NEC EXPRESSCLUSTER
System Construction Guide

NEC EXPRESSCLUSTER for Windows Ver 7.0

NEC EXPRESSCLUSTER for Windows Ver7.0

System Design (BASIC/Shared Disk)

Second Edition 05/26/2004

Revision History

Revision	Revision Date	Revision Page	Description
First	06/30/2003		New manual.
Second	05/26/2004	3	FastSync Option and Exchange2000 Support Kit R1.0 support-related descriptions deleted.

PREFACE

The "NEC EXPRESSCLUSTER/System Construction Guide" is intended for administrators who will construct cluster systems and for system engineers or maintenance personnel who will provide user support for these systems.

AUXILIARY INFORMATION

[Concerning the OS upgrade]

If an invalid procedure is used when upgrading the operating system of the cluster server, a failover may occur unexpectedly, or in the worst case, the system may be damaged.

This also holds for the application of service packs.

NEC EXPRESSCLUSTER/System Construction Guide Categories

The NEC EXPRESSCLUSTER/System Construction Guides are divided into separate volumes as follows according to the situations in which they are to be used and according to readers. If you are to construct a cluster system for the first time, please read "Introduction" first.

[Setup Card]
     This is a reference card attached to the product, which describes the product configuration, operating environment, and so on.

[Introduction]
     This is an introduction for someone who is to design and construct a cluster system for the first time.

[System Design (BASIC/Shared Disk, BASIC/Mirror Disk]
     This manual describes information necessary to install NEC EXPRESSCLUSTER and construct a system.
     It explains necessary information and precautions concerning system construction.
     There are two volumes, one for Shared Disk type and the other for Mirror Disk type.

[System Design (Application)]
     This manual is a collection of knowhows for using advanced functions of NEC EXPRESSCLUSTER   that are not described in System Design (BASIC).

[Cluster Installation and Configuration Guide (Shared Disk, Mirror Disk)]
     This manual clearly explains environment settings to be made after installing NEC EXPRESSCLUSTER   in the order that you will follow to build a cluster system.
     There are two volumes, one for Shared Disk type and the other for Mirror Disk type.

 [Operation/Maintenance]
     This is a document that describes information required for operating a cluster system, how to deal with system
       failure, and lists of errors.

[GUI Reference]
     This manual explains required concepts concerning NEC EXPRESSCLUSTER operation and describes how to operate NEC EXPRESSCLUSTER by using a graphical user interface (GUI) from a management client or a server.

[Command Reference]
     This is a reference of commands that can be put in NEC EXPRESSCLUSTER scripts and maintenance commands that can be executed from the server's command prompt.

[For PP]
     Every PP mentioned in this document can be executed in cooperation with NEC EXPRESSCLUSTER.
     It explains settings, script writings, and some notes that are necessary to operate the PPs and NEC EXPRESSCLUSTER together.
     Read carefully the respective PP edition for the PP you would like to use.

[Notes]
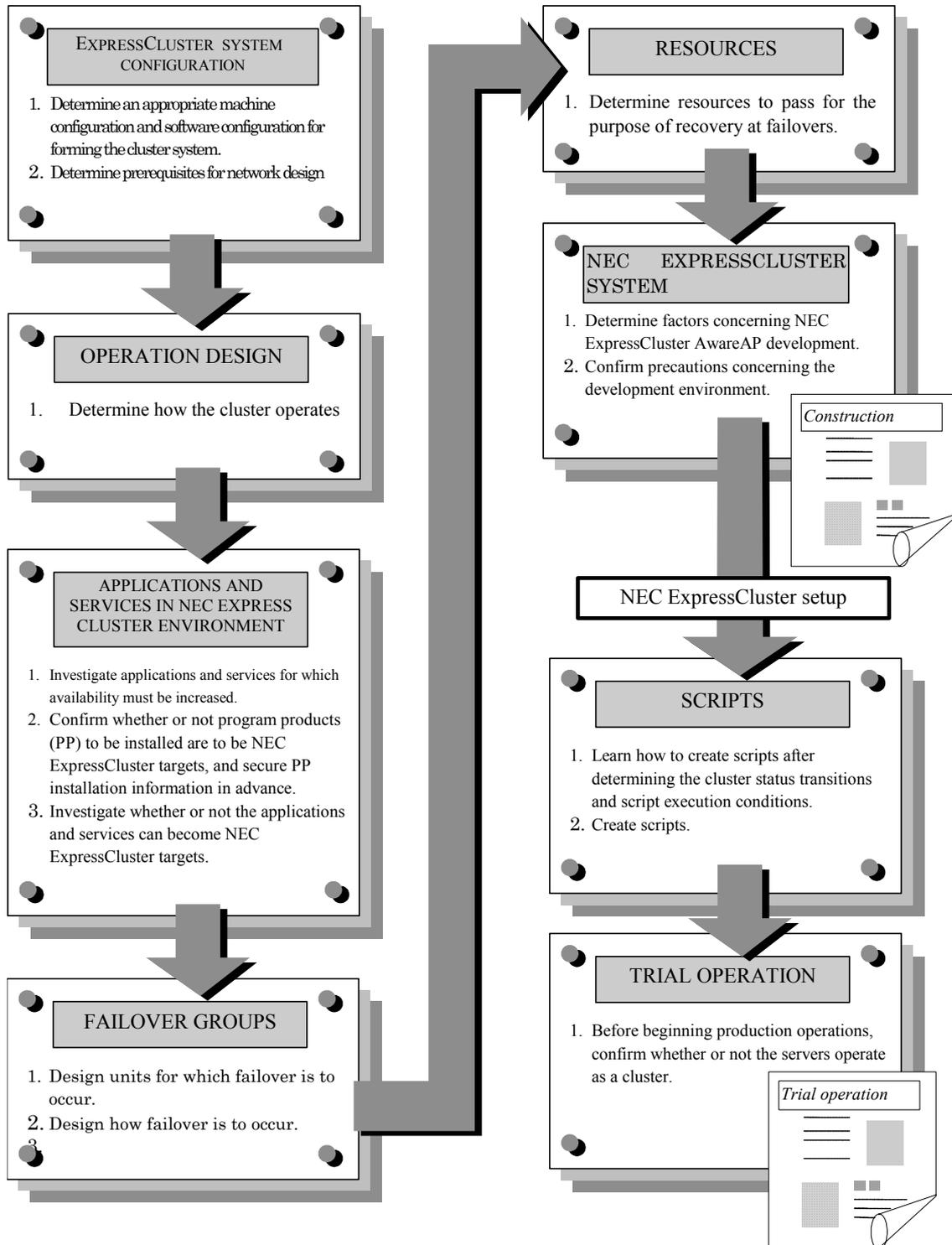   This reference describes items to be remembered in constructing, operating, and trouble shooting a cluster system. Read when necessary.

CONTENTS

1   NEC EXPRESSCLUSTER SYSTEM DESIGN PROCEDURE

EXPRESSCLUSTER SYSTEM CONFIGURATION

1. Determine an appropriate machine configuration and software configuration for forming the cluster system.
2. Determine prerequisites for network design

OPERATION DESIGN

1. Determine how the cluster operates

APPLICATIONS AND SERVICES IN NEC EXPRESS CLUSTER ENVIRONMENT

1. Investigate applications and services for which availability must be increased.
2. Confirm whether or not program products (PP) to be installed are to be NEC ExpressCluster targets, and secure PP installation information in advance.
3. Investigate whether or not the applications and services can become NEC ExpressCluster targets.

FAILOVER GROUPS

1. Design units for which failover is to occur.
2. Design how failover is to occur.
3.

RESOURCES

1. Determine resources to pass for the purpose of recovery at failovers.

NEC EXPRESSCLUSTER SYSTEM

1. Determine factors concerning NEC ExpressCluster AwareAP development.
2. Confirm precautions concerning the development environment.

Construction

NEC ExpressCluster setup

SCRIPTS

1. Learn how to create scripts after determining the cluster status transitions and script execution conditions.
2. Create scripts.

TRIAL OPERATION

1. Before beginning production operations, confirm whether or not the servers operate as a cluster.

Trial operation

2 NEC EXPRESSCLUSTER SYSTEM CONFIGURATION

2.1 Hardware Configuration

Management client

The interconnect uses Ethernet.

Clients

LAN

NEC EXPRESSCLUSTER Server
NEC EXPRESSCLUSTER Server
NEC EXPRESSCLUSTER Server
NEC EXPRESSCLUSTER Server

NEC Express5800/130Dp
NEC Express5800/130Dp
NEC Express5800/130Dp
NEC Express5800/130Dp

Shared disk

Up to sixteen servers can be added.

Different models of the target machines can be connected.
For more information about target EXPRESSCLUSTER machines and models, check the product notices and sales materials.

If you want to install the dual port feature in the disk array unit and dual disk path, the dual disk path SW is required.

There must be a sufficient amount of memory, virtual memory, and CPUs for the application programs and services to run in the event of failover.

2.1.1 Cluster server

* NEC EXPRESSCLUSTER SE should be 2-server configuration for shared disk operations. You can configure NEC EXPRESSCLUSTER EE/SX with 2 to 16 servers.
* You can connect different models of a machine. For information about NEC EXPRESSCLUSTER target machines and models, check the product notices and sales materials.
* Sufficient free space of memory, virtual memory and CPUs are required for application programs and services to run in the event of failover.
* Restrictions of interconnect are
    + Up to 16 interconnects are allowed per cluster system, there should be at least 2 interconnect.
    + You cannot use a primary interconnect as public LAN.

2.1.2 Management client

* A machine where NEC EXPRESSCLUSTER Manager is installed is referred to as Management Client.
* Either of Windows 98, Windows Me, Windows NT 4.0, Windows 2000, Windows Server 2003 or Windows XP should be operable on the management client.

2.1.3 Shared Disk

* As necessary, You must install the management tool for array disk used in the cluster system. Confirm this by setup card etc.
* If the data replication function of NEC iStorage series is used, Replication Volume (RV) cannot be used as shared disk of NEC EXPRESSCLUSTER.
    RV needs to be configured so that NEC EXPRESSCLUSTER server is unable to connect to RV, or only one server can connect to RV.

2.1.4 Interconnect

* Ethernet of 100BASE-TX or higher is used. (ex. NEC 100BASE-TX adapter set (N8104-85) or 1000BASE-SX adapter set (N8104-84))

2.2 Software Configuration

The following illustration is an example of the software configuration to install NEC EXPRESSCLUSTER.

Install the TCP/IP protocol.

Management client

Use TCP/IP for the communication protocol between clients and servers.

Clients

LAN

NEC EXPRESS SERVER NEC ExpressCluster NEC EXPRESS SERVER NEC ExpressCluster NEC EXPRESS SERVER NEC ExpressCluster NEC EXPRESS SERVER NEC EXPRESSCLUST
Server Server Server Server

Install TCP/IP and SNMP service on the server.

Shared disk

If a client uses NetBEUI or NetBIOS for communications, they must be installed on all servers.

The same revision of an application specified in a script must exist at all servers set in the failover policy and they must be handled in the same way at all of the servers.

Set up the shared disk so that the same partition can be accessed by the same drive letter from all servers.

2.2.1　　Operating environment

* NEC EXPRESSCLUSTER Manager of NEC EXPRESSCLUSTER Ver.6.05 (for example, ESS RL2000/09(including RUR application)) or later can manage cluster systems that were configured with NEC EXPRESSCLUSTER Ver 5.0 or later editions.
* NEC EXPRESSCLUSTER Enterprise Edition Manager of NEC EXPRESSCLUSTER Ver.6.02 (for example, ESS RL2000/06) or earlier can manage cluster systems that were configured with NEC EXPRESSCLUSTER Enterprise Edition.
 Likewise, NEC EXPRESSCLUSTER Standard Edition Manager of NEC EXPRESSCLUSTER Ver.6.02 (for example, ESS RL2000/06) or earlier can manage cluster systems that were configured with NEC EXPRESSCLUSTER Standard Edition.
* NEC EXPRESSCLUSTER Ver 7.0 Manager can manage cluster systems that were configured with NEC EXPRESSCLUSTER Ver 5.0 or later.　It cannot manage cluster systems that were configured with NEC EXPRESSCLUSTER Ver 4.2 or earlier.
 NEC EXPRESSCLUSTER Manager of Ver 4.2 or earlier cannot manage cluster systems that were configured with NEC EXPRESSCLUSTER Ver 5.0 or later.
* NEC EXPRESSCLUSTER EE or NEC EXPRESSCLUSTER SX can configure a cluster system with up to 16 servers.
 NEC EXPRESSCLUSTER SE or NEC EXPRESSCLUSTER LE can provide 2-server configuration.
* NEC EXPRESSCLUSTER Server should be installed.
* If you want to configure a load sharing type cluster system of load distribution with Oracle Parallel Server, NEC EXPRESSCLUSTER/Oracle Parallel Server support option is required.

2.2.2　　Cluster server

2.2.2.1　　　Precautions about Shared Disk

* Make settings so that you can use the same drive letters to access the same partition from servers that have a failover policy.
* You can create up to 32 partitions per disk.
* On a shared disk, FT disk's functions such as stripe set, volume set, mirroring, and parity-protected stripe set aren't supported.
* It is not allowed on Windows 2000 or Windows Server 2003 to use a dynamic disk as shared disk. Don't upgrade a basic disk to a dynamic disk. GPT-type disks are not available, either.
* Restrictions of shared disk are as follows;
  + Up to 16 cluster partitions are allowed per cluster system.
  + Up to 26 switched disk partitions (including system disks) are allowed per cluster system.
  + The cluster/shared partitions are not displayed in the server's resource tree.　If connection to a cluster or shared partition fails, that server will turn to yellow and a server status "Having some resources that connection is failed" will appear.

2.2.2.2 Precautions about network environment

* You need to install TCP/IP and SNMP service.
* If a client communicates with server through NetBEUI or NetBIOS, you should install them on the server, too.   (There are some precautions to use NETBEUI protocol.   See Section 6.3.6 "Precautions concerning the use of virtual computer names" for details.)
* Restrictions of IP addresses are as follows;
    + Up to 16 addresses are allowed per server (except for floating and virtual IP addresses).
    + One server cannot have more than one IP address that belongs to the same network address.
      Also, nesting relations as shown below aren't allowed:
        IP address: 10.1.1.10, Subnet mask: 255..255.0.0
        IP address: 10.1.2.10, Subnet mask: 255..255.255.0

2.2.2.3　Precautions about cluster settings

* Restrictions of cluster names are as follows;
  + Only single-byte alphanumeric uppercase and lowercase characters, hyphens (-), and underscores (_) are allowed.
  + Case-insensitive.
  + Up to 15 characters (15 bytes).
  + The name must be unique for each cluster system.
* Restrictions of server names are as follows;
  + The same restrictions applicable as computer names on the OS.
  + Case-insensitive.
  + Up to 15 characters (15 bytes).
* Restrictions of failover groups are as follows;
  + Up to 64 groups are allowed per cluster system.
  + Restrictions of failover group names are as follows;
    = Only single-byte alphanumeric uppercase and lowercase characters, hyphens (-), and underscores (_) are allowed.
    = Case-insensitive.
    = Up to 15 characters (15 bytes).
    = The name must be unique within a cluster system.
    = DOS I/O device names such as PRN are not allowed.　(Even though you can specify it as failover group name, it will fail to start.)
* Restrictions of cluster passwords are as follows;
  + Up to 15 bytes.
* Restrictions of user accounts are as follows;
  + Up to 16 user accounts are allowed per cluster system.
  + You can use only the user names and passwords that were set by the user manager.
  + User names can be up to 31 bytes.
  + Passwords can be up to 31 bytes.
* Restrictions of floating IP addresses are as follows;
  + Up to 64 addresses per cluster system.
  + You should assign an unoccupied host address within the same network address as the server's Public-LAN.
* Restrictions of virtual computer names are as follows;
  + Up to 64 virtual computer names per cluster system.　However, if a virtual computer name is associated with a floating IP address, it will be less than 64.　The maximum number will be (64 – NICs that are bound to WINS clients).
  + Allowed characters for virtual computer name are the same as those allowed as computer name in OS.
  + Up to 15 bytes.
  + Server names that exist within the network are not allowed.
  + The name must be unique within the network.

2.2.3　Management client

* TCP/IP must be installed.
* A single NEC EXPRESSCLUSTER Manager can manage up to 128 cluster systems.
* Up to 32 NEC EXPRESSCLUSTER Managers can access a server within cluster system.
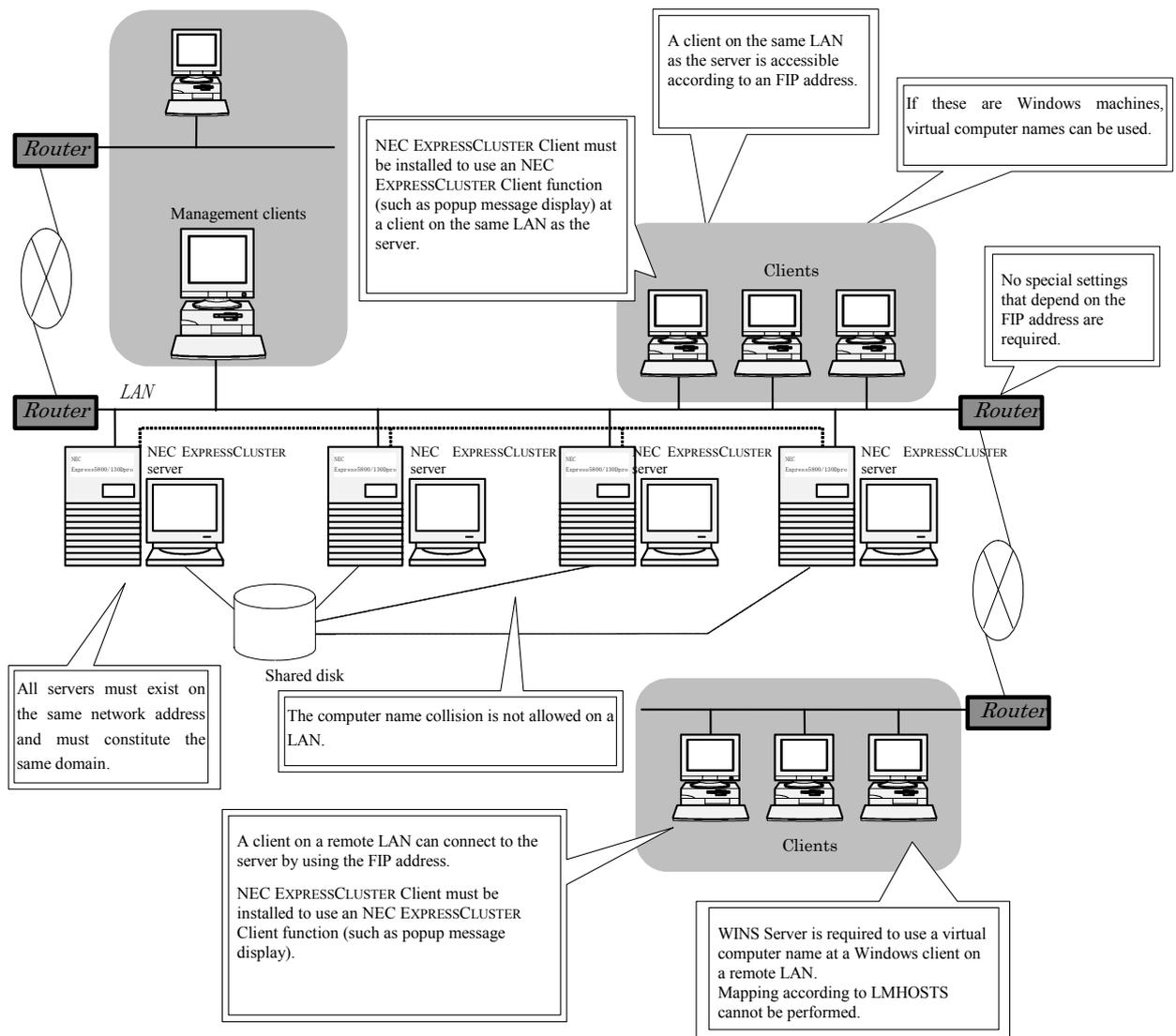
2.3   Network Design

2.3.1   Network overview

Network can use floating IP (FIP) addresses or virtual IP addresses.   FIP is recommended unless the LAN connection should be duplicated.

(This chapter describes only FIP.   See "*NEC EXPRESSCLUSTER System Construction Guide / System Design (Advanced)*" for details of settings with virtual IP address.)

Floating IP addresses differ from virtual IP addresses as follows.

	Floating IP Addresses	Virtual IP Addresses
Available IP addresses	Same network addresses as the public LAN	Different network addresses from the public LAN or interconnect
Switching method	The MAC address in the ARP table is switched according to an ARP broadcast from the server	The gateway address in the routing table is switched according to a RIP broadcast from the server and the NEC EXPRESSCLUSTER Client.
Client operating system	Any	OSes on clients on the same LAN is limited to Windows 98, Windows Me, Windows NT, Windows 2000, Windows XP, Windows Server 2003 and UNIX. Any OS at a remote LAN.
Router setting	Unnecessary	RIP must be interpreted.
NEC EXPRESSCLUSTER Client	Unnecessary	Required at a client on the same LAN.
LAN multiplexing	Not allowed	Allowed
Potential resources	List of IP addresses that were set by using floating IP settings	-
Use between servers	Allowed (however, only for the public LAN)	Allowed (supports route switching in the event of disconnections)
Use of Oracle Parallel Server	Allowed	Allowed

A client on the same LAN as the server is accessible according to an FIP address.

If these are Windows machines, virtual computer names can be used.

NEC EXPRESSCLUSTER Client must be installed to use an NEC EXPRESSCLUSTER Client function (such as popup message display) at a client on the same LAN as the server.

Management clients

No special settings that depend on the FIP address are required.

Clients

LAN

NEC EXPRESSCLUSTER server

NEC EXPRESSCLUSTER server

NEC EXPRESSCLUSTER server

NEC EXPRESSCLUSTER server

All servers must exist on the same network address and must constitute the same domain.

Shared disk

The computer name collision is not allowed on a LAN.

A client on a remote LAN can connect to the server by using the FIP address.

NEC EXPRESSCLUSTER Client must be installed to use an NEC EXPRESSCLUSTER Client function (such as popup message display).

Clients

WINS Server is required to use a virtual computer name at a Windows client on a remote LAN.
Mapping according to LMHOSTS cannot be performed.

Using Floating IP Addresses

See Section "6.2 Floating IP address" for details of floating IP addresses (FIP).

See Section 6.3.2 "Usage from remote LAN" for details of WINS servers to use virtual computer names on a remote LAN.

See "*NEC EXPRESSCLUSTER System Construction Guide / System Design (Advanced)*" for details of network designing to use virtual IP addresses.

2.3.2   Cluster server
* A cluster consists of 2 to 16 servers.   They must exist in the same LAN and must constitute the same domain with the same network address.
* The interconnect LAN and public-LAN must be different network addresses.   The IP addresses of the interconnect LAN can be private IP addresses.

2.3.2.1   Windows 2000, Windows Server 2003
You can construct an NEC EXPRESSCLUSTER system in the Active Directory environment of Windows 2000 or Window Server 2003.
* NEC EXPRESSCLUSTER servers need to belong to the same domain.
* It isn't advisable to use an NEC EXPRESSCLUSTER server as a domain controller.

2.3.3   Management Client
* NEC EXPRESSCLUSTER Manager must be installed.
* It need not belong to the same domain nor exist on the same LAN as the cluster servers.
* If there is a firewall between NEC EXPRESSCLUSTER Server and Manager, NEC EXPRESSCLUSTER Manager's behaviors are not assured.

2.3.4   Client
* All Windows clients on the same LAN as the server or on a remote LAN can be accessed according to an FIP address.   Also, even a client using a non-Windows operating system can be accessed according to an FIP address as long as ARP is supported.[1]

* Regardless of the LAN configuration, NEC EXPRESSCLUSTER Client must be installed to use an NEC EXPRESSCLUSTER Client function (such as popup message output or application execution using the API) on a Windows 98,Windows ME, Windows NT4.0, Windows 2000, Windows XP or Windows Server 2003 client.

* You need to install an NEC EXPRESSCLUSTER client to use a virtual computer name on Windows 2000 Windows Me, or Windows XP client.
  If you associated a virtual computer name with an FIP address (see Section "6.3.1.3 Associate a virtual computer name and a floating IP address" for details), you don't need to install an NEC EXPRESSCLUSTER client.

* To use virtual computer names on a Windows 98 client, you may need to install a CLUSTERRPO client.   See Section "6.3.8 To use virtual computer names on Windows 98" for details.
  If you associated a virtual computer name with an FIP address (see Section "6.3.1.3 Associate a virtual computer name and a floating IP address" for details), you don't need to install an NEC EXPRESSCLUSTER client.

* The cluster servers and clients need not belong to the same domain.   However, a client that uses security functions provided by OSes may need to belong to the same domain as the servers.

---

[1] Not all non-Windows hosts are assured to behave normally.   Verify yourself before starting operations.

## 2.3.5　Router

* No special settings are required for FIP addresses.

2.3.6　Duplication of LAN connection

* Dynamic routing is not applicable for floating IP addresses in a duplicate LAN connection configuration while it is applicable for virtual IP addresses.　If dynamic routing is required, use the virtual IP addresses.
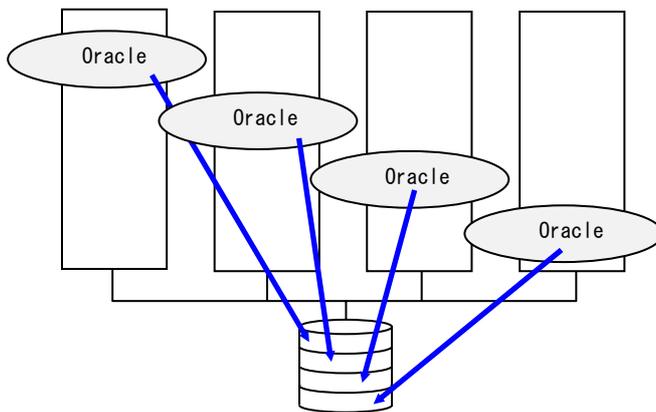
3   OPERATION DESIGN

NEC EXPRESSCLUSTER supports following three operational forms.

* Uni-directional standby cluster
  Only one application can run within the entire cluster system.

COBOL 85 AP
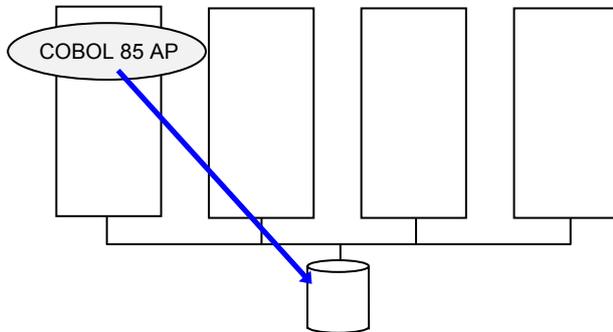
* The same application's multi-directional standby cluster
  The same applications can run on multiple servers in a cluster system.
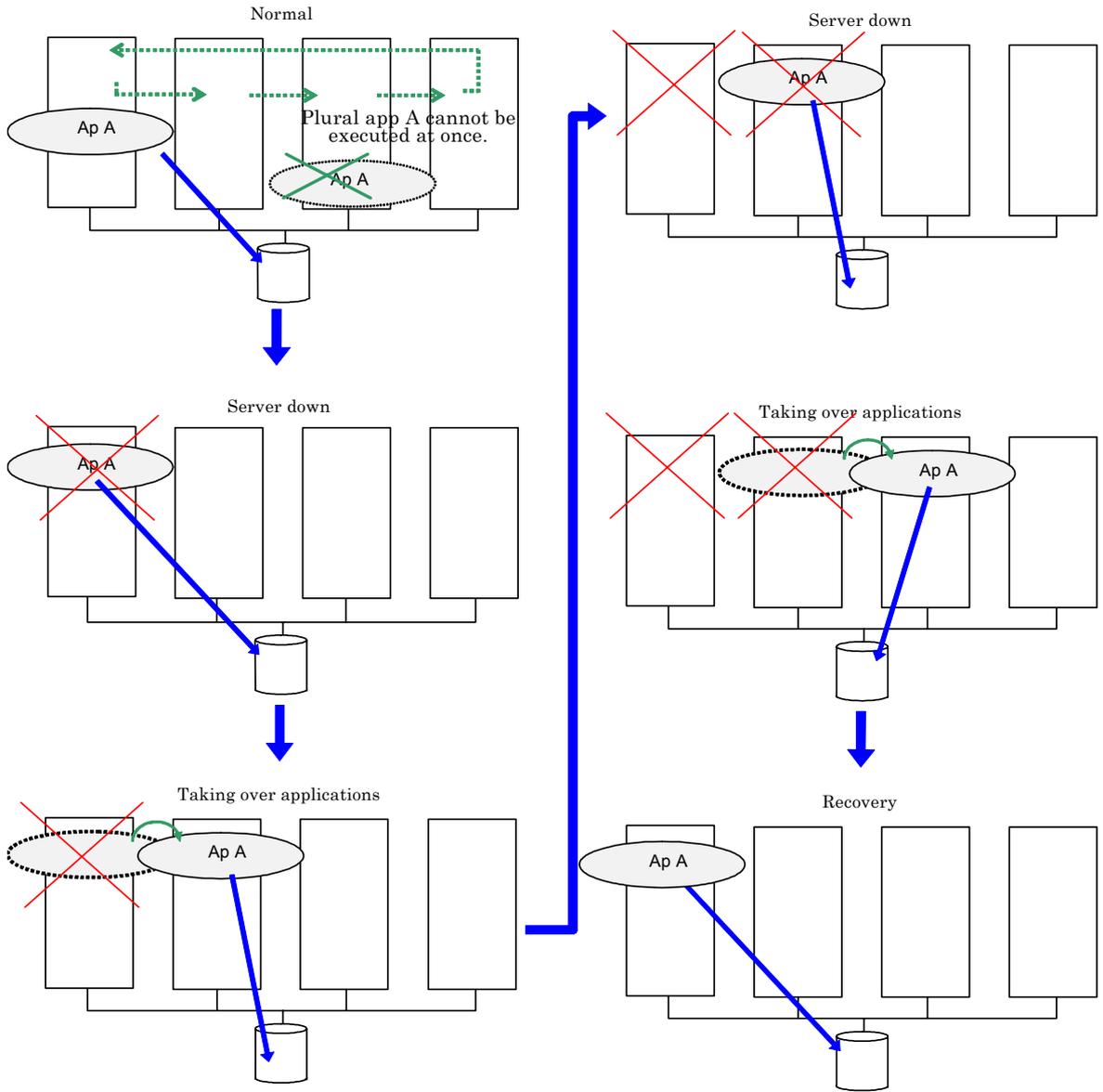
Oracle
Oracle
Oracle
Oracle

* Parallel cluster
  Each server shares the processes of same application in this cluster system (if any server goes down, the operations will be degenerated).
 Oracle Parallel Server is required.

Oracle Parallel Server
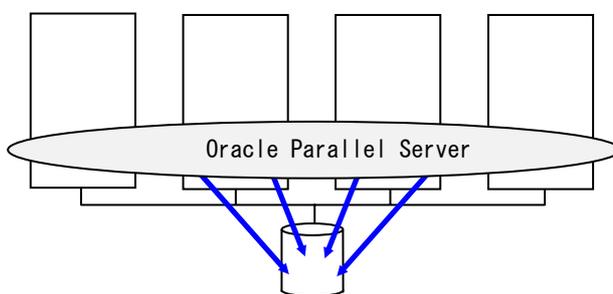
3.1   Uni-directional Standby Cluster

In uni-directionsl standby cluster system, only one failover group is allowed for a application.

Normal

Ap A

Plural app A cannot be
executed at once.

Ap A

Server down

Ap A

Server down

Ap A

Taking over applications

Ap A

Taking over applications

Ap A

Recovery

Ap A

Configuration sample: In this sample, COBOL85 AP, IIS, Oracle8 and Exchange Server of unidirectional standby cluster are combined.
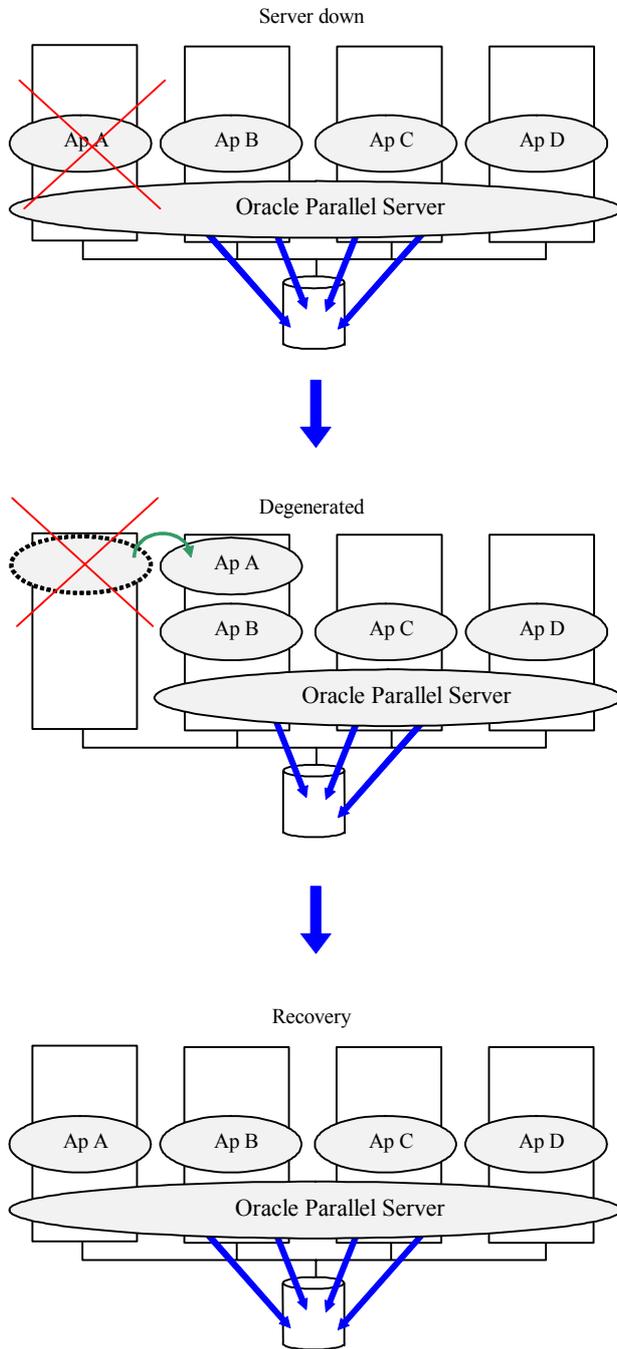
COBOL 85
AP

IIS

Oracle8

Exchange
Server
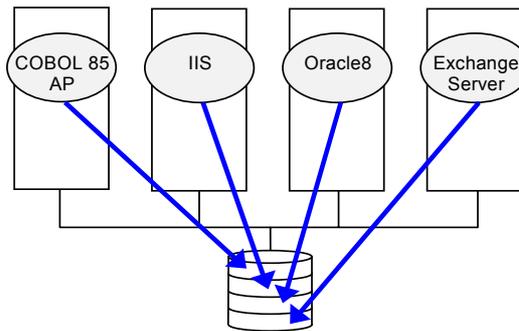
3.2   The Same Application's Multi-Directional Standby Cluster

An application can run on multiple servers simultaneously in this cluster system.

Normal

Server down

Server down

Taking over ap C

Taking over ap A

Recovery
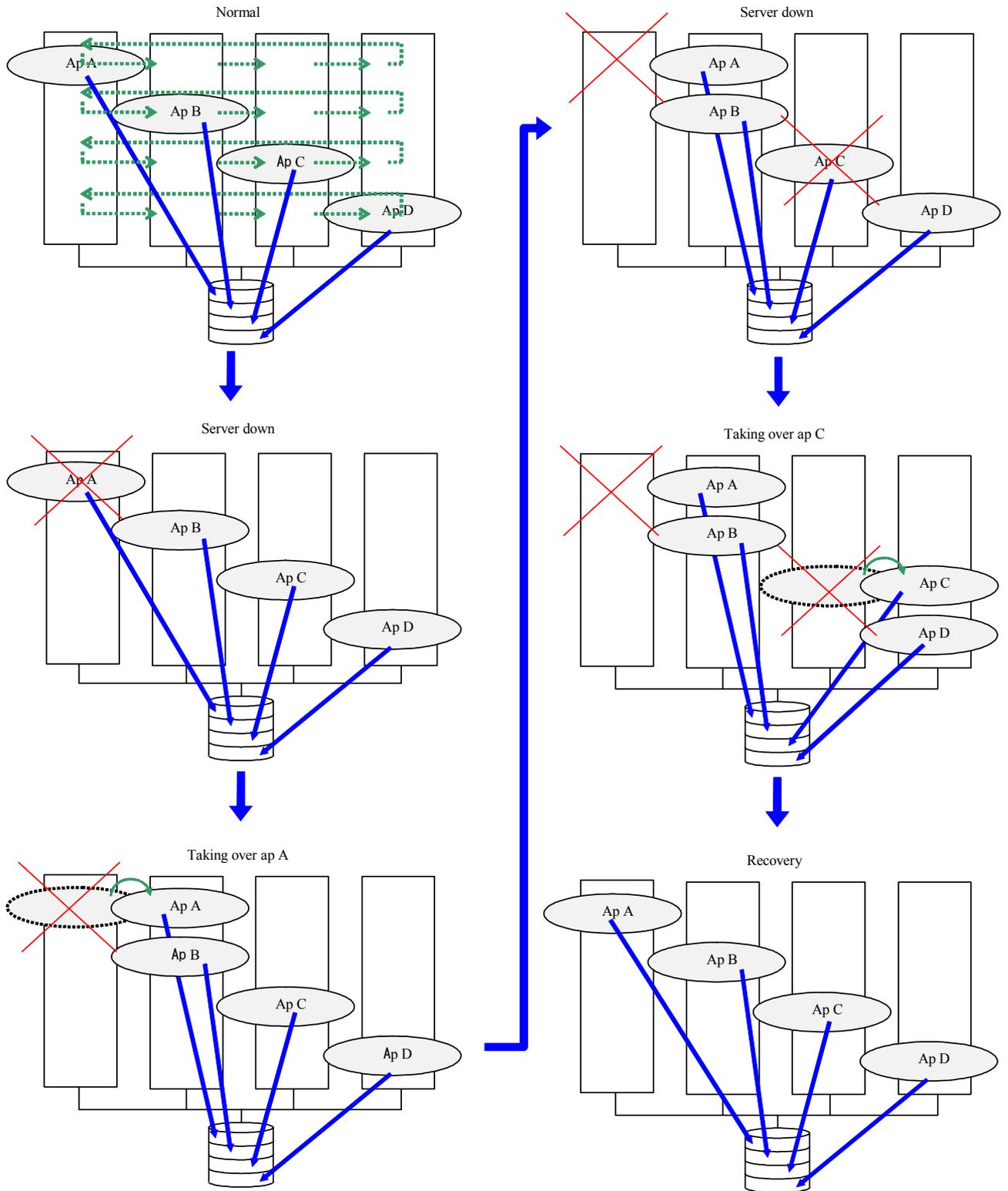
3.3   Parallel Cluster

Each server shares processes in parallel cluster system.   If any of servers goes down, the operations will be degenerated.   NEC EXPRESSCLUSTER supports cluster operations that use Oracle Parallel Server.

Server down

Ap A   Ap B   Ap C   Ap D

Oracle Parallel Server

Degenerated

Ap A

Ap B   Ap C   Ap D

Oracle Parallel Server

Recovery

Ap A   Ap B   Ap C   Ap D

Oracle Parallel Server

4 APPLICATIONS AND SERVICES IN AN NEC EXPRESSCLUSTER ENVIRONMENT

This chapter describes matters that should be carefully considered concerning applications and services that can operate in an NEC EXPRESSCLUSTER environment.

4.1 Determining Applications

When NEC EXPRESSCLUSTER is installed, you must first determine applications and services for which availability must be improved. Also, you must ascertain whether or not the applications and services that you found are appropriate for operating in an NEC EXPRESSCLUSTER environment.

Investigate the items described in the following sections to determine whether or not the selected applications and services can be applied in an NEC EXPRESSCLUSTER cluster.

4.2 Applications and Services in an NEC EXPRESSCLUSTER Environment

4.2.1 Server application

Each standby type has different notes for applications as described in Section 4.2.2.

* Uni-directional standby [Active - Standby]   Notes: 1 2 3 ~~4~~ 5
  An application runs on a server in this cluster system.

* Multi-directional standby [Active - Standby] Notes: 1 2 3 4 5
  An application runs on multiple servers in this cluster system.

* Co-resident behavior   Notes: ~~1 2 3 4~~ 5
  Applications are not the object of failover by cluster system, they coexist in this cluster system.
  This type is such as NEC ESMPRO/Server Agent and InocuLAN.

4.2.2 Precautions about server applications

(1) Restoring data after errors

If an application was just updating a file when an error occurred, the file may contain corrupted data when the standby system accesses the file.

The same is true also when a non-cluster system (a single server system) goes down and is rebooted. Basically, applications should be ready to handle this kind of errors. Scripts should be able to restore this type of errors on cluster systems without human interventions.

NEC EXPRESSCLUSTER's failover will not perform CHKDSK for this type of errors. Instead, NTFS Log file service restores journals automatically.

(2) Exiting applications

Before NEC EXPRESSCLUSTER stops and moves application groups (online failback), it unmounts the file systems which they use. Therefore, applications should issue a command to stop accessing all files on shared disks.

Typically, a stop script issues the command to exit the application. Be careful if the stop command does not synchronize with exiting application, you should pay attentions. (For example, the ARMSLEEP command is issued to wait for some time.)

(3) Where to store data

NEC EXPRESSCLUSTER can pass the following two types of data between servers.

- = Data on shared disk
- = Registry key to be synchronized

You should separate data to be passed and not to be passed between servers so that applications to distinguish it.

| Data type	(Sample)	Where to store
Data to be passed	(such as user data)	Shared disk registry
Data not to be passed	(such as programs, settings)	Server's local disk

(4) Multiple application groups

It should be assumed that multiple application groups of the same application program run on a server on multi-directional standby system (in case of degeneration due to errors).

Applications receive resources by either of following methods and should be able to run multiple application groups on a server.

- Invoking multiple instance
Invoke a new instance (process). Multiple applications should be executable.

- Restarting the application
Stop a running application and restart it to make added application available.

By restarting the application, data is passed.

- Adding dynamically
Resources are added to a running application dynamically or by instructions from script.

Dynamically adding data to a running application, data is passed.
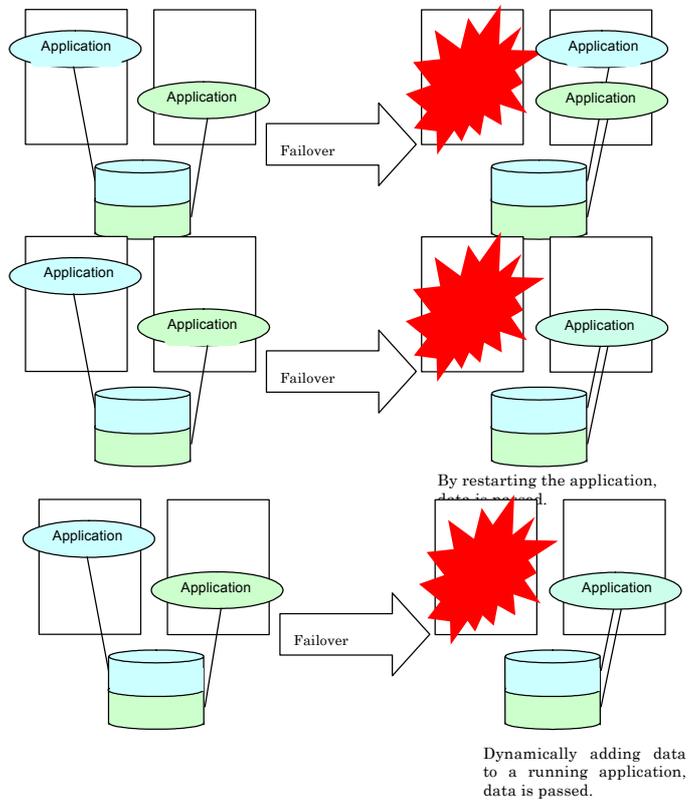
(5)   Mutual interference with applications (compatibility)

Applications or NEC EXPRESSCLUSTER may not function properly in some cases due to interferences between NEC EXPRESSCLUSTER and OS functions required for the former's functions and behaviors.

* I/O filtering

    NEC EXPRESSCLUSTER performs I/O filtering to switch disk partitions between servers. Therefore, it returns an error of "Not Ready" for I/O requests for inactive partitions.

* Applications are not allowed to access inactive mirror disks, i.e., mirror disks they do not have access rights.

    Usually, applications can assume that the mirror disk partitions are already accessible to them at the point of startup if they are started from cluster scripts.

* Access to mirror disks by applications

    Co-resident running applications do not receive notifications about stop of application groups. Therefore, if such an application is accessing a mirror disk partition used by an application group, the application will fail to unmount the disk after the group has stopped.

    Some applications such as system monitoring service access all disk partitions on regular basis. To prevent the above problem, such applications should have functions to limit the scope of partitions to be monitored.

* Multi-home environment and transfer of IP addresses

    Typically in cluster system, a server has multiple IP addresses and they (such as floating IP address) are transferred between servers.

    Problems will occur if an application assumes that gethostbyname of Winsock returns only one IP address.   In this case, you can resolve this kind of problems by changing the order of the returned IP addresses with Winsock Wrapper attached to NEC EXPRESSCLUSTER.

4.2.3　Precautions to use applications

This section describes information on server applications.
There may be more notes or information other than described here on products below.

(1)　Not supported by NEC EXPRESSCLUSTER

- Norton AntiVirus 4.0
　　This product is not usable in NEC EXPRESSCLUSTER environment because I/O to shared disks and mirror disks may be blocked (details about other version are unknown).

- Virus Buster
　　Disk isolation failures are reported.　The cause of failures was Virus Buster's occupation on shared and mirror disks at cluster shutdown, group's transfer or stop.

　　Use the following anti-virus software as described in "*NEC EXPRESSCLUSTER System Construction Guide / [For PP]*".
　　InocuLAN
　　ServerProtect

- ARCserve Disaster Recovery Option
　　OS installation in shared disks at restoration is reported.

(2)　Supported by NEC EXPRESSCLUSTER

- ARCserve Image Option
　　RAW partitions cannot be backed up.

4.2.4   Client applications

When a cluster system detects an error in the active server and a failover occurs, the connection between the client and server is temporarily disconnected.   Therefore, client applications should be able to reconnect to servers by either of following methods to proceed.

* Auto-reconnection

  Automatic reconnection to server by communication time-out or NEC EXPRESSCLUSTER client API
* Manual reconnection

  Manual reconnection to server by users by recognizing the failover by some means such as NEC EXPRESSCLUSTER Client pop-up message
* Client Applications' restart or Client PC's restart

  This method works when the connection with server is disconnected or you should restart the client application.

4.2.5   Sample countermeasures on cluster side

Problems	Actions	Numbers corresponding to those of notes in Section 4.2.2.
If a data file is unsuccessfully updated, applications cannot run properly on the standby system.	Modify the program.	(1)
To keep accessing shared disks for a certain period of time even after a service stopped.	Use the ARMSLEEP command in the stop script to wait.	(2)
Cannot access the registry on the standby system because a local server account is used, instead of domain account, to control access the registry information you want to pass between servers.	Use a domain account to control access.	(3)
Cannot start the same service multiple times simultaneously on a server.	Restart the service to pass shared data after failover in bidirectional standby operations.	(4)
I/Os to the entire disk are blocked if Norton AntiVirus receives "Not Ready" as reply of I/O request to a partition.	Use a different virus checker. (InocuLAN and ServerProtect are already evaluated.)	(5)

4.3   Determining the Application Format

Determine the application format based on the information presented throughout Chapter4.
* Which application or service is to be started up and when should it be started?
* What processing is required when the application or service is started up or when failover occurs?
* What data should be placed on the shared disk?
* What information should be placed in the synchronization target registry keys?


In addition, the operations should include,
* Regular backups of the shared disks

5   FAILOVER GROUP

A failover group, which is a collection of resources required to execute an application within a cluster system, is the unit for which failover occurs.

A failover group has a failover group name, cluster resources, and attributes.

Up to 64 failover groups can be created for one cluster system.
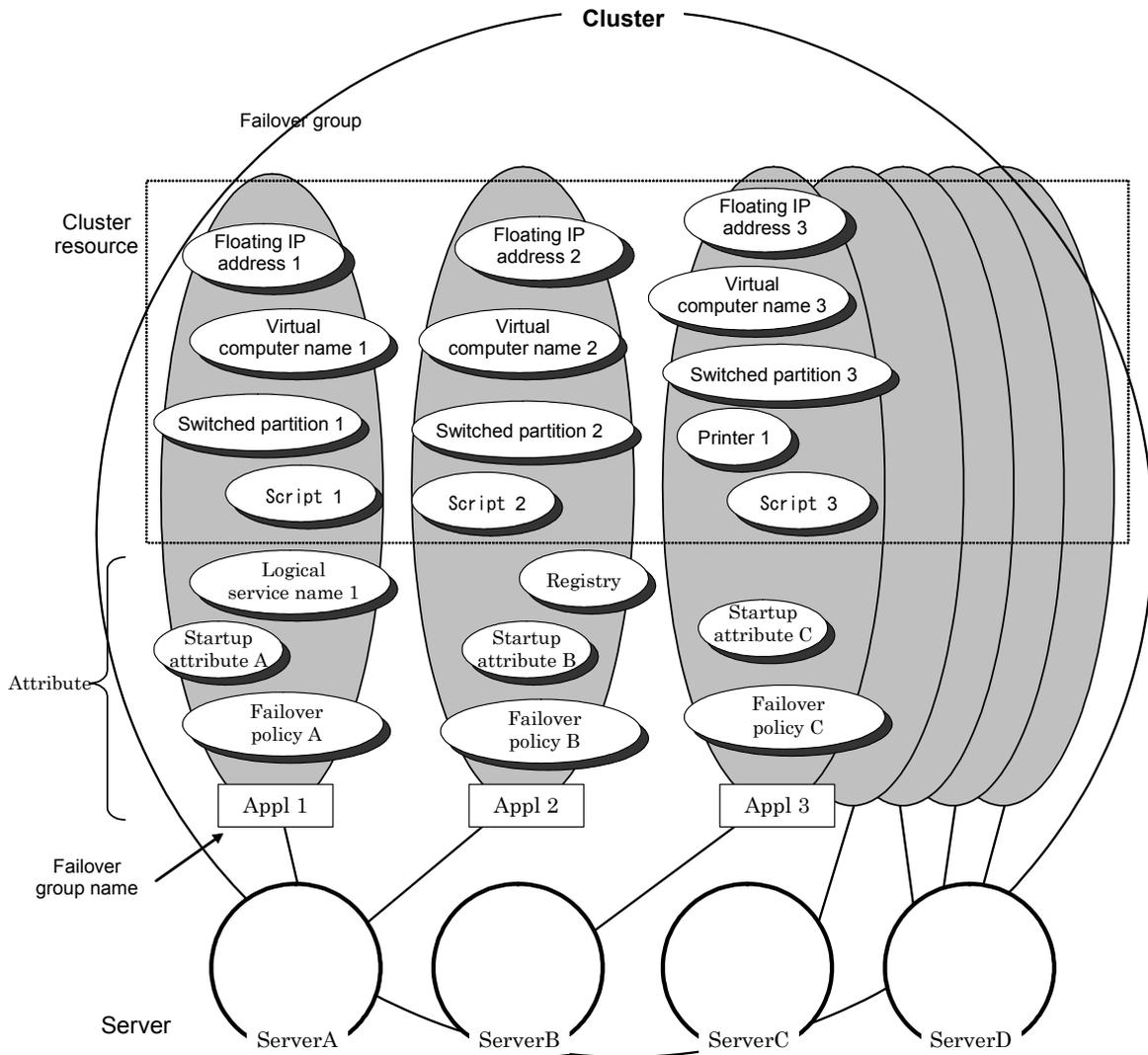
The cluster resources of each failover group are processed as an individual unified group. That is, when a failover occurs in Appl1, which has the switched partition 1 and floating IP address 1, both of the switched partition 1 and floating IP address 1 fail over (failover does not occur only for the switched partition 1).

Also, Switched partition 1 is not included in other failover groups (such as App 2).

5.1　　Cluster Resources

A failover group can possess the following cluster resources.
See Chapter "6 RESOURCES" for details.

* Switched Partition
* Floating IP address
* Virtual IP address
* Virtual Computer Name
* Printer
* LSW (To be connected to a line switching unit, Ver. 3.0 compatible operations only)
* Script
* Resource monitoring

5.2　　Attributes

A failover group possesses the following attributes.

* Failover group name
* Startup Attribute
* Failover Attribute
* Failback Attributes
* Failover policy
* Logical Service Name
* Synchronization target registry key

5.2.1　　Failover group name

This is the name of a failover group.
Since the NEC EXPRESSCLUSTER Client API performs monitoring or obtains the status in terms of individual failover groups, the cluster name and failover group names are used as keys.
Restrictions are as follows;

* Valid characters are one-byte alphanumeric uppercase or lowercase, hyphens (-), and underscores (_) only.
* Case-insensitive
* Up to 15 characters (15 bytes).
* Failover group names should be unique within a cluster system.
* A DOS input/output device name such as PRN cannot be used.
  (It can be set as the failover group name.　However, a failover group of such a name will fail to start.

5.2.2　　Startup attribute

The startup attribute specifies whether the failover group is to be started up automatically or manually. In automatic startup, NEC EXPRESSCLUSTER starts it up together with the cluster.　In manual startup, users start it up from the NEC EXPRESSCLUSTER Manager.

* Automatic startup
  The failover group is started up automatically by NEC EXPRESSCLUSTER.
  When the cluster is started up, the failover group is automatically started up (active status).
* Manual startup
  The failover group is not started up from NEC EXPRESSCLUSTER.　Users start it up from the NEC EXPRESSCLUSTER Manager.
  When the cluster is started up, the failover group is not started up yet (inactive status).　It is subsequently started up (active status) according to a user operation from the NEC EXPRESSCLUSTER Manager.

5.2.3 Failover attribute

Specify rules for determining failover destinations.

You can choose "Normal" or "Exclusive" as failover destination. In normal, the top priority server is always the failover destination. In exclusive, a server where no group has been started up is the failover destination.

* Normal
  Failover is automatically performed by NEC EXPRESSCLUSTER. The failover destination is always the server having the highest priority. Multiple groups may be started up at the same server. Offline failback is available.
* Exclusive
  Failover is automatically performed by NEC EXPRESSCLUSTER. The failover destination is always the server having the highest priority among the servers where no group of exclusive attribute has been started up. If no such a server can be found, failover is not performed. Multiple groups of exclusive attribute may not be started up at the same server. Offline failback is not available.
  It is not recommended to specify "Exclusive" in 2-server cluster systems (mostly, 2-server cluster systems do not fail over).

5.2.4 Failback attributes

You specify whether to automatically fail back to the top priority server specified in the failover policy when it returns to normal status.

Failback attribute is either of
* YES(= Fail back automatically)
* No
Default value is "No".

5.2.5 Failover policy

The failover policy is a list of available servers in case of failover, and the failover priorities among these servers.

Restrictions are as follows;
* Failover destination is the top priority server among servers that are operating as cluster.

The operations differ as described below according to the failover policy at the time a failover occurs.

(Explanation of symbols within the figure)

| Server Status	Description
O	Normal status (operating normally as a cluster)
X	Stopped status (stopped due to a cluster shutdown)
Xd	Down status (stopped because a server went down)
Od	Restarted after down status (isolated from the cluster)

Failover group	Failover policy		
	Priority 1 server	Priority 2 server	Priority 3 server
A	Server 1	Server 2	Server 3
B	Server 2	Server 3	Server 1
C	Server 3	Server 1	Server 2

When the failover attribute is Normal
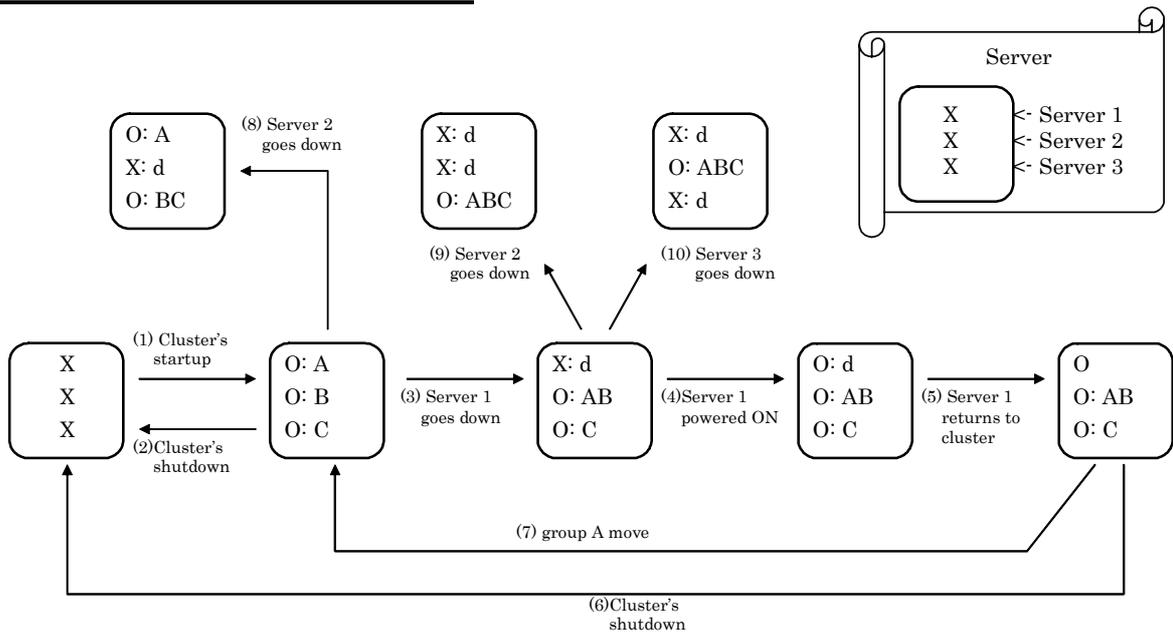
(1) Cluster startup
(2) Cluster shutdown
(3) Server 1 goes down: Failover is performed to the server having the next highest priority
(4) Turn on Server 1.
(5) Restore Server 1 to cluster
(6) Cluster shutdown: Off-line failback is performed.
(7) Failover group A is moved
(8) Server 2 goes down: Failover is performed to the server having the next highest priority
(9) Server 2 goes down: Failover is performed to the server having the next highest priority
(10)   Server 3 goes down: Failover is performed to the server having the next highest priority

When the failover attribute is Exclusive
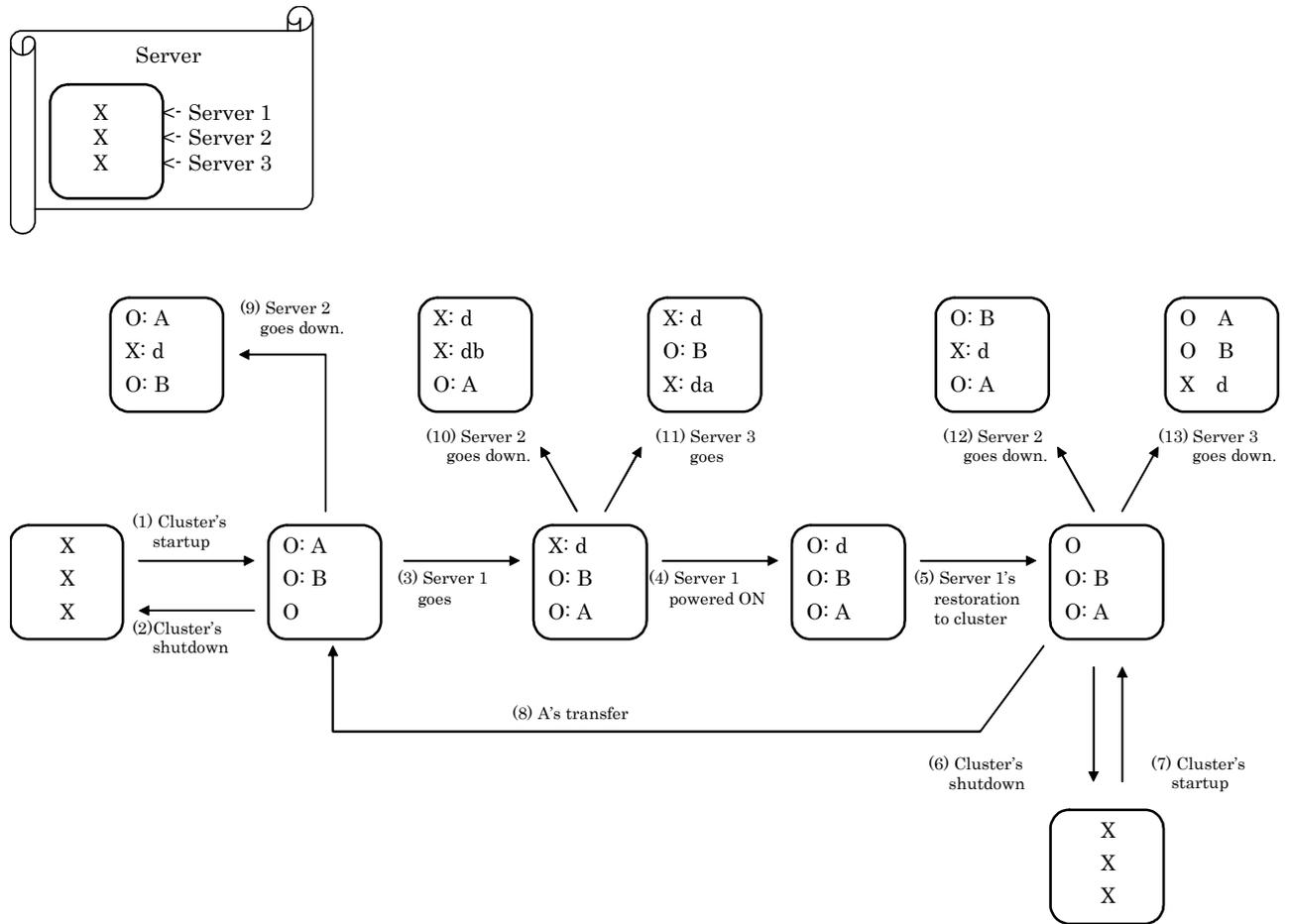
a and b mean stopped status

- （1） Cluster startup
- （2） Cluster shutdown
- （3） Server 1 goes down: Failover is performed to the server having the next highest priority
- （4） Turn on Server 1
- （5） Restore Server 1 to cluster
- （6） Cluster shutdown
- （7） Cluster startup: Offline failback is not performed
- （8） Failover group A is moved
- （9） Server 2 goes down: Failover is performed to the server having the next highest priority
- （10） Server 2 goes down: Failover is not performed
- （11） Server 3 goes down: Failover is not performed
- （12） Server 2 goes down: Failover is performed to the server having the next highest priority
- （13） Server 3 goes down: Failover is performed to the server having the next highest priority

If failover groups having different failover attributes coexist within a cluster, restrictions are as follows;

<Cluster startup>
*   First, determine the server to start up for the normal failover group. A normal failover group is started up on the server having the highest priority of its failover policy.
*   Next, determine the server to start up for the exclusive failover group.
    First, write off the servers with higher priorities than the server where the last cluster shutdown occurred. Next, investigate servers from the highest priority server to the lowest. Third, find servers on which any failover groups are not started up yet.   The found servers are the candidates for the failover group to start up.
    After all candidate servers of exclusive failover group are found, each candidate server chooses one failover group and starts it up. At this time, precedence among the failover groups to a server is based on the priority of failover on the server. When the priority is the same, precedence is determined according to alphabetical order of the failover group names.
    This is repeated until there is no server that no failover group is activated.

<Server failure>
*   First, determine the server to start up for the normal failover group. A normal failover group is moved to the server having the highest priority.   At this time, it is ignored even if a failover group for which the failover attribute is exclusive has been already activated.
*   Next, determine the server for the exclusive failover group to go to in case of failover.   As for each failover group to be failed over, how to determine it is the same way as cluster startup.   However, all servers that are working in the normal state are looked up.

5.2.6　Synchronization target registry key

　　You can set the registry keys to be synchronized when failover occurs.　The registry under a specified key can be synchronized with the failover destination server when failover occurs.　The failover destination is the server that has been set in the failover policy.

　　The synchronization process is described below.

(1)　If a synchronization target registry key has been set for a failover group, the registry key is always monitored while the group is running.

(2)　When a change in the registry key is recognized, the registry under that registry key is saved to the hive file and stored on the local disk.　Also, it is distributed to the failover destination server.

(3)　The server that received the distributed information maintains it on the local disk and restores it for the relevant key before starting up the script when failover occurs.

* Precautions when synchronizing the Registry
  + Registry synchronization may fail in the following situation.
    = Failover occurs when the synchronization target registry key is open at the failover destination server (since hive file expansion fails).
  + Set only the minimum required registry keys for synchronization target registry keys.
* Registry keys to be synchronized
  Keys under the following two registry keys are available as registry key.
  Therefore, not all the registries can be synchronized.
  + Any key under HKEY_USERS
  + Any key under HKEY_LOCAL_MACHINE
    Excluding,
    Keys under HKEY_LOCAL_MACHINE¥SOFTWARE¥NEC¥ESMARM
    HKEY_LOCAL_MACHINE¥SOFTWARE¥NEC
    HKEY_LOCAL_MACHINE¥SOFTWARE
    HKEY_LOCAL_MACHINE
* Restrictions of synchronization target registry keys are as follows;
  + Up to 16 synchronization target registry keys are allowed per failover group.
  + Restrictions of registry key names are as follows;
    = Available characters are compliant with the OS registry key specifications.
    = Up to 259 bytes.

5.2.7　Applications/services

　　Cluster-supporting applications and services are started up on the other server with scripts if failover occurs or a failover group is moved.　Therefore, it is prerequisite that the same revisions of applications and services exist on all servers on the failover policy and all servers can handle them the same way. Also, they should be able to collect data to be passed between servers on shared disks.

　　Applications and services that can run under NEC EXPRESSCLUSTER should meet other conditions also.　See Chapter 4 "APPLICATIONS AND SERVICES IN AN NEC EXPRESSCLUSTER ENVIRONMENT" for details.

　　See the same chapter for prerequisites of client application, too.　See the chapters about relevant PPs in "*NEC EXPRESSCLUSTER System Construction Guide / [For PP]*" as well as Chapter 4 "APPLICATIONS AND SERVICES IN AN NEC EXPRESSCLUSTER ENVIRONMENT" for details of NEC EXPRESSCLUSTER -supporting PPs.

5.3 Main Causes of Failover

The main causes of failover are listed below.

* Server shutdown
* Power supply failure
* Operating system (OS) panic
* Operating system (OS) complete stall
* NEC EXPRESSCLUSTER Server problem
* Fault of an application or service that was started up by an NEC EXPRESSCLUSTER command (ARMLOAD) from a script
    + The application or the service continues until it is terminated with an NEC EXPRESSCLUSTER command called ARMKILL.
    + The fault is either of the following states.
        = Loss of process for applications
        = Stop (SERVICE_STOPPED) or abnormally termination of service for services
    + The following options can be specified for ARMLOAD.
        = Whether or not to monitor the application or service
        = Threshold of Restart count
        = Interval until the Restart count is cleared to zero
        = Whether to restart the script or application only.
        = The behavior when the threshold value is exceeded; server shutdown or failover
* Registration of a specific event in the event log
    + ESMPRO/Server Agent can be used to trigger a failover when a specific event is registered.
    + The command for shutting down a server (ARMDOWN) and the command for causing failover of a failover group (ARMFORVER), which are provided by NEC EXPRESSCLUSTER, can be written in scripts so that ESMPRO/Server Agent can monitor a specific event and failover occurs when that event occurs.
* Detection of problem in the Public LAN or resources being monitored by the resource monitoring or ARMRSP command
* Detection of stall (no response for a certain period of time) of software being monitored by NEC EXPRESSCLUSTER watching option (for value).

6   RESOURCES
6.1   Disk Resources

6.1.1   Switched partition

A switched partition is an accessible partition to the primary server and is placed on the shared disk connected to multiple servers forming a cluster.

Switching is performed according to the failover policy for each failover group.   By saving the data required for an application in advance in the switched partition, it is automatically handed over when a failover occurs or the failover group is moved.

Make sure the file system of the switched partition is NTFS.   Also, assign the same drive letter to the same area at all servers.

See "*NEC EXPRESSCLUSTER System Construction Guide / Cluster Generation Guide (Shared disk)*" for details to set up the Switched Partitions.
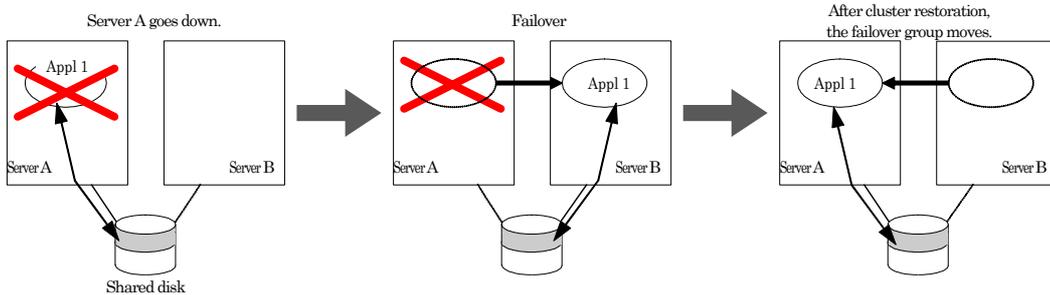
6.1.2　　Shared partition

A shared partition is on a shared disk connected to multiple servers that form a cluster, and accessible simultaneously.

* **Disk connection of shared partitions is supported only for parallel cluster systems that use Oracle Parallel Server**

By constructing the database on the shared partition, it can be passed automatically when failover occurs or a failover group is moved.

A shared partition must be a RAW partition.　　Do not format it.

Also, assign cluster characters to a shared partition from the *NEC EXPRESSCLUSTER disk administrator*.　　Do not assign a drive letter to it.　　For the cluster characters, assign the same character string to the same area at all servers.　　Up to 13 characters are allowed for cluster character strings.

See "*NEC EXPRESSCLUSTER System Construction Guide/ Cluster Installation and Configuration Guide*" for details to assign cluster letters.
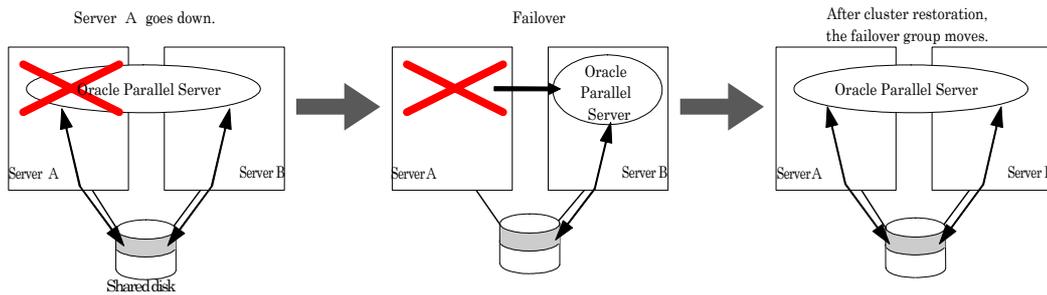
Terms are differentiated hereinafter as described in the table blow.

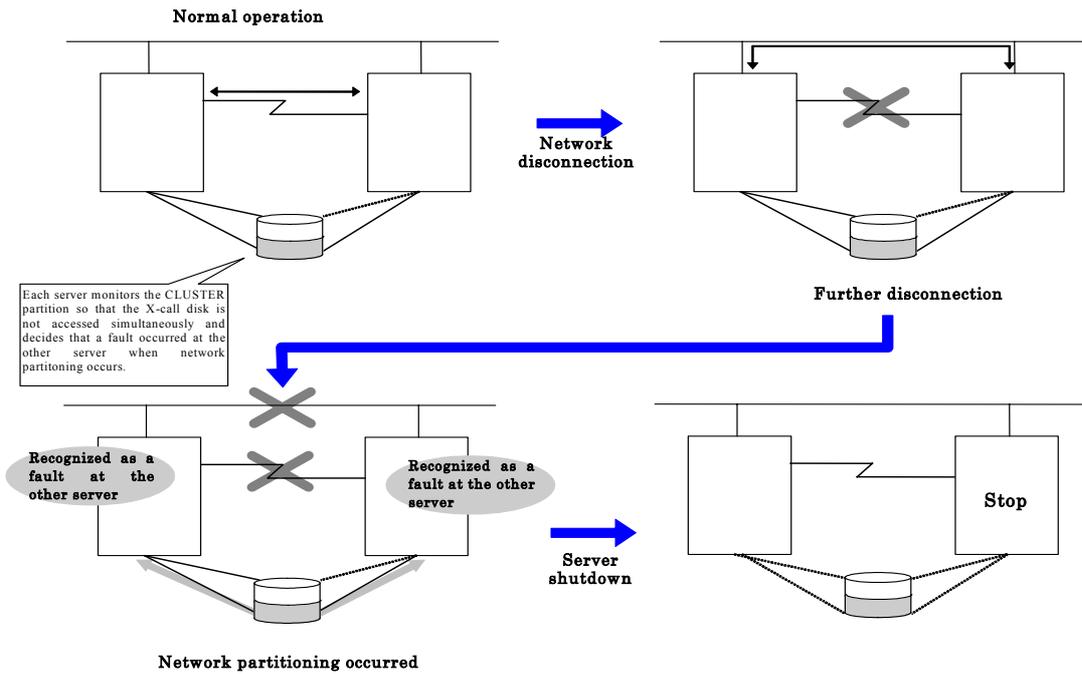|	Locations to assign icons	Abbreviations used herein
[Windows 2000] Disk management	[Start]-> [Programs] -> [Administrative Tools] -> [Computer Management]	Disk administrator
[Windows　Server 2003] Disk management	[Start] -> [Administrative Tools] -> [Computer Management]	
NEC EXPRESSCLUSTER Disk administrator	[Start] -> [Programs] -> [ActiveRecoveryManager server]	*NEC EXPRESSCLUSTER Disk administrator*

<Note> Do not start Disk administrator or *NEC EXPRESSCLUSTER Disk administrator* on servers simultaneously.

6.1.3 CLUSTER partition

The dedicated partition that NEC EXPRESSCLUSTER Server uses to detect network petitions[2] is called "CLUSTER Partition". At least one CLUSTER partition is required for each cluster. It is recommended to assign a CLUSTER partition per shared disk.

The CLUSTER partition must be a RAW partition. Do not format it.

Also, you should assign cluster characters from the *NEC EXPRESSCLUSTER Disk Administrator*. Do not assign a drive letter. Cluster characters representing a CLUSTER Partition starts with "###NEC_NP" and is up to 13 characters. The cluster characters should be the same character string to the same area at all servers.

Normal operation

Each server monitors the CLUSTER partition so that the X-call disk is not accessed simultaneously and decides that a fault occurred at the other server when network partitoning occurs.

Network disconnection

Further disconnection

Recognized as a fault at the other server

Recognized as a fault at the other server

Server shutdown

Stop

Network partitioning occurred

---

[2] Disconnections of heartbeats on all interconnects due to LAN errors.
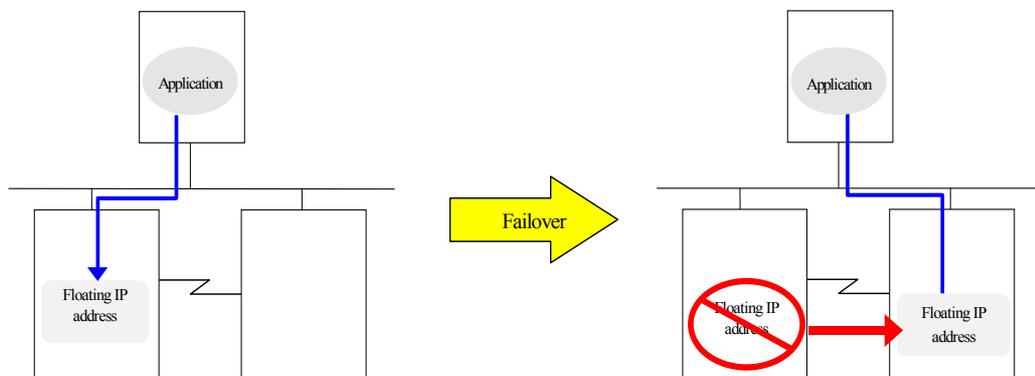
6.2　　Floating IP address

　　The floating IP address is a new mode of virtual IP address that is supported from Ver4.2.　See "*NEC EXPRESSCLUSTER System Construction Guide / System Design (Advanced)*" for details of conventional virtual IP address.

　　Unless the LAN connection should be duplicated, it is recommended to use the floating IP (FIP) address mode.　This section describes its details.
　　Oracle Parallel Server (OPS) Ver8.1.5 and higher support FIP.　FIP is not available for earlier versions.　Use virtual IP addresses for earlier versions of OPS.

　　Client Applications can connect to cluster servers with Floating IP address.　FIP is usable also to connect servers.　When a floating IP address is used, the client need not be aware of the switching of the connection destination server even if failover occurs or a failover group is moved.

　　Floating IP addresses are available on the same LAN and from a remote LAN.　It can even be used from a non-Windows operating system as long as the operating system supports the ARP protocol.

6.2.1　　Assigning addresses

　　An IP address assigned for a floating IP address must satisfy the following condition.

```
*   It is within the same network address as the LAN to which the cluster servers belong, and
    it is a host address that is not being used
```

　　Reserve the required number (generally, equivalent to the number of failover groups) of IP addresses satisfying this condition.
　　Since this IP address does not differ from an ordinary host address, it can be a global IP addresses such as the Internet.

6.2.2   Environment setup

To use a floating IP address, the following settings are required.

+ Assign an IP address to a failover group with NEC EXPRESSCLUSTER Manager

After the cluster is generated, select [Add failover group] - [Set resources] - [IP address] - [Floating IP address] from NEC EXPRESSCLUSTER Manager and then select the IP addresses to be used.

To change a floating IP address value, select [Failover group properties] - [Set resources] - [IP address] and then add a new FIP address and delete the old FIP address.

Floating IP addresses depend on LAN adapter driver registry at the Public-LAN. Before changing a LAN adapter model or updating a LAN adapter driver, you should stop NEC EXPRESSCLUSTER. Change the startup mode of related services to manual, and select [Cluster Shutdown] – [Reboot] and stop NEC EXPRESSCLUSTER.

6.2.3   Routing

Since the routing for the network IP addresses to be used for a server also controls routing of floating IP addresses, no special routing control for floating IP addresses is necessary.

6.2.4   Association with virtual computer name

If an FIP satisfies certain conditions, it should be associated with a virtual computer name.
See Section "6.3.1.3 Associate a virtual computer name and a floating IP address" for details.

6.2.5   Conditions to use

Floating IP addresses are accessible to the machines listed below.

- Cluster servers
- Other servers in the same cluster, and servers in other cluster systems
- Clients on the same LAN as the cluster server and on remote LAN

In addition, floating IP addresses are available on non-Windows machines as long as the following conditions are satisfied.[3]

- The communication protocol is TCP/IP
- ARP protocol is supported

The floating IP address mechanism will function without problems even on a LAN in which a switching HUB is used.
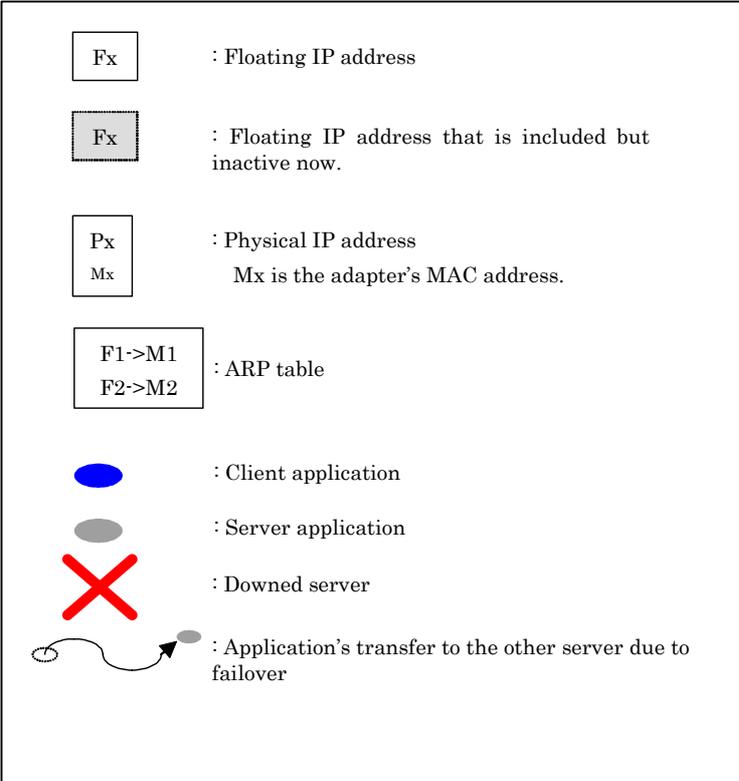When a server goes down, a connected TCP/IP connection is disconnected.
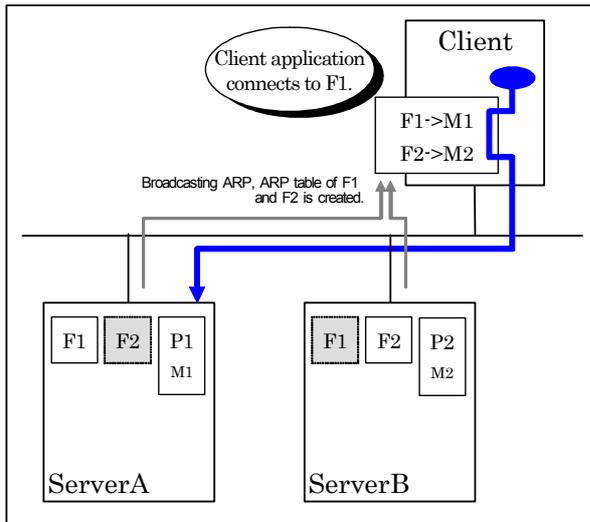Floating IP addresses are accessible to machines on a remote LAN, too.

---

[3] Connections are not assured on all machines or architectures. Assess non-Windows OSes yourself before using.

6.2.6　　Connections format by floating IP address

This section describes types of connections by FIP address.　The symbols used in the figures are defined as follows.

Fx	：Floating IP address
Fx	：Floating IP address that is included but inactive now.
Px Mx	：Physical IP address 　　Mx is the adapter's MAC address.
F1->M1 F2->M2	：ARP table
⬤	：Client application
⬤	：Server application
✕	：Downed server
	：Application's transfer to the other server due to failover

(1)   **Connection from client to server**

Connection format
* Client specifies the server's IP address to connect.

Connection method
* FIP addresses are used as connection destination.
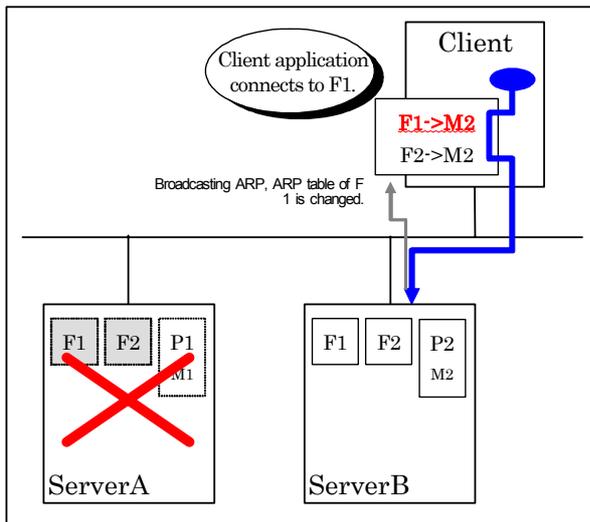
Operation when failover occurs
* The ARP table of client about FIP addresses is changed in the event of failover.
  The client can reconnect to the server with the same FIP address.

*   If clients use FIP addresses to connect to servers, they do not need to know the change of connected server in the event of failover.

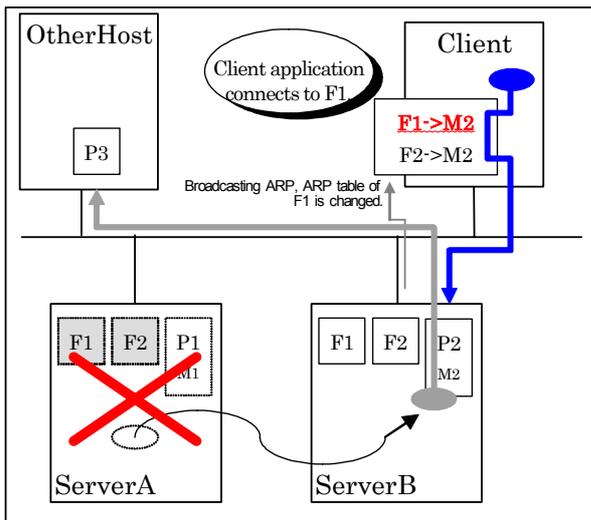**(2)   Connection to other host due to reception of request from client**
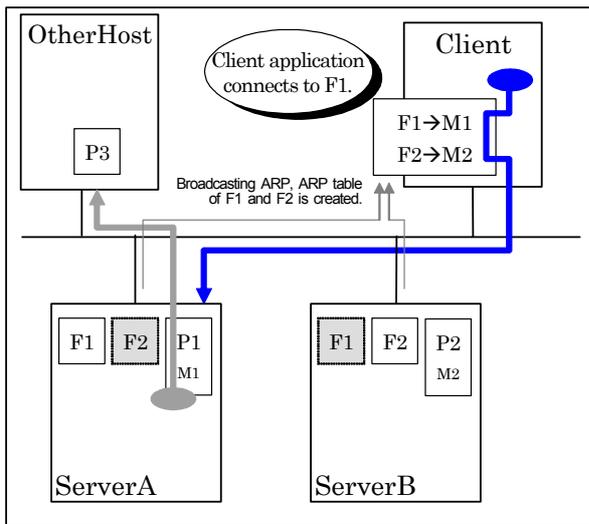
Connection format
* The client application connects to the server application. The server application receives a request from the client application to connects to another host, and sends back the result to the client application.

Connection method
* The client application connects to the server application with FIP address.
* When the server application connects to another host due to a request from the client, a physical IP address is used.
* You should make settings in advance of the host connected from the server application so that it accepts requests from the physical IP address of either server.
* The server application and client application are connected with FIP address.   And the server application and other host are connected by using a physical IP address.
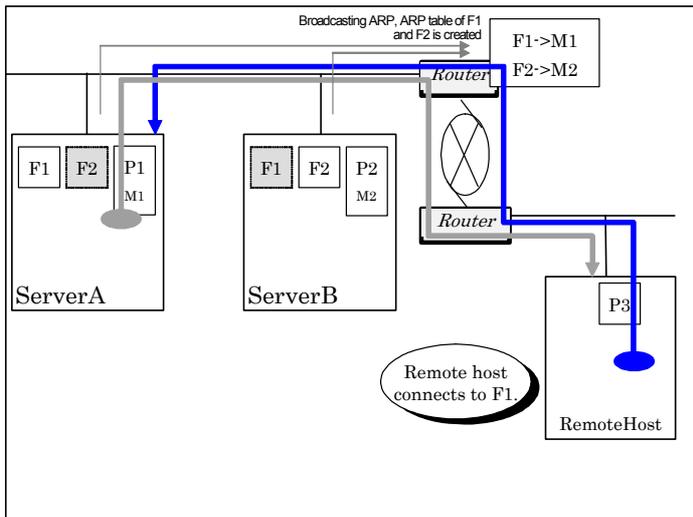

Operation when failover occurs
* The ARP table of client about FIP addresses is changed in the event of failover.   The client can reconnect to the server with the same FIP address.   Upon the request from the client, the server application at the failover destination connects to the other host.

*   Physical IP addresses should be used to connect servers with other hosts.   It is not required to explicitly bind an FIP address to connect servers with other hosts.

**(3)   Connection to a Non-Windows host on a remote network**

Connection format
* The client application connects to the host on a remote network (referred to as a remote host hereinafter). Also, the remote host connects to the server application.

Connection method
* Set up the remote host so that it can accept connection requests from the physical IP address of either server.
* The server application connects to the remote host with a physical IP address.
* The remote host uses an FIP address to connect to the server application.
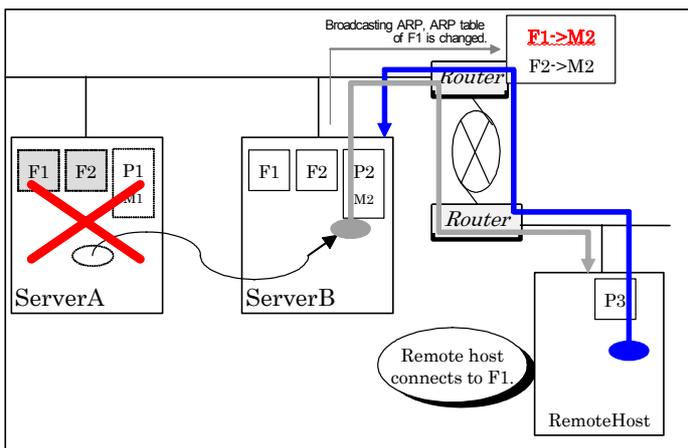
Operation when failover occurs
* The ARP table about FIP addresses is changed on the router on the cluster server side LAN in the event of failover. The remote host can reconnect to a new server with the same FIP address. Also, the failover destination server can reconnect to the remote host.

6.3 Virtual Computer Name

A client application can connect to a cluster server by using a virtual computer name. Also, virtual computer names can be used for connections between cluster servers. When virtual computer names are used, the client need not be aware of the switching of the connection destination server even if failover occurs or a failover group is moved.

Virtual computer names are available only to Windows machine clients to connect to a cluster server.
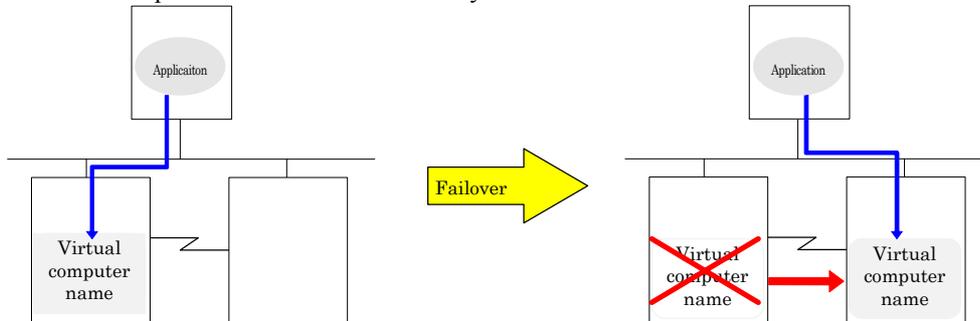
6.3.1 Assigning virtual computer name

6.3.1.1 Determining virtual computer names

A computer name assigned for a virtual computer name must satisfy the following conditions.

* The name should be different from the server name.
* The name should be different from the computer names of machines connected to the same network segment.
* The name can be up to 15 characters.
* Only alphanumeric characters and hyphen (-) are allowed.
* It is case-insensitive (however, supplied upper and lowercases are maintained directly).

6.3.1.2 Setting virtual computer names

Set the virtual computer name by selecting [Add failover group] - [Set resources] or [Failover group properties] - [Set resources] from NEC EXPRESSCLUSTER Manager.
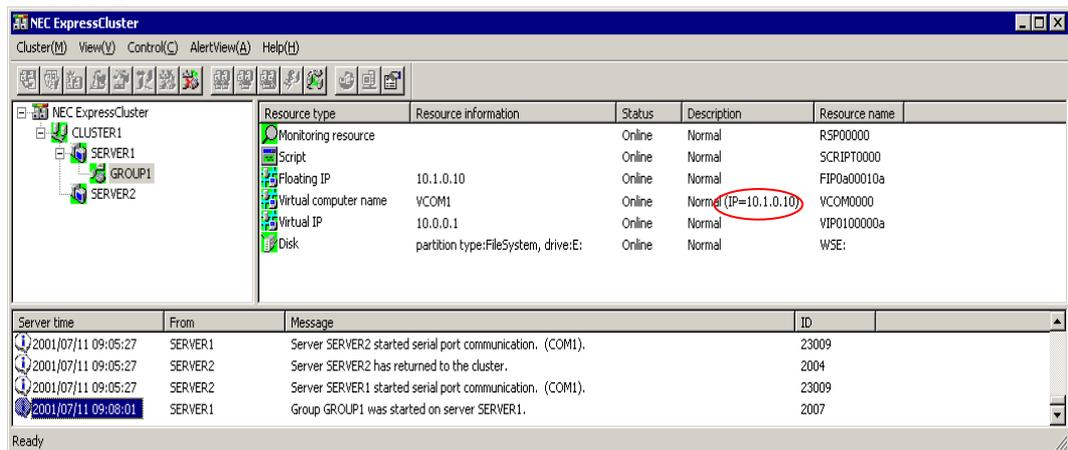
6.3.1.3　　　Associate a virtual computer name and a floating IP address

If the following conditions are met, NEC EXPRESSCLUSTER associates a virtual computer name with a floating IP (FIP) address:

* Only one virtual computer name exists within a failover group where FIP(s) is registered:

    (1)　　If one virtual computer name and one FIP address exist within the same failover group, NEC EXPRESSCLUSTER associates them automatically.

    (2)　　If one virtual computer name and several FIP addresses are registered within the same failover group, NEC EXPRESSCLUSTER associates the virtual name with one of these FIP addresses automatically.　You cannot specify an FIP address to associate with (When you start a group, you don't know which address NEC EXPRESSCLUSTER will choose.). However, it is assured that NEC EXPRESSCLUSTER uses the same FIP addresses at the time of failover.
    If FIP addresses on a different network address are registered, you should cancel associations with FIP addresses.　See Section "6.3.1.4 Cancel/reset an association of a virtual computer name with an FIP address." for details.

If a virtual computer name is associated with an FIP address, the NEC EXPRESSCLUSTER Manager displays the associated IP address in the "Description" section of the virtual computer name.　The Manager displays the IP addresses only when the virtual computer name is in normal operations or when activation failed.

If a virtual computer name isn't associated with an FIP address, the NEC EXPRESSCLUSTER Manager doesn't display an IP address in the "Description" section of the virtual computer name.

There are the following benefits if a virtual computer name and an FIP address are associated:
* A pair of virtual computer name and FIP address can be described in a client's LMHOSTS file.

Once a virtual computer name and FIP address are associated, you can lift some of the restrictions concerning the virtual computer name as discussed from the following sections.

6.3.1.4   Cancel/reset an association of a virtual computer name with an FIP address.

Use the ARMVCSET command to cancel/reset an association of a virtual computer name with an FIP address. See "*NEC EXPRESSCLUSTER System Construction Guide / Command Reference*" for details of commands.

6.3.1.5   Number of allowed virtual computer names

One cluster system can use up to 64 virtual computer names.
However, if a virtual computer name is associated with a floating IP address, it will be less than 64. The maximum number will be (64 – NICs that are bound to WINS clients).

6.3.2   Usage from remote LAN

6.3.2.1   If associated with FIP address

If a virtual computer name is associated with an FIP address, a client's LMHOSTS can contain an entry of virtual computer name and FIP address to use the virtual computer name from a remote LAN.
Virtual computer names can work in conjunction with WINS. (see Section "6.3.2.3 How to make settings of WINS server" for details).

6.3.2.2   If not associated with FIP address

If a virtual computer name is not associated with an FIP address, it should work with WINS to make virtual computer names available from remote LAN (see Section "6.3.2.3 How to make settings of WINS server" for details). It is not available from remote server in conjunction with LMHOSTS.

6.3.2.3 How to make settings of WINS server

To make settings of a WINS server, take the following steps:

* To set up a WINS server on two cluster servers:
  (1) Install a WINS server on two cluster servers. (When a dialog asks whether to restart the server after the setup, select [No].
  (2) [For Windows 2000]
       - Select [Control Panel]-[Network and Dial-up Connections] and open the [Properties] of icon representing the public LAN.
       - Select [Internet Protocol (TCP/IP)] and click [Properties].
       - Click [Advanced] and select [WINS] tab.
       - Add public LAN's IP addresses for the two servers in the WINS address fields. (You don't have to differentiate priorities between primary and secondary.)

       [For Windows Server 2003]
       - Open the [Properties] of local area connections representing public LAN from the [Control Panel]-[Network Connections].
       - Select the [Internet protocol (TCP/IP)] and click the [Properties].
       - Click the [Advanced] and select the [WINS] tab.
       - Specify IP addresses of public LAN in WINS address for two servers (you do not need to differentiate their priorities).

  (3) Make the same settings on both servers.
  (4) Perform cluster shutdown and reboot.
  (5) Likewise, add public LAN's addresses for the two servers on the client on remote LAN.

* To set up a WINS server on a server that isn't part of the cluster:
  (1) Install a WINS server that is independent from cluster servers.
  (2) [For Windows 2000]
       - Select [Control Panel]-[Network and Dial-up Connections] and open the [Properties] of icon representing the public LAN.
       - Select [Internet Protocol (TCP/IP)] and click [Properties].
       - Click [Advanced] and select [WINS] tab.
       - Add an IP address for the WINS server in the WINS address field.

       [For Windows Server 2003]
       - Open the [Properties] of local area connections representing public LAN from the [Control Panel]-[Network Connections].
       - Select the [Internet protocol (TCP/IP)] and click the [Properties].
       - Click the [Advanced] and select the [WINS] tab.
       - Specify the IP address of WINS server in WINS address.

  (3) Make the same settings on all cluster servers.
  (4) Perform cluster shutdown and reboot.
  (5) Likewise, add public LAN's addresses for the two servers on the client on remote LAN.

6.3.3   Requirements for using virtual computer names

Virtual computer names are accessible to the machines listed below.   The virtual computer name mechanism will function without problems even on a LAN at which a switching HUB is used.

(1)   Cluster server

To Use virtual computer names on Windows 2000 or Windows Server 2003, see Section 6.3.7 "To use virtual computer names on Windows 2000, Windows XP" for details.

(2)   Client

If the following conditions are satisfied, virtual computer names are available.
* The OS is either Windows Server 2003, Windows 2000, Windows NT, Windows 98, Windows Me, Windows XP.
* The communication protocol between servers and clients is TCP/IP or NetBEUI.   (There are some notes and restrictions for NetBEUI.   See Section 6.3.6 "Precautions concerning the use of virtual computer names" for details.)

If you want to use virtual computer names on either Windows Server 2003, Windows 2000, Windows XP, Windows 98 or Windows Me, see Section 6.3.7 "To use virtual computer names on Windows 2000, Windows XP", 6.3.8 To use virtual computer names on Windows 98" or 6.3.9 "To use virtual computer names on Windows Me" for details.

Make the following settings for Windows3.1 clients to connect to network drive with virtual computer name.
However, the client cannot access the network drive across a router.
* Install the NetBEUI protocol on NEC EXPRESSCLUSTER Server.   (There are some notes and restrictions for NETBEUI protocol.   See Section 6.3.6 "Precautions concerning the use of virtual computer names" for details.)
* Enable the NetBEUI protocol on LAN Manager of client.
(This can coexist with the TCP/IP protocol.)

(3)   WINS Client protocol

Each NEC EXPRESSCLUSTER Server should bind the WINS Client protocol to one network adapter only (typically bound to public LAN).

6.3.4   Available services with virtual computer names

The following 5 services are usable;
* Name resolution of TCP/IP (Computer name -> IP address)
* Network drive connection
* Network printer connection (there are some restrictions)
* Named pipe
* RPC(Named pipe, TCP/IP)

6.3.5   Unavailable services with virtual computer names

The following two services cannot be used.
* Mail slot
* RPC(NetBIOS)

6.3.6   Precautions concerning the use of virtual computer names

* Network printers cannot be connected from Control Panel.   Use the net use command.
    net use command: This command is used to connect or disconnect the computer to/ from shared
                         resources, or display information related to computer connections.   This is
                         used to control network reconnections at the next logon.
                         See Help of OS for details of commands.

  If an FIP address is not associated with a virtual computer name, there are following notes and
restrictions.
* It may take several minutes to be able to reconnect after failover due to downed server.   You can
  work out this problem by installing NEC EXPRESSCLUSTER Client.
* It may take several minutes to display virtual computer names on network computer after the cluster
  started.
* You cannot describe virtual computer names in LMHOSTS.

  If an FIP address is associated with a virtual computer name, there are following notes and restrictions.
* NetBEUI protocol is not usable.   You should cancel the association with FIP addresses to use
  NetBEUI protocol.   See Section 6.3.1.4 "Cancel/reset an association of a virtual computer name
 with an FIP address." for details.

There are the following notes and restrictions if the cluster servers run on Windows Server 2003 or Windows 2000

* Verify the followings and make settings if necessary before registering virtual computer names to WINS server on remote network.

Make following settings for binding of Interconnect and mirror connect.
- Select [Control Panel] - [Network and Dial-up Connections].
- Select [Advanced] - [Advanced Settings] - [Adapters and Bindings] tab.
- Put the public LAN (the network adapter where the address of WINS Server is registered) first on the Connections list.

Below is the image of [Adapters and Bindings].
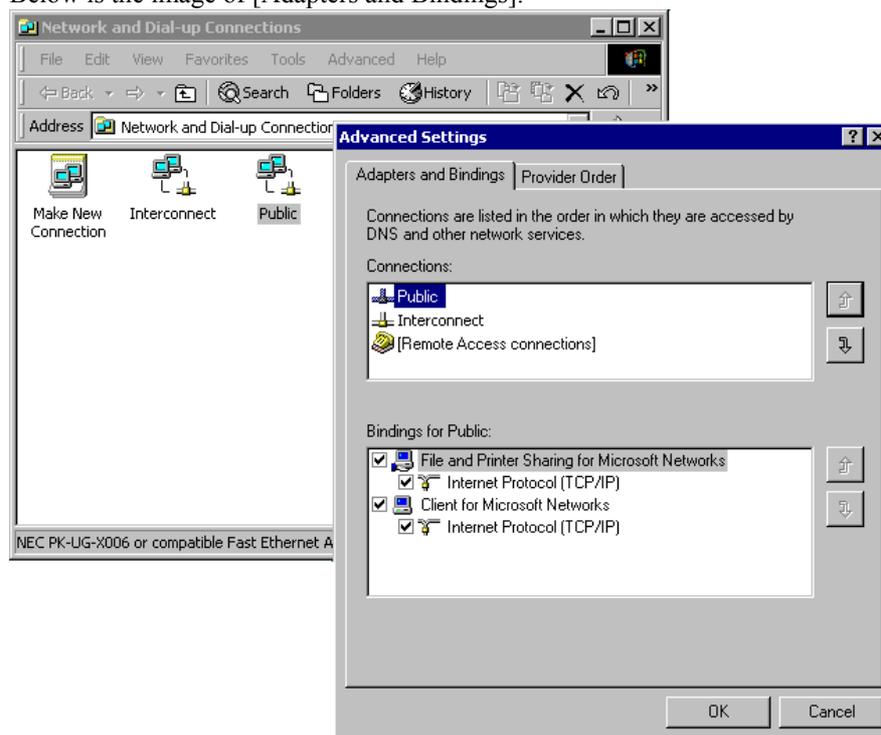
6.3.7 To use virtual computer names on Windows 2000, Windows XP or Windows Server 2003

Virtual computer names that meet the following condition are not available on cluster servers running on Windows 2000, Windows Server 2003 or Windows XP.

* A virtual computer name is registered in the cluster where its server is included. And the virtual computer name is active on other servers.

The following settings are required to use virtual computer names on Windows 2000, Windows Server 2003 or Windows XP.

If you associated virtual computer names with FIP addresses, you do not need to make the following settings. In that case, the following restrictions are lifted.

* Install NEC EXPRESSCLUSTER Client and make settings for the Client.
* If you have made settings to use DNS server and the server works with WINS, failover switching will be disabled while cached information on virtual computer names remain on the DNS server. Shorten the cache retention period for WINS records on the DNS server to about one second.

6.3.8 To use virtual computer names on Windows 98

If you associated virtual computer names with FIP addresses, this restriction is lifted.

To use virtual computer names on a Windows 98 client, you may need to install the CLUSTERRPO Client.

If its version number shown in [System] on the [Control Panel] is "Windows 98 Second Edition 4.10.2222 A", you should install the CLUSTERRPO Client.

6.3.9 To use virtual computer names on Windows Me

If you associated virtual computer names with FIP addresses, this restriction is lifted.

To use virtual computer names on Windows Me, you need to install the CLUSTERRPO Client.
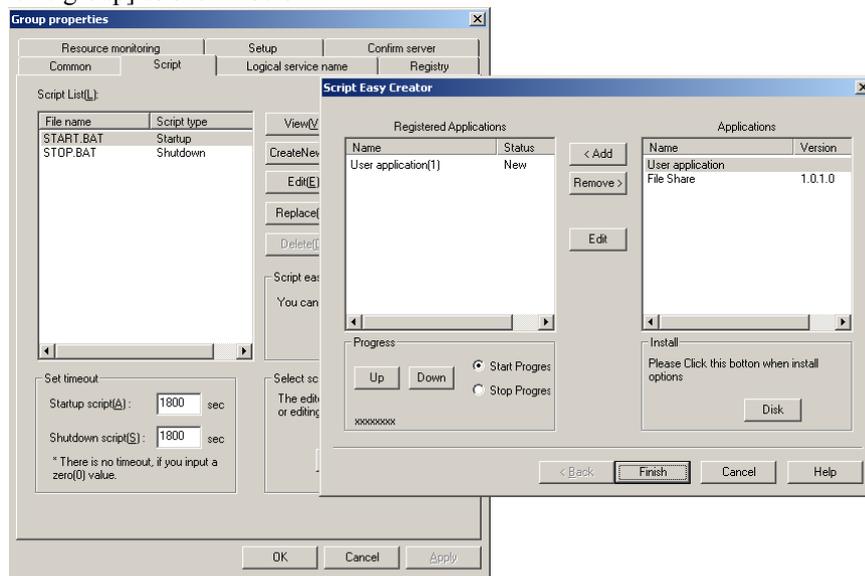
6.4　　Script

Scripts control clustered applications and services in NEC EXPRESSCLUSTER.　　NEC EXPRESSCLUSTER manages the scripts. Scripts are used to start, stop or fail over applications and/or services, and to move failover groups and to restore clusters.

Scripts have the same format as Windows batch files.　Therefore, you can describe processes best suited to each application or service.　By describing NEC EXPRESSCLUSTER commands in scripts, more productive functions can be provided.

"Script easy creation" provides a simple way to make scripts.

"Script easy creation" is started from [Group] – [Properties] of NEC EXPRESSCLUSTER Manager or [Create group] as shown below.
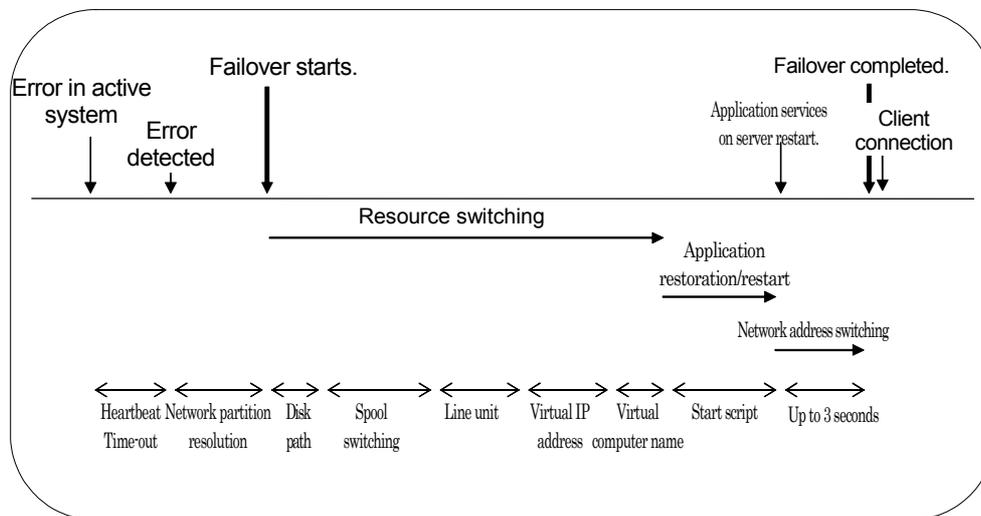

Separately released options are required to use the "Script easy creation".　The file sharing is a standard function.
See "*NEC EXPRESSCLUSTER System Construction Guide / GUI Reference*" for details of operations.

# 7 APPENDIX

7.1 Switching Time When a Server Goes Down

* Heartbeat timeout
  This is the period of time until a fault is detected by the standby server when it occurs at the primary server.   This time can be changed.
  The factory setting is 30 seconds; 3 seconds (fixed) x 10 times (default value).
* Network partition resolution
  It will take either of longer time from the below 1 or 2 to resolve a network partition problem.
  1. Approximately once or twice of heartbeat timeout are required to resolve a network partitioning problem. (When the 30-second default value is set for the heartbeat timeout, it takes roughly 30 to 60 seconds.)
  2. It takes about once or twice of disk I/O waiting time.   (If it is set to the default, 80 seconds, it takes roughly 80 to 160 seconds.)
* Resource switching (The time is just a rule of thumb.)
  If there are multiple resources to fail over, calculate the time by multiplying the following times by the number of resources.
  + Disk path switching(shared disk)
    It takes about 1 minute.
  + Spool switching
    It takes about 5 minutes.
  + Line unit switching
    It takes about 3 minutes.
  + Virtual IP address switching
    It takes about 6 minutes.
  + Floating IP address switching
    It takes about 6 minutes.
  + Virtual Computer Name switching
    It takes about 1 minute.[4]

  + Resource monitoring
    It takes about 1 minute.
* Start script execution time
  This time includes the startup times of applications and services, rollback time of database and so on.

---

[4]Switching takes about 1 minute in an ideal case.   However, it takes 9 minutes in typical setup.   The time required to activate virtual computer name at failover depends on the number of installed network adapters.

You can predict the rollback time to some extent by adjusting the checkpoint interval.   See documents of each database for details.

# 7.2   Diskcheck from Standby System

With the SCSI bus separated, it is possible to confirm the status of SCSI bus of standby server while the application is running on the active server, and increase the system availability.
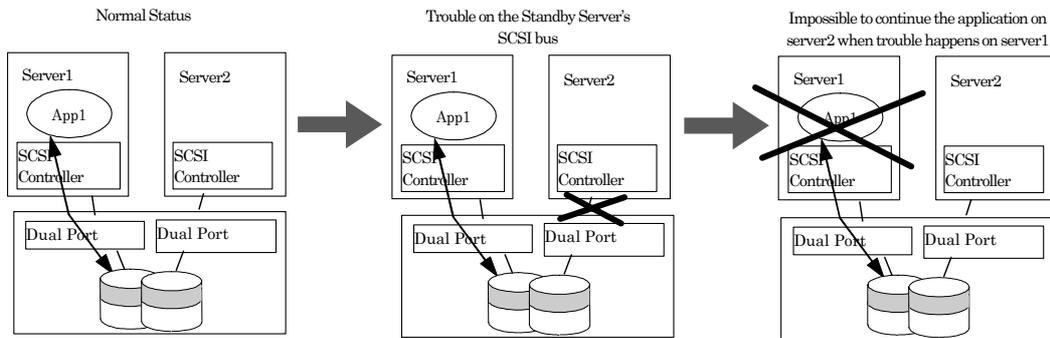
If NEC SCSI shared disk device (N8541-18,N8590-20,N8590-28, N8590-29) or such is used on an one-way standby cluster, the status of SCSI bus on a standby server can be monitored. It is possible to detect a SCSI bus problem on a standby server with NEC EXPRESSCLUSTER's commands and resolve the problem promptly to increase availability.

（1）   Make a failover group on a standby server.
（2）   Put the following line in the start script.
    ex.) ARMLOAD WatchID /R 9 /H 1 /FOV ARMRSP /D
（3）   Put the following line in the stop script.
     ARMKILL WatchID

When a failover happened on the failover group above, it means a trouble happened on the SCSI controller or such, not on the shared disk device.
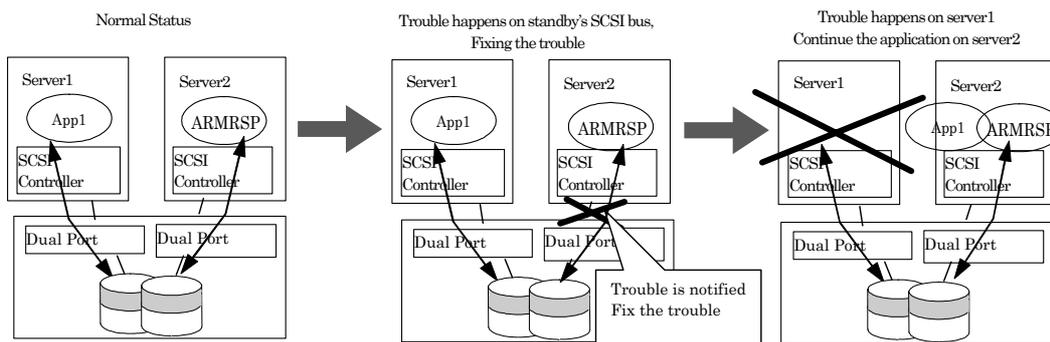
7.3 Monitoring Stall of OS

You can detect partial stall of OS by monitoring accessibility of share names with an NEC EXPRESSCLUSTER command (ARMWHSHR).   Detection of stall can be notified to NEC EXPRESSCLUSTER Manager and printed in event logs so that you can take necessary actions earlier, which improves the system's availability.   See "*NEC EXPRESSCLUSTER System Construction Guide / Command Reference*" for details of NEC EXPRESSCLUSTER commands.

Procedures to make settings to monitor partial stall of OS are as follows;

(1) Register a user account that has Administrator rights with NEC EXPRESSCLUSTER Manager.

(2) Create a failover group per server to monitor partial OS install.
* How to set a failover group
1. Deselect all of [Resource type of monitoring] on [Resource monitoring] tab.
2. Add a server (the failover group's server) only on [Failover policy]-[Startup of group server].

(3) Add the ARMLOAD command line in the start script so that it will be executed when ARMS_EVENT starts to start the ARMWHSRW command.

(4) Add the ARMKILL command line in the stop script so that it will be executed when ARMS_EVENT starts to stop the ARMLOAD command described in the start script.

Sample configuration
The following configuration and script samples take a 2-server cluster system as example.
Both servers monitor share names on the other side to monitor partial stall of OS each other.

* Configuration sample
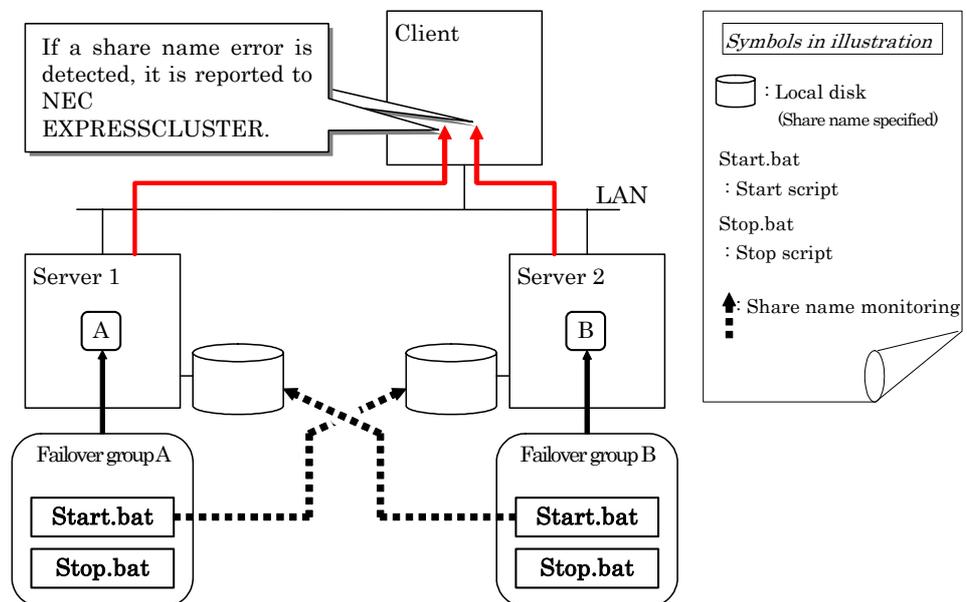
- Server information

Items	Server 1	Server 2
Server name	server1	server2
Public LAN IP address	10.10.10.1	10.10.10.2
Share name	(1)share1	(1)share2
	(2)share3	

* Script samples
  (1) Start.bat for Server 1

```
IF "%ARMS_EVENT%" == "START" GOTO NORMAL
GOTO EXIT
:NORMAL
ARMLOAD W1 /U Administrator ARMWHSHR ¥¥server2¥share2 10.10.10.2
:EXIT
EXIT
```

  (2) Stop.bat for Server 1

```
ARMKILL W1
EXIT
```

  (3) Start.bat for Server 2

```
IF "%ARMS_EVENT%" == "START" GOTO NORMAL
GOTO EXIT
:NORMAL
ARMLOAD W2 /U Administrator ARMWHSHR ¥¥server1¥share1 10.10.10.1
ARMLOAD W3 /U Administrator ARMWHSHR ¥¥server1¥share3 10.10.10.1
:EXIT
EXIT
```

  (4) Stop.bat for Server 2
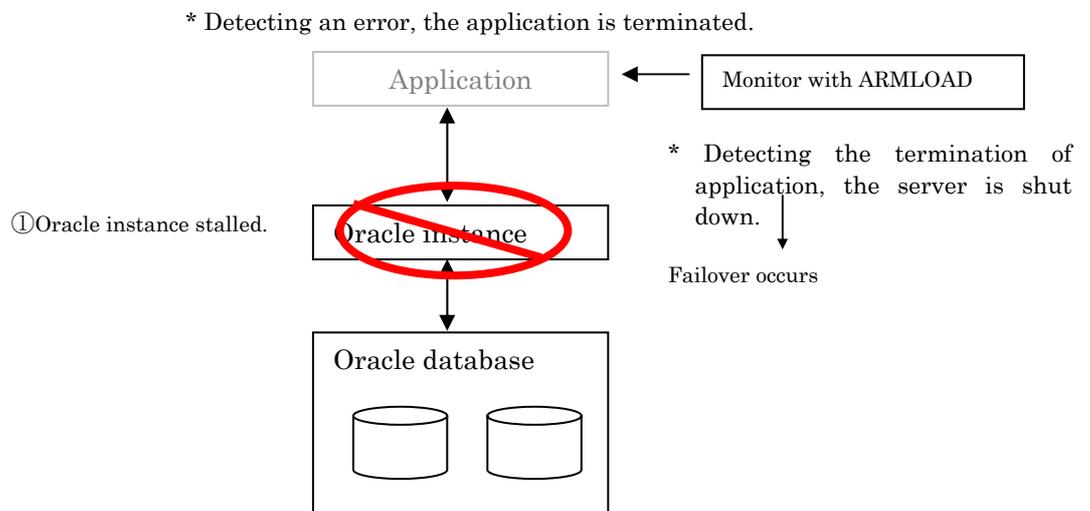
```
ARMKILL W2
ARMKILL W3
EXIT
```

For 3-server or more-server cluster systems (server names: A, B, C....n), you should describe scripts so that Server A monitors all other servers than itself, Server B monitors all other servers than itself ...Server n   monitors all other servers than itself.

7.4　　Countermeasures against Application Errors

NEC EXPRESSCLUSTER commands monitor application programs' services.　However, they cannot trigger failovers when an application program returns errors if its services themselves are normal. You can make failover occur in this case by creating an application program to monitor results of such application programs.

The following sample takes stall of Oracle instance as example.

(1)　Make an application program that accesses Oracle database on regular basis and quits itself if it detects an error on Oracle such as stall of Oracle instance.
(2)　Start the above application with the ARMLOAD command from a script.　The command monitors it, and if it is terminated, the command shuts down the server, which triggers a failover.

* Detecting an error, the application is terminated.

Application　←　Monitor with ARMLOAD

* Detecting the termination of application, the server is shut down.

①Oracle instance stalled.　　Oracle instance

Failover occurs

Oracle database

Command sample:
　　ARMLOAD WatchID /R 2 /H 1 AppName
　　*AppName represents an application program to detect the above Oracle instance errors.

See "*NEC EXPRESSCLUSTER System Construction Guide / Command Reference*" for details of NEC EXPRESSCLUSTER commands.

7.5   Working Together with NEC ESMPRO/Server Agent

Using the event log monitoring function of NEC ESMPRO/Server Agent, event logs representing system errors can trigger failover. Before starting the event log monitoring, you should set up NEC ESMPRO/Server Agent on cluster servers.

See the help and manuals of NEC ESMPRO/Server Agent for details.