Automation Technology for Virtualized Systems

● Takashi Fujiwara ● Yoshinobu Hibi ● Shigehiro Yoshikawa
● Hiroshi Yazawa

(Manuscript received May 31, 2007)

There has been a recent trend toward consolidating burgeoning systems onto blade servers in order to optimize IT infrastructures, maximize system investment efficiency, and reduce server power consumption. This trend heightens the demand to increase the efficiency of system integration and operation. Moreover, in line with the increased performance of multi-core CPUs and advances made in virtualization technology, more systems are being consolidated on blade servers. This consolidation entails making configuration changes, while balancing overall system resources and maintaining stable operation. Consequently, such consolidation is imposing a heavier workload on system managers. To address this problem, Fujitsu developed the Systemwalker Resource Coordinator Virtual server Edition (SRC VE) to reduce the system manager workload, and dramatically increase the manageability and availability of blade server systems. This paper describes the conventional technologies of SRC and introduces the following technologies of the new SRC VE package: integrated server management, Storage Area Network (SAN) virtualization management, and server orchestration for automatic resource assignments and changes.

1. Introduction

There has been a recent trend towards consolidating burgeoning systems onto blade servers in order to optimize IT infrastructures, maximize system investment efficiency, and reduce server power consumption. This trend also heightens the demand to increase the efficiency of system integration and operation.

Given the increased performance of multi-core CPUs and advances made in virtualization technology, more systems are now being consolidated on blade servers. This consolidation entails making configuration changes, while balancing overall system resources and maintaining stable operation. Consequently, such consolidation is imposing a heavier workload on system managers.

To address this problem, Fujitsu developed Systemwalker Resource Coordinator Virtual server Edition (SRC VE)[1] for the virtualization of blade server systems, by enhancing Systemwalker Resource Coordinator (SRC). SRC VE is now commercially available.

This new SRC VE package reduces the workload of system mangers relative to server consolidation, and significantly improves the manageability and availability of blade server systems by utilizing such automation technologies as automatic server recovery for system operation. This package is positioned as a core software product for infrastructure optimization.

This paper describes the basic technologies of SRC VE.[2]

2. Functions of SRC VE

SRC VE operates as server virtualization management software to provide the following benefits:

2.1 Unified, visualized control of physical and virtual blade servers

SRC VE provides unified operation management views so that the system administrator can collectively manage an environment including both physical and virtual servers. In other words, physical and virtual servers can be monitored and operated using the same operation procedures. Without having to consider the differences between physical and virtual servers, users can employ the same operation procedures for identifying virtual servers affected by hardware errors, and handling such routine jobs as starting and stopping servers, and making backup copies of servers.

2.2 Reduction of server management workload by enhancing the Storage Area Network (SAN) of blade servers

Fujitsu developed the SAN connection virtualization function in collaboration with Emulex Corporation, a major manufacturer of fibre-channel devices. With the SAN connection virtualization function, the SAN configuration need not be changed when adding or replacing a blade server. This means the server manager does not need to negotiate with SAN managers, and thus server management workload is reduced. This virtualization function can be applied to both Fujitsu's "ETERNUS" disk array and the existing storage environments of other manufacturers.

2.3 Highly available blade server systems at low cost

SRC VE provides highly available blade server systems at low cost regardless of a physical or virtual environment, by utilizing server virtualization technology and SAN environment virtualization technology as follows:

- SRC VE automatically continues operations from error detection to job startup on a spare server without operator intervention. Since this automation function is provided as standard, neither the purchase of additional software products nor a complicated scripts program is needed.
- In conjunction with VMware HA (highly available function of VMware), one spare blade server can be shared between SRC VE and VMware HA in a mixed physical-virtual environment. Therefore, a more economical, highly available system ($N + 1$ standby system) can be configured than with other server management software products.

3. Technologies of SRC VE

SRC VE is developed based on technologies of the existing SRC. This section describes the conventional SRC technologies and introduces the new technologies of SRC VE.

3.1 Technologies of SRC

SRC is software that manages all system resources including servers, storage, and networks. SRC features both an element management function and system orchestration function. The element management function consists of integrated server management for managing various servers, integrated storage management for managing various storage systems, and integrated network management for managing various networks. The system orchestration function is used for total system management (**Figure 1**).

Each type of management above operates in conjunction with system orchestration and provides the following unified management processing for each server, storage, and network.

1) Integrated server management

Integrated server management systematically manages the servers and software products on those servers.

Regardless of the type of server, integrated server management automatically collects configuration and event information, and visualizes server information (e.g., type of cabinet, hardware model, and whether an error has occurred) and software information (e.g., name and version of the operating system) on one operation management screen. Integrated server management manages the operating system and applications as a master image (boot disk), installs and configures servers, and makes backup copies of servers.

2) Integrated storage management

Integrated storage management systematically manages the storage devices and access routes from a given server to those storage devices.

Regardless of the type of device, integrated storage management automatically collects and visualizes configuration and event information, and access routes. Integrated storage management also assigns disks to servers and changes routes.

3) Integrated network management

Integrated network management manages network devices and topology (i.e., network connections between servers).

Regardless of the type of device, integrated network management automatically collects and visualizes configuration, event, and topology information, and changes the topology.

4) System orchestration

In conjunction with each type of management above, system orchestration automates a series of management tasks in various situations, while maintaining consistency with various independent system resources (such as servers, storage, and networks). These management tasks entail a variety of procedures, such as operations for various resources during initial installation or system expansion, making a backup copy of the server, and performing error recovery.

3.2 Technologies of SRC VE

SRC VE is a server virtualization management software package developed by separating integrated storage management and integrated network management from SRC, and enhancing the virtualization function so that both SAN and LAN can be managed on the server side. Therefore, SRC VE can be flexibly combined with existing storage and network resources.

SRC VE consists of integrated server management (for managing both physical and virtual servers), SAN virtualization management (for virtualizing the connection paths to servers and storage), LAN virtualization management (for virtualizing the connection paths between servers and networks), and server orchestration (for managing servers including LAN and SAN connections) (**Figure 2**).

3.2.1 Integrated server management

Integrated server management systematically manages servers in a mixed physical-virtual environment, and manages the software products on the servers (**Figure 3**).

Integrated server management automatically collects configuration and event information from the server virtualization software (e.g., VMware), and displays physical and virtual server information in the same format on one operation management screen. That is, such server information as cabinet, model, and error

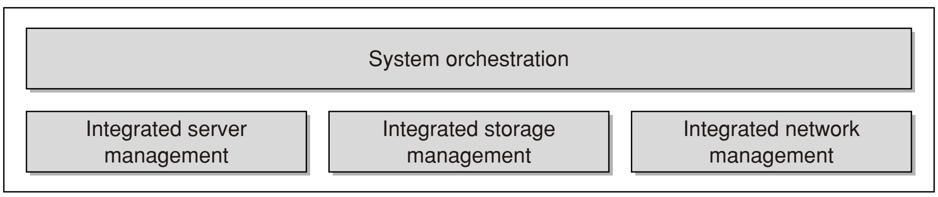


Figure 1
Architectural overview of SRC.

detection information is displayed in Status-view, Tree-view, or List-view format. Moreover, such software information as the name and version of the operating system is also displayed.

Since server installation, settings, and making a backup copy are performed with a master image (boot disk) that encapsulates the type of server, the user need not consider whether the server is physical or virtual.

Status-view displays the total number of all managed servers and the number of servers classified by server status (error, alarm, stop, or start state). Therefore, the total system status can be checked at a glance in Status-view.

Tree-view uses icons in displaying the hierarchical relation between the server cabinet, physical servers, and virtual servers. The virtual servers that may be influenced by hardware errors can therefore be found in Tree-view (**Figure 4**).

After a virtual or physical server is selected in Tree-view, an operation can be selected from the menu. In this case, instructions to operate a virtual server can be issued via the Web interface of VMware, and instructions to operate a physical server can be issued via the management interface of the physical server. Thus, the server manager can perform such routine operations as starting and stopping a server, and making a backup copy of the server without having to consider the type of server (physical or virtual).

List-view displays the setting status of automatic switching for shared spare servers, etc. (**Figure 5**).

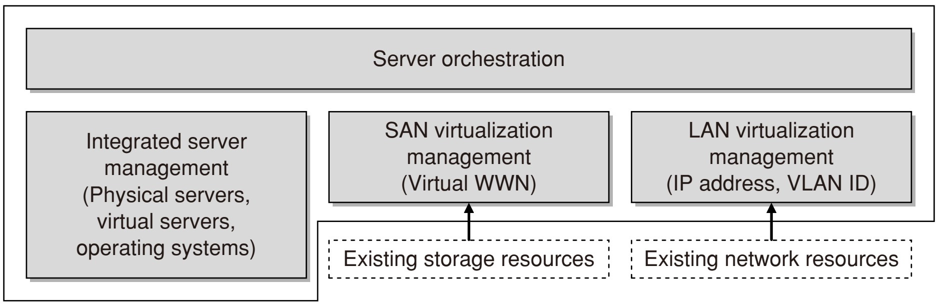


Figure 2
Architectural overview of SRC VE.

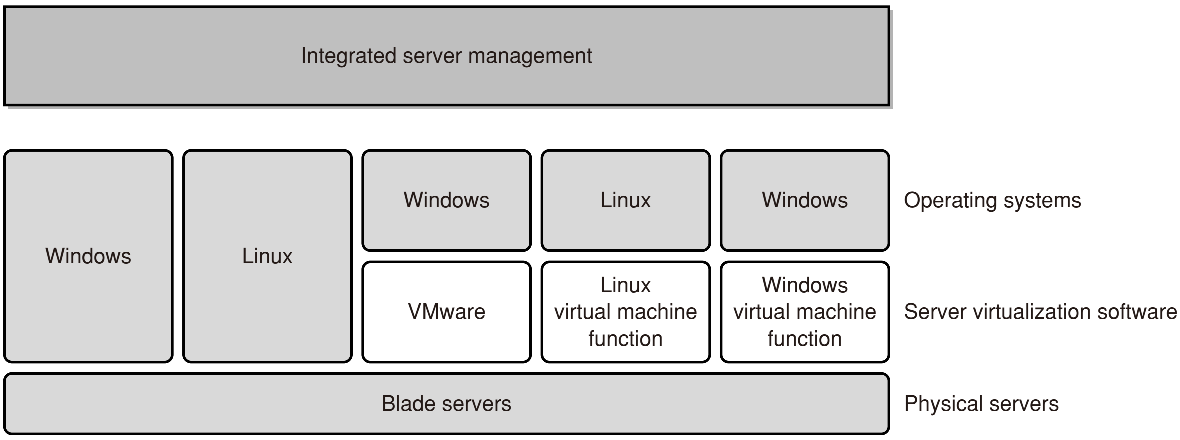


Figure 3
Resources managed by integrated server management.

3.2.2 SAN virtualization management

SAN virtualization is technology that hides the server-SAN physical relation on the server side so that servers can be separated from SAN. This eliminates the need to change the SAN configuration when adding or changing a server.

Fujitsu developed SAN virtualization management technology for managing fibre-channel adapters — the parts for connecting servers to SAN. This management assigns virtual World Wide Names (WWNs), instead of conventional fixed addresses, to the adapters.

WWN is a fixed address assigned to a fibre-channel adapter. SAN uses WWNs (via a

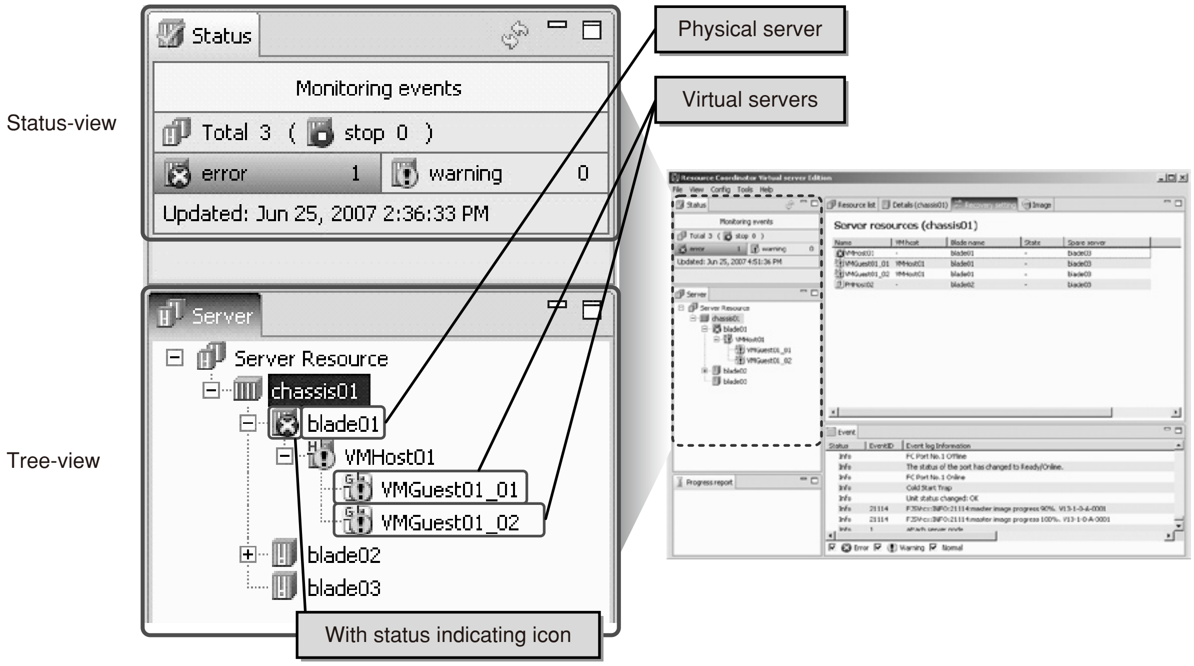


Figure 4
Example of SRC VE GUI screen (Status-view and Tree-view).

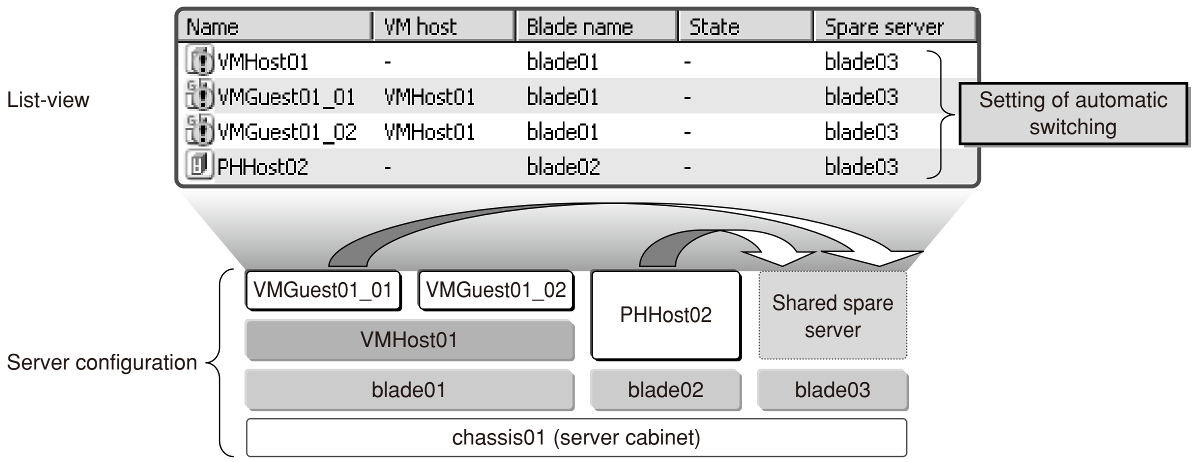


Figure 5
Example of SRC VE GUI screen (List-view).

switching device or storage device) to distinguish between servers.

For example, when a job on a faulty server in a conventional system is to be restarted on another server, WWN is changed from A to B. Therefore, the WWN set on the SAN side must be changed in the conventional system [**Figure 6 (a)**]. Since the virtual WWN is automatically inherited in SRC VE, the WWN setting on the SAN side need not be changed at all [**Figure 6 (b)**].

Combining SRC VE with the "SAN connection virtualization option" of the PRIMERGY blade server enables the use of this function. This option supplies a unique, virtual WWN address as Fujitsu's vendor ID. SRC VE assigns or changes virtual WWNs (**Figure 7**).

3.2.3 Server orchestration

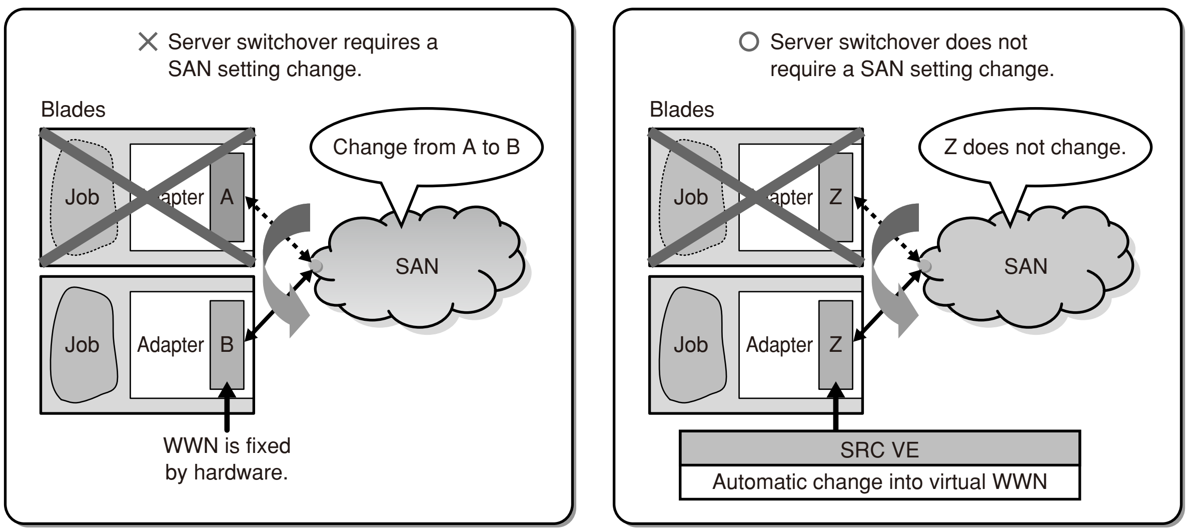
In conjunction with integrated server management, SAN virtualization management, and LAN virtualization management, server orchestration automates server assignment and changes operations, while maintaining consistency among the resources of all servers (Figure 2).

When the resource information necessary for a job is defined in an abstracted system design created by the user, server orchestration selects appropriate resources that match the system design from among the share-managed resources, and efficiently assigns the resources selected.

Both system design and resource information are managed in a unified manner by using the Resource Control eXtensible Markup Language (RCXML) developed by Fujitsu (**Figure 8**).

The system design can be defined by using either of two methods. One is the table format (worksheet) method, which is appropriate for defining large-scale systems. The other is the GUI method, which is appropriate for defining small-scale systems.

Use of the table format method allows system design to be done in the office before going to the data center where actual hardware devices are installed. The newly created system design definitions can be saved in a file, and then read by a hardware device at the data center. Thus, to complete the system settings, users can efficiently perform operations without making



(a) Conventional system

(b) SRC VE

Figure 6
Mechanism of WWN and virtualized WWN setting.

any mistakes (**Figure 9**).

Use of the GUI method allows a user without special skills to easily create a system design. For example, a spare server to be used in case of a server error can be completely defined by simply selecting the target current server and the spare server (**Figure 10**).

Based on the system design, the process from the detection of a server failure to the restart of operation on a spare server can be fully automated, and thereby create a highly available system that does not impose a heavy workload on the manager. The following six operations are automated in concrete terms (**Figure 11**):

1) Detect error

The routine monitoring results and Simple Network Management Protocol (SNMP) trap posted from a physical server are analyzed to detect errors.

2) Stop server

An error-detected server is identified according to server information in the SRC VE

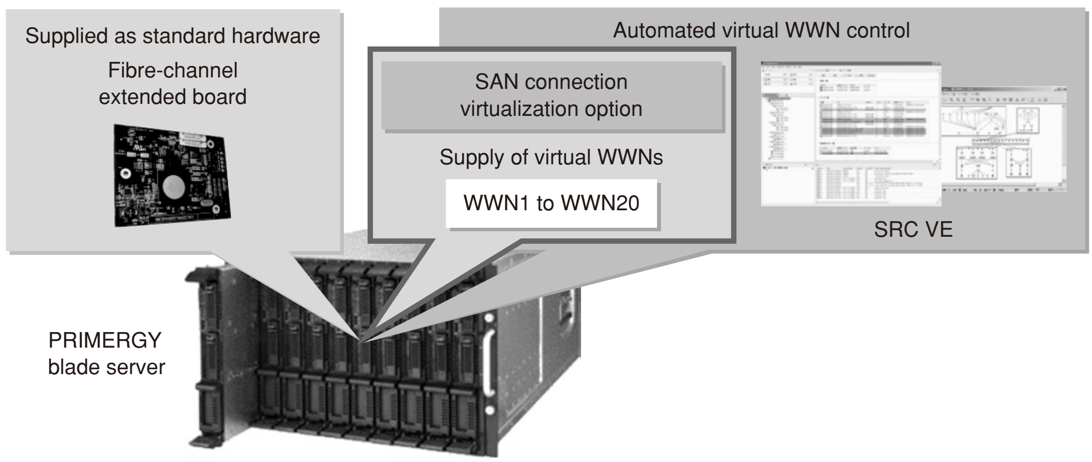


Figure 7
Components of SAN virtualization management.

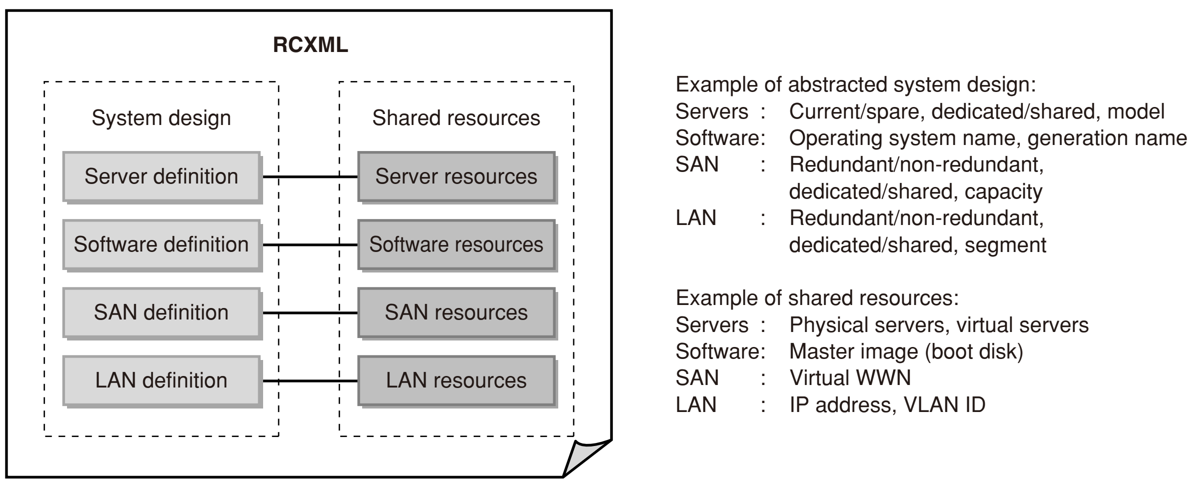


Figure 8
Implementation overview of system design and shared resources.

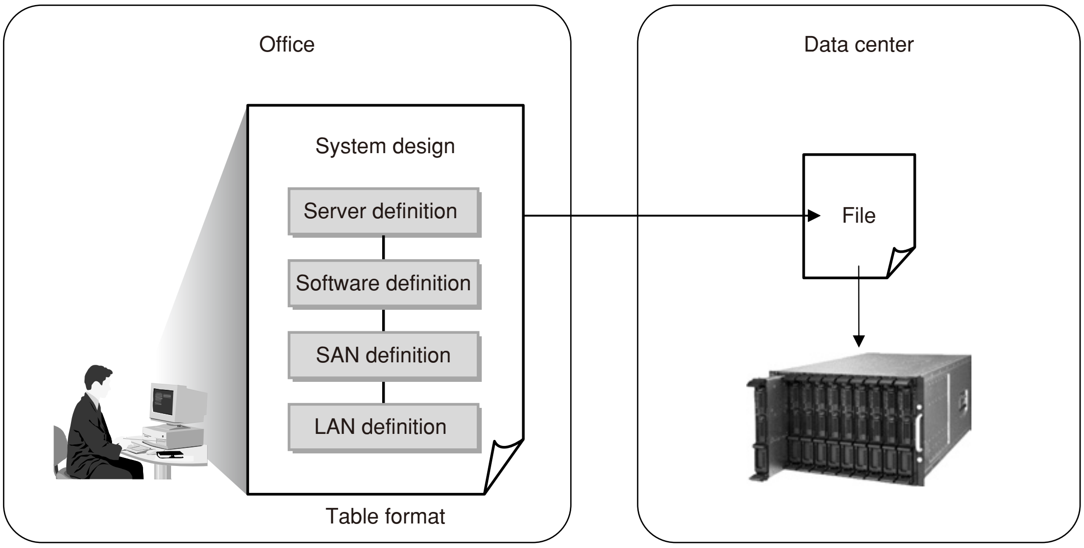


Figure 9
Example of creating system design definitions (worksheet method).

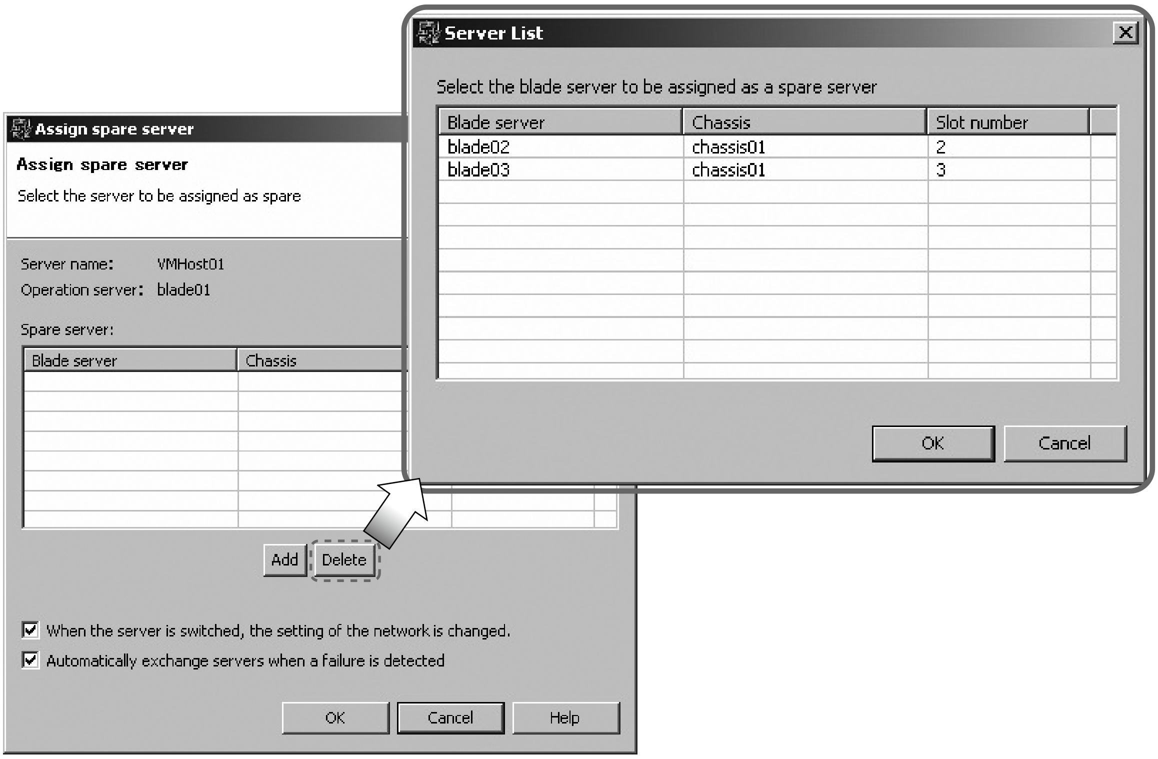


Figure 10
Example of creating system design definition (GUI method).

database, and then stopped.

3) Select spare server

A spare server is selected according to the system design definitions, with a check that other software such as VMware HA has not used the spare server for server switchover.

4) Set storage

The disk of the error-detected server is switched to the spare server by assigning the error-detected server's virtual WWN to the spare server according to SAN information on the shared resources.

5) Set network

The LAN connection is switched from the error-detected server to the spare server by assigning the error-detected server's network setting VLAN ID to the LAN switch port where the spare server is connected according to LAN information on the shared resources.

6) Restart job

The power of the spare server is turned on.

The HA function of other server virtualization software such as VMware HA only manages virtual servers. Therefore, when both virtual and physical servers are contained in an environment where SRC VE is not installed, spare servers for both the virtual and physical servers must be prepared independently for normal server switchover [**Figure 12 (a)**].

These server switchover conditions differ when SRC VE operates in conjunction with

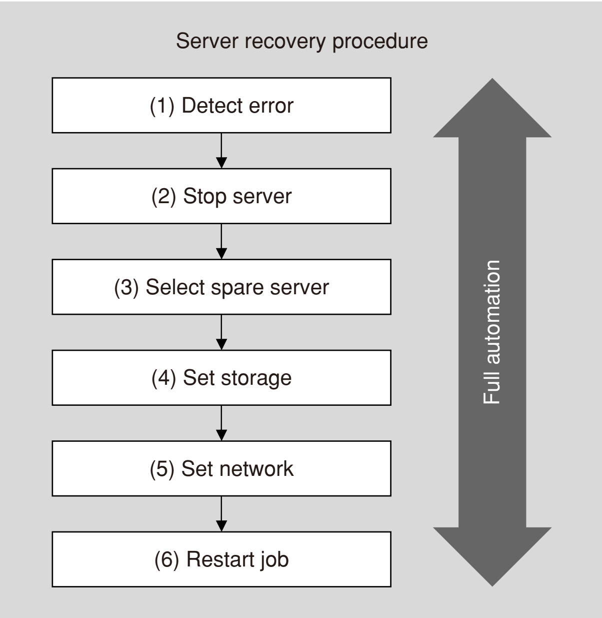
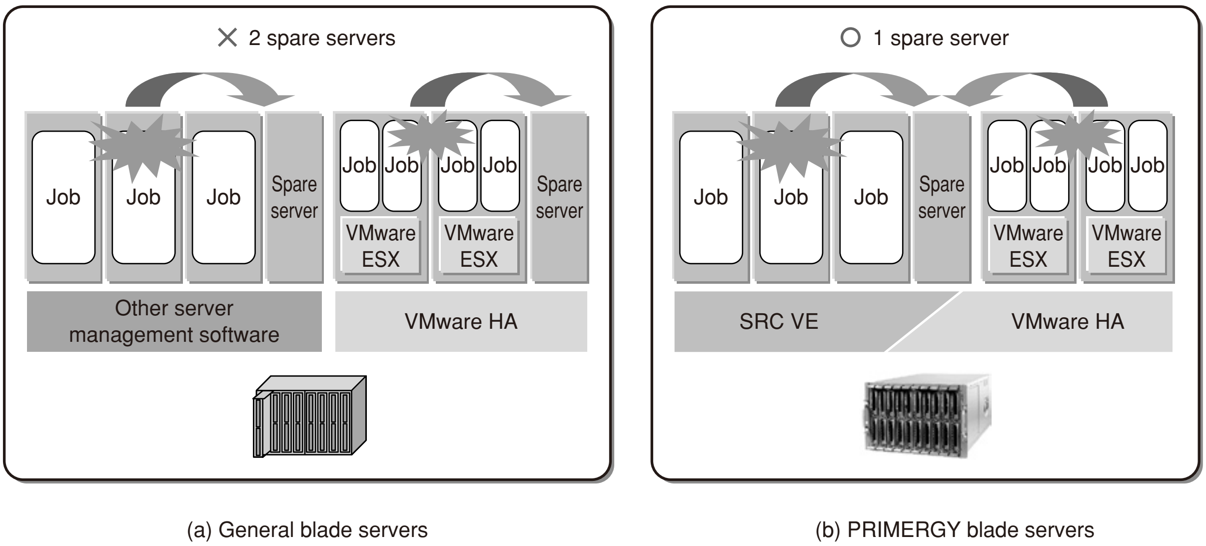


Figure 11
Automated procedure for server recovery based on system design.



(a) General blade servers

(b) PRIMERGY blade servers

Figure 12
Difference between spare server usage in mixed physical-virtual server environment.

VMware. If an error occurs on a server where a virtual server operates, VMware HA switches the error-detected server to the spare server. If an error similarly occurs on a physical server, SRC VE switches the error-detected server to the same spare server. Thus, physical and virtual servers can be used more efficiently by sharing a spare server [**Figure 12 (b)**].

SRC VE also considers energy efficiency with a power-saving mode for making spare servers wait in the power-off state.

4. Conclusion

This paper introduced the basic technologies of SRC VE server virtualization management software.

Major operating systems are now being planned to provide the virtualization function as standard. Upon hearing these plans, we anticipate a growing corporate demand for virtualization in consolidating servers and improving resource efficiency in the future.

Fujitsu will promote virtualization technologies for the open system environment by adopting industry standards and cooperating with related vendors and communities. Of course, Fujitsu will continue developing virtualization technologies with focus on the advantages to customers equipped with SRC VE, and to satisfy existing information system needs in the market.

References

1) Fujitsu: Introduction to Systemwalker Resource Coordinator Virtual server Edition V13.2 (December 2007). *http://www.fujitsu.com/downloads/PRMQST/documents/presen/RCVE_v13.2_Introduction.pdf*
2) R. Fichera et al: Organic IT in Japan, Forrester Research, November 21, 2005. *http://www.forrester.com/Research/Document/Excerpt/0,7211,38224,00.html*

Takashi Fujiwara, *Fujitsu Ltd.*
Mr. Fujiwara received the B.S. degree in Electrical Engineering and Electronics from Aoyama Gakuin University, Tokyo Japan in 1985. He joined Fujitsu Ltd., Kawasaki, Japan in 1985, where he was engaged in research and development of server clustering software for open systems. He is currently engaged in development of PRIMECLUSTER and also in research of autonomic technology for Systemwalker Resource Coordinator.

Shigehiro Yoshikawa, *Fujitsu Ltd.*
Mr. Yoshikawa received the B.S. and M.S. degrees in Computer Science from Tsukuba University, Ibaraki, Japan in 2000 and 2002, respectively. He joined Fujitsu Ltd., Yokohama, Japan in 2002, where he was engaged in research and development of autonomic system software. He is currently engaged in research and development of Systemwalker Resource Coordinator.

Yoshinobu Hibi, *Fujitsu Ltd.*
Mr. Hibi received the B.S. degree in Mathematics from Tokyo Metropolitan University, Tokyo, Japan in 1986. He joined Fujitsu Ltd., Kawasaki, Japan in 1986, where he was engaged in research and development of open system software. He is currently engaged in research and development of Systemwalker Resource Coordinator.

Hiroshi Yazawa, *Fujitsu Ltd.*
Mr. Yazawa received the B.S. degree in Electrical and Computer Engineering from Yokohama National University, Yokohama, Japan in 1992. He joined Fujitsu Ltd., Yokohama, Japan in 1992, where he was engaged in research and development of applications for UNIX workstations. He is currently engaged in research and development of Systemwalker Resource Coordinator.