

# Disaster Recovery Feature of Symfoware DBMS

● Teruyuki Goto

*(Manuscript received December 26, 2006)*

The demands for stable operation of corporate information systems continue to grow, and in recent years there has been an urgent demand for protections against natural disasters, terrorism, and other destabilizing events. To meet these demands, Fujitsu offers its Symfoware Disaster Recovery system, which is based on the Symfoware high-reliability database management system (DBMS) and incorporates Fujitsu's expertise in mainframes, new hardware, and other core technology areas. When a disaster occurs at a customer's main center, this system switches operation to a previously prepared remote standby center. By using this system, customers can protect their data and quickly resume business operations when a disaster occurs. This paper describes Symfoware Disaster Recovery.

## 1. Introduction

Corporate information systems have a close relationship with the economy and with social activities, and requirements for stable operation are increasing year by year. Particularly, in recent years there has been an urgent demand for countermeasures against disasters such as earthquakes and terrorism, which can inhibit stable operation. On the other hand, only a few corporations have implemented comprehensive business continuity planning (BCP) disaster recovery measures. The dearth of optimal solutions and few actual examples of implementation can be given as reasons for this current situation.

In this paper, we introduce our optimal disaster recovery solution system for databases, which are the lifelines of companies. This system—Symfoware Disaster Recovery—combines

Fujitsu's proven technologies with newly-developed technologies.<sup>1,2)</sup>

## 2. Increasing demand for disaster recovery ability

In this section, we describe the necessity of stable operation and the factors that inhibit it.

### 2.1 Stable operation as a company's top priority

Current corporate information systems comprise various systems that perform integrated services on a reciprocal basis. Therefore, if one of these systems is interrupted, many companies and consumers will be affected. If an important system is interrupted, not only is there direct damage such as loss of business opportunities, but the interruption also leads to a drop in public trust. Against this background, stable operation

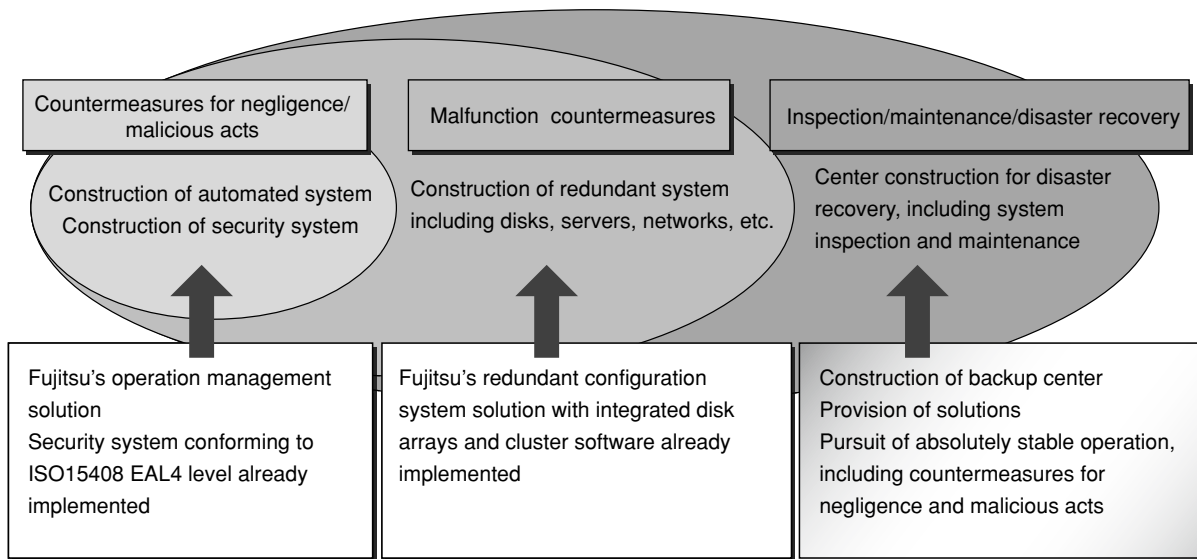


Figure 1  
Measures toward stable operation of Symfoware.

has become a high-priority capital investment issue for most companies, and there is a growing demand for absolutely stable operation of databases that manage important data.<sup>note 1)</sup>

## 2.2 Factors that inhibit stable operation and their countermeasures

There are three main groups of factors that inhibit stable operation, and optimal countermeasures for these factors that ascertain importance and urgency are required. The first group are problems caused by negligence or malicious acts such as human error and computer viruses. The second group are malfunctions or interruptions caused by hardware malfunctions and power outages. The third group are interruptions due to disasters such as earthquakes, fires, and terrorism. Moreover, in addition to these causes of unplanned interruptions, preventive measures for planned interruptions such as inspections required by

note 1) A database manages the data that forms the framework of a corporate information system. An initiative that minimizes data loss and recovery time even when the building housing the database is destroyed by a disaster becomes particularly important.

law and hardware/software upgrades are also important.

Our Symfoware database management system (DBMS) has autonomic functions such as automatic reorganization, a cluster function that can provide switching within seconds, and security/automated operations. It can provide solutions to problems caused by negligence, malicious acts, and malfunctions and has already been incorporated into many corporate systems. The next stage in the development of our Symfoware DBMS is to realize the highest level of stable operation by providing a Symfoware Disaster Recovery system that uses our proven mainframe technology and the latest storage technology.

Measures for the stable operation of Symfoware, including planned interruptions for maintenance, are shown in **Figure 1**.

## 3. Current status of disaster recovery

The current approach to maintaining business continuity is to construct remote centers and then switch operation to these centers when a disaster occurs. This section describes this

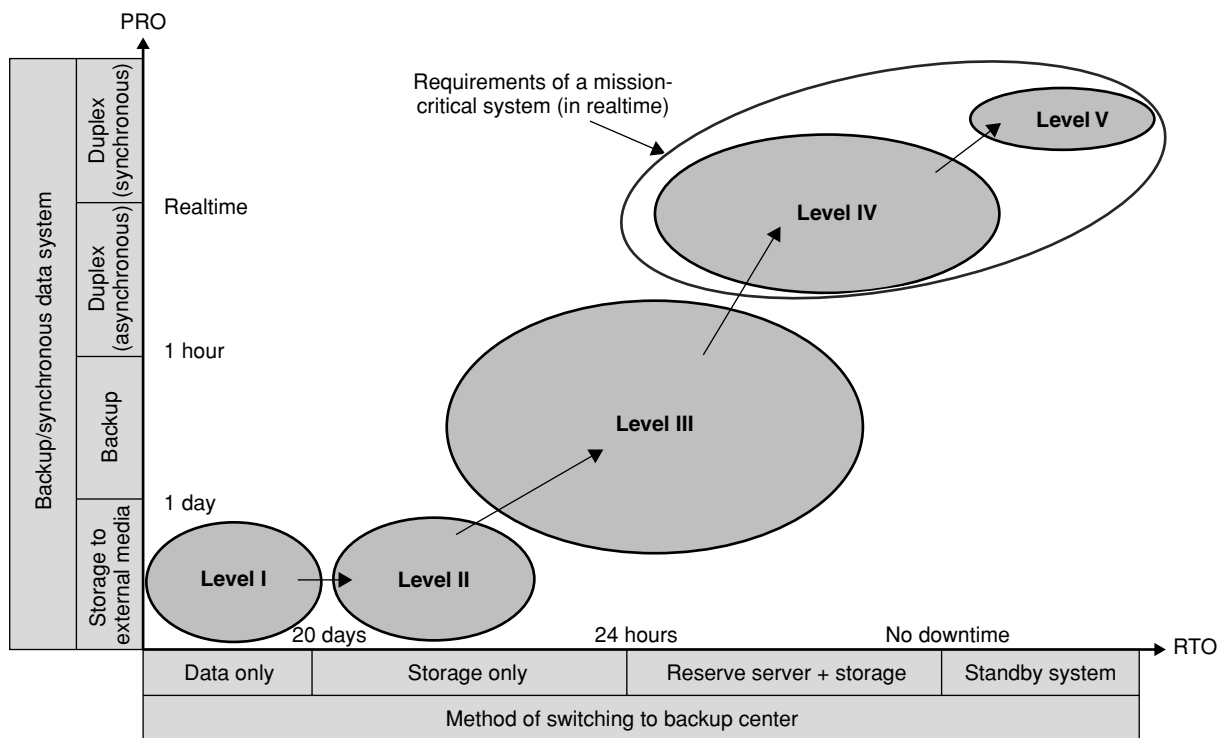


Figure 2 Requirements and systems analyzed from restoration service level.

approach.

### 3.1 Rapid resumption of business operations by switching data centers

On-site backup systems have been widely used as effective countermeasures against system failure. However, if we consider the extent of the destruction caused by a major disaster, this method will not be effective. Therefore, technology and measures to rapidly restore a database environment at a remote center and resume business operations by switching centers have become essential.<sup>note 2)</sup> The methods used to implement a remote center based on the restoration service level (recovery point objective [RPO] and recovery time objective [RTO]<sup>note 3)</sup>) are shown below.

note 2) After the September 11 terrorist attacks in New York, some companies had backup centers in New Jersey that enabled them to restart business operations the same day, which affirms the usefulness of such backup centers.

The requirements and systems analyzed from the restoration service level are shown in **Figure 2**.

Implementation examples are as follows:

- 1) Level I  
Periodically back up data to magnetic tapes and store them in a fireproof vault at a remote location.
- 2) Level II  
Use a network-attached storage (NAS) device or other network backup product and store the data backups at a remote location.
- 3) Level III  
Use mirroring software and synchronize the remote center database with the main center database.

note 3) RPO and RTO are data recovery metrics. In the case of RPO, the question to ask is "How much data loss can be tolerated?" while in the case of RTO, the question is "How quickly can business operations be restored?"

4) Level IV

Use DBMS replication to synchronize the remote center database with the main center database.

5) Level V

Use the remote copy function of the storage system and synchronize the remote center database with the main center database in realtime.

The time required to resume business operations is reduced as the level moves up from I to V, but the costs increase. Using Level I as an example, if one day's worth of backup data is transported by truck, the additional capital investment is minimal. However, up to one day's worth of data could be lost if something happens to the main center. Furthermore, depending on how that data is recovered, several days to several weeks are required to resume business operations.

Which of these methods should be used depends on the values chosen for RPO and RTO. That is, the measures that should be taken and the investment expenditures differ greatly according to the designed objective. For this reason, it becomes very important to set a policy that ascertains importance and urgency. For example, if the loss of one day's data is permissible, then transporting that data by truck is an adequate backup method. However, in today's mission-critical systems, disaster recovery measures of Level IV or higher that provide realtime functions are being demanded in order to reduce RPO and RTO. For this reason, the preparing of a standby database in realtime at a remote center and the switching of centers when the main center is hit by a disaster will become necessary to enable rapid resumption of business operations. The rapid resumption of business operations by switching centers is shown in **Figure 3** (hereafter the centers are called the main center and the remote standby center).

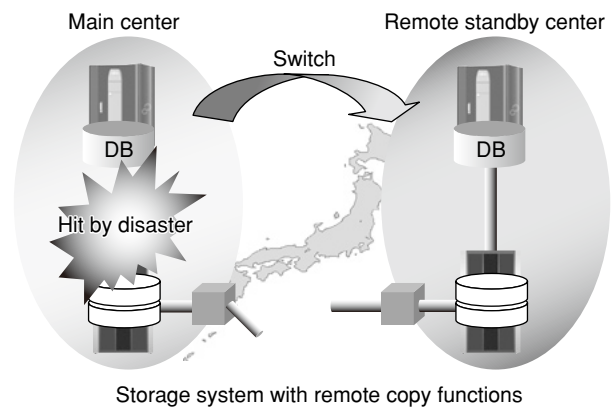


Figure 3  
Rapid resumption of business operations by switching centers.

### 3.2 Technology for attaining disaster recovery requirements

To implement disaster recovery at Level IV or higher, technology for synchronizing the database's updated results (difference log) or the database itself in realtime is required. The following two methods are representative of the methods currently being used.

1) Database mirroring using remote copy

The entire database system environment is assigned to storage disks for remote copying, and the main center database is mirrored to the remote standby center database.

2) DBMS log shipping

The difference logs generated at the main center are reflected to the remote standby center's database. Proprietary DBMS real-time transfer functions are used to assure equivalency between the main and remote standby centers.

Currently, there is no way to link the databases of centers in realtime in the mission-critical business domain. This is because the following problems will occur if either of the two methods described above is used for a large-scale database system.

- Database mirroring that uses remote copying of the storage system cannot assure safe, flexible center operations, including

operations for assuring database integrity, because copying is not performed in database management units. For example, both centers must be synchronized before database maintenance operations can be performed. Moreover, because the entire database is mirrored, the volume transferred to the remote center is large, resulting in increased line costs. In particular, with reorganized operation, the database that was reorganized becomes the object of transfer and operation of a large-scale database system that contains terabytes of data cannot be performed.

- DBMS log shipping uses server resources. Because real-time transfer is performed in small units, if a large volume of traffic is generated, the throughput cannot be maintained, resulting in a drop in real-time performance.

#### 4. Symfoware initiatives

With Symfoware Disaster Recovery, we are targeting absolutely stable operation, including that of large-scale systems. For this reason, log shipping is used because of the small transfer volume and superior database operability. Furthermore, by providing and integrating the features of the storage function and database function, we have also realized the advantages provided by database mirroring, for example, less time required to switch centers and ease of use. In the next section, we describe our initiatives toward stable operation of Symfoware Disaster Recovery.

#### 5. Initiatives toward high-load systems

The following initiatives realize the stable operation of database systems, including high-load systems.

##### 5.1 Real-time transfer using storage mirroring technology

With Symfoware Disaster Recovery, a rerun log pipeline (RLP) is created in log group units (database log environment units) with respect to the business operation. An RLP consists of a rerun log management file (RLM) that stores log management data and multiple rerun log cycle files (RLCs) that store the database difference logs. These files are assigned to the remote copy storage section and are mirrored in realtime between the centers. The log environment configuration is shown in **Figure 4**.

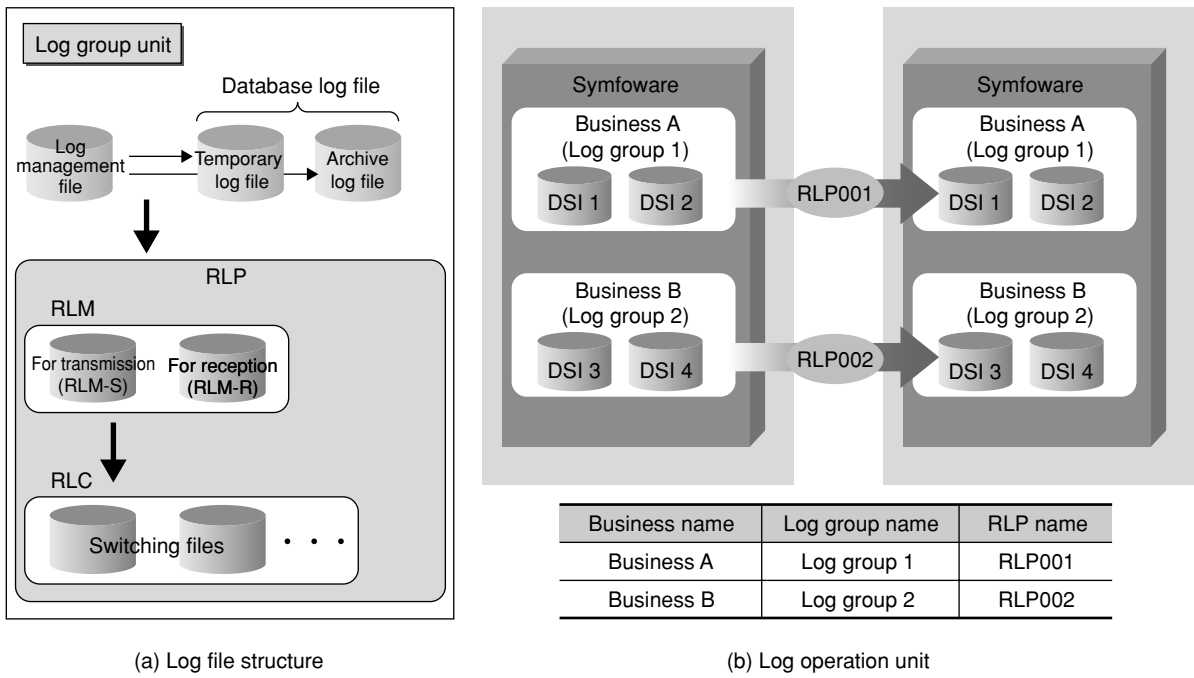
Moreover, when the main center's database is updated, logs are acquired by the database recovery log acquisition processing extension and the logs are written to the RLCs. When the volume of logs in an RLC reaches a fixed amount, log writing is switched to the next RLC. At the same time that logs are acquired, the RLC for which switching has been completed is loaded to the remote standby center and reflected to the database. This mechanism maintains equivalency between the centers and provides preparations for center switching if the main center is hit by a disaster. Log transportation by RLC switching is shown in **Figure 5**.

##### 5.2 Dedicated logs for reducing TCO (line costs)

In Symfoware Disaster Recovery, only the difference logs that were used to update the database are copied. Furthermore, our proprietary technology assures that the difference logs for acquisition are only those that represent the minimum data required for reflection (indexes and so forth are not acquired). Therefore, compared with database mirroring, the copy target is reduced to 1/8 or less<sup>note 4)</sup> of the original size, resulting in a big reduction of the storage volume and line costs.

---

note 4) This figure was obtained in an in-house comparison test.



note) DSI (Data Structure Instance) is a unit of table storage.

Figure 4  
Log environment configuration.

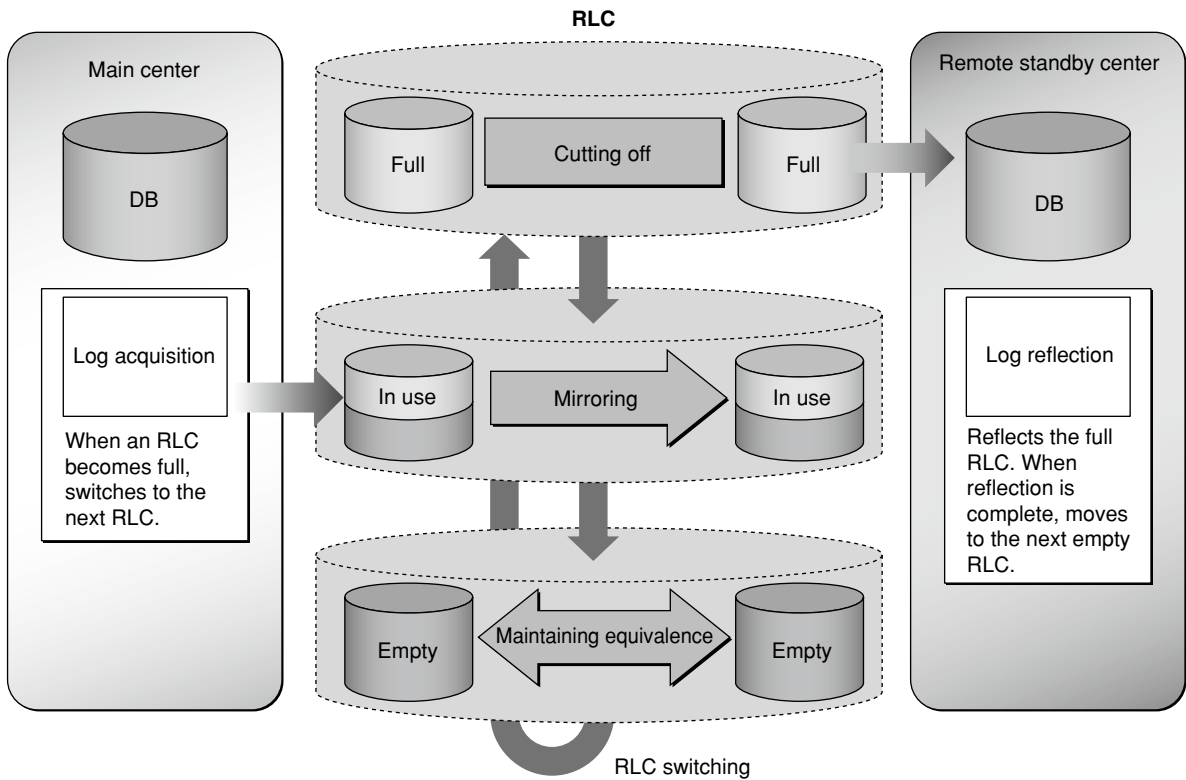


Figure 5  
Log transportation by switching RLCs.

Table 1  
Mode selection.

Selected mode	Transaction	Storage	Setting guidelines
Synchronous	Synchronous	Synchronous	Select this setting to give maximum priority to preventing data loss
Asynchronous	Asynchronous	Asynchronous	Select this setting to give maximum priority to online performance

### 5.3 Rapid resumption of business operations due to a reduction in the scope of recovery

To rapidly resume business operations through log shipping, it is important to reduce the volume of logs that have not yet been reflected.

- 1) Dedicated log files enable high-speed center switching

Because the difference logs are reflected to the remote standby center when switching is performed from one RLC to the next, if there is no backlog when the centers are switched, only the difference logs in the final RLC are not reflected. Therefore, when targeting the rapid resumption of business operations, RLC capacity becomes important. With Symfoware Disaster Recovery, the RLC-dedicated log files can be flexibly designed to match the RTO requirements. Furthermore, the RLC switching timing can also be changed dynamically. In planned center switching, the volume of difference logs that have not been reflected can be reduced even further and the business interruption time shortened.

- 2) Selection of synchronous or asynchronous mode as required

With Symfoware Disaster Recovery, either synchronous or asynchronous mode can be selected for the secure writing of the transaction difference logs and the storage transfer method. This option enables the optimal design for business performance and RTO requirements.

The mode selections are shown in **Table 1**.

### 5.4 Rapid resumption of business operations due to a reduction in recovery operations

When a disaster occurs, it is necessary to check the business results that have arrived at the remote standby center and establish database integrity in order to maintain business integrity prior to the resumption of business operations.<sup>note 5)</sup> These operations must be performed according to the type of business involved. This is normally not an easy task, and the extent to which it is performed efficiently is important for the rapid resumption of business operations. Symfoware Disaster Recovery realizes rapid resumption of business operations by reducing the scope of recovery when business has been hit by a disaster and also by using a function that supports actions performed by the user to maintain business integrity. An illustration of the operation performed when a disaster occurs is shown in **Figure 6**.

- 1) Maximum use of the logs that are being written

When a disaster occurs, database difference logs that have not been reflected remain in the RLP. With Symfoware Disaster Recovery, the data loss reduction function identifies the secure transaction difference logs from among the remaining difference logs, including those being written to the RLC, and enables their reflection

note 5) With Symfoware Database Recovery, transaction integrity is assured as standard. Maintaining business and database integrities becomes particularly important for businesses that handle multiple transactions such as flow tasks.

to the database.

2) Disaster support function for rapid resumption of business operations

Symfoware Disaster Recovery provides a difference log inspection function and a database recovery function to support recovery operations such as database and business integrity matching performed by the user. The difference log inspection function can inspect the arrival status of updated results at the remote

standby database based on the difference log and transaction identification information. If the recovery point at the main center is periodically input at the main center as a routine task, the database recovery function can quickly restore the database to the recovery point status that was selected when the disaster occurred. The database recovery function is shown in **Figure 7**. These functions enable efficient implementation of restore operations by the user and a large-scale reduction of the time required to resume business operations.

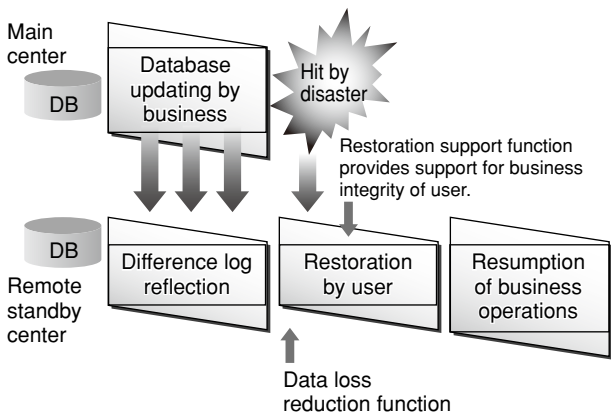


Figure 6 Operation when disaster occurs.

5.5 Assured performance even for high-traffic business applications

In the case of database mirroring, the entire database environment must be assigned to storage disks that enable remote copying, which affects the existing business [**Figure 8(a)**]. The I/O costs of these disks are high when compared with local disks. Symfoware Disaster Recovery copies only the dedicated log files that are not direct business application updates. Furthermore, the I/O costs are reduced and the existing business assured because the size of the

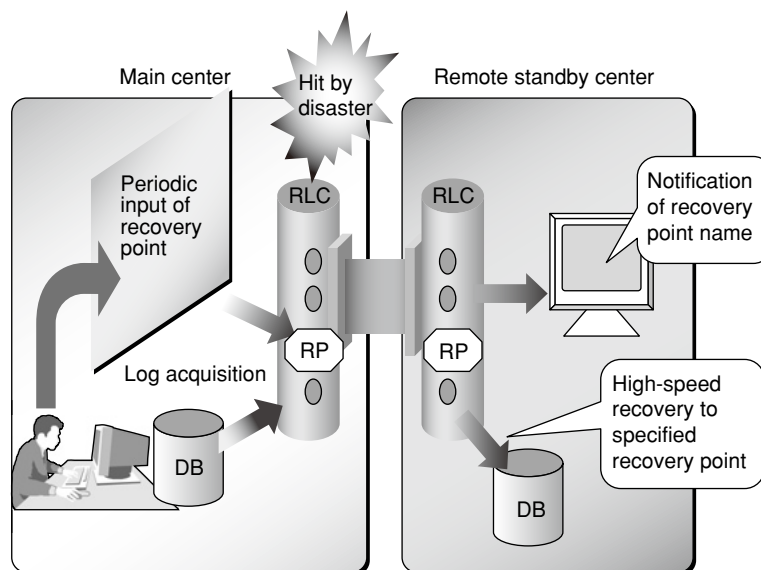


Figure 7 Recovering a database to recovery point using recovery function.



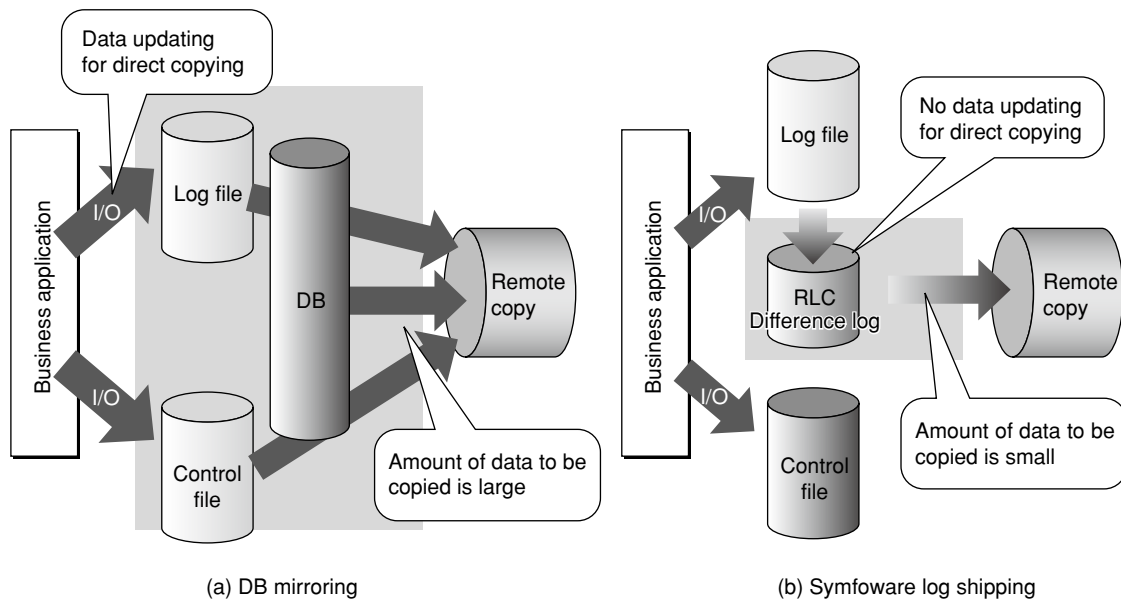


Figure 8  
Log acquisition that minimizes effect on business operations.

difference logs is minimized as described earlier in this paper. Log acquisition with the minimum effect on business is shown in **Figure 8 (b)**.

Furthermore, thanks to our high-performance log acquisition and log reflection technologies, we have realized a system that does not even affect high-traffic business applications.

#### 1) High-performance log acquisition

If asynchronous mode has been selected, log acquisition is completed simply by transferring log blocks to the dedicated I/O cache in the same way as database recovery logs are written. The log blocks transferred to the I/O cache have almost no effect on business performance because they are written to difference log files in the asynchronous transfer mode. If downtime has occurred, integration with the database down recovery operation assures the generation of difference logs. Furthermore, due to line load and other factors, even if an unforeseen situation such as an exhausted buffer occurs, only the disaster recovery operation itself is stopped and online business operations can be continued.

#### 2) High-performance log reflection

The use of high-speed technology such as

internal paths for direct reflection, a dedicated memory cache, and multiplexing enables the logs generated at the main center to be reflected without any backlog.

## 6. Initiatives toward large-scale database operations

Symfoware Disaster Recovery fuses proven operation technology that has been cultivated for high-level mainframe requirements with the latest open technology to realize stable operation for large-scale database systems.

### 6.1 Proven operation technology

In the 1990s, Fujitsu was the first Japanese company to build remote backup centers for online databases that are important social infrastructures. In particular, in large-scale systems for financial institutions, our AIM<sup>note 6)</sup> system has produced many excellent results.

For example, a major financial institution

note 6) AIM is an abbreviation for Advanced Information Manager.

This is a database/data communication (DB/DC) management system for Fujitsu mainframes.

customer that integrated operations by remote disk transfer in realtime with a remote site located 100km away was able to resume business operations in two hours with the loss of only a few seconds of data. Moreover, the backup center structure is being utilized effectively as a remote standby center in cases such as when maintenance operations are being performed at the main center.

## 6.2 Symfoware Disaster Recovery initiatives

The following initiatives enabled us to realize safe and flexible operation of database systems, including those of large-scale database systems.

### 6.2.1 Flexible operation by using logical logs

Symfoware Disaster Recovery provides flexibility in the physical construction between centers. As a result, by using logical reflection of the acquired difference logs, reorganization and recovery operations can be performed independently for each center, so there is no need to interrupt operations simultaneously at both centers.

### 6.2.2 Center reversion operation enabled by log shipping

When a disaster occurs, after a switch has been made from the main center to the remote standby center and business continuity has been assured, operations can be reverted to the main center. If, for example, the effect of the disaster is limited to network damage, the original main center can be set momentarily as the remote standby center after network operation has been restored. Next, center reversion is performed and the original center configuration can be reestablished. In the same way as for center switching, center reversion is realized by using log-shipping technology. The business interruption time can be kept to a minimum by minimizing the number of difference logs

required for reflection at center switching. Moreover, even if the asynchronous transfer mode has been used to link to the remote center, center switching can be performed easily with almost no data loss. The flow of operations is described below.

- 1) When a disaster occurs, difference log acquisition operations are started at the center to which business operations have been switched (hereafter called the former remote standby center).
- 2) After operations have been restored at the center where the disaster occurred (hereafter called the former main center), according to the database status, recovery to the status when the difference log acquisition was started in 1) is performed. This is done by copying all of the data or by using a specified recovery point.
- 3) As required, the difference log is reflected to the database at the former main center after it has been restored from the former remote standby center. During the difference log reflection, business operations are continued at the former remote standby center, and when center reversion is performed, the reflection of the difference logs from the last RLC assures equivalency between the centers.

### 6.2.3 Operation enabled by automatic linkage through the storage system

The remote standby center is normally established at a separate location from the main center. However, to keep the database ordinarily updated, operations are required at both centers. Moreover, if there is a long-term network malfunction or a power outage at one of the centers, operation that is independent of the affected center becomes necessary. This requirement increases the costs of operating both centers. With Symfoware Disaster Recovery, by using the storage system's remote copy function for automatically linking the current status

of the database difference logs generated at the main center with the current status of the update processing of the conserved database that is performed at the remote standby center, operability between the centers is enhanced and the TCO reduced.

#### 6.2.4 Message linkage function for linked operation between centers

As a part of database administration, database maintenance is periodically required for configuration changes and expansion. When this kind of maintenance is performed, the maintenance procedures performed at the main center must be synchronized with those performed at the remote standby center. This requires a mechanism and operation that enable synchronized timing. With Symfoware Disaster Recovery, we provide a message linkage function for the complete assurance of database integrity between the centers as an answer to this requirement.

The message linkage function synchronizes the operation timing by automatically informing the remote standby center of the operational information generated at the main center while

assuring the order relationship with the difference logs. By using the remote copy function in the same way as with the difference logs to transfer messages, we have realized a method of assuring the order relationship with the difference logs.

The message linkage function is shown in **Figure 9**.

## 7. Developments toward new solutions

Symfoware Disaster Recovery can use the remote standby center database reproduction technology to provide solutions other than disaster recovery solutions. In this section, we describe the following functions.

- 1) Business continuity during periodic maintenance
- 2) Using the remote standby center in normal operation
- 3) Database snapshot
- 4) Purpose-specific maintenance center

### 7.1 Business continuity during periodic maintenance

Even during lengthy maintenance operations such as facility inspections required

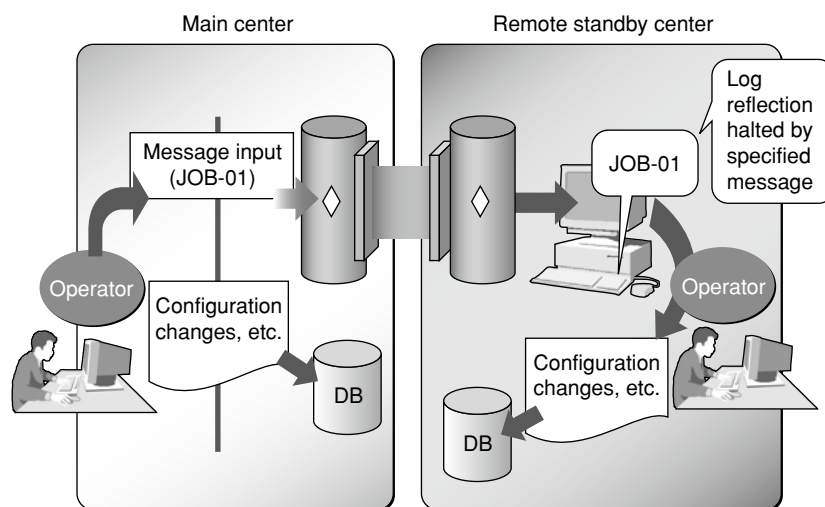


Figure 9  
Message linkage function.

by law, hardware preventive maintenance, and software upgrades, Symfoware Disaster Recovery assures business continuity through the use of log-shipping technology. In the case of conventional systems, if a remote standby center has not been established, business operations are interrupted during maintenance.

Moreover, if a storage system is used to perform database mirroring to the remote standby center and the maintenance operations cause the contents of the main center database to be reorganized, all of the reorganized disk area has to be mirrored. Because of the line load and other factors, it is necessary to interrupt business operations in the case of large-scale database systems. However, with Symfoware Disaster Recovery, the use of logical log-shipping technology for maintenance operations such as reorganization enables operations to be performed independently for each center. Because a DBMS manages log transportation by using a storage system, center switching can be easily performed.

Maintenance measures differ considerably from disaster recovery measures from the viewpoint of operating on a planned basis. Normally, difference logs are mirrored in real-time in the same way as for disaster recovery measures, and the difference logs accumulated at the remote standby center are reflected as required. When center switching is performed, log acquisition from the main center is interrupted, and center switching is performed after all the remaining logs in the remote standby center have been reflected (to assure equivalency between the centers). Therefore, there is no loss of data.

An example of planned switching is shown in **Figure 10**.

## 7.2 Using the remote standby center in normal operation

Requirements for the effective use of the remote standby center are extremely high, not just in the case of a disaster, but also for

the expansion of services. With Symfoware Disaster Recovery, browsing tasks can always be performed at the remote standby center. While online business operations are being performed at the main center, effective use of the remote standby center for browsing tasks is enabled.

In the case of conventional systems, to use the remote standby center for browsing tasks, the online business operations have to be interrupted momentarily and the remote standby center database restarted (when database mirroring is used for storage). Costly log acquisition and reflection processing is also required when other companies' log-shipping technologies are used. With Symfoware Disaster Recovery, if a key or other information is set in advance in the acquired difference logs, logical reflection can be performed at high speed, thereby enabling the remote standby center to be used for browsing tasks.

## 7.3 Database snapshot

With financial systems and other similar systems, batched interest processing (assigning of interest) is required in daily operations. The load that this processing places on online business operations must be considered, and the creation of a database snapshot on a different server to perform batch processing is a method that is often used. In the case of conventional systems, a database copy function is used to create a database snapshot on a different server each time batch processing is performed. Symfoware Disaster Recovery employs log-shipping technology, and therefore the remote standby center can be used as a batch server. When this is done, it is no longer necessary to create a new database snapshot for each batch processing operation and the remote standby center can be readied for the next batch processing operation as an add-on. The creation of a database snapshot at a remote standby center is shown in **Figure 11**.

- 1) A freeze point is input at the main center

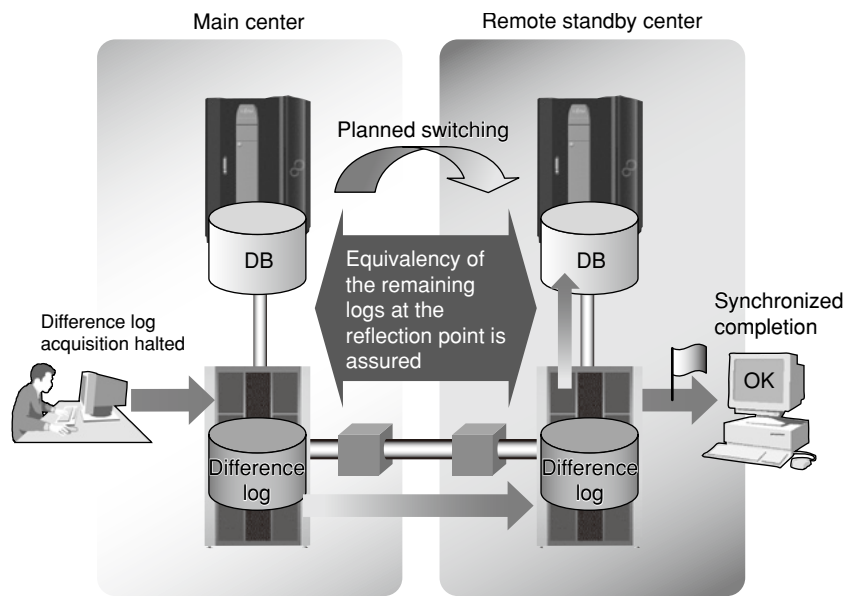


Figure 10  
Centers are switched in advance.

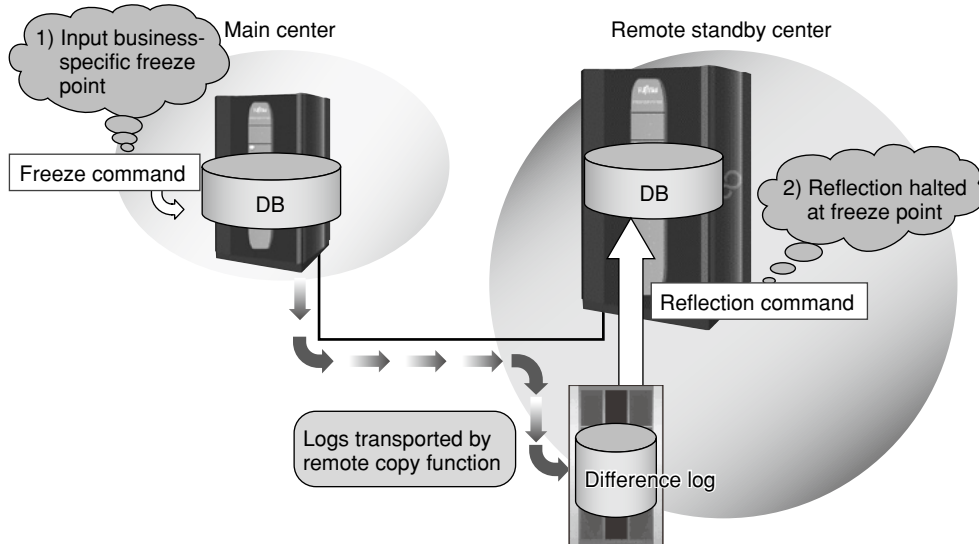


Figure 11  
Database snapshot is made for remote standby center.

at the required timing for the static status. At this time, the input freeze point is also automatically transmitted to the remote standby center.

- 2) Reflection of the difference logs is automatically interrupted at the transmitted freeze point, and a message indicating that batch processing can be started is output (completion of the creation of the database snapshot).

#### 7.4 Purpose-specific maintenance center

To implement the disaster recovery measures required for absolutely stable operation, there are cases where the establishment of a third or even a fourth remote standby center is required. Symfoware Disaster Recovery provides for the construction of multiple remote standby centers for each main center. Multiple remote standby centers are established not only for disaster recovery but also for maintenance measures and batched business transactions. Therefore, Symfoware Disaster Recovery makes it possible to prepare for an unexpected situation and also make effective use of remote standby centers for other purposes.

### 8. Conclusion

Symfoware is a high-reliability database management system that is designed to provide customers with solid security. It was developed in keeping with our development goals of unrivalled stable operation and laborsaving application.



**Teruyuki Goto, Fujitsu Ltd.**  
Mr. Goto received the B.S. degree in Mathematics from Okayama University of Science, Okayama, Japan in 1990. He joined Fujitsu Ltd., Kawasaki, Japan in 1990, where he has been engaged in research and development of replication technology for database management systems.

For example, to provide stable operation, we have included functions such as switching within seconds, a cluster system for cluster scale-out, instant backup, and operation at the world's highest speed. Our system has shown proven results in a wide range of customer applications. Moreover, we have realized an easy-to-use database environment that provides autonomic functions such as automatic reorganization and automated grid technology for laborsaving operation. Furthermore, it is provided with a long-term warranty and is highly evaluated by a large number of customers.

Symfoware Disaster Recovery combines the Symfoware high-reliability technology with Fujitsu's proven mainframe technology, new hardware technology, and other core technology areas to realize absolutely stable operation through the construction of backup centers that can handle high-end databases.

Furthermore, we are providing preventative maintenance measures, a database snapshot function, an operation function that is integrated with the storage system, and laborsaving operation through synergy with Fujitsu's TRIOLE IT infrastructure. Through these technologies, we have realized high-level business continuity for today's global corporate management.

### References

- 1) M. Erbschloe: Guide to Disaster Recovery. Course Technology Ptr (Sd), 2003.
- 2) M. Nagata: Distributed Database System. (in Japanese), Japan, Sobunkan, 1989.