Vertically Integrated Database System Combining Advanced Hardware with Database Technologies

● Yosuke Miyauchi   ● Ryohei Takahashi   ● Mitsuhiro Kinomura

To survive in a highly competitive, rapidly changing business world, an enterprise needs information and communications technology (ICT) systems that can respond flexibly and quickly to the market. However, as ICT systems become increasingly complex, the need to reduce their management and operating costs increases. Databases, moreover, are ICT systems having high-reliability and high-performance requirements. The FUJITSU Integrated System HA Database Ready product integrates database technologies cultivated by Fujitsu over many years with advanced hardware. Providing "high-reliability and high-performance," "open-standards," and "ready-to-use" features, HA Database Ready can significantly reduce the costs of deploying, designing, and operating a database infrastructure. This paper introduces a fully redundant system that uses advanced hardware to support high-reliability and high-performance operation and Fujitsu's Smart Software Technology for performing a full range of tasks from system deployment to system operation, monitoring, and troubleshooting using a simple graphical user interface (GUI).

1. Introduction

In the face of severe and fast-changing competition, business managers have a keen interest in building up their business through information and communications technology (ICT), which enables them to respond in a flexible and rapid manner to diverse needs such as creating a new line of business or accelerating the management process. However, a complex and massive set of ICT assets incurs costs, and finding a means of reducing them is becoming an important issue that cannot be ignored if the company wants to improve its competitiveness.

FUJITSU Integrated System HA Database Ready[1] is a database system that can eliminate complexity in existing ICT assets and meet the need for both reducing maintenance and operation costs and expanding business. The value that HA Database Ready can bring to a customer is threefold: high performance by optimally combining hardware and software, reduction of setup costs by enabling a database system to be easily set up in a short period of time compared to past systems, and high operability though a dedicated graphical user interface (GUI).

This paper begins by introducing the HA Database Ready product concept. It then describes its high-performance and high-reliability features, the open-standards interface, and the measures taken to simplify implementation and operation.

2. Product concept

One of the objectives of HA Database Ready is to reduce the maintenance and management costs of database systems, which tend to increase on a yearly basis. HA Database Ready is a vertically integrated system that can reduce total costs while maintaining the performance and reliability expected of a database system. The configuration decided on for HA Database Ready is based on the results and know-how that Fujitsu has accumulated over many years in the solutions and services business—high performance and high reliability are key features of this configuration.

Furthermore, in the area of operability, all tasks from "implementation" to "operation" of the database system have been simplified to the utmost in HA Database Ready. This simplified operation reduces the cost of setting up and operating a database system.

FUJITSU Sci. Tech. J., Vol. 51, No. 2, pp. 55–61 (April 2015)

55

In addition, the design of HA Database Ready places importance on the rapid incorporation of the systems that are needed to create a new line of business or speed-up management tasks. Up to now, Fujitsu has been providing the FUJITSU Software Symfoware Server[2] as a high-performance, high-reliability database for a wide variety of customers, and the emphasis has been on large-scale, social-infrastructure ICT systems that demand high performance and high reliability as in stock exchange systems.[3] Today, however, customers seek assistance in creating new value, and to this end, importance is coming to be placed on combining existing ICT systems with solutions and services such as cloud computing and big data. Under these conditions, database systems, in addition to supporting existing ICT systems, are coming to play a major role as an infrastructure for flexibly and rapidly expanding new services through novel combinations with existing systems.

Consequently, in addition to Symfoware Server, it was decided to adopt a database that incorporates the high-performance, high-reliability technologies of Symfoware Server in the PostgreSQL open-source database. Using PostgreSQL in this way makes it easier to work with other types of software and provides a common application development environment across the entire system. Additionally, combining the high-performance and high-reliability technologies of Symfoware Server and PostgreSQL can realize an open-standards database system that can be used for mission-critical applications.

The following sections describe the feature technologies of HA Database Ready.

3. High performance and high reliability

3.1 High transaction performance

In general, the performance issue that is most likely to arise in a database is I/O bottleneck. In particular, for systems managing medium-scale to large-scale volumes of data, the amount of data that can be loaded into cache memory is only a small part of total data, so situations may occur in which most of the time taken up by Structured Query Language (SQL) execution involves I/O. This problem has been conventionally dealt with by performing complicated tuning in accordance with the customer's performance requirements.

HA Database Ready eliminates this need for tuning by mounting PCIe solid-state drives (SSDs) capable of high-speed I/O in Fujitsu's high-performance FUJITSU Server PRIMERGY PC server.[4] An SSD is a storage device containing flash memory chips used for recording data. It features exceptionally high-speed data reading/writing compared to a hard disk drive (HDD). In fact, an SSD has no seek time or search time as in the case of an HDD, so SSD random-access performance is about the same as sequential-access performance.

In addition, each PCIe SSD mounted in HA Database Ready connects directly to the PCIe bus. This scheme eliminates the I/O bottleneck and maximizes CPU performance, enabling HA Database Ready to achieve exceptionally high transaction throughput. However, simply mounting a PCIe SSD is not sufficient for greatly increasing the performance of online transaction processing (OLTP) in the case of huge volumes of data. HA Database Ready therefore incorporates the following optimization methods for achieving ultra-high-speed OLTP by making efficient and maximum use of PCIe SSD performance.

1) Placing only resources causing I/O bottlenecks on a PCIe SSD

In database processing, a huge amount of writing to data files, transaction logs, etc. can take place, and this can cause I/O bottleneck. To resolve this problem, we place all database resources on PCIe SSDs. As a result, large-volume transaction processing and batch processing can be executed at high speeds even in the case of a low cache-hit ratio.

2) Tuning to maximize PCIe SSD random-access performance

The database is tuned so that an optimizer can create an access plan that achieves equivalent costs for sequential access and random access. The result is highly efficient SQL execution when using PCIe SSDs.

In addition, I/O access to PCIe SSDs can be made more efficient by adjusting the data block size and kernel parameters so as to optimize both buffer consumption and the frequency of I/O access.

These optimization methods enable HA Database Ready to achieve a throughput about 20 times that of a conventional database using HDDs (in-house comparison), as shown in **Figure 1**.
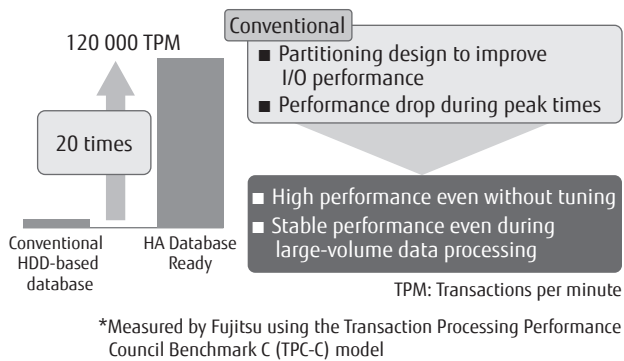
Figure 1
HA Database Ready features.

Figure 2
Data triplication by mirroring and automatic backup.

3.2 High-reliability database system

HA Database Ready incorporates as standard a redundant configuration for both hardware and software to achieve a high-reliability database system.

1) Completely redundant hardware

HA Database Ready duplicates a wide range of hardware including server units, storage units, and network devices, thereby avoiding a single point of failure (i.e., the condition in which a failure at one location causes the entire system to fail). This redundant configuration supports business continuity by automatically switching to a secondary server in the event of an abnormality such as a hardware failure.

2) Data triplication

The data on the primary server is continuously replicated on the secondary server by using the mirroring technology[5] of Symfoware Server, thereby achieving a completely redundant configuration in an active standby system. This enables business to continue by switching to the secondary server after several seconds in the event of a fault in the primary server. In addition, database resources are automatically backed up onto a storage unit to triplicate system data (**Figure 2**). Consequently, if both the primary server and secondary server should fail, this backup data can be used to return the database to its latest state through a simple operation.

In the above ways, HA Database Ready protects business data by supporting business continuity in case of failure.
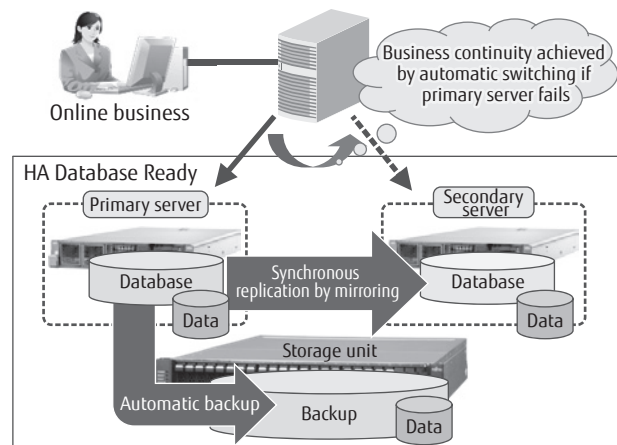
4. Open standards
4.1 Ecosystem access through PostgreSQL

PostgreSQL[6] is open-source software (OSS) developed by a community consisting of about 40 000 members worldwide. An ecosystem linking various types of OSS and independent software vendors has already been formed around PostgreSQL. In addition to the existing Symfoware interface, HA Database Ready uses PostgreSQL as a database interface so that the ecosystem formed around PostgreSQL can be used. Selecting the PostgreSQL interface means that applications currently used with PostgreSQL can be directly run with HA Database Ready without modification. Also, it means a lower learning cost when developing business systems, and it results in systems that easily accommodate users compared with past systems.

Application servers, information application products, and development tools that can use PostgreSQL are many and varied, as shown by the examples listed below. The PostgreSQL community has introduced many PostgreSQL-related products.[6]

1) Application servers
- JBoss Application Server
- Glassfish Server
- Interstage Application Server

2) Information application products
- ASTERIA WARP
- Jaspersoft Business Intelligence Suite
- Oracle Business Intelligence
- Pentaho Open BI Suite

- RapidMiner
3) Development tools
- SI Object Browser
- Eclipse

In short, using the PostgreSQL interface of HA Database Ready in this way means that a user has access to techniques and know-how shared by other users around the world. In addition, using various types of upper-level middleware compatible with PostgreSQL can make system development even more efficient.

4.2 Enhanced operability and reliability through PostgreSQL

HA Database Ready is the only database appliance that applies PostgreSQL to mission-critical systems. Users can therefore construct an enterprise-oriented database system by simply applying their know-how of PostgreSQL.

For systems that fall into the mission-critical category, high operability and high reliability must be achieved to ensure that operations continue uninterrupted. Fujitsu has introduced Symfoware Server into many areas that require mission-critical systems such as finance, medicine, distribution, and social infrastructures. By incorporating the technologies and know-how that it has accumulated through database development and system construction in PostgreSQL, Fujitsu has achieved a database system featuring both high operability and high reliability as demanded by mission-critical systems.

1) Mirroring technology and smart recovery technology

As described in the previous section, HA Database Ready duplicates the database through the mirroring technology of Symfoware Server. It also uses the smart recovery technology of Symfoware Server to automatically determine the type of recovery process that should be taken at the time of an abnormality or failure on the basis of abnormality/failure patterns, thus enabling recovery to be executed using a one-click operation.

With these mirroring and smart recovery technologies, HA Database Ready can easily achieve a high-reliability system that would be troublesome to achieve with only PostgreSQL.

2) Storage management technology

Fujitsu has developed Extended Storage Manager (ESM) for PostgreSQL[7] as a follow-up to its existing

Symfoware Server storage management technology. HA Database Ready uses this ESM for PostgreSQL storage management in the form of a transaction-log duplication function and a data-file corruption-checking function to raise the reliability of PostgreSQL with respect to data.

- Transaction-log duplication function

This function writes the transaction log not only onto the primary and secondary servers but also into a backup area on the storage unit at the time of a data update. It can also recover the latest version of this log at the time of a server disk failure or a double failure of the primary/secondary servers during which time PostgreSQL archive log operations cannot be performed.

- Data-file corruption-checking function

This function stores page-by-page checksum values[note] and compares these values with data to check for corrupted data. If such data is detected, it suspends processing and reports an error. It also uses proprietary technology to minimize the overhead incurred by these operations and to prevent the spread of data corruption at the time of an abnormality.

3) Enhanced security

The encryption function of HA Database Ready uses the 256-bit key of the Advanced Encryption Standard (AES)—the strongest of its key options—as its encryption algorithm. It also links with Intel AES New Instructions (AES-NI) incorporated in HA Database Ready as new encryption-related technology to carry out data encryption and decryption without degrading performance.

The encryption function of PostgreSQL has also been extended to protect stored data by transparent data encryption. Data stored in the database is encrypted when writing the data file and decrypted when reading it. HA Database Ready automatically performs these processes to protect stored data. In this way, encryption/decryption of stored data can be executed without users or applications having to be aware of it.

5. Simplified implementation and operation
5.1 Simplified implementation

With HA Database Ready, hardware connections and settings and OS and software installation/tuning

---

note) Values calculated to check data reliability.

are already completed at the time of shipping. This eliminates the need for integration work (design/implementation/testing) on the customer side. The following gives some examples of settings that are made beforehand.

- Inter-device time synchronization
- HA Database Ready internal monitoring
- Network redundancy
- Backup/Recovery
- Log rotation
- Database mirroring
- Performance tuning

Applying power to the system has also been simplified. After installation, only power to the primary server needs be turned on to start up the entire HA Database Ready system. The setup is completed by using the Appliance Manager (operations management GUI) wizard to enter the management ID, IP address, and database information (interface, character code, and port number).

Simplifying implementation in this way significantly shortens the period from installation to database use, which in the past could take a couple of months. Now, the database can be used on the day it is installed.

5.2 Simplified operation

Appliance Manager is equipped with a function for monitoring the operation of HA Database Ready. It represents the state of HA Database Ready operation by displaying one of three simple icons corresponding to normal, warning, and abnormal conditions (**Figure 3**). Through the following types of monitoring, Appliance Manager changes the displayed icon from normal to warning or abnormal upon detecting an abnormality and displays an appropriate response method.

1) Monitoring related to database switching/halting
- Database process survival status
- Hardware device abnormality
- Network state abnormality
2) Monitoring related to database resources
- Back up execution results
- Archive log capacity
- Data storage capacity

5.3 One-click recovery

HA Database Ready is equipped with "one-click recovery": a recovery procedure is automatically selected and executed in accordance with the abnormal event that has occurred. The procedure shown in **Figure 4** is executed for any kind of hardware failure. The recovery functions support recovery of both the hardware settings and the database.

1) Hardware settings recovery

HA Database Ready duplicates all hardware settings on servers and in storage to simplify recovery after a hardware failure. Once the work for replacing a hardware unit has been completed, clicking the recovery button automatically applies these stored settings, thereby eliminating the need to reconfigure the new hardware. In the case of PCIe SSDs, for example, necessary partitions are automatically created during the recovery process, and, in the case of a network that has turned into a single-system operation due to a
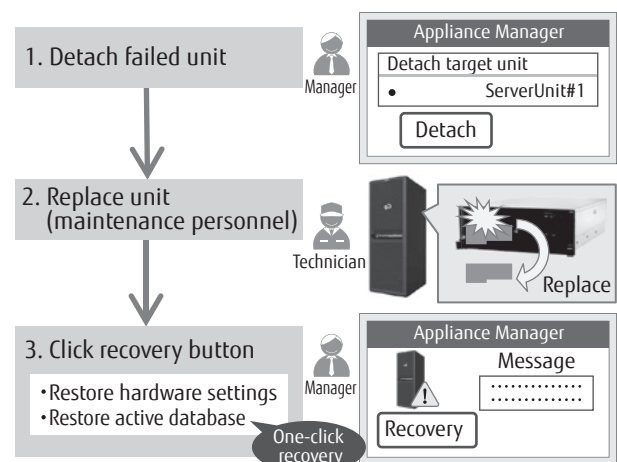

Figure 3
Display of operating state by Appliance Manager.

Figure 4
HA Database Ready recovery procedure.

Table 1
Database recovery patterns.

Event	Recovery procedure	
	Symfoware interface	PostgreSQL interface
Active database is corrupted	• Restore data from backup area to active database	• Restore from active backup database to corrupted/damaged database
Backup area is corrupted	• Recreate backup area • Redo backup	• Recreate backup area • Redo backup
Both databases are corrupted	• Restore data from backup area to both databases	• Restore data from backup area to one database • Copy data in restored database to other database
One database and backup area are corrupted	• Recreate backup area • Redo backup • Restore data from backup area to one database	• Recreate backup area • Redo backup • Copy data from active database to damaged database

disconnected network cable, a redundant configuration is automatically reestablished.

2)   Database recovery

If either of the databases shown in the database configuration of Figure 2 should become corrupted, that database will be recovered using one of the four patterns listed in **Table 1**.

Once the system manager clicks the recovery button, HA Database Ready automatically determines the recovery pattern matching the event that has just occurred and recovers the database using the method appropriate for that pattern. There is therefore no need for the system manager to create and test a recovery procedure corresponding to the type of database failure, which means a great reduction in operating costs.

6.   Conclusion

This paper described FUJITSU Integrated System HA Database Ready, a database appliance that integrates database technology and advanced hardware on the basis of the results and know-how that Fujitsu has accumulated over many years in the solutions and services business.

HA Database Ready achieves high performance by optimally arranging resources on high-performance PCIe SSDs and achieves high reliability by duplicating hardware in a redundant configuration. Furthermore, its use of PostgreSQL as an open-standards database makes it easier to interface and link with other products and promotes the formation of an ecosystem. In addition, the Appliance Manager operations management GUI of HA Database Ready simplifies system setup, visualizes current operating conditions, and simplifies operations by enabling recovery from any kind of

abnormality by using a one-click operation.

Fujitsu's customers make use of the functions and technologies of HA Database Ready to meet their medium-scale to large-scale system needs. Customers are also adopting HA Database Ready for mission-critical areas.

Going forward, the plan is to meet both the need for reducing maintenance and operation costs and the need for expanding business by providing a database consolidation scheme that incorporates cutting-edge, high-performance hardware and extracts maximum hardware performance. Fujitsu is committed to pursuing state-of-the-art, high-performance and high-reliability technologies to develop products that are increasingly easier for its customers to use.

References
1)   Fujitsu: Fujitsu Enhances Lineup of Vertically Integrated Database Systems.
http://www.fujitsu.com/global/about/resources/news/press-releases/2014/0228-02.html
2)   Fujitsu: Symfoware Server.
http://www.fujitsu.com/global/products/software/middleware/database/symfoware/products/symfoware-server/
3)   Fujitsu: TSE Launches Next-Generation "arrowhead" Trading System.
http://www.fujitsu.com/global/about/resources/news/press-releases/2010/0108-01.html
4)   Fujitsu: FUJITSU Server PRIMERGY.
http://www.fujitsu.com/global/products/computing/servers/primergy/
5)   Fujitsu: FUJITSU Software Symfoware.
http://www.fujitsu.com/global/products/software/middleware/database/symfoware/
6)   PGECons: PostgreSQL Enterprise Consortium.
https://www.pgecons.org/en/about/

7) Fujitsu: Database PowerGres Plus (in Japanese).
   http://software.fujitsu.com/jp/powergresplus/

Yosuke Miyauchi
Fujitsu Ltd.
Mr. Miyauchi is engaged in the development of HA Database Ready.

Mitsuhiro Kinomura
Fujitsu Ltd.
Mr. Kinomura is engaged in the development of HA Database Ready.

Ryohei Takahashi
Fujitsu Ltd.
Mr. Takahashi is engaged in the development of HA Database Ready.

FUJITSU Sci. Tech. J., Vol. 51, No. 2 (April 2015)

61