Express5800/Scalable HA Server Achieving High Reliability and Scalability

NASU Yasuyuki

Abstract

Locating the target data as close as possible to the CPU is of importance in the high-speed processing of a large amount of big data. The Express5800/Scalable HA Server series are servers optimized for use as big data processing platforms capable of using a large-capacity memory of up to 2 TB. Building a high cost-efficiency system using a high-speed PCI Express SSD is attracting recent market attention. The Express5800/Scalable HA Server series products are suitable for such a system because they support multiple I/O slots. This paper introduces their excellent advantages and discusses the actual examples in which their features may be usefully applied.

Keywords

server, memory, SSD, PCI Express, in-memory, database

1. Introduction

Some technologies are attracting attention for use in efficient, high-speed analyses of the big data that continues to multiply. Among them, the parallel distributed processing and in-memory processing are the two technologies that are receiving the most attention. The representative product applying the parallel distributed processing technology is the open-source middleware “Hadoop.” This product implements high-speed big data processing by finely dividing a large amount of data, applying distributed parallel processing and aggregating the results. On the other hand, in-memory processing enables data updating and analysis at ultrahigh speeds by the short term placing of part or all of the processing target data in the main memory. When the amount of processed data is too large to be placed wholly in the main memory, a high efficiency system may be built by utilizing newly developed high-performance storage devices such as a PCI Express SSD (Solid State Drive).

2. Outline of the Express5800/Scalable HA Server Series

The Express5800/A1080a (**Photo 1**) is the high-end model of the Express5800/Scalable HA Server series. It is a high-performance server capable of simultaneous parallel processing of 160 threads by integrating up to 8 CPUs of the Intel Xeon processor E7-8800/4800 product family that have up to 10 cores per CPU. With a large-capacity memory of up to 2 TB, and 14 PCI Express slots, it is a server optimized for use as a big data processing platform that is capable of processing a large amount of data with high efficiency.

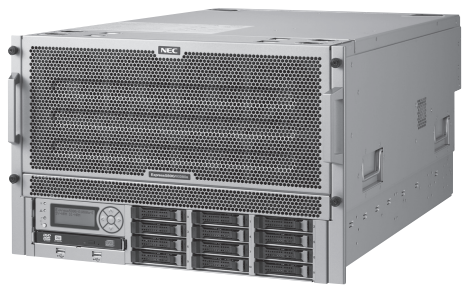

Photo 1 External view of the Express5800/A1080a and the Express5800/A1040a.

The Express5800/A1080a is available in three variations. These are the A1080a-S with a node for up to 4 CPUs, the A1080a-D with two nodes for up to 4 CPUs and the A1080a-E with a node for 8 CPUs. As shown in **Fig. 1** , these various types may be exchanged or extended easily as required. For example, if the required processing capability is small, the A1080a-S may be first introduced in order to reduce the initial cost and subsequently its performance may be improved to that of the A1080a-E if an improved processing performance is required.

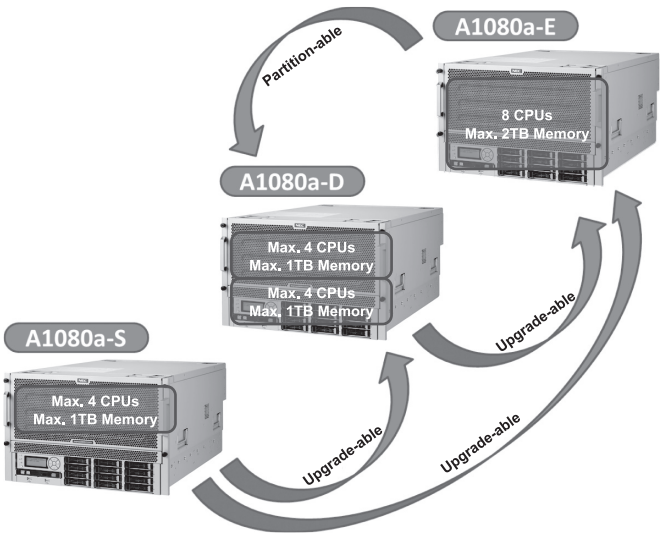

Fig. 1 Diagram of the model change.

Photo 2 External view of the Express5800/A1020a.

"Long-term maintenance-guaranteed models [*1] " are also available, with which the servers may be used securely for up to ten years by concluding a long-term maintenance service agreement. Extension of the server life will reduce the high cost of system updating that would accompany a replacement of servers.

In addition, the Express5800/Scalable HA Server series also includes other products for dealing with various mission-critical tasks. The Express5800/A1040a incorporates up to 4 CPUs and a 1 TB large-capacity memory with an 80-thread simultaneous parallel processing capability (Photo 1). The Express5800/A1020a [*1] (**Photo 2**) incorporates up to 2 CPUs of the Xeon processor E5-2600 product family that have up to 8 cores per CPU with up to 384 GB memory and a 32-thread simultaneous parallel processing capability.

The Express5800/A1080a and Express5800/A1040a achieve high reliability thanks to the EXPRESSSCOPE engine SP2 that features an excellent fault analysis capability/diagnosis function. They also benefit from the outstanding design and the rigorous shipment inspections based on our high-reliability parts selection criteria that result from our long years of mainframe development.

For use with the Express5800/Scalable HA Server series, we offer the "Enterprise Linux with Dependable Support" to support construction of a Linux platform for mission-critical tasks. Together with the "MC SCOPE," a platform middleware product that enables detection and investigation into the causes of system faults, and "Linux Dependable Support." This is a Linux support service related to technology and operation that implements a highly accessible Linux-based system. [*2]

2.1 Express5800/A1080a Architecture

The Express5800/A1080a-E connects eight CPUs via a high-speed interface called the QPI (QuickPath Interconnect) to achieve high performance and scalability. The QPI interface is also used to connect two CPUs and one IOH (I/O Hub) (**Fig. 2**).

Each CPU has 16 DDR3 DIMM slots that are connected via four MBs (Memory Buffers,) so up to 128 DDR3 DIMM slots may be used. 128 16 GB DIMMs are used to deal with the large-capacity memory of up to 2 TB.

Each IOH has 3 or 4 PCI Express 2.0 slots that can mount up to 14 PCI Express cards. Among the 14 PCI Express slots, two of them support 16 lanes.

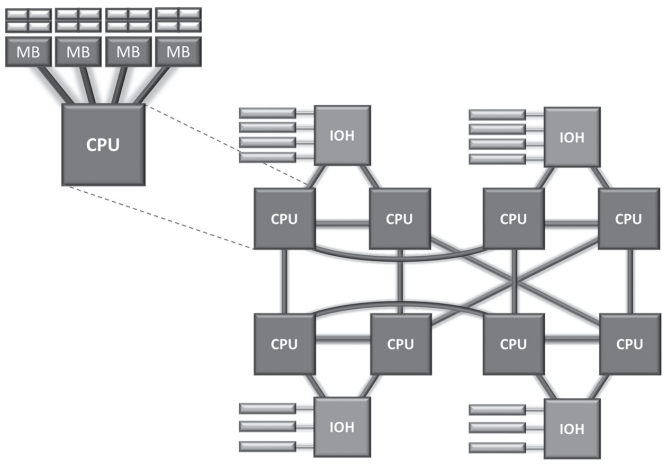

Fig. 2 Block diagram of the Express5800/A1080a-E.

[*1] At present, this model is available only in Japan.
[*2] At present, this service is available only in Japan.

2.2 Capacity Optimization (COPT) Function

The product lineup of the Express5800/A1080a series includes a model that supports the COPT (Capacity Optimization) function. This function makes it possible to introduce reserve CPU cores in the disabled (unavailable) status and to enable each of them individually whenever necessary during operation. This system not only facilitates countermeasures to deal with increased loads but also makes it possible to eliminate unnecessary investment if the software in use requires the same number of licenses as the number of CPU cores used. The number of software licenses can be as small as being equal to the number of CPU cores, even if the CPUs need to be increased in order to increase the memory capacity or to use many PCI Express slots.

2.3 RAS Function of Express5800/A1080a and Express5800/A1040a

The RAS (Reliability, Availability and Serviceability) of the main memory is important for the secure use of a large-capacity memory.

These servers support the DDDC (Double Device Data Correction) function that can correct errors while continuing the system operation even if the two DRAM chips forming the main memory fail simultaneously.

They also support the demand scrubbing function that reads the main memory, corrects the detected correctable errors and writes the corrected data back in the main memory. A patrol scrubbing function reads the entire data in the main memory in a certain period and corrects any detected correctable errors before they degrade into uncorrectable errors.

In addition, when the Windows Server 2008 R2 is used, these servers can access the MCA (Machine Check Architecture) Recovery function. This function reduces the potential for system failure due to uncorrectable errors being detected by the hardware during writing the cache memory data back into the main memory and during patrol scrubbing of the main memory. This is done by recovering or restarting the process related to the data in the error locations in association with the OS.

3. Actual Examples of Use with Big Data

The Express5800/Scalable HA Server series inherits the design concepts of NEC supercomputers and mainframes that have evolved over long years. It can thereby offer excellent performance together with high reliability and availability. These servers are adopted in many enterprise based mission-critical systems due to their excellent characteristics. When they are combined with the big data processing software introduced in the following subsections, they can also build big data platforms with high cost efficiencies.

3.1 InfoFrame TAM (Table Access Method)

The InfoFrame TAM (Table Access Method) is an in-memory database management software product that processes a large amount of transactions in real time by placing the processing target data in memory. With the Express5800/A1080 series the COPT function is used, the Express5800/A1080a incorporates the COPT function and can therefore optimize the customer's investment because the customer is required to purchase the same number of InfoFrame TAM licenses as the number of necessary CPU cores regardless of the mounted memory capacity.

This software was introduced with scalable HA servers with PCI Express slots that can be expanded up to 14 slots together with Enterprise Linux with Dependable Support in the construction of the next-generation information distribution platform of QUICK Corp., which is Japan's biggest financial information vender. As a result, we were able to implement a mission-critical service platform for the high-speed distribution of a large amount of information in the order of milliseconds.

3.2 InfoFrame DataBooster

The InfoFrame DataBooster is a software product that performs ultrahigh-speed data processing by means of an optimized data structure, processing algorithms and efficient internal parallel processing. As it executes processing after storing data in the memory, the use of a scalable HA server that can handle a large-capacity memory of up to 2 TB is secure, even if the data quantity is increased subsequently. Additional CPUs may sometimes be required in order to increase the memory capacity. However, if the Express5800/A1080a is used, the memory capacity can be increased without adding core licenses of the InfoFrame DataBooster, thanks to the built-in COPT function.

3.3 SAP HANA

The SAP HANA (High-Performance Analytic Appliance) is

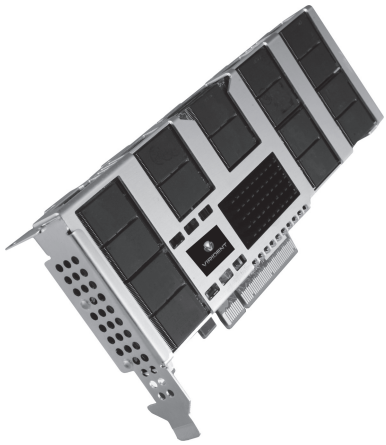

Photo 3 Virident Systems "FlashMAX."

an in-memory software product of SAP AG. Its featured high speed is made possible by compression of the databases of both column and row structures as well as by provision of those database with an in-memory processing. At the SAP Global Competence Center in Germany, NEC engineers verified the Express5800/A1080a and Express5800/A1040a in cooperation with SAP and SUSE engineers, which subsequently obtained approval as servers for the SAP HANA.

In marketing the appliance servers for the SAP HANA, we adopted the SLC-type FlashMAX, of Virident Systems, Inc., USA (**Photo 3** ,) a flash storage that has excellent write performance, in order to enable maximum performance.

The FlashMAX is a PCIe SSD. Unlike the SSDs connected via the same interface as the HDD such as the SAS and SATA, it is connected to the server by PCI Express to achieve high write/read throughputs.

Since the Express5800/A1080a has up to 14 PCI Express slots, a sufficient number of network interface cards and storage interface cards can be mounted, even when several FlashMAX cards are in use. This makes it possible to build a real-time analysis platform that matches the needs of the era of big data.

4. Conclusion

In the above, we introduced the Express5800/Scalable HA Server series that features high scalability optimized for use in big data processing platforms.

The Express5800/Scalable HA Server series functions not only as a big data processing platform for use in processing big data. The excellent reliability and availability also ensure that the series will continue to evolve as servers that can be used securely in platforms for supporting mission-critical tasks and in private cloud systems.

Reference

1) NEC's Appliance Server for SAP HANA(R) Certified by SAP
http://www.nec.com/en/press/201203/global_20120309_01.html

Author's Profile

NASU Yasuyuki
Manager, Product Planning
IT Hardware Operations Unit

The details about this paper can be seen at the following.

Related URL:
Scalable Enterprise Servers:
http://www.nec.com/en/global/prod/express/scalable/index.html
Scalable Enterprise Servers:
http://www.necam.com/Servers/Enterprise/
SAP Certified NEC Express5800 Servers:
http://www.nec.com/en/global/prod/express/related/sap_certified.html

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.7 No.2 Big Data

Vol.7 No.2
September, 2012

Special Issue TOP

Remarks for Special Issue on Big Data
NEC IT Infrastructure Transforms Big Data into New Value

◇ Papers for Special Issue

Big data processing platforms
Ultrahigh-Speed Data Analysis Platform “InfoFrame DWH Appliance”
UNIVERGE PF Series: Controlling Communication Flow with SDN Technology
InfoFrame Table Access Method for Real-Time Processing of Big Data
InfoFrame DataBooster for High-speed Processing of Big Data
“InfoFrame Relational Store,” a New Scale-Out Database for Big Data
Express5800/Scalable HA Server Achieving High Reliability and Scalability
OSS Hadoop Use in Big Data Processing

Big data processing infrastructure
Large-Capacity, High-Reliability Grid Storage: iStorage HS Series (HYDRAstor)

Data analysis platforms
“Information Assessment System” Supporting the Organization and Utilization of Data Stored on File Servers
Extremely-Large-Scale Biometric Authentication System - Its Practical Implementation
MasterScope: Features and Experimental Applications of System Invariant Analysis Technology

Information collection platforms
M2M and Big Data to Realize the Smart City
Development of Ultrahigh-Sensitivity Vibration Sensor Technology for Minute Vibration Detection, Its Applications

Advanced technologies to support big data processing
Key-Value Store “MD-HBase” Enables Multi-Dimensional Range Queries
Example-based Super Resolution to Achieve Fine Magnification of Low-Resolution Images
Text Analysis Technology for Big Data Utilization
The Most Advanced Data Mining of the Big Data Era
Scalable Processing of Geo-tagged Data in the Cloud
Blockmon: Flexible and High-Performance Big Data Stream Analytics Platform and its Use Cases

◇ General Papers

“A Community Development Support System” Using Digital Terrestrial TV