NEC Enterprise Server
Express5800/1000 Series

Empowered by Innovation

NEC

Express5800/1000 Series Guide (1)

Powered by Intel® Itanium® 2 Processor

Express5800/1000 Series

1320Xe/1160Xe/1080Xe RAS Technology

Three keys to success in real-time business

Reliability
Availability
Serviceability

intel Itanium 2 inside

intel

Enterprise Server Express5800/1000 Series Equipped with Powerful RAS Features to Support Mission-critical Systems

The demands on the enterprise server are increasing daily—and so are the costs of system failure. That's why real-time businesses need more than just increased processing capabilities. They also need to reinforce those capabilities with powerful RAS (Reliability, Availability and Serviceability) features.

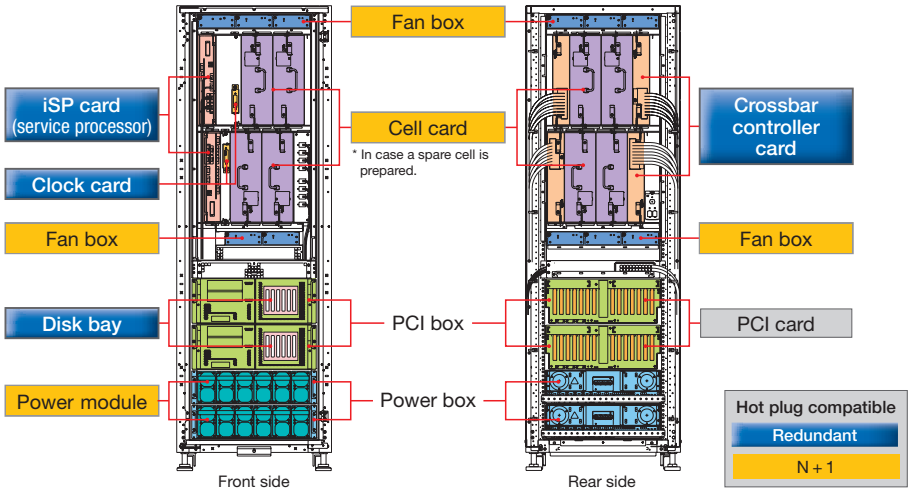
The Express5800/1000 Series (AsAmA*) fulfills the needs of IT system infrastructures in five ways, providing:

- **Ample processing power** delivered by Intel® Itanium® 2 processors.
- **An architecture that permits hardware partitioning.** The ability to divide the server into multiple nodes means that each node can function as an independent server with an independent operating system, resulting in greater reliability. These servers come with Red Hat Enterprise Linux AS v.4, and Windows Server 2003 so administrators can tailor the system to match the needs of current business environments.
- **An optimized NEC chipset** designed explicitly to maximize the performance of the Intel® Itanium® 2 processors. This chipset incorporates expertise and features refined during many years of building NEC mainframes and supercomputers.
- **An autonomic service processor/manager** that takes care of the server whether the maintenance staff is there or not, providing platform troubleshooting features such as failure monitoring, failure log recording, and log analysis. This processor also provides server management elements such as system boot, initial diagnosis and configuration management of hardware, power control, and partition control.
- **Design modularization and redundancy** to prevent the entire system from going down and make simple online maintenance possible.

But these are just the highlights of the Express5800/1000 Series Enterprise Servers. Let's look more closely at the careful engineering that pays off for the real-time business.

* AsAmA: A development code name for the NEC enterprise server platform, including Express5800/1320Xe, 1160Xe, and 1080Xe for Express5800/1000 Series.

Figure 1: Configuration sample for the Express5800/1000 Series (Express5800/1320Xe)



Ample Processing Power

Enterprise-class Performance, Availability and Reliability Delivered by the Intel® Itanium® 2 Processor

The Intel® Itanium® 2 processor was designed expressly for large-scale, mission-critical business environments and systems that perform large-volume operations. That is why more than 50 Global 100 companies* rely on this processor, and why it is at the heart of 46 of the world's Top 500 supercomputers**.

These demanding facilities prefer the Intel® Itanium® 2 processor for three reasons:

More parallel instructions—simply

The new EPIC (Explicitly Parallel Instruction Computing) architecture of the Intel® Itanium® 2 processor adds more parallel instructions to improve performance without complicating the electric circuit. Conventional parallel processing requires a preliminary judgment about which instructions can be executed in parallel, which in turn requires more complex circuitry. EPIC technology simply recognizes these instructions from the start.

Automatic recovery from errors

MCA (Machine Check Architecture) provides a 3-stage troubleshooting mechanism that passes on any unresolved errors to the next stage for handling:

Hardware CPU and chipset Stage 1 error handling → **Firmware** Stage 2 error handling → **Operating System** Stage 3 error handling

If any errors still remain after the third stage, the operating system then runs recovery procedures based on the error report and error log it has received. If a critical error has occurred, the system resets automatically to significantly lower the risk of server system failure.

ECC/parity protection for high data integrity

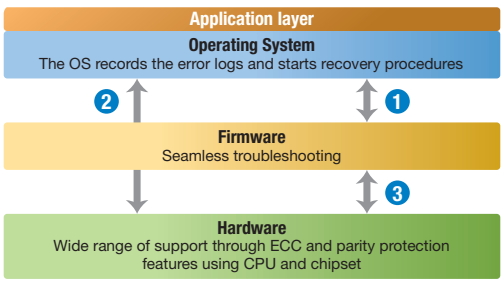
System failure can occur when alpha rays or cosmic rays that randomly irradiate the transistor change its state, causing faults in the semiconductor memory and related software errors. This type of error is difficult to avoid even with strict IT controls that employ repeated verification. Just as damaging are hardware errors involving memory and failures of lines and devices. To ensure high performance and reliability, the platform must detect and resolve both software and hardware errors, and must work to ensure data integrity.

The Intel® Itanium® 2 processor accomplishes this with ECC (Error Checking and Correction) and parity protection features. These methods apply to the main data path, cache, TLB (Translation Look-aside Buffer) and other areas, with the following specific effects:

ECC	Parity Protection
L3 cache	L2 instruction cache
L3 TAG	L2 TAG
L2 D cache	L1 data/instruction cache
Corrects 1-bit errors	L1 data/instruction TAG
Detects 2-bit hardware errors	TLB

When errors are detected, the firmware reloads the correct data retained in the lower storage layer.

Figure 2: MCA: Machine Check Architecture



1. Firmware and OS help correct complex platform errors and restore the platform.
2. Error logs and OS report flows provide all details.
3. Wide range of hardware error detection and correction applies to all areas of the main data structure.

* Intel research result, as of August 2005.
** World's supercomputer Top 500 list (www.top500.org), as of November 2005.

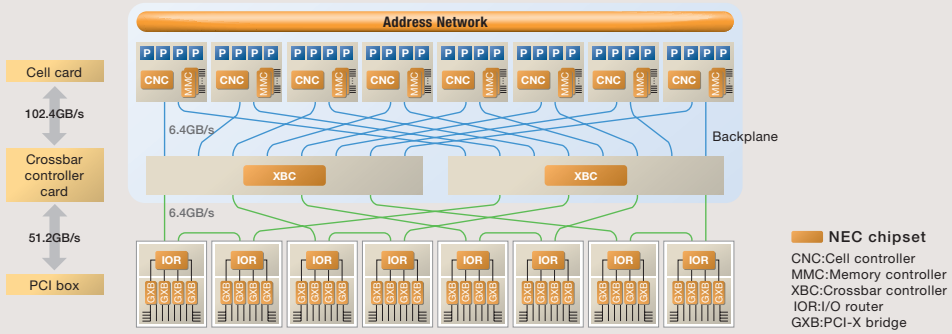
Architecture Permits Hardware Partitioning

Partitions Enable Independent Operations under Multiple OSs for Increased RAS

A modular architecture makes it possible to divide the Express5800 server into independent partitions, each capable of running under a separate operating system. A partition can support up to eight cells, with as many as four CPUs allocated for each cell. Partitions occur on the cell and PCI box boundaries, and are separated by hardware so a software error does not affect the security or availability of another partition. This modularity makes it possible to isolate workloads, improves the ability to withstand a crash of the entire system, and increases the efficiency of the hardware resource utilization.

In addition, the ability to run separate partitions under different operating systems permits building a flexible system to adjust to changes in the business environment. For example, increasing the use rate of the system environment and optimizing system resource allotment per application can have a profound effect on efficiency and general performance of the server.

Figure 3: Express5800/1320Xe architecture



Optimized NEC Chipset

In-House, high-performance chipset equipped with powerful RAS features to maximize processor performance

The Express5800/1000 Series* servers use a chipset designed specifically by NEC with one purpose in mind: to get the highest possible performance from these servers. NEC was able to apply enormous expertise derived from years of designing powerful mainframes and supercomputers to create a chipset with features such as partial chipset reduction, multiple error checkers, advance error checker, and acquisition of extensive error analysis data.

* Express5800/1320Xe, 1160Xe, and 1080Xe models.

ECC/parity protection features strengthen data integrity through function to retransmit error data

ECC/parity protection features come preloaded with the Intel® Itanium® 2 processor, so these features are also loaded with the Express5800/1000 Series chipset. The ECC is first set on the main data path of the chipset; then, when 1-bit errors are detected, the hardware corrects the data. To strengthen data integrity, chipset interfaces support the multi-bit error detection function and the error data retransmitting function.

Error check and correction accuracy also benefit from use of SDDC (Single Device Data Correction) memory, memory scrubbing and other functions in the memory area. When correctable errors occur frequently, this memory page is dynamically deallocated by OS.*

* Supported OS are Windows Server 2003.

Redundancy and high-speed interface realize high reliability and performance

The NEC chipset incorporates strict error check and correction functions and a redundant data path to minimize the risk of system failure. Interconnect devices between the cell card, crossbar controller card and PCI box increase the data throughput by employing the high bandwidth of 6.4 GB/s per path. High reliability also results from main data path redundancy and the error check function.

Memory RAS features permit prompt detection of memory error, prompt recovery, and reduced occurrence of errors

Even when faced with memory chip failures, memory-related RAS features ensure the server can rapidly detect memory errors, reduce multi-bit errors, and operate without interruption. These features include:

- Memory scanning runs on all the memories that are onboard during bootup and detects errors before OS boot. This function isolates errors when they occur, which prevents any effects on actual business processes.
- Patrol & scrubbing checks memory content regularly (every few milliseconds) during operation without affecting performance, writes corrected data when an error is detected, and reports the error. This patrol & scrubbing function is effective in detecting errors quickly and reducing multi-bit errors.
- SDDC memory is a memory system loaded with several DRAM chips that can correct errors at the chip level. If a memory chip fails, the error can be corrected immediately to enable uninterrupted operation. If more than a prescribed number of correctable errors are detected, the memory page supported by the OS is isolated without interrupting operations. This prevents the failure spreading to other areas.*

* Supported OS are Windows Server 2003.

Partial reduction of chipset prevents multi-partition failure by reducing only error area

Partial reduction of the chipset means the Express5800/1000 Series chipset can prevent failure from spreading inside the chipset, but can also perform data path redundancy and error detection and correction.

Each LSI in the chipset consists of smaller sub-units. A single partition includes several connected sub-units. When an error occurs in an LSI sub-unit, that sub-unit is reduced to isolate only the partition to which the sub-unit belongs, thus preventing the damage from spreading to other partitions.

For example, assume the partition is divided as shown in Figure 4 and an error occurs in the crossbar controller. Only Partition ① is isolated, leaving other partitions unaffected and able to continue business operations. In addition, after the defective subunit is reduced, the failed Partition ① can automatically reboot itself without intervention from an administrator, so business operations recover immediately.

Collection and analysis of high-volume log data enable identification of location and cause of failure for prompt recovery

In order to identify the location and cause of the failure so the server can recover, it is first necessary to collect and analyze error data. NEC's substantial experience with mainframe technologies yields yet another benefit, for the Express5800/1000 Series chipset has extensive capabilities to collect and analyze error logs that are unique to these servers.

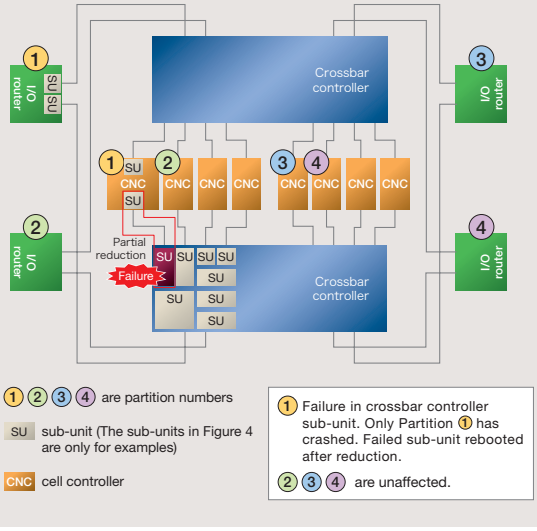
The chipset contains multiple error checkers (failure detection circuits) to detect errors without fail. An EIF (Error Indicating Flip-flop) retains the error check results, and a dedicated error analysis tracer retains passing data content. The chipset buffer content, which contains high volumes of passing data, can also provide log data.

These extensive error analysis results make it quick and simple to identify the exact location and cause of a failure. The error log data detected and collected by the error checker also makes it possible to identify the location of the failure through the service processor diagnosis function (built-in diagnosis) below, enabling simple and prompt identification and correction of the failure.

Testing the error checker in advance prevents defective error detection

An especially powerful RAS feature contained in the Express5800/1000 Series chipset is the capability to test the error checker after every system bootup. This diagnostic function prevents any failures in error detection during actual business operations.

Figure 4: Example of partial reduction inside chipset



Autonomic Service Processor/Manager

Server Troubleshooting/Configuration Functions at Core of RAS

Automated operation and recovery when failure occurs

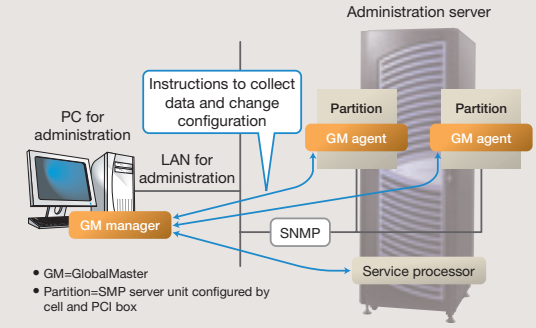
The service processor is the heart of RAS functionality for the Express5800/1000 Series*. This autonomic processor works over a dedicated connection to perform hardware initialization, configuration, diagnosis and power control management, plus platform error handling through monitoring and logging.

NEC's SystemGlobe GlobalMaster software works in cooperation with the service processor to perform these functions. This joint effort provides flexible and autonomic configuration control, such as virtualization, autonomic recovery, and autonomic adjustment. For example, the administrator can set up an optional recovery policy. The system can then automatically recover failed parts and—based on the administrator's recovery policy—optimize for an increased load.

A single user interface makes it simple to manage partitions.

* Express5800/1320Xe, 1160Xe, and 1080Xe models.

Figure 5: Autonomic configuration control



Extraction of log information/FRU analysis/reduction permits automatic recovery

When a failure occurs, the service processor uses a built-in diagnostic process to examine the error log provided by the chipset. With this information, it can identify the location of the failure and related FRU (Field Replaceable Unit). The system then reports as FRU information the location in the rack and the module, replaceable units and false rate. The service processor determines the reduced unit based on FRU information, detaches the unit, and reboots automatically. A report of the failure then automatically goes to the OS and the administrator, who can quickly replace any faulty units by referring to the FRU information.

The service processor also implements advanced troubleshooting using MCA, in collaboration with the chipset and firmware.

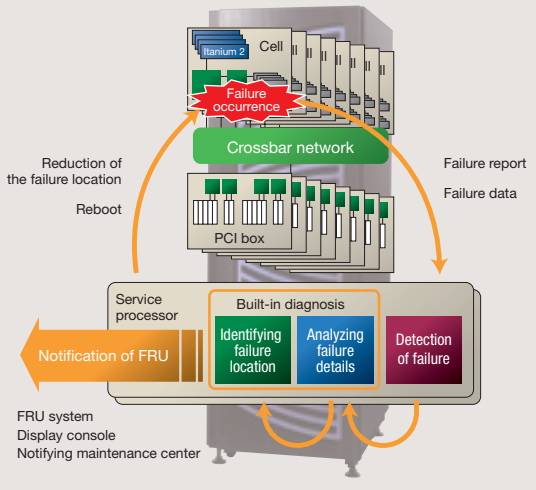
Console management, remote maintenance achieve flexible management

Console management enables serial console redirection and display of boot phase details for each partition. Remote maintenance features enable failure reports to be sent to, and logs to be collected from, a remote location. These elements offer flexible operation and management of RAS capabilities.

The service processor also supports a redundant configuration.* This can prevent RAS features from being disabled if a failure occurs in the service processor. A redundant service processor is achieved through the master-slave configuration, with the switch between master and slave occurring within a few seconds of a failure.

* 1320Xe and 1160Xe support this feature.

Figure 6: Service processor: Detection/analysis/identification of failure based on log data



Design Modularization and Redundancy

Redundancy and Hot Plug Support Module Prevent System Failure, Permit Simple Online Maintenance

Modularization/redundancy minimizes system crashes from single-point failures

Most Express5800/1000 Series* hardware components, including the clock module and power module, are redundant (2N or N+1). This minimizes the risk of single-point failures to prevent entire system crashes and ensures prompt recovery. In addition, most hardware resources consist of modules that support a hot plug. This is a feature that enables online maintenance and prompt, easy recovery to redundant configurations.

Moreover, a dual-power module system enables continual operation in case of any power failure. Failures can range from commercial power failures due to disasters or other causes to distribution board or UPS (Uninterruptible Power Supply) failure.

* Express5800/1320Xe, 1160Xe, and 1080Xe models.

Clock module redundancy/partitioning minimize crashes from clock failures

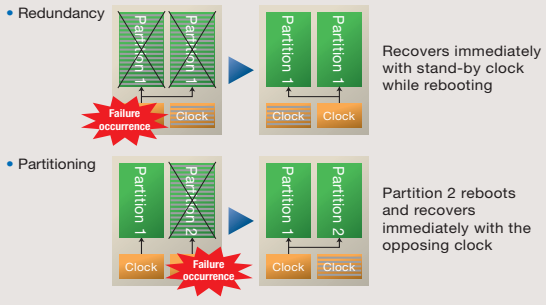
The Express5800/1000 Series employs redundant clock cards that supply the clock signal.* During regular operation, only one clock card is used while another stands by.** If a clock card failure occurs, upon rebooting the system switches to the stand-by clock card for immediate recovery.

When the system is operating in partitioning mode, however, both clock cards are used, with each supplying a clock signal to different partitions. In this case, the failure is limited to only the partition with the faulty clock card.

For a cluster configuration, reliability improves when different clock cards provide signals to different configured cluster partitions. In partitioning mode, the partition that belongs to the side of the failed clock card is shut down, but is restored by the operating clock card while rebooting.

* 1320Xe and 1160Xe support this feature.
** 1320Xe supports this feature.

Figure 7: Redundancy and partitioning of clock card



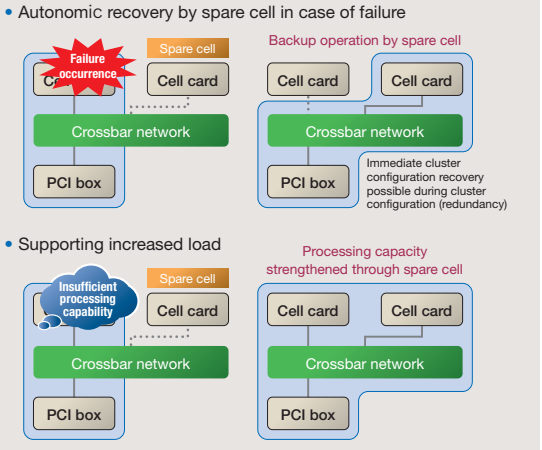
Floating I/O and spare cell prevent cell card failures from affecting I/O area

The Express5800/1000 Series employs floating I/O to create flexible combinations of cell card and I/O box connections. Support for spare cells further enhances this feature.

Floating I/O makes it possible to separate or connect multiple cell cards and PCI boxes with PCI cards. The connections occur over a crossbar network (switch) configured by multiple crossbar controllers. Because these devices can be combined in a flexible manner, any cell card failures that occur do not affect I/O. Figure 8 illustrates the features of this capability, showing how the failed cell card is isolated and a spare cell boots the system automatically in cooperation with SystemGlobe GlobalMaster.

The cooperative functions with SystemGlobe GlobalMaster extend beyond recovery to enable flexible rearrangement of hardware resources through changes in load.

Figure 8: Application of spare cells



Express5800/1000 Series Products

Mid-range Server

Powerful utility model excels in availability and expandability

1080Xe (Max. 8 CPUs)

- Itanium® 2 processor (1.6 GHz / 9 MB, 1.6 GHz / 6 MB, 1.5 GHz / 4 MB). Max. 8 processors
- Max. 64 GB memory
- Max. 1.46 TB built-in disk*
- 26 PCI-X slots*
- Memory supports ECC protection, SDDC (Single Device Data Correction) memory and memory check architecture (Windows Server 2003).
- NEC original crossbar switch with low latency and high capability of data transfer.
- Realizes near-uniform high speed memory access using NEC chipset and crossbar switch.
- Uses floating I/O technology via cell and I/O through crossbar switch to enable flexible troubleshooting for cell failure.
- Fully utilizes the power of the Itanium® 2 processor.
- 8U size of EIA rack
- Windows, Linux support

* When the extension PCI box is used.

High-end Server

Powerful high-end model consolidating NEC's latest technologies excels in availability and expandability

1160Xe (Max. 16 CPUs)

- Itanium® 2 processor (1.6 GHz / 9 MB, 1.6 GHz / 6 MB, 1.5 GHz / 4 MB). Max. 16 processors
- Max. 256 GB memory
- Max. 3.5 TB built-in disk*
- 56 PCI-X slots*
- Memory supports ECC protection, SDDC (Single Device Data Correction) memory and memory check architecture (Windows Server 2003).
- NEC original crossbar switch with low latency and high capability of data transfer.
- Realizes near-uniform high speed memory access using NEC chipset and crossbar switch.
- Uses floating I/O technology via cell and I/O through crossbar switch to enable flexible troubleshooting for cell failure.
- Fully exploits the power of the Itanium® 2 processor, optimal for mission-critical businesses.
- Windows, Linux support

* When the extension PCI box is used.

1320Xe (Max. 32 CPUs)

- Itanium® 2 processor (1.6 GHz / 9 MB, 1.6 GHz / 6 MB, 1.5 GHz / 4 MB). Max. 32 processors
- Max. 512 GB memory
- Max. 7 TB built-in disk*
- 112 PCI-X slots*
- Memory supports ECC protection, SDDC (Single Device Data Correction) memory and memory check architecture (Windows Server 2003).
- NEC original crossbar switch with low latency and high capability of data transfer.
- Realizes near-uniform high speed memory access using NEC chipset and crossbar switch.
- Uses floating I/O technology via cell and I/O through crossbar switch to enable flexible troubleshooting for cell failure.
- Fully exploits the power of the Itanium® 2 processor, optimal for mission-critical businesses.
- Windows, Linux support

* When the extension PCI box is used.

Safety notes

Please read carefully before use and observe the cautions and prohibitions in the instruction, installation, planning, operations and other manuals. Incorrect usage may cause fire, electric shock, or injury.

If this product (including the software) comes under the regulations of Foreign Exchange and Foreign Trade Law as a regulated article or other item, observe the procedures (such as application for export permission) required by the Japanese government when taking the product out of Japan.

The colors of the products in this catalogue may be slightly different from the actual colors. Specifications are subject to change without prior notice for the purpose of improving the product.

For inquiries, please contact the following NEC office:
NEC Solutions (America), Inc
2890 Scott Blvd.
Santa Clara, CA 95050

As of October 2005