents
22	4.0.0-1	The function to create a rule for exclusive attribute groups has been added.
23	4.0.0-1	Internal communication has been improved to save TCP port usage.
24	4.0.0-1	The list of files for log collection has been revised.
25	4.0.0-1	Difference Bitmap Size to save differential data for mirror disk and hybrid disk resource is tunable.
26	4.0.1-1	The newly released kernel is now supported.
27	4.0.1-1	When HTTPS is unavailable in WebManager due to incorrect settings, messages are output to syslog and alert log.

Corrected information

Modification has been performed on the following minor versions.

Critical level:
- L: Operation may stop. Data destruction or mirror inconsistency may occur. Setup may not be executable.
- M: Operation stop should be planned for recovery. The system may stop if duplicated with another fault.
- S: A matter of displaying messages. Recovery can be made without stopping the system.

No.	Version in which the problem has been solved / Version in which the problem occurred	Phenomenon	Level	Occurrence condition/ Occurrence frequency	Cause
1	4.0.1-1 / 4.0.0-1	Two fixed-term licenses of the same product may be enabled.	S	This problem occurs on rare occasions if the following two operations are performed simultaneously. ● An unused license in stock is automatically enabled when the license expires. ● A new license is registered by the command for registering a license.	There was a flaw in performing exclusive control when operating license information.
2	4.0.1-1 / 4.0.0-1	The clpgrp command fails to start a group.	S	In a configuration where exclusive rules are set, this problem occurs when the clpgrp command is executed without specifying the name of the group to be started.	There was a flaw in the process when the group name is omitted.
3	4.0.1-1 / 4.0.0-1	In a configuration where CPU license and VM node license are mixed, a warning message appears, indicating that CPU licenses are insufficient.	S	This problem occurs when CPU license and VM node license are mix.	There was a flaw in counting licenses.

No.	Version in which the problem has been solved / Version in which the problem occurred	Phenomenon	Level	Occurrence condition/ Occurrence frequency	Cause
4	4.0.1-1 / 4.0.0-1	In Azure DNS monitor resources, even if the DNS server on Azure runs properly, it may be judged to be an error.	S	If all the following conditions are met, this problem inevitably occurs: ● [Check Name Resolution] is set to ON. ● When the version of Azure CLI is between 2.0.30 and 2.0.32 (this problem does not occur when the version is 2.0.29 or earlier, or 2.0.33 or later).	Since tab characters were included in the list of DNS servers acquired by the version of Azure CLI, an analysis for output results of Azure CLI failed.
5	4.0.1-1 / 4.0.0-1	In Azure DNS monitor resources, even if some of the DNS servers on Azure run properly, it may be judged to be an error.	S	If all the following conditions are met, this problem inevitably occurs: ● When [Check Name Resolution] is set to ON. ● The first DNS server on the list of the DNS servers acquired by Azure CLI does not run properly (The other DNS servers run properly.).	There was a flaw in confirming the soundness of DNS server.
6	4.0.1-1 / 4.0.0-1	In Azure DNS monitor resource, even if it fails to acquire the list of the DNS servers on Azure, it is not judged to be an error.	S	If all the following conditions are met, this problem inevitably occurs: ● When [Check Name Resolution] is set to ON. ● Azure CLI fails to acquire the list of the DNS servers.	There was a flaw in judging whether it is normal or abnormal.
7	4.0.1-1 / 4.0.0-1	When using the JVM monitor resources, memory leak may occur in the Java VM to be monitored.	M	This problem may occur under the following condition: ● [Monitor the number of Active Threads] on [Thread] tab in [Tuning] properties on [Monitor (special)] tab is set to on.	When extending Java API being used, classes which are not released in Scavenge GC may be accumulated.

No.	Version in which the problem has been solved / Version in which the problem occurred	Phenomenon	Level	Occurrence condition/ Occurrence frequency	Cause
8	4.0.1-1 / 4.0.0-1	Memory leak may occur In Java process of JVM monitor resources.	M	If all the following conditions are met, this problem may occur: ● All the settings in the [Tuning] properties on the [Monitor (special)] tab are set to OFF. ● More than one JVM monitor resource are created.	There was a flaw in disconnecting Java VM to be monitored.
9	4.0.1-1 / 4.0.0-1	The JVM statistics log (jramemory.stat) is output, even if the following parameters are set to OFF in JVM monitor resources. ● [Monitor (special)] tab – [Tuning] properties – [Memory] tab – [Memory Heap Memory Rate] ● [Memory (special)] tab – [Tuning] properties – [Memory] tab – [Monitor Non-Heap Memory Rate]	S	If all the following conditions are met, this problem inevitably occurs: ● [Oracle Java (usage monitoring)] is selected for [JVM type] on the [Monitor (special)] tab. ● [Monitor Heap Memory Rate] on the [Memory] tab in the [Tuning] properties on the [Monitor (special)] tab is set to OFF. ● [Monitor Non-Heap Memory Rate] on the [Memory] tab in the [Tuning] properties on the [Monitor (special)] tab is set to OFF.	There was a flaw in deciding whether or not to output the JVM statistics log.

Chapter 5 Notes and Restrictions

This chapter provides information on known problems and how to troubleshoot the problems.
This chapter covers:

Designing a system configuration

Hardware selection, option products license arrangement, system configuration, and shared disk configuration are introduced in this section.

Function list and necessary license

The following option products are necessary as many as the number of servers.

Those resources and monitor resources for which the necessary licenses are not registered are not on the resource list of the Builder (online version).

Necessary function	Necessary license
Mirror disk resource	EXPRESSCLUSTER X Replicator 4.0 *1
Hybrid disk resource	EXPRESSCLUSTER X Replicator DR 4.0 *2
Oracle monitor resource	EXPRESSCLUSTER X Database Agent 4.0
DB2 monitor resource	EXPRESSCLUSTER X Database Agent 4.0
PostgreSQL monitor resource	EXPRESSCLUSTER X Database Agent 4.0
MySQL monitor resource	EXPRESSCLUSTER X Database Agent 4.0
Sybase monitor resource	EXPRESSCLUSTER X Database Agent 4.0
SQL Server monitor resource	EXPRESSCLUSTER X Database Agent 4.0
ODBC monitor resource	EXPRESSCLUSTER X Database Agent 4.0
Samba monitor resource	EXPRESSCLUSTER X File Server Agent 4.0
nfs monitor resource	EXPRESSCLUSTER X File Server Agent 4.0
http monitor resource	EXPRESSCLUSTER X Internet Server Agent 4.0
smtp monitor resource	EXPRESSCLUSTER X Internet Server Agent 4.0
pop3 monitor resource	EXPRESSCLUSTER X Internet Server Agent 4.0
imap4 monitor resource	EXPRESSCLUSTER X Internet Server Agent 4.0
ftp monitor resource	EXPRESSCLUSTER X Internet Server Agent 4.0
Tuxedo monitor resource	EXPRESSCLUSTER X Application Server Agent 4.0
Weblogic monitor resource	EXPRESSCLUSTER X Application Server Agent 4.0
Websphere monitor resource	EXPRESSCLUSTER X Application Server Agent 4.0
WebOTX monitor resource	EXPRESSCLUSTER X Application Server Agent 4.0
JVM monitor resource	EXPRESSCLUSTER X Java Resource Agent 4.0
System monitor resource	EXPRESSCLUSTER X System Resource Agent 4.0
Mail report actions	EXPRESSCLUSTER X Alert Service 4.0
Network Warning Light status	EXPRESSCLUSTER X Alert Service 4.0

*1 When configuring data mirror form, product **Replicator** must be purchased.
*2 When configuring mirror between shared disk, product **Replicator DR** must be purchased.

Hardware requirements for mirror disks

- Linux md stripe set, volume set, mirroring, and stripe set with parity cannot be used for either mirror disk resource cluster partitions or data partitions.
- Linux LVM volumes can be used for both cluster partitions and data partitions.
  For SuSE, however, LVM and MultiPath volumes cannot be used for data partitions. (This is because for SuSE, ReadOnly or ReadWrite control over these volumes cannot be performed by EXPRESSCLUSTER.)
- Mirror disk resource cannot be made as a target of a Linux md stripe set, volume set, mirroring, and stripe set with parity.
- Mirror partitions (data partition and cluster partition) to use a mirror disk resource.
- There are two ways to allocate mirror partitions:
  - Allocate a mirror partition (data partition and cluster partition) on the disk where the operating system (such as root partition and swap partition) resides.
  - Reserve (or add) a disk (or LUN) not used by the operating system and allocate a mirror partition on the disk.
- Consider the following when allocating mirror partitions:
  - When maintainability and performance are important:
    - It is recommended to have a mirror disk that is not used by the OS.
  - When LUN cannot be added due to hardware RAID specification or when changing LUN configuration is difficult in hardware RAID pre-install model:
    - Allocate a mirror partition on the same disk where the operating system resides.
- When multiple mirror disk resources are used, it is recommended to prepare (adding) a disk per mirror disk resource. Allocating multiple mirror disk resources on the same disk may result in degraded performance and it may take a while to complete mirror recovery due to disk access performance on Linux operating system.
- Disks used for mirroring must be the same in all servers.
  - Disk interface

    Mirror disks on both servers and disks where mirror partition is allocated should be of the same disk interface

    Example

Combination	server1	server2
OK	SCSI	SCSI
OK	IDE	IDE
NG	IDE	SCSI

  - Disk type

    Mirror disks on both servers and disks where mirror partition is allocated should be of the same disk type

    Example

Combination	server1	server2
OK	HDD	HDD
OK	SSD	SSD
NG	HDD	SSD

  - Sector size

Mirror disks on both servers and disks where mirror partition is allocated should be of the same sector size

Example

Combination	server1	server2
OK	512B	512B
OK	4KB	4KB
NG	512B	4KB

- Notes when the geometries of the disks used as mirror disks differ between the servers.

  The partition size allocated by the `fdisk` command is aligned by the number of blocks (units) per cylinder. Allocate a data partition considering the relationship between data partition size and direction for initial mirror configuration to be as indicated below:

  Source server ≤ Destination server

  "Source server" refers to the server where the failover group that a mirror disk resource belongs has a higher priority in failover policy. "Destination server" refers to the server where the failover group that a mirror disk resource belongs has a lower priority in failover policy.

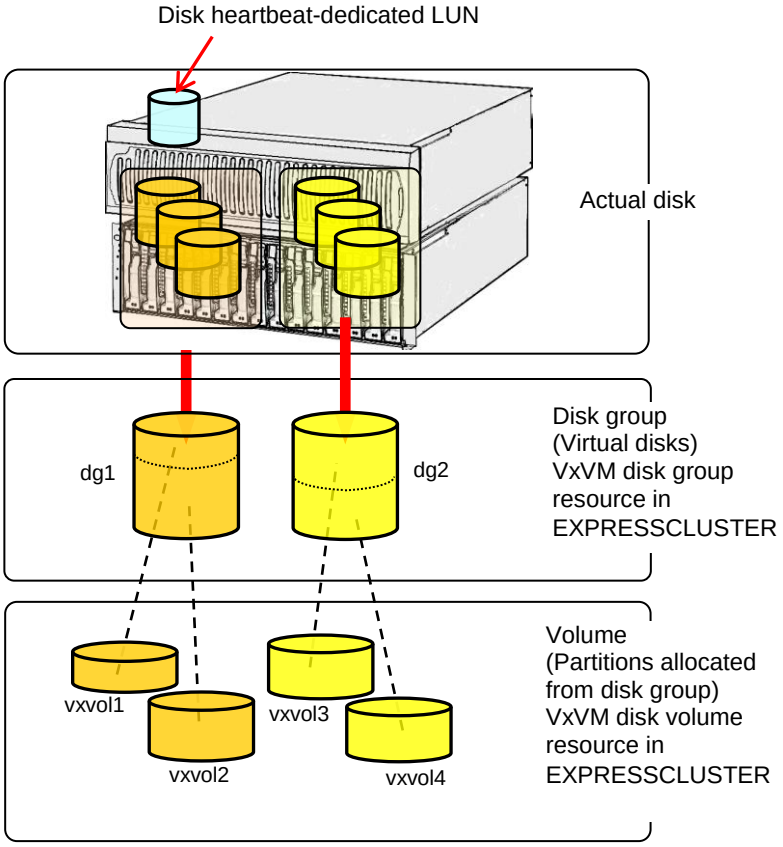
  Make sure that the data partition sizes do not cross over 32GiB, 64GiB, 96GiB, and so on (multiples of 32GiB) on the source server and the destination server. For sizes that cross over multiples of 32GiB, initial mirror construction may fail. Be careful, therefore, to secure data partitions of similar sizes.

  Example)

Combination	Data partition size		Description
	On server 1	On server 2	
OK	30GiB	31GiB	OK because both are in the range of 0 to 32GiB.
OK	50GiB	60GiB	OK because both are in the range of 32GiB to 64GiB.
NG	30GiB	39GiB	Error because they are crossing over 32GiB.
NG	60GiB	70GiB	Error because they are crossing over 64GiB.

Hardware requirements for shared disks

- When a Linux LVM stripe set, volume set, mirroring, or stripe set with parity is used:
  - EXPRESSCLUSTER cannot control ReadOnly/ReadWrite of the partition configured for the disk resource.
- When you use VxVM or LVM, a LUN that is not controlled by VxVM or LVM is required on a shared disk for the disk heartbeat of EXPRESSCLUSTER. You should bear this in your mind when configuring LUN on the shared disk.
- When you use LVM features, use the disk resource (disk type: “lvm”) and the volume manager resource.



Hardware requirements for hybrid disks

- Disks to be used as a hybrid disk resource do not support a Linux md stripe set, volume set, mirroring, and stripe set with parity.
- Linux LVM volumes can be used for both cluster partitions and data partitions.
  For SuSE, however, LVM and MultiPath volumes cannot be used for data partitions. (This is because for SuSE, ReadOnly or ReadWrite control over these volumes cannot be performed by EXPRESSCLUSTER.)
- Hybrid disk resource cannot be made as a target of a Linux md stripe set, volume set, mirroring, and stripe set with parity.
- Hybrid partitions (data partition and cluster partition) are required to use a hybrid disk resource.
- When a disk for hybrid disk is allocated in the shared disk, a partition for disk heartbeat resource between servers sharing the shared disk device is required.
- The following are the two ways to allocate partitions when a disk for hybrid disk is allocated from a disk which is not a shared disk:
  - Allocate hybrid partitions (data partition and cluster partition) on the disk where the operating system (such as root partition and swap partition) resides.
  - Reserve (or add) a disk (or LUN) not used by the operating system and allocate a hybrid partition on the disk.
- Consider the following when allocating hybrid partitions:
  - When maintainability and performance are important:
    - It is recommended to have a hybrid disk that is not used by the OS.
  - When LUN cannot be added due to hardware RAID specification or when changing LUN configuration is difficult in hardware RAID pre-install model:
    - Allocate a hybrid partition on the same disk where the operating system resides.
- When multiple hybrid disk resources are used, it is recommended to prepare (add) a LUN per hybrid disk resource. Allocating multiple hybrid disk resources on the same disk may result in degraded in performance and it may take a while to complete mirror recovery due to disk access performance on Linux operating system.

	Device for which hybrid disk resource is allocated	
Type of required partition	Shared disk device	Non-shared disk device
Data partition	Required	Required
Cluster partition	Required	Required
Partition for disk heart beat	Required	Not Required
Allocation on the same disk (LUN) as where the OS is	-	Possible

- Notes when the geometries of the disks used as hybrid disks differ between the servers.

  Allocate a data partition considering the relationship between data partition size and direction for initial mirror configuration to be as indicated below:

  Source server ≤ Destination server

  "Source server" refers to the server with a higher priority in failover policy in the failover group where the hybrid disk resource belongs. "Destination server" refers to the server with a lower priority in failover policy in the failover group where the hybrid disk resource belongs has.

  Make sure that the data partition sizes do not cross over 32GiB, 64GiB, 96GiB, and so on (multiples of 32GiB) on the source server and the destination server. For sizes that cross over

multiples of 32GiB, initial mirror construction may fail. Be careful, therefore, to secure data partitions of similar sizes.

Example)

Combination	Data partition size		Description
	On server 1	On server 2	
OK	30GiB	31GiB	OK because both are in the range of 0 to 32GiB.
OK	50GiB	60GiB	OK because both are in the range of 32GiB to 64GiB.
NG	30GiB	39GiB	Error because they are crossing over 32GiB.
NG	60GiB	70GiB	Error because they are crossing over 64GiB.

IPv6 environment

The following function cannot be used in an IPv6 environment:

- BMC heartbeat resource

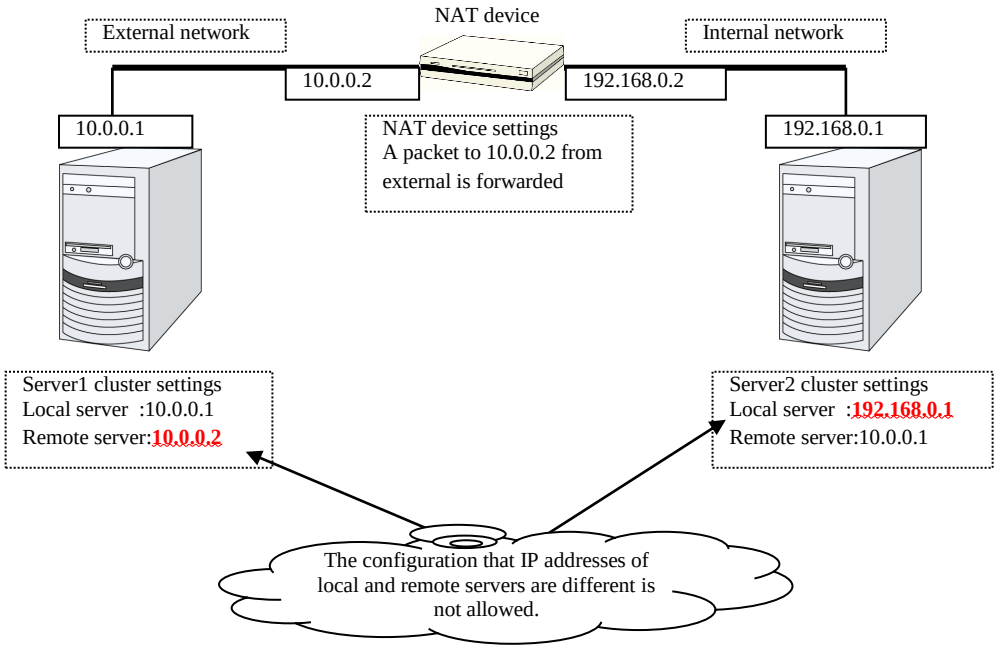
The following functions cannot use link-local addresses:

- LAN heartbeat resource
- Kernel mode LAN heartbeat resource
- Mirror disk connect
- PING network partition resolution resource
- FIP resource
- VIP resource

Network configuration

The cluster configuration cannot be configured or operated in an environment, such as NAT, where an IP address of a local server is different from that of a remote server.

Example of network configuration



Execute Script before Final Action setting for monitor resource recovery action

EXPRESSCLUSTER version 3.1.0-1 and later supports the execution of a script before reactivation and before failover.

The same script is executed in either case. Therefore, if **Execute Script before Final Action** is set with a version earlier than 3.1.0-1, editing of the script file may be required.

For the additional script configuration needed to execute the script before reactivation and before failover, the script file must be edited to assign processing to each recovery action.

For the assignment of processing for a recovery action, see "Recovery/pre-recovery action script" in Chapter 5, "Monitor resource details" in the *Reference Guide*.

NIC Link Up/Down monitor resource

Some NIC boards and drivers do not support required `ioctl()`.

The propriety of a NIC Link Up/Down monitor resource of operation can be checked by the `ethtool` command which each distributor offers.

```
ethtool eth0
Settings for eth0:
     Supported ports: [ TP ]
     Supported link modes:   10baseT/Half 10baseT/Full
                             100baseT/Half 100baseT/Full
                             1000baseT/Full
     Supports auto-negotiation: Yes
     Advertised link modes:  10baseT/Half 10baseT/Full
                             100baseT/Half 100baseT/Full
                             1000baseT/Full
     Advertised auto-negotiation: Yes
     Speed: 1000Mb/s
     Duplex: Full
     Port: Twisted Pair
     PHYAD: 0
     Transceiver: internal
     Auto-negotiation: on
     Supports Wake-on: umbg
     Wake-on: g
     Current message level: 0x00000007 (7)
     Link detected: yes
```

- When the LAN cable link status ("Link detected: yes") is not displayed as the result of the `ethtool` command:
  - It is highly likely that NIC Link Up/Down monitor resource of EXPRESSCLUSTER is not operable. Use IP monitor resource instead.

- When the LAN cable link status ("Link detected: yes") is displayed as the result of the `ethtool` command:
  - In most cases NIC Link Up/Down monitor resource of EXPRESSCLUSTER can be operated, but sometimes it cannot be operated.
  - Particularly in the following hardware, NIC Link Up/Down monitor resource of EXPRESSCLUSTER may not be operated. Use IP monitor resource instead.
  - When hardware is installed between the actual LAN connector and NIC chip such as a blade server
  - When the monitored NIC is in a bonding environment, check whether the MII Polling Interval is set to 0 or higher.

To check if NIC Link Up/Down monitor resource can be used by using EXPRESSCLUSTER on an actual machine, follow the steps below to check the operation.

1. Register NIC Link Up/Down monitor resource with the configuration information. Select **No Operation** for the configuration of recovery operation of NIC Link Up/Down monitor resource upon failure detection.
2. Start the cluster.

3. Check the status of NIC Link Up/Down monitor resource.
   If the status of NIC Link Up/Down monitor resource is abnormal while LAN cable link status is normal, NIC Link Up/Down monitor resource cannot be operated.
4. If NIC Link Up/Down monitor resource status becomes abnormal when LAN cable link status is made abnormal status (link down status), NIC Link Up/Down monitor resource cannot be operated.
   If the status remains to be normal, NIC Link Up/Down monitor resource cannot be operated.

Write function of the mirror disk resource and hybrid disk resource

- A mirror disk and a hybrid disk resource write data in the disk of its own server and the disk of the remote server via network. Reading of data is done only from the disk on own server.
- Writing functions shows poor performance in mirroring when compared to writing to a single server because of the reason provided above. For a system that requires through-put as high as single server, use a shared disk.

Not outputting syslog to the mirror disk resource or the hybrid disk resource

Do not set directories or subdirectories which mounted the mirror disk resource or the hybrid disk resource as syslog output destination directories.

When the mirror disk connection is disconnected, the I/O to the mirror partition may stop until the disconnection is detected. The system may become abnormal because of the syslog output stoppage at this time.

When outputting syslog to the mirror disk resource or the hybrid disk resource is necessary, consider the followings.

- Use bonding as a way of path redundancy of the mirror disk connection.
- Adjust the user-mode monitoring timeout value or the mirror related timeout values.

Notes when terminating the mirror disk resource or the hybrid disk resource

- In case that processes which access to the directories, subdirectories and files which mounted the mirror disk resource or the hybrid disk resource exist, terminate the accesses to each disk resource by using ending script or other methods at deactivation of each disk resource like when shutdown or failover.
  Depending on the settings of each disk resource, action at abnormity detection when unmounting (forcibly terminate processes while each disk resource is being accessed) may occur, or recovery action at deactivation failure caused by unmount failure (OS shutdown or other actions) may be executed.
- In case that a massive amount of accesses to directories, subdirectories or files which mounted the mirror disk resource or hybrid disk resource are executed, it may take much time before the cache of the file systems is written out to the disks when unmounting at disk resource deactivation.
  At times like this, set the timeout interval of unmount longer enough so that the writing to the disks will successfully complete.
- For the details of this setting,
  see Chapter 4, "Group resource details" in Reference Guide,
  Settings Tab or **Mirror Disk Resource Tuning Properties** or **Unmount Tab** in **Details Tab** in "Understanding mirror disk resources" or "Understanding mirror disk resources".

Data consistency among multiple asynchronous mirror disks

In mirror disk or hybrid disk with asynchronous mode, writing data to the data partition of the active server is performed in the same order as the data partition of the standby server.

This writing order is guaranteed except during the initial mirror disk configuration or recovery (copy) period after suspending mirroring the disks. The data consistency among the files on the standby data partition is guaranteed.

However, the writing order is not guaranteed among multiple mirror disk resources and hybrid disk resources. For example, if a file gets older than the other and files that cannot maintain the data consistency are distributed to multiple asynchronous mirror disks, an application may not run properly when it fails over due to server failure.

For this reason, be sure to place these files on the same asynchronous mirror disk or hybrid disk.

Mirror data reference at the synchronization destination if mirror synchronization is interrupted

If mirror synchronization is interrupted for a mirror disk or a hybrid disk in the mirror synchronization state, using the mirror disk helper or the `clpmdctrl` / `clphdctrl` command (with the `--break` / `-b` / `--nosync` option specified), the file system and application data may be abnormal if the mirror disk on the server on the mirror synchronization destination (copy destination) is made accessible by performing forced activation (removing the access restriction) or forced mirror recovery.

This occurs because if mirror synchronization is interrupted on the server on the mirror synchronization source (server on which the resources are activated) leading to an inconsistent state in which there are portions that can be synchronized with the synchronization destination and portions that cannot be synchronized such as; for example, when an application is writing to a mirror disk area, part of the data and so on will be retained in the cache and so on (memory) of the OS, but not yet actually written to the mirror disk, or may be in the process of being written.

If you want to perform access in a state in which consistency with the mirror disk on the mirror synchronization destination (standby server) is ensured, secure a rest point on the mirror synchronization source (active server on which the resources are activated) first and then interrupt mirror synchronization. Alternatively, secure a rest point by deactivating. (With an application end, access to the mirror area ends, and by unmounting the mirror disk, the cache and so on of the OS are all written to the mirror disk.)

For an example of securing a rest point, refer to the "EXPRESSCLUSTER X PP Guide (Schedule Mirror)," provided with the StartupKit.

Similarly, if mirror recovery is interrupted for a mirror disk or a hybrid disk that is in the middle of mirror recovery (mirror resynchronization), the file system and application data may be abnormal if the mirror disk on the mirror synchronization destination is accessed by performing forced activation (removing the access restriction) or forced mirror recovery.
This also occurs because mirror recovery is interrupted in an inconsistent state in which there are portions that can be synchronized but also portions that cannot.

O_DIRECT for mirror or hybrid disk resources

Do not use the O_DIRECT flag of the `open()` system call for mirror or hybrid disk resources. Examples include the Oracle parameter `filesystemio_options = setall`.

Do not specify the disk monitor O_DIRECT mode for mirror or hybrid disk resources.

Initial mirror construction time for mirror or hybrid disk resources

The time that takes to construct the initial mirror is different between ext2/ext3/ext4 and other file systems.

Mirror or hybrid disk connect

- When using redundant mirror or hybrid disk connect, both version of IP address are needed to be the same.
- All the IP addresses used by mirror disk connect must be set to IPv4 or IPv6.

JVM monitor resources

- Up to 25 Java VMs can be monitored concurrently. The Java VMs that can be monitored concurrently are those which are uniquely identified by the Builder (with **Identifier** in the **Monitor (special)** tab).
- Connections between Java VMs and Java Resource Agent do not support SSL.
- It may not be possible to detect thread deadlocks. This is a known problem in Java VM. For details, refer to "Bug ID: 6380127" in the Oracle Bug Database.
- The JVM monitor resources can monitor only the Java VMs on the server on which the JVM monitor resources are running.
- The JVM monitor resources can monitor only one JBoss server instance per server.
- The Java installation path setting made by the Builder (with **Java Installation Path** in the **JVM monitor** tab in **Cluster Properties**) is shared by the servers in the cluster. The version and update of Java VM used for JVM monitoring must be the same on every server in the cluster.
- The management port number setting made by the Builder (with **Management Port** in the **Connection Setting** dialog box opened from the **JVM monitor** tab in **Cluster Properties**) is shared by all the servers in the cluster.
- Application monitoring is disabled when an application to be monitored on the IA32 version is running on an x86_64 version OS.
- If a large value such as 3,000 or more is specified as the maximum Java heap size by the Builder (by using Maximum Java Heap Size on the **JVM monitor** tab in **Cluster Properties**), The JVM monitor resources will fail to start up. The maximum heap size differs depending on the environment, so be sure to specify a value based on the capacity of the mounted system memory.
- Using SingleServerSafe is recommended if you want to use the target Java VM load calculation function of the coordination load balancer. It’s supported only by Red Hat Enterprise Linux.
- If "-XX:+UseG1GC" is added as a startup option of the target Java VM, the settings on the **Memory** tab on the **Monitor(special)** tab in **Properties** of JVM monitor resources cannot be monitored before Java 7.
  It’s possible to monitor by choosing **Oracle Java (usage monitoring)** in **JVM Type** on the **Monitor(special)** tab after Java 8.

Mail reporting

The mail reporting function is not supported by STARTTLS and SSL.

Requirements for network warning light

- When using “DN-1000S” or “DN-1500GL,” do not set your password for the warning light.
- To play an audio file as a warning, you must register the audio file to a network warning light supporting audio file playback.
  For details about how to register an audio file, see the manual of the network warning light you want to use.
- Set up a network warning light so that a server in a cluster is permitted to execute the `rsh` command to that warning light.

Installing operating system

Notes on parameters to be determined when installing an operating system, allocating resources, and naming rules are described in this section.

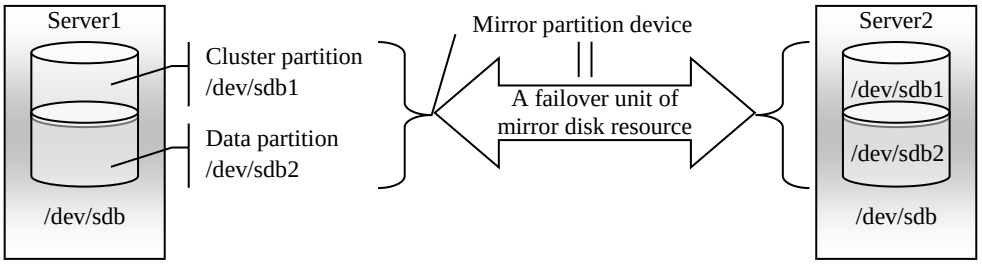
/opt/nec/clusterpro file system

It is recommended to use a file system that has journaling functions to improve tolerance for system failure. File systems such as ext3, ext4, JFS, ReiserFS, XFS are available for a journaling file system supported by Linux (kernel version 2.6 or later). If a file system that is not capable of journaling is used, run an interactive command (fsck the root file system) when rebooting from server or OS stop (i.e. normal shutdown could not be done.)

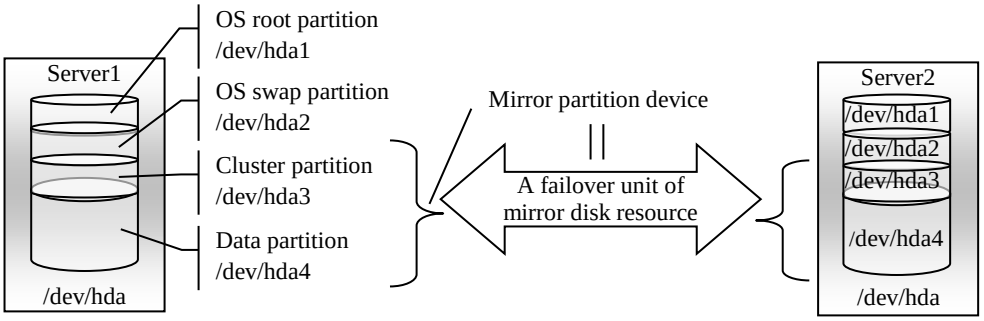
Mirror disks

- Disk partition

  Example: When adding one SCSI disk to each of both servers and making a pair of mirrored disks:



  Example: When using free space of IDE disks of both servers, where the OS is stored, and making a pair of mirrored disks:



  - Mirror partition device refers to cluster partition and data partition.
  - Allocate cluster partition and data partition on each server as a pair.
  - It is possible to allocate a mirror partition (cluster partition and data partition) on the disk where the operating system resides (such as root partition and swap partition.).
    - When maintainability and performance are important:

      It is recommended to have a mirror disk that is not used by the operating system (such as root partition and swap partition.)

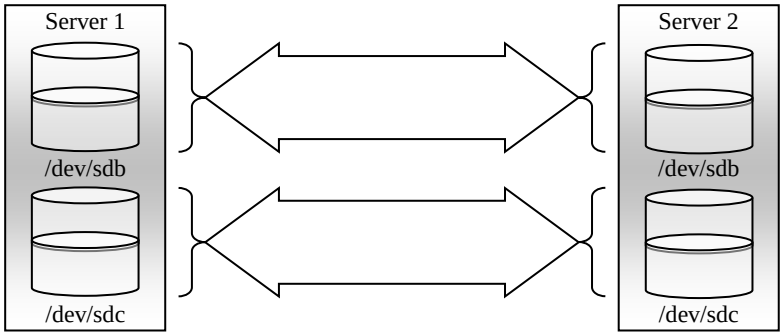
- When LUN cannot be added due to hardware RAID specification: or
  When changing LUN configuration is difficult in hardware RAID pre-install model:

  It is possible to allocate a mirror partition (cluster partition and data partition) on the disk where the operating system resides (such as root partition and swap partition.)

◆ Disk configurations

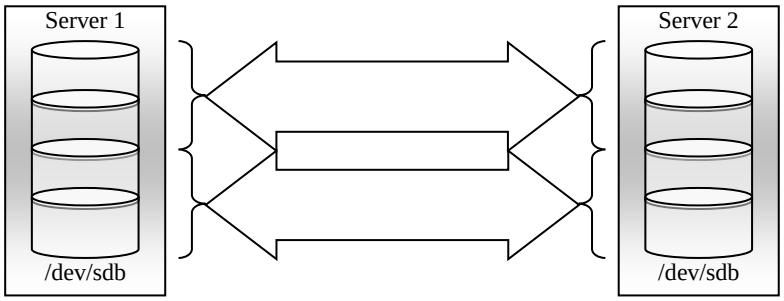
Multiple disks can be used as mirror disks on a single server. Or, you can allocate multiple mirror partitions on a single disk.

Example: When adding two SCSI disks to each of both servers and making two pairs of mirrored disks:



- Allocate two partitions, cluster partition and data partition, as a pair on each disk.
- Use of the data partition as the first disk and the cluster partition as the second disk is not permitted.

Example: When adding one SCSI disk to each of both servers and making two mirror partitions:



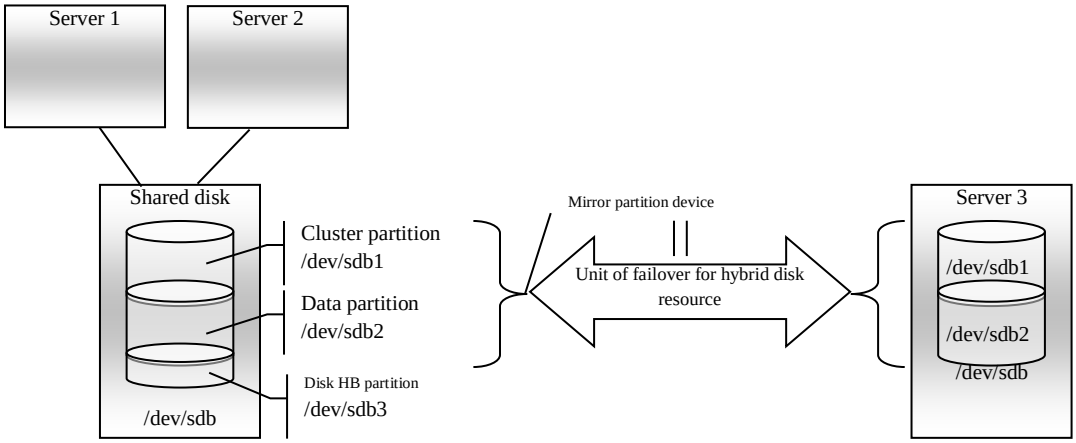
◆ A disk does not support a Linux md stripe set, volume set, mirroring, and stripe set with parity.

Hybrid disks

- Disk partition

  Disks that are shared or not shared (server with built-in disk, external disk chassis not shared by servers etc.) can be used.

  Example) When two servers use a shared disk and the third server uses a built-in disk in the server:



  - Mirror partition device is a device EXPRESSCLUSTER mirroring driver provides in the upper.
  - Allocate cluster partition and data partition on each server as a pair.
  - When a disk that is not shared (e.g. server with a built-in disk, external disk chassis that is not shared among servers) is used, it is possible to allocate mirror partitions (cluster partition and data partition) on the disk where the operating system resides (such as root partition and swap partition.).
    - When maintainability and performance are important:

      It is recommended to have a mirror disk that is not used by the operating system (such as root partition and swap partition.)

    - When LUN cannot be added due to hardware RAID specification: or
      When changing LUN configuration is difficult in hardware RAID pre-install model:

      It is possible to allocate mirror partitions (cluster partition and data partition) on the disk where the operating system resides (such as root partition and swap partition.)
  - When a hybrid disk is allocated in a shared disk device, allocate a partition for the disk heart beat resource between servers sharing the shared disk device.
  - A disk does not support a Linux md stripe set, volume set, mirroring, and stripe set with parity.

Dependent library

- libxml2

Install libxml2 when installing the operating system.

Dependent driver

- softdog

This driver is necessary when softdog is used to monitor user-mode monitor resource.

Configure a loadable module. Static driver cannot be used.

The major number of Mirror driver

Use mirror driver's major number 218. Do not use major number 218 for other device drivers.

The major number of Kernel mode LAN heartbeat and keepalive drivers

- Use major number 10, minor number 240 for kernel mode LAN heartbeat driver.
- Use major number 10, minor number 241 for keepalive driver.

Make sure to check that other drivers are not using major and minor numbers described above.

Partition for RAW monitoring of disk monitor resources

Allocate a partition for monitoring when setting up RAW monitoring of disk monitor resources. The partition size should be 10MB.

SELinux settings

- Configure permissive or disabled for the SELinux settings.
- If you set enforcing, communication required in EXPRESSCLUSTER may not be achieved.

NetworkManager settings

If the NetworkManager service is running in a Red Hat Enterprise Linux 6 environment, an unintended behavior (such as detouring the communication path, or disappearance of the network interface) may occur upon disconnection of the network. It is recommended to set NetworkManager to stop the service.

LVM metadata daemon settings

- When controlling or monitoring the LVM by using the volume manager resource or volume manager monitor resource in an environment of Red Hat Enterprise Linux 7 or later, the LVM metadata daemon must be disabled.

  The procedure to disable the metadata daemon is as follows:

  (1) Execute the following command to stop the LVM metadata daemon.

  ```
  # systemctl stop lvm2-lvmetad.service
  ```

  (2) Edit `/etc/lvm/lvm.conf` to set the value of `use_lvmetad` to 0.

Before installing EXPRESSCLUSTER

Notes after installing an operating system, when configuring OS and disks are described in this section.

Communication port number

In EXPRESSCLUSTER, the following port numbers are used. You can change the port number by using the Builder.

Make sure not to access the following port numbers from a program other than EXPRESSCLUSTER.

Configure to be able to access the port number below when setting a firewall on a server.

For an AWS environment, configure to able to access the following port numbers in the security group setting in addition to the firewall setting.

Server to Server					
Loopback in servers					
From			**To**		**Used for**
Server	Automatic allocation[1]	→	Server	29001/TCP	Internal communication
Server	Automatic allocation	→	Server	29002/TCP	Data transfer
Server	Automatic allocation	→	Server	29002/UDP	Heartbeat
Server	Automatic allocation	→	Server	29003/UDP	Alert synchronization
Server	Automatic allocation	→	Server	29004/TCP	Communication between mirror agents
Server	Automatic allocation	→	Server	29006/UDP	Heartbeat (kernel mode)
Server	Automatic allocation	→	Server	XXXX[2]/TCP	Mirror disk resource data synchronization
Server	Automatic allocation	→	Server	XXXX[3]/TCP	Communication between mirror drivers
Server	Automatic allocation	→	Server	XXXX[4]/TCP	Communication between mirror drivers
Server	icmp	→	Server	icmp	keepalive between mirror drivers, duplication check for FIP/VIP resource and mirror agent
Server	Automatic allocation	→	Server	XXXX[5]/UDP	Internal log communication

Cluster WebUI / WebManager to Server					
From			**To**		**Used for**
Cluster WebUI WebManager	Automatic allocation	→	Server	29003/TCP	http communication

Server connected to the Integrated WebManager to Target server					
From			**To**		**Used for**
Server connected to the Integrated WebManager	Automatic allocation	→	Server	29003/TCP	http communication
Server to be managed by the Integrated WebManager	29003	→	Client	29010/UDP	UDP communication

Others					
From			**To**		**Used for**
Server	Automatic allocation	→	Network warning light	See the manual for each product.	Network warning light control
Server	Automatic allocation	→	Management LAN of server BMC	623/UDP	BMC control (Forced stop / Chassis lamp association)
Management LAN of server BMC	Automatic allocation	→	Server	162/UDP	Monitoring target of the external linkage monitor configured for BMC linkage
Management LAN of server BMC	Automatic allocation	→	Management LAN of server BMC	5570/UDP	BMC HB communication
Server	icmp	→	Monitoring target	icmp	IP monitor
Server	icmp	→	NFS server	icmp	Checking if NFS server is active by NAS resource
Server	icmp	→	Monitoring target	icmp	Monitoring target of Ping method network partition resolution resource
Server	Automatic allocation	→	Server	Management port number set by the Builder[6]	JVM monitor
Server	Automatic allocation	→	Monitoring target	Connection port number set by the Builder[6]	JVM monitor
Server	Automatic allocation	→	Server	Load balancer linkage management port number set by the Builder[6]	JVM monitor
Server	Automatic allocation	→	BIG-IP LTM	Communication port number set by the Builder[6]	JVM monitor
Server	Automatic allocation	→	Server	Probe port set by the Builder[7]	Azure probe port resource
Server	Automatic allocation	→	AWS region endpoint	443/tcp[8]	AWS elastic ip resource AWS virtual ip

Others					
From			**To**		**Used for**
					resource
					AWS DNS resource
					AWS elastic ip monitor resource
					AWS virtual ip monitor resource
					AWS AZ monitor resource
					AWS DNS monitor resource
Server	Automatic allocation	→	Azure endpoint	443/tcp[9]	Azure DNS resource
Server	Automatic allocation	→	Azure authoritative name server	53/udp	Azure DNS monitor resource

1. In automatic allocation, a port number not being used at a given time is allocated.
2. This is a port number used per mirror disk resource or hybrid disk resource and is set when creating mirror disk resource or hybrid disk resource. A port number 29051 is set by default. When you add a mirror disk resource or hybrid disk resource, this value is automatically incremented by 1. To change the value, click **Details** tab in the **[md] Resource Properties** or the **[hd] Resource Properties** dialog box of the Builder. For more information, refer to Chapter 4, "Group resource details" in the *Reference Guide*.
3. This is a port number used per mirror disk resource or hybrid disk resource and is set when creating mirror disk resource or hybrid disk resource. A port number 29031 is set by default. When you add a mirror disk resource or a hybrid disk resource, this value is automatically incremented by 1. To change the value, click **Details** tab in the **[md] Resource Properties** or the **[hd] Resource Properties** dialog box of the Builder. For more information, refer to Chapter 4, "Group resource details" in the *Reference Guide*.
4. This is a port number used per mirror disk resource or hybrid disk resource and is set when creating mirror disk resource or hybrid disk resource. A port number 29071 is set by default. When you add a mirror disk resource or hybrid disk resource this value is automatically incremented by 1. To change the value, click **Details** tab in the **[md] Resource Properties** or the **[hd] Resource Properties** dialog box of the Builder. For more information, refer to Chapter 4, "Group resource details" in the *Reference Guide*.
5. Select **UDP** for the **Communication Method for Internal Logs** in the **Port No. (Log)** tab in **Cluster Properties**. Use the port number configured in Port No. Communication port is not used for the default log communication method **UNIX Domain**.
6. The JVM monitor resource uses the following four port numbers.
   - A management port number is a port number that the JVM monitor resource internally uses. To set this number, use the **Connection Setting** dialog box opened from the **JVM monitor** tab in **Cluster Properties** of the Builder. For details, refer to Chapter 2, "Function of the Builder" in the *Reference Guide*.
   - A connection port number is used to establish a connection to the target Java VM (WebLogic Server or WebOTX). To set this number, use the **Monitor (special)** tab in **Properties** of the Builder for the corresponding JVM monitor resource. For details, refer to Chapter 6, "Monitor resource details" in the *Reference Guide*.
   - A load balancer linkage management port number is used for load balancer linkage. When load balancer linkage is not used, this number does not need to be set. To set the

number, use opened from the **JVM monitor** tab in **Cluster Properties** of the Builder. For details, refer to Chapter 2, "Function of the Builder" in the *Reference Guide*.

- A communication port number is used to accomplish load balancer linkage with BIG-IP LTM. When load balancer linkage is not used, this number does not need to be set. To set the number, use the **Load Balancer Linkage Settings** dialog box opened from the **JVM monitor** tab in **Cluster Properties** of the Builder. For details, refer to Chapter 2, "Function of the Builder" in the *Reference Guide*.

7. Port number used by the Microsoft Azure load balancer for the alive monitoring of each server.
8. The AWS elastic ip resource, AWS virtual ip resource, AWS DNS resource, AWS elastic ip monitor resource, AWS virtual ip monitor resource, AWS AZ monitor resource, and AWS DNS monitor resource run the AWS CLI. The above port numbers are used by the AWS CLI.
9. The Azure DNS resource runs the Azure CLI. The above port numbers are used by the Azure CLI.

Changing the range of automatic allocation for the communication port numbers

- The range of automatic allocation for the communication port numbers managed by the OS might overlap the communication port numbers used by EXPRESSCLUSTER.
- Change the OS settings to avoid duplication when the range of automatic allocation for the communication numbers managed by OS and the communication numbers used by EXPRESSCLUSTER are duplicated.

Examples of checking and displaying OS setting conditions.

The range of automatic allocation for the communication port numbers depends on the distribution.

```
# cat /proc/sys/net/ipv4/ip_local_port_range
1024   65000
```

This is the condition to be assigned for the range from 1024 to 65000 when the application requests automatic allocation for the communication port numbers to the OS.

```
# cat /proc/sys/net/ipv4/ip_local_port_range
32768   61000
```

This is the condition to be assigned for the range from 32768 to 61000 when the application requests automatic allocation for the communication port numbers to the OS.

Examples of OS settings change

Add the line below to `/etc/sysctl.conf`. (When changing to the range from 30000 to 65000)

```
net.ipv4.ip_local_port_range = 30000  65000
```

This setting takes effect after the OS is restarted.

After changing `/etc/sysctl.conf`, you can reflect the change instantly by executing the command below.

```
# sysctl -p
```

Avoiding insufficient ports

If a lot of servers and resources are used for EXPRESSCLUSTER, the number of temporary ports used for internal communications by EXPRESSCLUSTER may be insufficient and the servers may not work properly as the cluster server.

Adjust the range of port number and the time before a temporary port is released as needed.

Clock synchronization

In a cluster system, it is recommended to synchronize multiple server clocks regularly. Synchronize server clocks by using `ntp`.

NIC device name

Because of the `ifconfig` command specification, when the NIC device name is shortened, the length of the NIC device name which EXPRESSCLUSTER can handle depends on it.

Shared disk

- When you continue using the data on the shared disk at times such as server reinstallation, do not allocate a partition or create a file system.
- The data on the shared disk gets deleted if you allocate a partition or create a file system.
- EXPRESSCLUSTER controls the file systems on the shared disk. Do not include the file systems on the shared disk to `/etc/fstab` in operating system.
  (If the entry to is required /etc/fstab, please use the noauto option is not used ignore option.)
- See the *Installation and Configuration Guide* for steps for shared disk configuration.

Mirror disk

- Set a management partition for mirror disk resource (cluster partition) and a partition for mirror disk resource (data partition).
- EXPRESSCLUSTER controls the file systems on mirror disks. Do not set the file systems on the mirror disks to `/etc/fstab` in operating system.
  (Do not enter a mirror partition device, mirror mount point, cluster partition, or data partition in `/etc/fstab` of the operating system.)
  (Do not enter `/etc/fstab` even with the `ignore` option specified.
  If you enter `/etc/fstab` with the `ignore` option specified, the entry will be ignored when `mount` is executed, but an error may subsequently occur when `fsck` is executed.)
  (Entering `/etc/fstab` with the `noauto` option specified is not recommended, either, because it may lead to an inadvertent manual mount or result in some application being mounted.)
- See the *Installation and Configuration Guide* for steps for mirror disk configuration.

Hybrid disk

- Configure the management partition (cluster partition) for hybrid disk resource and the partition used for hybrid disk resource (data partition).
- When a hybrid disk is allocated in the shared disk device, allocate the partition for the disk heart beat resource between servers sharing the shared disk device.
- EXPRESSCLUSTER controls the file systems on the hybrid disk. Do not include the file systems on the hybrid disk to `/etc/fstab` in operating system.
  (Do not enter a mirror partition device, mirror mount point, cluster partition, or data partition in `/etc/fstab` of the operating system.)
  (Do not enter `/etc/fstab` even with the `ignore` option specified.
  If you enter `/etc/fstab` with the `ignore` option specified, the entry will be ignored when `mount` is executed, but an error may subsequently occur when `fsck` is executed.)
  (Entering `/etc/fstab` with the `noauto` option specified is not recommended, either, because it may lead to an inadvertent manual mount or result in some application being mounted.)
- See the *Installation and Configuration Guide* for steps for hybrid disk configuration.

- When using this EXPRESSCLUSTER version, a file system must be manually created in a data partition used by a hybrid disk resource. For details about what to do when a file system is not created in advance, see “Settings after configuring hardware” in Chapter 1 “ Determining a system configuration” of the *Installation and Configuration Guide*.

If using ext4 with a mirror disk resource or a hybrid disk resource

- If ext4 is used as a file system with a mirror disk resource or a hybrid disk resource, and a disk used in the past is reused and configured (some unnecessary data remains in the disk), copying may take time more than the disk usage amount when full mirror recovery (copying between mirror disk servers) is performed.

  To avoid this, initialize the data partition with the `mkfs` command with the following options specified beforehand, before configuring a cluster (after allocating the data partition for the mirror disk resource or hybrid disk resource).

  When OS is RHEL7, Asianux Server 7 or Ubuntu:
  `mkfs -t ext4 -O -64bit,-uninit_bg {data_partition_device_name}`

  When OS is besides RHEL7, Asianux Server 7 and Ubuntu(RHEL6, etc...):
  `mkfs -t ext4 -O -uninit_bg {data_partition_device_name}`

  Operation above-mentioned is needed in case of the following one of conditions.

  - When [Execute initial mkfs] is off on setting of mirror disk resources.
  - When hybrid disk resource will be used.

- When using ext4 as a file system with a mirror disk resource or a hybrid disk resource, 64bit option of ext4 to correspond more than 16TB isn't being supported. Therefore, in case of Red Hat Enterprise Linux 7, Asianux Server 7 or Ubuntu, when using `mkfs` manually for a mirror disk, hybrid disk or their data partitions, please invalidate an option 64 bits.

  Further, the option designation which invalidates this is needed because an option becomes effective 64 bits by default in Red Hat Enterprise Linux 7 and Asianux Server 7. To be judged automatically by default in Ubuntu, please do invalidated option designation. In RHEL6, the option designation to invalidate is unnecessary because this option is invalidated.

  When OS is Red Hat Enterprise Linux 7, Asianux Server 7 or Ubuntu:
  `mkfs -t ext4 -O -64bit,-uninit_bg {data_partition_device_name}`

  When OS is besides Red Hat Enterprise Linux 7, Asianux Server 7 and Ubuntu(Red Hat Enterprise Linux 6, etc...):
  `mkfs -t ext4 -O -uninit_bg {data_partition_device_name}`

  Further, when an option becomes effective 64 bits in ext4, initial mirror configuration and full mirror recovery will be an error, and the following message is recorded in SYSLOG.

```
kernel: [I] <type: liscal><event: 271> NMPx FS type is EXT4
(64bit=ON, desc_size=xx).

kernel: [I] <type: liscal><event: 270> NMP1 this FS type (EXT4
with 64bit option) is not supported for high speed full copy.
```

Adjusting OS startup time

It is necessary to configure the time from power-on of each node in the cluster to the server operating system startup to be longer than the following:

- The time from power-on of the shared disks to the point they become available.
- Heartbeat timeout time

See the *Installation and Configuration Guide* for configuration steps.

Verifying the network settings

- The network used by Interconnect or Mirror disk connect is checked. It checks by all the servers in a cluster.
- See the *Installation and Configuration Guide* for configuration steps.

OpenIPMI

- The following functions use OpenIPMI.
  - Final Action at Activation Failure / Deactivation Failure
  - Monitor resource action upon failure
  - User-mode monitor
  - Shutdown monitor
  - Forcibly stopping a physical machine
  - Chassis Identify
- OpenIPMI do not come with EXPRESSCLUSTER. You need to download and install the rpm packages for OpenIPMI.
- Check whether or not your server (hardware) supports OpenIPMI in advance.
- Note that even if the machine complies with `ipmi` standard as hardware, OpenIPMI may not run if you actually try to run them.
- If you are using a software program for server monitoring provided by a server vendor, do not choose `ipmi` as a monitoring method for user-mode monitor resource and shutdown stall monitor. Because these software programs for server monitoring and OpenIPMI both use BMC (Baseboard Management Controller) on the server, a conflict occurs preventing successful monitoring.

User-mode monitor resource, shutdown monitoring (monitoring method: softdog)

- When softdog is selected as a monitoring method, use the soft dog driver.
  Make sure not to start the features that use the softdog driver except EXPRESSCLUSTER. Examples of such features are as follows:
  - Heartbeat feature that comes with OS
  - i8xx_tco driver
  - iTCO_WDT driver
  - watchdog feature and shutdown monitoring feature of systemd
- When `softdog` is selected as a monitoring method, make sure to set heartbeat that comes with OS not to start.
- When it sets `softdog` in a monitor method in SUSE LINUX 11, it is impossible to use with an i8xx_tco driver. When an i8xx_tco driver is unnecessary, make it the setting that i8xx_tco is not loaded.
- For Red Hat Enterprise Linux 6, when `softdog` is selected as a monitoring method, `softdog` cannot be used together with the iTCO_WDT driver. If the iTCO_WDT driver is not used, specify not to load iTCO_WDT.

Log collection

- The designated function of the generation of the syslog does not work by a log collection function in SUSE LINUX. The reason is because the suffixes of the syslog are different. Please change setting of rotate of the syslog as follows to use the appointment of the generation of the syslog of the log collection function.
- Please comment out “compress” and “date ext” of the `/etc/logrotate.d/syslog` file.
- When the total log size exceeds 2GB on each server, log collection may fail.

nsupdate and nslookup

- The following functions use `nsupdate` and `nslookup`.
  - Dynamic DNS resource of group resource (ddns)
  - Dynamic DNS monitor resource of monitor resource (ddnsw)
- EXPRESSCLUSTER does not include `nsupdate` and `nslookup`. Therefore, install the rmp files of `nsupdate` and `nslookup`, in addition to the EXPRESSCLUSTER installation.
- NEC does not support the items below regarding `nsupdate` and `nslookup`. Use `nsupdate` and `nslookup` at your own risk.
  - Inquiries about `nsupdate` and `nslookup`
  - Guaranteed operations of `nsupdate` and `nslookup`
  - Malfunction of `nsupdate` or `nslookup` or failure caused by such a malfunction
  - Inquiries about support of `nsupdate` and `nslookup` on each server

FTP monitor resources

- If a banner message to be registered to the FTP server or a message to be displayed at connection is long or consists of multiple lines, a monitor error may occur. When monitoring by the FTP monitor resource, do not register a banner message or connection message.

Notes on using Red Hat Enterprise Linux 7

- In mail reporting function takes advantage of the [mail] command of OS provides. Because the minimum composition is [mail] command is not installed, please execute one of the following.
  - Select the [**SMTP**] by the **Mail Method** on the **Alert Service** tab of **Cluster Properties**.
  - Installing mailx.

Notes on using Ubuntu

- To execute EXPRESSCLUSTER-related commands, execute them as the root user.
- Only a Websphere monitor resource is supported in Application Server Agent. This is because other Application Server isn't supporting Ubuntu.
- In mail reporting function takes advantage of the [mail] command of OS provides. Because the minimum composition is [mail] command is not installed, please execute one of the following.
  - Select the [**SMTP**] by the **Mail Method** on the **Alert Service** tab of **Cluster Properties**.
  - Installing mailutils.
- Information acquisition by SNMP cannot be used.

Time synchronization in the AWS environtment

AWS CLI is executed at the time of activation/deactivation/monitoring for AWS elastic ip resources, AWS virtual ip resoruces, AWS DNS resoruces, AWS elastic ip monitor resources, AWS virtual ip monitor resources, and AWS DNS monitor resources. If the date is not correctly set to an instance, AWS CLI may fail and the message saying “Failed in the AWS CLI command.” may be displayed due to the specification of AWS.
In such a case, correct the date and time of the instance by using a server such as an NTP server. For details, refer to “Setting the Time for Your Linux Instance” (https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/set-time.html)

IAM settings in the AWS environtment

This section describes the settings of IAM (Identity & Access Management) in AWS environment.

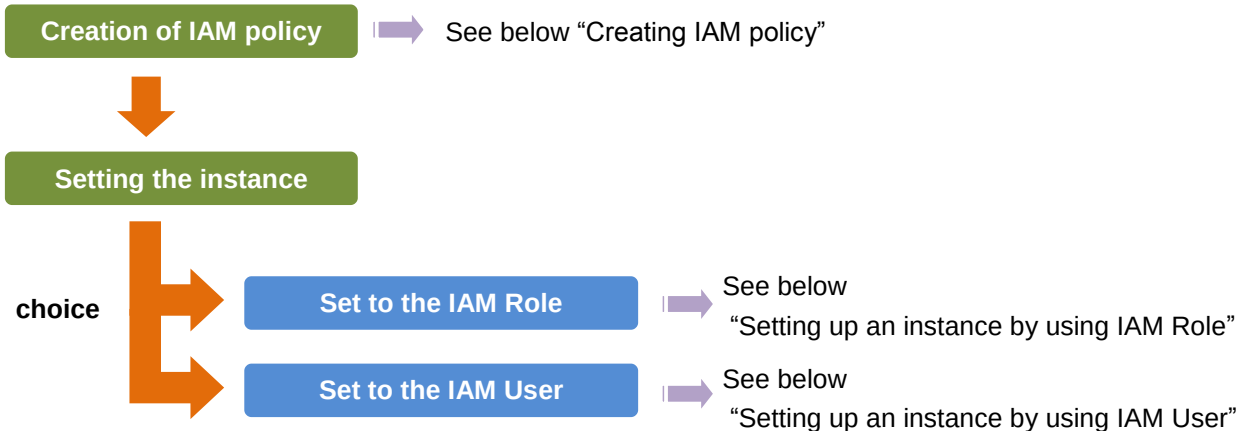
Resources and monitor resources such as AWS virtual ip resouces execute AWS CLi internally. To run AWS CLI successfully, you need to set up IAM in advance.

You can give access permissions to AWS CLI by using IAM Role or IAM User. IAM Role method offers a high-level of security because you do not have to store AWS access key ID and AWS secret access key in an instance. Therefore, it is recommended to use IAM Role basically.

Advantages and disadvantages of the two methods are as follows:

	Advantages	Disadvantages
IAM Role	- This method is more secure th using IAM user - The procedure for maintaining key information simple.	You cannot modify access permissions for each instance because IAM Role is immutable.
IAM User	You can set access permissions for each instance later.	The risk of key information leakage is high. The procedure for maintaining key information is complicated.

The procedure of setting IAM is shown below.



Creating IAM policy

Create a policy that describes access permissions for the actions to the services such as EC2 and S3 of AWS. The actions required for AWS-related resources and monitor resources to execute AWS CLI are as follows:

The necessary policies are subject to change.

◆ AWS virtual ip resource / AWS virtual ip monitor resource

Action	Description
ec2:Describe*	This is required when obtaining information of VPC, route table and network interfaces.
ec2:ReplaceRoute	This is required when updating the route table.

◆ AWS elastic ip resource /AWS elastic ip monitor resource

Action	Description
ec2:Describe*	This is required when obtaining information of EIP and network interfaces.
ec2:AssociateAddress	This is required when associating EIP with ENI.
ec2:DisassociateAddress	This is required when disassociating EIP from ENI.

◆ AWS AZ monitor resource

Action	Description
ec2:Describe*	This is required when obtaining information of the availability zone.

◆ AWS DNS resource / AWS DNS monitor resource

Action	Description
route53:ChangeResourceRecordSets	This is required when a resource record set is added or deleted or when the resource record set configuration is updated.
route53:ListResourceRecordSets	This is required when obtaining information of a resource record set.

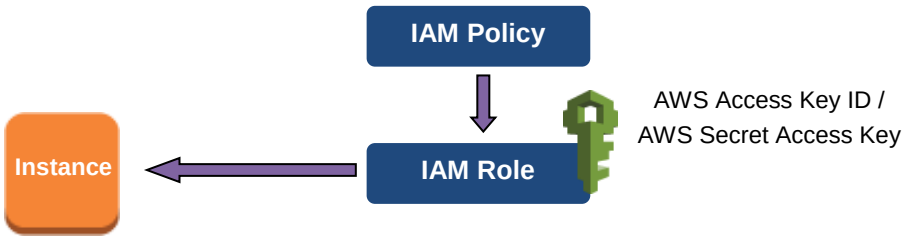
The example of a custom policy as shown below permits actions used by all the AWS-related resources and monitor resources.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "ec2:Describe*",
                "ec2:ReplaceRoute",
                "ec2:AssociateAddress",
                "ec2:DisassociateAddress" ,
                "route53:ChangeResourceRecordSets",
                "route53:ListResourceRecordSets"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

You can create a custom policy from [Policies] - [Create Policy] in IAM Management Console

Setting up an instance by using IAM Role

In this method, you can execute execute AWS CLI after creating IAM Role and associate it with an instance.



1) Create the IAM Role and attach the IAM Policy to the role.

   You can create the IAM Role from [Roles] - [Create New Role] in IAM Management Console

2) When creating an instance, specify the IAM Role you created to **IAM Role**. **(**You cannot assign the IAM Role after the instance has been created.)

3) Log on to the instance.

4) Install Python.

   Install Python required by EXPRESSCLUSTER. First, confirm that Python has been installed on the machine. If not, install it by using the command such as the yum command. The installation path of the python command must be one of the following:

   /sbin, /bin, /usr/sbin, /usr/bin

5) Execute the pip command from the shell to install AWS CLI.

```
$ pip install awscli
```

   For details about the pip command, refer to the following:

   https://pip.pypa.io/en/latest/

   For the AWS CLI installation path, select any of the following:

   /sbin, /bin, /usr/sbin, /usr/bin, /usr/local/bin

   For details on how to set up AWS CLI, refer to the following web page.

   http://docs.aws.amazon.com/cli/latest/userguide/installing.html

   (If EXPRESSCLUSTER has been installed when you install Pyhon or AWS CLI, restart OS before operating EXPRESSCLUSTER.)

6) Execute the command from the shell as shown below

```
$ sudo aws configure
```

   Input the information required to execute AWS CLI in response to the prompt. Do not input AWS access key ID and AWS secret access key.

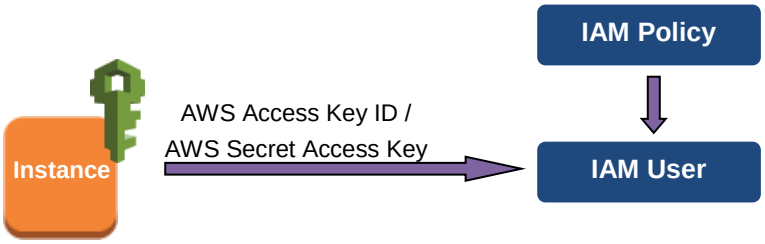
```
AWS Access Key ID [None]: (Just press Enter key)
AWS Secret Access Key [None]: (Just press Enter key)
Default region name [None]: <default region name>
Default output format [None]: text
```

   For "Default output format", other format than "text" may be specified.

   If you input wrong information, delete the entire /root/.aws directory and execute the step described above.

Setting up an instance by using IAM user

In this method, you can execute execute AWS CLI after creating the IAM User and storing its access key ID and secret access key in the instance. You do not have to assign the IAM Role to the instance when creating the instance.



1) Create the IAM User and attach the IAM Policy to the role.

   You can create the IAM user in [Users] - [Create New Users] of IAM Management Console

2) Log on to the instance.

3) Install Python.

   Install Python required by EXPRESSCLUSTER. First, confirm that Python has been installed on the machine. If not, install it by using the command such as the yum command.

   The installation path of the python command must be one of the following:

   /sbin, /bin, /usr/sbin, /usr/bin

4) Execute the pip command from the shell to install AWS CLI.

```
$ pip install awscli
```

   For details about the pip command, refer to the following:

   https://pip.pypa.io/en/latest/

   For the AWS CLI installation path, select any of the following:

   /sbin, /bin, /usr/sbin, /usr/bin, /usr/local/bin

   For details on how to set up AWS CLI, refer to the following web page.

   http://docs.aws.amazon.com/cli/latest/userguide/installing.html

   (If EXPRESSCLUSTER has been installed when you install Pyhon or AWS CLI, restart OS before operating EXPRESSCLUSTER.)

5) Execute the command from the shell as shown below

```
$ sudo aws configure
```

   Input the information required to execute AWS CLI in response to the prompt. Obtain AWS access key ID and AWS secret access key from IAM user detail screen to input.

```
AWS Access Key ID [None]: <AWS access key>
AWS Secret Access Key [None]: <AWS secret access key>
Default region name [None]: <default region name >
Default output format [None]: text
```

   If you input wrong information, delete the entire /root/.aws directory and execute the step described above.

Azure probe port resources

- In the setting of the Microsoft Azure load balancer, do not assign a load balancing rule for TCP and UDP to users of one health probe mixedly.

  If both rules are required, prepare two health probes for different ports and assign them so that the load balancing rules for TCP and UDP are different.

  In that case, prepare two Azure probe port resources and specify the port of each health probe as **Probeport**.

Azure DNS resources

- For the procedures to install Azure CLI and create a service principal, refer to the *EXPRESSCLUSTER X 4.0 HA Cluster Configuration Guide for Microsoft Azure (Linux)*.
- The Azure CLI and Python must be installed because the Azure DNS resource uses them. Python is supplied with an OS such as Red Hat Enterprise Linux and Cent OS. For details about the Azure CLI, refer to the following website:

  Microsoft Azure document:

  `https://docs.microsoft.com/en-us/azure/`
- The Azure DNS service must be installed because the Azure DNS resource uses it. For details about Azure DNS, refer to the following website:

  Azure DNS: `https://azure.microsoft.com/en-us/services/dns/`
- To set up EXPRESSCLUSTER to work with Microsoft Azure, a Microsoft Azure organizational account is required. An account other than the organizational account cannot be used because an interactive login is required when executing the Azure CLI.
- It is necessary to create a service principal with Azure CLI.

  The Azure DNS resource logs in to Microsoft Azure and performs the DNS zone registration. The Azure DNS resource uses Azure login based on service principal when logging in Microsoft Azure.
  For details about a service principal and procedure, refer to the following websites:
  Log in with Azure CLI 2.0:
  `https://docs.microsoft.com/en-us/cli/azure/authenticate-azure-cli?view=azure-cli-latest`
  Create an Azure service principal with Azure CLI 2.0:
  `https://docs.microsoft.com/en-us/cli/azure/create-an-azure-service-principal-azure-cli?view=azure-cli-latest`

  When changing the role of the created service principal from the default role “Contributor” to another role, select the role that can access all of the following operations as the Actions properties.
  If the role is changed to one that does not meet this condition, starting the Azure DNS resource fails due to an error.

  For Azure CLI 1.0:
  ```
  Microsoft.Network/dnsZones/read
  Microsoft.Network/dnsZones/A/write
  Microsoft.Network/dnsZones/A/read
  Microsoft.Network/dnsZones/A/delete
  Microsoft.Network/dnsZones/NS/read
  ```

  For Azure CLI 2.0:
  ```
  Microsoft.Network/dnsZones/A/write
  ```

```
Microsoft.Network/dnsZones/A/delete
Microsoft.Network/dnsZones/NS/read
```

Samba monitor resources

- Samba monitor resources use SMB protocol version 1.0 for monitoring. If the SMB protocol version accepted by a Samba server is limited to SMB2.0 or later (for example, when 'server min protocol' is set to 'SMB2' in smb.conf), a monitoring error will occur. Therefore, enable SMB protocol version 1.0 when using the Samba monitor resource.
- When the SMB signature is enabled in the Samba server (for example, when 'client signing' is set to 'mandatory' in smb.conf), a monitoring error will occur. Therefore, disable the SMB signature.
- The Samba monitor resource uses NTLMv1 authentication for monitoring. Therefore, a monitoring error occurs if NTLMv1 authentication is disabled on the Samba server (for example, `lanman auth = no` and `ntlm auth = no` are set in `smb.conf`). To use the Samba monitor resource, enable NTLMv1 authentication. Note that NTLMv1 authentication is disabled by default for Samba 4.5.0 or later.

Notes when creating EXPRESSCLUSTER configuration data

Notes when creating a cluster configuration data and before configuring a cluster system is described in this section.

Directories and files in the location pointed to by the EXPRESSCLUSTER installation path

The directories and files in the location pointed to by the EXPRESSCLUSTER installation path must not be handled (edited, created, added, or deleted) by using any application or tool other than EXPRESSCLUSTER.
Any effect on the operation of a directory or file caused by using an application or tool other than EXPRESSCLUSTER will be outside the scope of NEC technical support.

Environment variable

The following processes cannot be executed in an environment in which more than 255 environment variables are set. When using the following function of resource, set the number of environmental variables less than 256.

- Group start/stop process
- Start/Stop script executed by EXEC resource when activating/deactivating
- Script executed by Custom monitor Resource when monitoring
- Script before final action after the group resource or the monitor resource error is detected
- Script to be executed before and after activating or deactivating a group resource
- The script for forced stop

Note:

The total number of environment variables set in the system and EXPRESSCLUSTER must be less than 256. About 30 environment variables are set in EXPRESSCLUSTER.

Force stop function, chassis identify lamp linkage

When using forced stop function or chassis identify lamp linkage, settings of BMC IP address, user name and password of each server are necessary. Use definitely the user name to which the password is set.

Server reset, server panic and power off

When EXPRESSCLUSTER performs “Server Reset”, “Server Panic,” or “Server power off”, servers are not shut down normally. Therefore, the following may occur.

- Damage to a mounted file system
- Loss of unsaved data
- Suspension of OS dump collection

“Server reset” or “Server panic” occurs in the following settings:

- Action at an error occurred when activating/inactivating group resources
  - Sysrq Panic

- Keepalive Reset
-Keepalive Panic
- BMC Reset
- BMC Power Off
- BMC Power Cycle
- BMC NMI
- I/O Fencing(High-End Server Option)

- Final action at detection of an error in monitor resource
  - Sysrq Panic
  - Keepalive Reset
  - Keepalive Panic
  - BMC Reset
  - BMC Power Off
  - BMC Power Cycle
  - BMC NMI
  - I/O Fencing(High-End Server Option)
- Action at detection of user-mode monitor timeout
  - Monitoring method softdog
  - Monitoring method ipmi
  - Monitoring method keepalive
  - Monitoring method ipmi(High-End Server Option)

Note: "Server panic" can be set only when the monitoring method is "keepalive."

- Shutdown stall mentoring
  - Monitoring method softdog
  - Monitoring method ipmi
  - Monitoring method keepalive
  - Monitoring method ipmi(High-End Server Option)

Note: "Server panic" can be set only when the monitoring method is "keepalive."

- Operation of Forced Stop
  - BMC reset
  - BMC power off
  - BMC cycle
  - BMC NMI
  - VMware vSphere power off

Final action for group resource deactivation error

If you select **No Operation** as the final action when a deactivation error is detected, the group does not stop but remains in the deactivation error status. Make sure not to set **No Operation** in the production environment.

Verifying raw device for VxVM

Check the raw device of the volume raw device in advance:

1. Import all disk groups which can be activated on one server and activate all volumes before installing EXPRESSCLUSTER.
2. Run the command below:

```
# raw -qa
/dev/raw/raw2:  bound to major 199, minor 2
/dev/raw/raw3:  bound to major 199, minor 3
```

(A) (B)

Example: Assuming the disk group name and volume name are:

- Disk group name: dg1
- Volume name under dg1: vol1, vol2

3. Run the command below:

```
# ls -l /dev/vx/dsk/dg1/
brw-------    1 root     root      199,   2 May 15 22:13 vol1
brw-------    1 root     root      199,   3 May 15 22:13 vol2
```

(C)

4. Confirm that major and minor numbers are identical between (B) and (C).

Never use these raw devices (A) as an EXPRESSCLUSTER disk heartbeat resource, raw resource, raw monitor resource, disk resource for which the disk type is not VxVM, or disk monitor resource for which the monitor method is not READ(VxVM).

Selecting mirror disk file system

Following is the currently supported file systems:

- ext3
- ext4
- xfs
- reiserfs
- jfs
- vxfs

ext4 operations are not proved for operating systems other than Red Hat Enterprise Linux 6.

Selecting hybrid disk file system

The following are the currently supported file systems:

- ext3
- ext4
- reiserfs

Setting of mirror or hybrid disk resource action

In a system that uses mirror or hybrid disks, do not set the monitoring resources final action to **Stop the cluster service**.

If only the cluster service is stopped while the mirror agent is active, hybrid disk control or collecting mirror disk status may fail.

Time to start a single serve when many mirror disks are defined.

If many mirror disk resources are defined and a short time is set to **Time to wait for the other servers to start up**, it may take time to start a mirror agent and mirror disk resources and monitor resources related to mirror disks may not start properly when a single server is started.

If such an event occurs when starting a single server, change the value set to the time to wait for synchronization to a large value (by selecting **Cluster Properties** - **Timeout** tab - **Server Sync Wait Time**).

RAW monitoring of disk monitor resources

- When raw monitoring of disk monitor resources is set up, partitions cannot be monitored if they have been or will possibly be mounted. These partitions cannot be monitored even if you set device name to “whole device” (device indicating the entire disks).
- Allocate a partition dedicated to monitoring and set up the partition to use the raw monitoring of disk monitor resources.

Delay warning rate

If the delay warning rate is set to 0 or 100, the following can be achieved:

- When 0 is set to the delay monitoring rate

  An alert for the delay warning is issued at every monitoring.
  By using this feature, you can calculate the polling time for the monitor resource at the time the server is heavily loaded, which will allow you to determine the time for monitoring time-out of a monitor resource.

- When 100 is set to the delay monitoring rate

  The delay warning will not be issued.

Be sure not to set a low value, such as 0%, except for a test operation.

Disk monitor resource (monitoring method TUR)

- You cannot use the TUR methods on a disk or disk interface (HBA) that does not support the Test Unit Ready (TUR) and `SG_IO` commands of SCSI. Even if your hardware supports these commands, consult the driver specifications because the driver may not support them.
- S-ATA disk interface may be recognized as IDE disk interface (hd) or SCSI disk interface (sd) by OS depending on disk controller type and distribution. When it is recognized as IDE interface, all TUR methods cannot be used. If it is recognized as SCSI disk interface, TUR (legacy) can be used. Note that TUR (generic) cannot be used.
- TUR methods burdens OS and disk load less compared to Read methods.
- In some cases, TUR methods may not be able to detect errors in I/O to the actual media.

WebManager reload interval

- Do not set the “Reload Interval” in the WebManager tab for less than 30 seconds.

LAN heartbeat settings

- As a minimum, you need to set either the LAN heartbeat resource or kernel mode LAN heartbeat resource.
- You need to set at least one LAN heartbeat resource. It is recommended to set two or more LAN heartbeat resources.
- It is recommended to set both LAN heartbeat resource and kernel mode LAN heartbeat resource together.

Kernel mode LAN heartbeat resource settings

- As a minimum, you need to set either the LAN heartbeat resource or kernel mode LAN heartbeat resource.
- It is recommended to use kernel mode LAN heartbeat resource for distribution kernel of which kernel mode LAN heartbeat can be used.

COM heartbeat resource settings

- It is recommended to use a COM heartbeat resource if your environments allows. This is because using COM heartbeat resource prevents activating both systems when the network is disconnected.

BMC heartbeat settings

- The hardware and firmware of the BMC must support BMC heartbeat. For available BMCs, see Chapter 3, "Servers supporting NX7700x series linkage" and "Servers supporting Express5800/A1080a and Express5800/A1040a series linkage" in the *Getting Started Guide*.

BMC monitor resource settings

- The hardware and firmware of the BMC must support BMC heartbeat. For available BMCs, see Chapter 3, "Servers supporting NX7700x series linkage" in the *Getting Started Guide*.

IP address for Integrated WebManager settings

- **Public LAN IP address** setting, EXPRESSCLUSTER X2.1 or before, is available in the Builder at **IP address for Integrated WebManager** which is on the **WebManager** tab of **Cluster Properties**.

Double-byte character set that can be used in script comments

- Scripts edited in Linux environment are dealt as EUC code, and scripts edited in Windows environment are dealt as Shift-JIS code. In case that other character codes are used, character corruption may occur depending on environment.

Failover exclusive attribute of virtual machine group

- The group set to a virtual machine group must not be set to the exclusive rule.

System monitor resource settings

- Pattern of detection by resource monitoring
  The System Resource Agent detects by using thresholds and monitoring duration time as parameters.
  The System Resource Agent collects the data (number of opened files, number of user processes, number of threads, used size of memory, CPU usage rate, and used size of virtual memory) on individual system resources continuously, and detects errors when data keeps exceeding a threshold for a certain time (specified as the duration time).

Message receive monitor resource settings

- Error notification to message receive monitor resources can be done in any of three ways: using the `clprexec` command, BMC linkage, or linkage with the server management infrastructure.
- To use the `clprexec` command, use the relevant file stored on the EXPRESSCLUSTER CD. Use this method according to the OS and architecture of the notification-source server. The notification-source server must be able to communicate with the notification-destination server.
- To use BMC linkage, the BMC hardware and firmware must support the linkage function. This method requires communication between the IP address for management of the BMC and the IP address of the OS.
- For the linkage with the server management infrastructure, see Chapter 9, "Linkage with Server Management Infrastructure" in the *Reference Guide*.

JVM monitor resource settings

- When the monitoring target is the WebLogic Server, the maximum values of the following JVM monitor resource settings may be limited due to the system environment (including the amount of installed memory):
  - **The number** under **Monitor the requests in Work Manager**
  - **Average** under **Monitor the requests in Work Manager**
  - **The number** of **Waiting Requests** under **Monitor the requests in Thread Pool**
  - **Average** of **Waiting Requests** under **Monitor the requests in Thread Pool**
  - **The number** of **Executing Requests** under **Monitor the requests in Thread Pool**
  - **Average** of **Executing Requests** under **Monitor the requests in Thread Pool**
- When the monitoring-target is a 64-bit JRockit JVM, the following parameters cannot be monitored because the maximum amount of memory acquired from the JRockit JVM is a negative value that disables the calculation of the memory usage rate:
  - **Total Usage** under **Monitor Heap Memory Rate**
  - **Nursery Space** under **Monitor Heap Memory Rate**
  - **Old Space** under **Monitor Heap Memory Rate**
  - **Total Usage** under **Monitor Non-Heap Memory Rate**
  - **Class Memory** under **Monitor Non-Heap Memory Rate**
- To use the JVM monitor resources, install the Java runtime environment (JRE) described in "Operation environment for JVM monitor" in Chapter 3, "Installation requirements for EXPRESSCLUSTER" You can use either the same JRE as that used by the monitoring target (WebLogic Server or WebOTX) or a different JRE.
- The monitor resource name must not include a blank.
- **Command**, which is intended to execute a command for a specific failure cause upon error detection, cannot be used together with the load balancer linkage function.

EXPRESSCLUSTER startup when using volume manager resources

- When EXPRESSCLUSTER starts up, the system startup may take some time because of the deactivation processing performed by the `vgchange` command if the volume manager is lvm or the deport processing if it is vxvm. If this presents a problem, edit the startup or stop script of the EXPRESSCLUSTER main body as shown below.

For an init.d environment, edit `/etc/init.d/clusterpro` as shown below.

```
#!/bin/sh
#
# Startup script for the EXPRESSCLUSTER daemon
#
                                        :
                                        :
# See how we were called.
case "$1" in
  start)
                                        :
                                        :
        # export all volmgr resource
#           clp_logwrite "$1" "clpvolmgrc start." init_main
#           ./clpvolmgrc -d > /dev/null 2>&1
#           retvolmgrc=$?
#           clp_logwrite "$1" "clpvolmgrc end.("$retvolmgrc")" init_main
                                        :
                                        :
```

For a systemd environment, edit `/opt/nec/clusterpro/etc/systemd/clusterpro.sh` as shown below.

```
#!/bin/sh
#
# Startup script for the EXPRESSCLUSTER daemon
#
                                        :
                                        :
# See how we were called.
case "$1" in
  start)
                                        :
                                        :
        # export all volmgr resource
#           clp_logwrite "$1" "clpvolmgrc start." init_main
#           ./clpvolmgrc -d > /dev/null 2>&1
#           retvolmgrc=$?
#           clp_logwrite "$1" "clpvolmgrc end.("$retvolmgrc")" init_main
                                        :
                                        :
```

Setting up AWS elastic ip resources

- Only a data mirror configuration is possible. A shared disk configuration and a hybrid configuration are not supported.
- IPv6 is not supported.
- In the AWS environment, floating IP resources, floating IP monitor resources, virtual IP resources, and virtual IP monitor resources cannot be used.
- Only ASCII characters is supported. Check that the character besides ASCII character isn't included in an execution result of the following command.
  `aws ec2 describe-addresses --allocation-ids <EIP ALLOCATION ID>`

Setting up AWS virtual ip resources

- Only a data mirror configuration is possible. A shared disk configuration and a hybrid configuration are not supported.
- IPv6 is not supported.
- In the AWS environment, floating IP resources, floating IP monitor resources, virtual IP resources, and virtual IP monitor resources cannot be used.
- Only ASCII characters is supported. Check that the character besides ASCII character isn't included in an execution result of the following command.
  `aws ec2 describe-vpcs --vpc-ids <VPC ID>`
  `aws ec2 describe-route-tables --filters Name=vpc-id,Values=<VPC ID>`
  `aws ec2 describe-network-interfaces --network-interface-ids <ENI ID>`
- AWS virtual IP resources cannot be used if access via a VPC peering connection is necessary. This is because it is assumed that an IP address to be used as a VIP is out of the VPC range and such an IP address is considered invalid in a VPC peering connection. If access via a VPC peering connection is necessary, use the AWS DNS resource that use Amazon Route 53.

Setting up AWS DNS resources

- Only a data mirror configuration is possible. A shared disk configuration and a hybrid disk configuration are not supported.
- IPv6 is not supported.
- In the AWS environment, floating IP resources, floating IP monitor resources, virtual IP resources, and virtual IP monitor resources cannot be used.

Setting up AWS DNS monitor resources

- The AWS DNS monitor resource runs the AWS CLI for monitoring. The AWS DNS monitor resource uses **AWS CLI timeout** set to the AWS DNS resource as the timeout of the AWS CLI execution.
- Immediately after the AWS DNS resource is activated, monitoring by the AWS DNS monitor resource may fail due to the following events. If monitoring failed, set **Wait Time to Start Monitoring** of the AWS DNS monitor resource longer than the time to reflect the changed DNS setting of Amazon Route 53 (https://aws.amazon.com/route53/faqs/).
  1. When the AWS DNS resource is activated, a resource record set is added or updated.
  2. If the AWS DNS monitor resource starts monitoring before the changed DNS setting

of Amazon Route 53 is applied, name resolution cannot be done and monitoring fails.
The AWS DNS monitor resource will continue to fail monitoring while a DNS resolver cache is enabled.

3. The changed DSN setting of Amazon Route 53 is applied.
4. Name resolution succeeds after the **TTL** valid period of the AWS DNS resource elapses. Then, the AWS DNS monitor resource succeeds monitoring.

Setting up Azure probe port resources

- Only a 2-node configuration is supported.
- Only a data mirror configuration is possible. A shared disk configuration and a hybrid configuration are not supported.
- IPv6 is not supported.
- In the Microsoft Azure environment, floating IP resources, floating IP monitor resources, virtual IP resources, and virtual IP monitor resources cannot be used.

Setting up Azure load balance monitor resources

- When a Azure load balance monitor resource error is detected, there is a possibility that switching of the active server and the stand-by server from Azure load balancer is not performed correctly. Therefore, in the Final Action of Azure load balance monitor resources and the recommended that you select Stop the cluster service and shutdown OS.

Setting up Azure DNS resources

- Only a data mirror configuration is possible. A shared disk configuration and a hybrid configuration are not supported.
- IPv6 is not supported.
- In the Microsoft Azure environment, floating IP resources, floating IP monitor resources, virtual IP resources, and virtual IP monitor resources cannot be used.

Notes on using an iSCSI device as a cluster resource

In an environment in which it takes some time for an iSCSI device to become available after an iSCSI service is started, a cluster may start before an iSCSI device becomes available.
In this case, add `sleep` to the startup or stop script of the mirror agent as follows.
For an `init.d` environment, add the following change. This change is not necessary for a `systemd` environment.

Example: When it takes 30 seconds until an iSCSI device becomes available after an iSCSI service is started

Add `sleep 30` to `/etc/init.d/clusterpro_md`.

```
                    :
                    :
case "$1" in
  start)
      sleep 30
      clp_filedel "$1" init_md
                    :
                    :
```

After starting operating EXPRESSCLUSTER

Notes on situations you may encounter after start operating EXPRESSCLUSTER are described in this section.

Error message in the load of the mirror driver in an environment such as udev

In the load of the mirror driver in an environment such as udev, logs like the following may be recorded into the message file:

```
kernel: [I] <type: liscal><event: 141> NMP1 device does not exist.
(liscal_make_request)
kernel: [I] <type: liscal><event: 141> - This message can be
recorded on udev environment when liscal is initializing NMPx.
kernel: [I] <type: liscal><event: 141> - Ignore this and following
messages 'Buffer I/O error on device NMPx' on udev environment.
kernel: Buffer I/O error on device NMP1, logical block 0
```

```
kernel: <liscal liscal_make_request> NMP1 device does not exist.
kernel: Buffer I/O error on device NMP1, logical block 112
```

This phenomenon is not abnormal.

When you want to prevent the output of the error message in the udev environment, add the following file in `/etc/udev/rules.d`.

Note, however, that error messages may be output even if the rule files are added in Red Hat Enterprise Linux 7 or Asianux Server 7.

filename: `50-liscal-udev.rules`

```
ACTION=="add", DEVPATH=="/block/NMP*", OPTIONS+="ignore_device"
ACTION=="add", DEVPATH=="/devices/virtual/block/NMP*", OPTIONS+="ignore_device"
```

Buffer I/O error log for the mirror partition device

If the mirror partition device is accessed when a mirror disk resource or hybrid disk resource is inactive, log messages such as the ones shown below are recorded in the messages file.

```
kernel: [W] <type: liscal><event: 144> NMPx I/O port has been
closed, mount(0), io(0). (PID=xxxxx)
kernel: [I] <type: liscal><event: 144> - This message can be
recorded on hotplug service starting when NMPx is not active.
kernel: [I] <type: liscal><event: 144> - This message can be
recorded by fsck command when NMPx becomes active.
kernel: [I] <type: liscal><event: 144> - Ignore this and following
messages 'Buffer I/O error on device NMPx' on such environment.
:
kernel: Buffer I/O error on device /dev/NMPx, logical block xxxx
kernel: [W] <type: liscal><event: 144> NMPx I/O port has been
closed, mount(0), io(0). (PID=xxxx)
:
kernel: [W] <type: liscal><event: 144> NMPx I/O port has been
closed, mount(0), io(0). (PID=xxxx)
```

```
kernel: <liscal liscal_make_request> NMPx I/O port is close,
mount(0), io(0).
kernel: Buffer I/O error on device /dev/NMPx, logical block xxxx
```

(Where *x* and *xxxx* each represent a given number.)

The possible causes of this phenomenon are described below.
(In the case of a hybrid disk resource, the term "mirror disk resource" should be replaced with "hybrid disk resource" hereinafter.)

- When the udev environment is responsible
  - In this case, when the mirror driver is loaded, the message "`kernel: Buffer I/O error on device /dev/NMPx, logical block xxxx`" is recorded together with the message "`kernel: [I] <type: liscal><event: 141>`".
  - These messages do not indicate any error and have no impact on the operation of EXPRESSCLUSTER.
  - For details, see "Error message in the load of the mirror driver in an environment such as udev" in this chapter.
- When an information collection command (`sosreport`, `sysreport`, `blkid`, etc.) of the operating system has been executed
  - In this case, these messages do not indicate any error and have no impact on the operation of EXPRESSCLUSTER.
  - When an information collection command provided by the operating system is executed, the devices recognized by the operating system are accessed. When this occurs, the inactive mirror disk is also accessed, resulting in the above messages being recorded.

- There is no way of suppressing these messages by using the settings of EXPRESSCLUSTER or other means.

◆ When the unmount of the mirror disk has timed out

- In this case, these messages are recorded together with the message that indicates that the unmount of the mirror disk resource has timed out.
- EXPRESSCLUSTER performs the “recovery operation for the detected deactivation error” of the mirror disk resource. It is also possible that there is inconsistency in the file system.
- For details, see “Cache swell by a massive I/O ” in this chapter.

◆ When the mirror partition device may be left mounted while the mirror disk is inactive

- In this case, the above messages are recorded after the following actions are taken.

  (1) After the mirror disk resource is activated, the user or an application (for example, NFS) specifies an additional mount in the mirror partition device (`/dev/NMPx`) or the mount point of the mirror disk resource.

  (2) Then, the mirror disk resource is deactivated without unmounting the mount point added in (1).

- While the operation of EXPRESSCLUSTER is not affected, it is possible that there is inconsistency in the file system.
- For details, see “When multiple mounts are specified for a resource like a mirror disk resource ” in this chapter.

◆ When multiple mirror disk resources are configured

- With some distributions, when two or more mirror disk resources are configured, the above messages may be output due to the behavior of fsck if the resources are active.
- For details, see “Messages written to syslog when multiple mirror disk resources or hybrid disk resources are used.”

◆ When the mirror disk resource is accessed by a certain application

- Besides the above cases, it is possible that a certain application has attempted to access the inactive mirror disk resource.
- When the mirror disk resource is not active, the operation of EXPRESSCLUSTER is not affected.

Cache swell by a massive I/O

- In case that a massive amount of write over the disk capability to the mirror disk resource or the hybrid disk resource are executed, even though the mirror connection is alive, the control from write may not return or memory allocation failure may occur.

  In case that a massive amount of I/O requests over transaction performance exist, and then the file system ensure a massive amount of cache and the cache or the memory for the user space (HIGHMEM zone) are insufficient, the memory for the kernel space (NORMAL zone) may be used.

  Change the settings so that the parameter will be changed at OS startup by using `sysctl` or other commands.

  ```
  /proc/sys/vm/lowmem_reserve_ratio
  ```

- In case that a massive amount of accesses to the mirror disk resource or the hybrid disk resource are executed, it may take much time before the cache of the file systems is written out to the disks when unmounting at disk resource deactivation.
  If, at this moment, the unmounting times out before the writing from the file system to the disks is completed, I/O error messages or unmount failure messages like those shown below may be recorded.

  In this case, change the unmount timeout length for the disk resource in question to an adequate value such that the writing to the disk will be normally completed.

Example 1:

```
expresscls: [I] <type: rc><event: 40> Stopping mdx resource has
started.

kernel: [I] <type: liscal><event: 193> NMPx close I/O   port OK.

kernel: [I] <type: liscal><event: 195> NMPx close mount port OK.

kernel: [I] <type: liscal><event: 144> NMPx I/O port has been closed,
mount(0), io(0).

kernel: [I] <type: liscal><event: 144> - This message can be recorded
on hotplug service starting when NMPx is not active.

kernel: [I] <type: liscal><event: 144> - This message can be recorded
by fsck command when NMPx becomes active.

kernel: [I] <type: liscal><event: 144> - Ignore this and following
messages 'Buffer I/O error on device NMPx' on such environment.

kernel: Buffer I/O error on device NMPx, logical block xxxx

kernel: [I] <type: liscal><event: 144> NMPx I/O port has been closed,
mount(0), io(0).

kernel: Buffer I/O error on device NMPx, logical block xxxx

:
```

Example 2:

```
expresscls: [I] <type: rc><event: 40> Stopping mdx resource has
started.

kernel: [I] <type: liscal><event: 148> NMPx holder 1. (before umount)

expresscls: [E] <type: md><event: 46> umount timeout. Make sure that
the length of Unmount Timeout is appropriate. (Device:mdx)

:

expresscls: [E] <type: md><event: 4> Failed to deactivate mirror
disk. Umount operation failed.(Device:mdx)

kernel: [I] <type: liscal><event: 148> NMPx holder 1. (after umount)

expresscls: [E] <type: rc><event: 42> Stopping mdx resource has
failed.(83 : System command timeout (umount, timeout=xxx))

:
```

When multiple mounts are specified for a resource like a mirror disk resource

- If, after activation of a mirror disk resource or hybrid disk resource, you have created an additional mount point in a different location by using the mount command for the mirror partition device (/dev/NMP*x*) or the mount point (or a part of the file hierarchy for the mount point), you must unmount that additional mount point before the disk resource is deactivated. If the deactivation is performed without the additional mount point being unmounted, the file system data remaining in memory may not be completely written out to the disks. As a result, the I/O to the disks is closed and the deactivation is completed although the data on the disks are incomplete.
  Because the file system will still try to continue writing to the disks even after the deactivation is completed, I/O error messages like those shown below may be recorded.
  After this, an attempt to stop the mirror agent, such as when stopping the server, will fail, since the mirror driver cannot be terminated. This may cause the server to restart.

Example:

```
expresscls: [I] <type: rc><event: 40> Stopping mdx resource has
started.

kernel: [I] <type: liscal><event: 148> NMP1 holder 1. (before umount)

kernel: [I] <type: liscal><event: 148> NMP1 holder 1. (after umount)

kernel: [I] <type: liscal><event: 193> NMPx close I/O   port OK.

kernel: [I] <type: liscal><event: 195> NMPx close mount port OK.

expresscls: [I] <type: rc><event: 41> Stopping mdx resource has
completed.


kernel: [I] <type: liscal><event: 144> NMPx I/O port has been closed,
mount(0), io(0).

kernel: [I] <type: liscal><event: 144> - This message can be recorded
on hotplug service starting when NMPx is not active.

kernel: [I] <type: liscal><event: 144> - This message can be recorded
by fsck command when NMPx becomes active.
```

```
kernel: [I] <type: liscal><event: 144> - Ignore this and following
messages 'Buffer I/O error on device NMPx' on such environment.
kernel: Buffer I/O error on device NMPx, logical block xxxxx
kernel: lost page write due to I/O error on NMPx
kernel: [I] <type: liscal><event: 144> NMPx I/O port has been closed,
mount(0), io(0).
kernel: Buffer I/O error on device NMPx, logical block xxxxx
kernel: lost page write due to I/O error on NMPx
:
```

Messages written to syslog when multiple mirror disk resources or hybrid disk resources are used

When more than two mirror disk resources or hybrid disk resources are configured on a cluster, the following messages may be written to the OS message files when the resources are activated.

This phenomenon may occur due to the behavior of the `fsck` command of some distributions (`fsck` accesses an unintended block device).

```
kernel: [I] <type: liscal><event: 144> NMPx I/O port has been
closed, mount(0), io(0).
kernel: [I] <type: liscal><event: 144> - This message can be
recorded by fsck command when NMPx becomes active.
kernel: [I] <type: liscal><event: 144> - This message can be
recorded on hotplug service starting when NMPx is not active.
kernel: [I] <type: liscal><event: 144> - Ignore this and following
messages 'Buffer I/O error on device NMPx' on such environment.
kernel: Buffer I/O error on device /dev/NMPx, logical block xxxx
```

```
kernel: <liscal liscal_make_request> NMPx I/O port is close,
mount(0), io(0).
kernel: Buffer I/O error on device /dev/NMPx, logical block xxxx
```

This is not a problem for EXPRESSCLUSTER. If this causes any problem such as heavy use of message files, change the following settings of mirror disk resources or hybrid disk resources.

- Select “Not Execute” on “fsck action before mount”

- Select “Execute” on “fsck Action When Mount Failed”

Messages displayed when loading a driver

When loading a mirror driver, messages like the following may be displayed at the console and/or syslog. However, this is not an error.

```
kernel: liscal: no version for "xxxxx" found: kernel tainted.
kernel: liscal: module license 'unspecified' taints kernel.
```

(Any character strings are set to *xxxxx*.)

And also, when loading the `clpka` or `clpkhb` driver, messages like the following may be displayed on the console and/or syslog. However, this is not an error.

```
kernel: clpkhb: no version for "xxxxx" found: kernel tainted.
kernel: clpkhb: module license 'unspecified' taints kernel.
```

```
kernel: clpka: no version for "xxxxx" found: kernel tainted.
kernel: clpka: module license 'unspecified' taints kernel.
```

(Any character strings are input into *xxxxx*.)

Messages displayed for the first I/O to mirror disk resources or hybrid disk resources

When reading/writing data from/to a mirror disk resource or hybrid disk resource for the first time after the resource was mounted, a message like the following may be displayed at the console and/or syslog. However, this is not an error.

```
kernel: JBD: barrier-based sync failed on NMPx - disabling barriers
```

(Any character strings are set to *x*.)

File operating utility on X-Window

Some of the file operating utilities (coping and moving files and directories via GUI) on X-Window perform the following:

- Checks if the block device is usable.
- Mounts the file system if there is any that can be mounted.

Make sure not to use file operating utility that perform above operations. They may cause problem to the operation of EXPRESSCLUSTER.

IPMI message

When you are using ipmi for user-mode monitor resources, the following kernel module warning log is recorded many times in the syslog.

```
modprobe: modprobe: Can`t locate module char-major-10-173
```

When you want to prevent this log from being recorded, rename /dev/ipmikcs.

Limitations during the recovery operation

Do not control the following commands, clusters and groups by the WebManager while recovery processing is changing (reactivation → failover → last operation), if a group resource is specified as a recovery target and when a monitor resource detects an error.

- Stop and suspend of a cluster
- Start, stop, moving of a group

If these operations are controlled at the transition to recovering due to an error detected by a monitor resource, the other group resources in the group may not be stopped.

Even if a monitor resource detects an error, it is possible to control the operations above after the last operation is performed.

Executable format file and script file not described in manuals

Executable format files and script files which are not described in Chapter 4, ”EXPRESSCLUSTER command reference” in the *Reference Guide* exist under the installation directory. Do not run these files on any system other than EXPRESSCLUSTER. The consequences of running these files will not be supported.

Executing fsck

- When `fsck` is specified to execute at activation of disk resources, mirror disk resources, or hybrid disk resources, `fsck` is executed when an ext2/ext3/ext4 file system is mounted. Executing `fsck` may take times depending on the size, usage or status of the file system, resulting that an `fsck` timeout occurs and mounting the file system fails.

  This is because `fsck` is executed in either of the following ways.

  (a) Only performing simplified journal check.
  Executing `fsck` does not take times.
  (b) Checking consistency of the entire file system.
  When the data saved by OS has not been checked for 180 days or more or the data will be checked after it is mounted around 30 times.
  In this case, executing `fsck` takes times depending the size or usage of the file system.

  Specify a time in safe for the `fsck` timeout of disk resources so that no timeout occurs.

- When `fsck` is specified not to execute at activation of disk resources, mirror disk resources, or hybrid disk resources, the warning described below may be displayed on the console and/or syslog when an ext2/ext3/ext4 file system is mounted more than the mount execution count set to OS that it is recommended to execute `fsck`.

```
EXT2-fs warning: xxxxx, running e2fsck is recommended.
```

Note: There are multiple patterns displayed in *xxxxx*.

It is recommended to execute `fsck` when this waning is displayed.

Follow the steps below to manually execute `fsck`.
Be sure to execute the following steps on the server where the disk resource in question has been activated.

(1) Deactivate a group to which the disk resource in question belongs by using a command such as `clpgrp`.

(2) Confirm that no disks have been mounted by using a command such as `mount` and `df`.

(3) Change the state of the disk from Read Only to Read Write by executing one of the following commands depending on the disk resource type.

Example for disk resources: A device name is `/dev/sbd5`

```
# clproset -w -d /dev/sbd5
/dev/sbd5 : success
```

Example for mirror disk resources: A resource name is `md1`.

```
# clpmdctrl --active -nomount md1
<md1@server1>: active successfully
```

Example for hybrid disk resources: A resource name is `hd1`.

```
# clphdctrl --active -nomount hd1
<hd1@server1>: active successfully
```

(4) Execute `fsck`.
(If you specify the device name for fsck execution in the case of a mirror disk resource or hybrid disk resource, specify the mirror partition device name (/dev/NMP*x*) corresponding to the resource.)

(5) Change the state of the disk from Read Write to Read Only by executing one of the following commands depending on the disk resource type.

Example for disk resources: A device name is `/dev/sbd5`.

```
# clproset -o -d /dev/sbd5
/dev/sbd5 : success
```

Example for mirror disk resources: A resource name is `md1`.

```
# clpmdctrl --deactive md1
<md1@server1>: active successfully
```

Example for hybrid disk resources: A resource name is `hd1`.

```
# clphdctrl --deactive -nomount hd1
<hd1@server1>: active successfully
```

(6) Activate a group to which the disk resource in question belongs by using a command such as `clpgrp`.

If you need to specify that the warning message is not output without executing fsck, for ext2/ext3/ext4, change the maximum mount count by using tune2fs. Be sure to execute this command on the server where the disk resource in question has been activated.

(1) Execute one of the following commands..

Example for disk resources: A device name is `/dev/sbd5`.

```
# tune2fs -c -l /dev/sbd5
tune2fs 1.42.9 (28-Dec-2013)
Setting maximal mount count to -1
```

Example for mirror disk resources: A resource name is `/dev/NMP1`.

```
# tune2fs -c -l /dev/NMP1
tune2fs 1.42.9 (28-Dec-2013)
Setting maximal mount count to -1
```

Example for hybrid disk resources: A resource name is `/dev/NMP1`.

```
# tune2fs -c -l /dev/NMP1
tune2fs 1.42.9 (28-Dec-2013)
Setting maximal mount count to -1
```

(2) Confirm that the maximum mount count has been changed.

Example: A device name is `/dev/sbd5`.

```
# tune2fs -l /dev/sbd5
tune2fs 1.42.9 (28-Dec-2013)
Filesystem volume name: <none>
     :
Maximum mount count: -1
     :
```

Messages when collecting logs

When collecting logs, the message described below is displayed at the console, but this is not an error. Logs are collected successfully.

```
hd#: bad special flag: 0x03
ip_tables: (C) 2000-2002 Netfilter core team
```

(“hd#” is replaced with the device name of IDE.)

```
kernel: Warning: /proc/ide/hd?/settings interface is obsolete, and
will be removed soon!
```

Failover and activation during mirror recovery

- When mirror recovery is in progress for a mirror disk resource or hybrid disk resource, a mirror disk resource or hybrid disk resource placed in the deactivated state cannot be activated.
  During mirror recovery, a failover group including the disk resource in question cannot be moved.
  If a failover occurs during mirror recovery, the copy destination server does not have the latest status, so a failover to the copy destination server or copy destination server group will fail.
  Even if an attempt to fail over a hybrid disk resource to a server in the same server group is made by actions for when a monitor resource detects an error, it will fail, too, since the current server is not changed.
  Note that, depending on the timing, when mirror recovery is completed during a failover, move, or activation, the operation may be successful.
- At the first mirror startup after configuration information registration and also at the first mirror startup after a mirror disk is replaced after a failure, the initial mirror configuration is performed.
  In the initial mirror configuration, disk copying (full mirror recovery) is performed from the active server to the mirror disk on the standby server immediately after mirror activation.
  Until this initial mirror configuration (full mirror recovery) is completed and the mirror enters the normal synchronization state, do not perform either failover to the standby server or group movement to the standby server.
  If a failover or group movement is performed during this disk copying, the standby server may be activated while the mirror disk of the standby server is still incomplete, causing the data that has not yet been copied to the standby server to be lost and thus causing mismatches to occur in the file system.

Cluster shutdown and reboot (mirror disk resource and hybrid disk resource)

When using a mirror disk resource or a hybrid disk resource, do not execute cluster shutdown or cluster shutdown reboot from the `clpstdn` command or the WebManager while a group is being activated.

A group cannot be deactivated while a group is being activated. Therefore, OS may be shut down in the state that mirror disk resource or hybrid disk resources is not deactivated successfully and a mirror break may occur.

Shutdown and reboot of individual server (mirror disk resource and hybrid disk resource)

When using a mirror disk and a hybrid disk resource, do not shut down the server or run the shutdown reboot command from the `clpdown` command or the WebManager while activating the group.

A group cannot be deactivated while a group is being activated. Therefore, OS may be shut down and a mirror break may occur in the state that mirror disk resources and hybrid disk resources are not deactivated successfully.

Scripts for starting/stopping EXPRESSCLUSTER services

For an init.d environment, an error occurs in the service startup and stop scripts in the following cases. For a systemd environment, an error does not occur.

- Before start operating EXPRESSCLUSTER
  When a server start up, the error occurs in the following starting scripts. There is no problem for the error because cluster configuration data has not uploaded.
  - `clusterpro_md`
- At following case, the script to terminate EXPRESSCLUSTER services may be executed in the wrong order.
  EXPRESSCLUSTER services may be terminated in the wrong order at OS shutdown if all of EXPRESSCLUSTER services are disabled. This problem is caused by failure in termination process for the service has been already disabled.
  As long as the system shutdown is executed by WebManager or `clpstdn` command, there is no problem even if the services is terminated in the wrong order. But, any other problem may not be happened by wrong order termination.

Service startup time

EXPRESSCLUSTER services might take a while to start up, depending on the wait processing at startup.

- `clusterpro_evt`
  Servers other than the master server wait up to two minutes for configuration data to be downloaded from the master server. Downloading usually finishes within several seconds if the master server is already operating. The master server does not have this wait process.
- `clusterpro_trn`
  There is no wait process. This process usually finishes within several seconds.
- `clusterpro_md`
  This service starts up only when the mirror or hybrid disk resources exist. The system waits up to one minute for the mirror agent to normally start up. This process usually finishes within several seconds.
- `clusterpro`
  Although there is no wait process, EXPRESSCLUSTER might take several tens of seconds to start up. This process usually finishes within several seconds.
- `clusterpro_webmgr`
  There is no wait process. This process usually finishes within several seconds.
- `clusterpro_alertsync`
  There is no wait process. This process usually finishes within several seconds.

In addition, the system waits for cluster activation synchronization after the EXPRESSCLUSTER daemon is started. By default, this wait time is five minutes.

For details, see Chapter 10, “The system maintenance information” in the Reference Guide.

Checking the service status in a systemd environment

For a `systemd` environment, the service status displayed by using the `systemctl` command may differ from the actual cluster status.

Use the `clpstat` command or Cluster WebUI / WebManager to check the cluster status.

Scripts in EXEC resources

EXEC resource scripts of group resources stored in the following location.

/opt/nec/clusterpro/scripts/*group-name*/*resource-name*/

The following cases, old EXEC resource scripts are not deleted automatically.

- When the EXEC resource is deleted or renamed
- When a group that belongs to the EXEC resource is deleted or renamed

Old EXEC resource scripts can be deleted when unnecessary.

Monitor resources that monitoring timing is “Active”

When monitor resources that monitoring timing is “Active” have suspended and resumed, the following restriction apply:

- In case stopping target resource after suspending monitor resource, monitor resource becomes suspended. As a result, monitoring restart cannot be executed.
- In case stopping or starting target resource after suspending monitor resource, monitoring by monitor resource starts when target resource starts.

Notes on the WebManager

- The information displayed on the WebManager does not necessarily show the latest status. If you want to get the latest information, click the **Reload** button.
- If the problems such as server shutdown occur while the WebManager is getting the information, acquiring information may fail and a part of object may not be displayed correctly. Wait for the next automatic update or click the **Reload** button to reacquire the latest information.
- When using a browser on Linux, a dialog box may be displayed behind the window managers depending on the combination of the managers. Change the window by pressing the **ALT** + **TAB** keys.
- Collecting logs of EXPRESSCLUSTER cannot be executed from two or more WebManager simultaneously.
- If the WebManager is operated in the state that it cannot communicate with the connection destination, it may take a while until the control returns.
- If you move the cursor out of the browser in the state that the mouse pointer is displayed as a wristwatch or hourglass, the cursor may be back to an arrow.
- When going through the proxy server, make the settings for the proxy server be able to relay the port number of the WebManager.
- When going through the reverse proxy server, the WebManager will not operate properly.
- When updating EXPRESSCLUSTER, close all running browsers. Clear the Java cache (not browser cache) and open browsers.
- When updating Java, close all running browsers. Clear the Java cache (not browser cache) and open browsers

Notes on the Builder (Config mode of Cluster Manager)

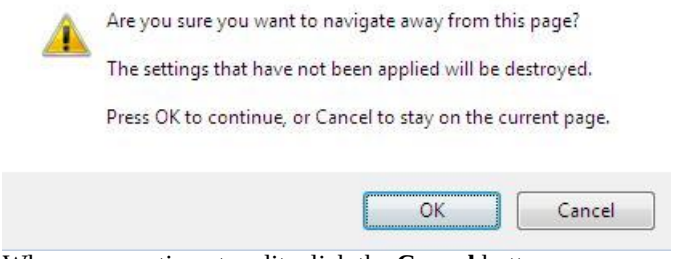
- EXPRESSCLUSTER does not have the compatibility of the cluster configuration data with the following products.

- Builder for Linux other than EXPRESSCLUSTER X 4.0 for Linux

◆ Cluster configuration data created using a later version of this product cannot be used with this product.

◆ Cluster configuration data of EXPRESSCLUSTER X1.0/2.0/2.1/3.0/3.1/3.2/3.3 /4.0 for Linux can be used with this product.
You can use such data by clicking **Import** from the **File** menu in the Builder.

◆ Closing the Web browser (by clicking **Exit** from the menu), the dialog box to confirm to save is displayed.

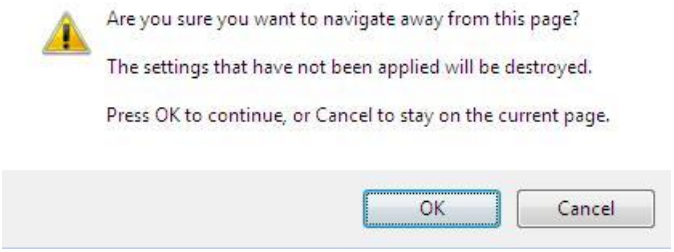


When you continue to edit, click the **Cancel** button.

Note:

This dialog box is not displayed if JavaScript is disabled.

◆ Reloading the Web browser (by selecting **Refresh** button from the menu or tool bar), the dialog box to confirm to save is displayed.



When you continue to edit, click the **Cancel** button.

Note:

This dialog box is not displayed if JavaScript is disabled.

◆ When creating the cluster configuration data using the Builder, do not enter the value starting with 0 on the text box. For example, if you want to set 10 seconds for a timeout value, enter "10" but not "010."

◆ When going through the reverse proxy server, the Builder will not operate properly.

Changing the partition size of mirror disks and hybrid disk resources

When changing the size of mirror partitions after the operation is started, see "Changing offset or size of a partition on mirror disk resource" in Chapter 10 "The system maintenance information" in the *Reference Guide*.

Changing kernel dump settings

If you are changing the kdump settings and "applying" them through "kernel dump configuration" (system-config-kdump) while the cluster is running on Red Hat Enterprise Linux 6 or the like, you may see the following error message output.

In this case, stop the cluster once (stop the mirror agent as well as the cluster when using a mirror disk resource or hybrid disk resource), and then retry the kernel dump configuration.

* The following *{driver_name}* indicates clpka, clpkhb, or liscal.

```
No module {driver_name} found for kernel {kernel_version}, aborting
```

Notes on floating IP and virtual IP resources

- Do not execute a network restart on a server on which floating IP resources or virtual IP resources are active. If the network is restarted, any IP addresses that have been added as floating IP resources or virtual IP resources are deleted.

Notes on system monitor resources

- To change a setting, the cluster must be suspended.
- System monitor resources do not support a delay warning for monitor resources.
- If the date or time setting on the OS is changed by the `data(1)` command or another method while a system monitor resource is operating, that system monitor resource may fail to operate normally.
  If you have changed the date or time setting on the OS, suspend and then resume the cluster.
- Set SELinux to either the permissive or disabled state.
  If SELinux is set to the enforcing state, the communication required for EXPRESSCLUSTER may be disabled.
- If the "system monitor" is not displayed in the **Type** field of the monitor resource definition dialog box, update the server information by selecting **Update Server Data** from the **File** menu in the Builder.
- Up to 64 disks that can be monitored by the disk resource monitor function at the same time.

Notes on JVM monitor resources

- When restarting the monitoring-target Java VM, suspend or shut down the cluster before restarting the Java VM.
- To change a setting, the cluster must be suspended.
- JVM monitor resources do not support a delay warning for monitor resources.

HTTP monitor resource

- The HTTP monitor resource uses any of the following OpenSSL shared library symbolic links:
  `libssl.so`
  `libssl.so.10` (OpenSSL 1.0 shared libraly)
  `libssl.so.6` (OpenSSL 0.9 shared libraly)

  The above symbolic links may not exist depending on the OS distribution or version, or the package installation status.
  If the above symbolic links cannot be found, the following error occurs in the HTTP monitor resource.

```
Detected an error in monitoring<Module Resource Name>. (1 :Can
not found library.  (libpath=libssl.so, errno=2))
```

For this reason, if the above error occurred, be sure to check whether the above symbolic links exit in `/usr/lib` or `/usr/lib64`.
If the above symbolic links do not exit, create the symbolic link `libssl.so`, as in the command example below.

Command example:

```
cd /usr/lib64                          # Move to /usr/lib64.
ln -s libssl.so.1.0.1e libssl.so       # Create a symbolic link.
```

Restoration from an AMI in an AWS environment

If the ENI ID of a primary network interface is set to the **ENI ID** of the AWS virtual ip resource and AWS elastic ip resource, the AWS virtual ip resource and AWS elastic ip resource setting is required to change when restoring data from an AMI.

If the ENI ID of a secondary network interface is set to the **ENI ID** of the AWS virtual ip resource and AWS elastic ip resource, it is unnecessary to set the AWS virtual ip resource and AWS elastic ip resource again because the same ENI ID is inherited by a detach/attach processing when restoring data from an AMI.

Notes when changing the EXPRESSCLUSTER configuration

The section describes what happens when the configuration is changed after starting to use EXPRESSCLUSTER in the cluster configuration.

Exclusive rule of group properties

When the exclusive attribute of the exclusive rule is changed, the change is applied by suspending and resuming the cluster.

When a group is added to the exclusion rule whose exclusive attribute is set to Absolute, multiple groups of **Absolute** exclusive start on the same server depending on the group startup status before suspending the cluster.

Exclusive control will be performed at the next group startup.

Dependency between resource properties

When the dependency between resources has been changed, the change is applied by suspending and resuming the cluster.

If a change in the dependency between resources that requires the resources to be stopped during application is made, the startup status of the resources after the resume may not reflect the changed dependency.

Dependency control will be performed at the next group startup.

Adding and deleting group resources

When you move a group resource to another group, follow the procedure shown below.

If this procedure is not followed, the cluter may not work normally.

Example) Moving fip1 (floating ip resource) from failover1 group to failover2 group

1. Delete fip1 from failover1.
2. Reflect the setting to the system.
3. Add fip1 to failover2.
4. Reflect the setting to the system.

Deleting disk resources

When a disk resource is deleted, the corresponding device is sometimes set to **Read Only**.

Change the status of the device to **Read Write** by using the clproset command.

Notes on Upgrading EXPRESSCLUSTER

This section describes the notes on upgrading EXPRESSCLUSTER from X 3.3 (internal version: 3.3.5-1) to X 4.0 after starting a cluster operation.

Management tool

The default management tool has been changed to Cluster WebUI. If you want to use the conventional WebManager as the management tool, click **WebManager** of Cluster WebUI or specify "`http://management IP address of management group` or `actual IP address:port number of the server in which` *`EXPRESSCLUSTER Server is installed`*`/main.htm`" in the address bar of a web browser.

Functions Removed in X 4.0

The following functions were used in X3.3 (internal version: 3.3.5-1), but have been removed in X 4.0.

- WebManager Mobile
- OracleAS monitor resource

Removed Parameters

Among the parameters that can be set by using the Builder, the X 3.3 (internal version: 3.3.5-1) compatible parameters in the following table have been removed in X 4.0

Cluster

Parameters	default values of the X 3.3
Cluster Properties	
Alert Service Tab	
Use Alert Extension	Off
WebManager Tab	
Enable WebManager Mobile Connection	Off
Web Manager Mobile Password	
Password for Operation	-
Password for Reference	-

JVM monitor resource

Parameters	default values of the X 3.3
JVM Monitor Resource Properties	
Monitor(special) Tab	
Memory Tab (when Oracle Java is selected for JVM Type)	
Monitor Virtual Memory Usage	2048 megabytes
Memory Tab (when Oracle JRockit is selected for JVM Type)	
Monitor Virtual Memory Usage	2048 megabytes
Memory Tab(when Oracle Java(usage monitoring) is selected for JVM Type)	
Monitor Virtual Memory Usage	2048 megabytes

Changed Default Values

Among the parameters that can be set by using the Builder, the following default values of the X 3.3 (internal version: 3.3.5-1) compatible parameters have been changed in X 4.0.

- The default value of the following parameters will be changed after upgrading EXPRESSCLUSTER from the previous version to the target version or later.
- If you want to keep using the "default values of the X 3.3", you have to change these parameters to this value after upgrading EXPRESSCLUSTER.
- If you have changed the parameters from "default values of the X 3.3", the setting values of these parameters will not be changed. Therefore you do not have to change these parameters.

Cluster

Parameters	default values of the X 3.3	default values of the X 4.0
Cluster Properties		
JVM monitor Tab		
Maximum Java Heap Size	7 megabytes	16 megabytes

Exec resource

Parameters	default values of the X 3.3	default values of the X 4.0
Exec Resource Properties		
Dependence Tab		
Follow the default dependence	On • floating IP resources • virtual IP resources • disk resources • mirror disk resources • hybrid disk resources • NAS resources • Dynamic DNS resource • Volume manager resource • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • floating IP resources • virtual IP resources • disk resources • mirror disk resources • hybrid disk resources • NAS resources • Dynamic DNS resource • Volume manager resource • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource

Disk resource

Parameters	default values of the X 3.3	default values of the X 4.0
Disk Resource Properties		
Dependence Tab		
Follow the default dependence	On • floating IP resources • virtual IP resources • Dynamic DNS resource • Volume manager resource • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • floating IP resources • virtual IP resources • Dynamic DNS resource • Volume manager resource • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource
Details Tab		
Disk Resource Tuning Properties		
Mount Tab		

Timeout	60 seconds	180 seconds
xfs_repair Tab (when xfs is selected for File System)		
xfs_repair Action When Mount Failed Execute	On	Off

NAS resource

Parameters	default values of the X 3.3	default values of the X 4.0
NAS Resource Properties		
Dependence Tab		
Follow the default dependence	On • floating IP resources • virtual IP resources • Dynamic DNS resources • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • floating IP resources • virtual IP resources • Dynamic DNS resources • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource

Mirror disk resource

Parameters	default values of the X 3.3	default values of the X 4.0
Mirror Disk Resource Properties		
Dependency Tab		
Follow the default dependence	On • floating IP resources • virtual IP resources • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • floating IP resources • virtual IP resources • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource
Details Tab		
Mirror Disk Resource Tuning Properties		
xfs_repair Tab (when xfs is selected for File System)		
xfs_repair Action When Mount Failed Execute	On	Off

Hybrid disk resource

Parameters	default values of the X 3.3	default values of the X 4.0
Hybrid Disk Resource Properties		
Dependency Tab		
Follow the default dependence	On • floating IP resources • virtual IP resources • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • floating IP resources • virtual IP resources • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource
Details Tab		
Hybrid Disk Resource Tuning Properties		
xfs_repair Tab (when xfs is selected for File System)		
xfs_repair Action When Mount Failed Execute	On	Off

.

Volume manager resource

Parameters	default values of the X 3.3	default values of the X 4.0
Volume Manager Resource Properties		
Dependency Tab		
Follow the default dependence	On • Floating IP resources • Virtual IP resources • Dynamic DNS resources • AWS elastic ip resource • AWS virtual ip resource • Azure probe port resource	On • Floating IP resources • Virtual IP resources • Dynamic DNS resources • AWS elastic ip resource • AWS virtual ip resource • AWS DNS resource • Azure probe port resource • Azure DNS resource

Virtual IP monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Virtual IP Monitor Resource Properties		
Monitor(common)		
Timeout	30 seconds	180 seconds

PID monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
PID Monitor Resource Properties		
Monitor(common)Tab		
Wait Time to Start Monitoring	0 seconds	3 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

User-mode monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
User-mode Monitor Resource Properties		
Monitor(special) Tab		
Method	softdog	keepalive

NIC Link Up/Down monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
NIC Link Up/Down Monitor Resource Properties		
Monitor(common) Tab		
Timeout	60 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

ARP monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0

ARP Monitor Resource Properties		
Monitor(common) Tab		
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

Dynamic DNS monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Dynamic DNS Monitor Resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds

Process name monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Process Monitor Resource Properties		
Monitor(common) tab		
Wait Time to Start Monitoring	0 seconds	3 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

DB2 monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
DB2 Monitor Resource Properties		
Monitor(special) Tab		
Password	ibmdb2	-
Library Path	/opt/IBM/db2/V8.2/lib/libdb2.so	/opt/ibm/db2/V11.1/lib64/libdb2.so

MySQL monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
MySQL Monitor Resource Properties		
Monitor(special) Tab		
Storage Engine	MyISAM	InnoDB
Library Path	/usr/lib/mysql/libmysqlclient.so.15	/usr/lib64/mysql/libmysqlclient.so.20

Oracle monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Oracle Monitor Resource Properties		
Monitor(special) Tab		
Password	change_on_install	-
Library Path	/opt/app/oracle/product/10.2.0/db_1/lib/libclntsh.so.10.1	/u01/app/oracle/product/12.2.0/dbhome_1/lib/libclntsh.so.12.1

PostgreSQL monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
PostgreSQL Monitor Resource Properties		

Monitor(special) Tab		
Library Path	/usr/lib/libpq.so.3.0	/opt/PostgreSQL/10/lib/libpq.so.5.10

Sybase monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Sybase Monitor Resource Properties		
Monitor(special) Tab		
Library Path	/opt/sybase/OCS-12_5/lib/libsybdb.so	/opt/sap/OCS-16_0/lib/libsybdb64.so

Tuxedo monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Tuxedo Monitor Resource Properties		
Monitor(special) Tab		
Library Path	/opt/bea/tuxedo8.1/lib/libtux.so	/home/Oracle/tuxedo/tuxedo12.1.3.0.0/lib/libtux.so

Weblogic monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Weblogic Monitor Resource Properties		
Monitor(special) Tab		
Domain Environment File	/opt/bea/weblogic81/samples/domains/examples/setExamplesEnv.sh	/home/Oracle/product/Oracle_Home/user_projects/domains/base_domain/bin/setDomainEnv.sh

JVM monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
JVM Monitor Resource Properties		
Monitor(common) Tab		
Timeout	120 seconds	180 seconds

Floating IP monitor resources

Parameters	default values of the X 3.3	default values of the X 4.0
Floating IP Monitor Resource Properties		
Monitor(common) Tab		
Timeout	60 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

AWS Elastic IP monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
AWS elastic ip Monitor Resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds

Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

AWS Virtual IP monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
AWS virtual ip Monitor Resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

AWS AZ monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
AWS AZ Monitor Resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

Azure probe port monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Azure probe port Monitor Resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

Azure load balance monitor resource

Parameters	default values of the X 3.3	default values of the X 4.0
Azure load balance monitor resource Properties		
Monitor(common) Tab		
Timeout	100 seconds	180 seconds
Do Not Retry at Timeout Occurrence	Off	On
Do not Execute Recovery Action at Timeout Occurrence	Off	On

Moved Parameters

Among the parameters that can be set by using the Builder, the locations of the following X 3.3 (internal version: 3.3.5-1) compatible parameters have been changed in X 4.0.

Parameter location in X 3.3	Parameter location in X 4.0
[Cluster Properties] - [Recovery Tab] - [Max Reboot Count]	[Cluster Properties] - [Extension Tab] - [Max Reboot Count]
[Cluster Properties] - [Recovery Tab] - [Max Reboot Count Reset Time]	[Cluster Properties] - [Extension Tab] - [Max Reboot Count Reset Time]
[Cluster Properties] - [Recovery Tab] - [Use Forced Stop]	[Cluster Properties] - [Extension Tab] - [Use Forced Stop]
[Cluster Properties] - [Recovery Tab] - [Forced Stop Action]	[Cluster Properties] - [Extension Tab] - [Forced Stop Action]
[Cluster Properties] - [Recovery Tab] - [Forced Stop Timeout]	[Cluster Properties] - [Extension Tab] - [Forced Stop Timeout]
[Cluster Properties] - [Recovery Tab] - [Virtual Machine Forced Stop Setting]	[Cluster Properties] - [Extension Tab] - [Virtual Machine Forced Stop Setting]
[Cluster Properties] - [Recovery Tab] - [Execute Script for Forced Stop]	[Cluster Properties] - [Extension Tab] - [Execute Script for Forced Stop]
[Cluster Properties] - [Power Saving Tab] - [Use CPU Frequency Control]	[Cluster Properties] - [Extension Tab] - [Use CPU Frequency Control]
[Cluster Properties] - [Auto Recovery Tab] - [Auto Return]	[Cluster Properties] - [Extension Tab] - [Auto Return]
[Cluster Properties] - [Exclusion Tab] - [Mount/Umount Exclusion]	[Cluster Properties] - [Extension Tab] - [Exclude Mount/Unmount Commands]
[Group Properties] - [Attribute Tab] - [Failover Exclusive Attribute]	[Group Common Properties] - [Exclusion Tab]

Chapter 6 Upgrading EXPRESSCLUSTER

This chapter provides information on how to upgrade EXPRESSCLUSTER.
This chapter covers:

How to upgrade from EXPRESSCLUSTER X3.0 or X3.1 or X3.2 or X3.3

How to upgrade from X3.0 or X3.1 or X3.2 or X3.3 to X4.0

Before starting the upgrade, read the following notes.

- If mirror disk resources or hybrid disk resources are set, cluster partitions require space of 1024 MB or larger. And also, executing full copy of mirror disk resources or hybrid disk resources is required.
- If mirror disk resources or hybrid disk resources are set, it is recommended to backup data in advance. For details of a backup procedure, refer to "Backup procedures" and "Restoration procedures" in Chapter 8, "Verifying operation" in the *Installation and Configuration Guide*.
- Upgrade the EXPRESSCLUSTER Server RPM as root user.

The following procedures explain how to upgrade from EXPRESSCLUSTER X 3.0, 3.1, 3.2 or 3.3 to EXPRESSCLUSTER X 4.0.

1. Before upgrading, confirm that the servers in the cluster and all the resources are in normal status by using WebManager or the command.
2. Save the current cluster configuration file with the Builder or `clpcfctrl` command. For details about saving the cluster configuration file with `clpcfctrl` command, refer to "Backing up the cluster configuration data" in Chapter 3, "EXPRESSCLUSTER command reference" in the *Reference Guide*.
3. Uninstall the EXPRESSCLUSTER Server from all the servers. For details, refer to "Uninstallation" in Chapter 10, "Uninstalling and reinstalling EXPRESSCLUSTER" in the *Installation and Configuration Guide*.
4. Install the EXPRESSCLUSTER Server on all the servers. For details, refer to "Setting up the EXPRESSCLUSTER Server" in Chapter 3, "Installing EXPRESSCLUSTER" in the *Installation and Configuration Guide*.
5. If mirror disk resources or hybrid disk resources are set, allocate cluster partition (The cluster partition should be 1024 MB or larger).
6. Import the cluster configuration file which was saved in the step 2 with the Builder. If the cluster partition is different from the configuration, modify the configuration. And regarding the groups which mirror disk resources or hybrid disk resources belong to, if **Startup Attribute** is **Auto Startup** on the **Attribute** tab of **Group Properties**, change it to **Manual Startup**.
7. If mirror disk resources are set, perform the following steps for each mirror disk resource.
   - Click **Tuning** on the **Details** tab of **Resouce Properties**. Then, **Mirror disk resource tuning properties** dialog box is displayed.
   - Uncheck **Execute the initial mirror construction** on **Mirror** tab of the **Mirror disk resource tuning properties** dialog box.
8. Upload the the cluster configuration data with the Builder.

   If mirror disk resources or hybrid disk resources are set, initialize the cluster partition of all mirror disk resources and hybrid disk resources as below on each server.

   For the mirror disk

   `clpmdinit --create force <mirror disk resource name>`

   For the hybrid disk

```
clphdinit --create force <hybrid disk resource name>
```

9. Start the cluster.
10. If mirror disk resources or hybrid disk resources are set, start Mirror Disk Helper and execute a full copy assuming that the server with the latest data is the copy source.
11. Start the groups and confirm that each resource starts normally.
12. If **Startup Attribute** or **Execute the initial mirror construction** was changed in the step 6 or 7, change back the setting and select **Apply the Configuration File** from the **File** menu to apply the cluster configuration data to the cluster.
13. This completes the procedure for upgrading the EXPRESSCLUSTER Server. Check that the servers are operating normally as the cluster by the clpstat command or Cluster WebUI / WebManager.

Appendix

- Appendix A Glossary
- Appendix B Index

Appendix A Glossary

Cluster partition	A partition on a mirror disk. Used for managing mirror disks. (Related term: Disk heartbeat partition)
Interconnect	A dedicated communication path for server-to-server communication in a cluster. (Related terms: Private LAN, Public LAN)
Virtual IP address	IP address used to configure a remote cluster.
Management client	Any machine that uses the Cluster WebUI / WebManager to access and manage a cluster system.
Startup attribute	A failover group attribute that determines whether a failover group should be started up automatically or manually when a cluster is started.
Shared disk	A disk that multiple servers can access.
Shared disk type cluster	A cluster system that uses one or more shared disks.
Switchable partition	A disk partition connected to multiple computers and is switchable among computers. (Related terms: Disk heartbeat partition)
Cluster system	Multiple computers are connected via a LAN (or other network) and behave as if it were a single system.
Cluster shutdown	To shut down an entire cluster system (all servers that configure a cluster system).
Active server	A server that is running for an application set. (Related term: Standby server)
Secondary server	A destination server where a failover group fails over to during normal operations. (Related term: Primary server)
Standby server	A server that is not an active server. (Related term: Active server)
Disk heartbeat partition	A partition used for heartbeat communication in a shared disk type cluster.
Data partition	A local disk that can be used as a shared disk for switchable partition. Data partition for mirror disks and hybrid disks. (Related term: Cluster partition)
Network partition	All heartbeat is lost and the network between servers is partitioned. (Related terms: Interconnect, Heartbeat)

Node	A server that is part of a cluster in a cluster system. In networking terminology, it refers to devices, including computers and routers, that can transmit, receive, or process signals.
Heartbeat	Signals that servers in a cluster send to each other to detect a failure in a cluster. (Related terms: Interconnect, Network partition)
Public LAN	A communication channel between clients and servers. (Related terms: Interconnect, Private LAN)
Failover	The process of a standby server taking over the group of resources that the active server previously was handling due to error detection.
Failback	A process of returning an application back to an active server after an application fails over to another server.
Failover group	A group of cluster resources and attributes required to execute an application.
Moving failover group	Moving an application from an active server to a standby server by a user.
Failover policy	A priority list of servers that a group can fail over to.
Private LAN	LAN in which only servers configured in a clustered system are connected. (Related terms: Interconnect, Public LAN)
Primary (server)	A server that is the main server for a failover group. (Related term: Secondary server)
Floating IP address	Clients can transparently switch one server from another when a failover occurs. Any unassigned IP address that has the same network address that a cluster server belongs to can be used as a floating address.
Master server	The server displayed at the top of **Master Server** in **Server Common Properties** of the Builder
Mirror disk connect	LAN used for data mirroring in mirror disks and hybrid disks. Mirror connect can be used with primary interconnect.
Mirror disk type cluster	A cluster system that does not use a shared disk. Local disks of the servers are mirrored.

Appendix B Index

M

N

O

R

S

T

U

V

W