RAS features of the Mission-Critical Converged Infrastructure

Reliability, Availability, and Serviceability (RAS) features of
HP Integrity Systems: Superdome 2, BL8x0c, and rx2800 i2

Technical White Paper

Table of contents

Executive Summary

Integrity Servers for Mission-Critical Computing

HP Integrity Servers have been leaders in the industry for providing mission-critical RAS since their inception. This latest generation of Integrity, based on HP's Converged Infrastructure, not only offers common infrastructure, components, and management from x86 to Superdome, but also extends the Integrity RAS legacy. The new line of HP Integrity servers consists of the BL860c i2, BL870c i2, BL890c i2 server blades; the rx2800 i2 rack-mount Integrity server; and the Integrity Superdome 2. All of these latest generation of Integrity servers are designed to maximize uptime and reduce operating costs.

HP's latest generation of Integrity Servers are all part of HP's Converged Infrastructure. This means that all of HP's servers from x86 to Superdome use the same racks, enclosures, and management environments, thus allowing HP to focus on the value-add mission-critical RAS features of Integrity.

Hot-swap n+1 redundancy for fans and power supplies; and single-bit detect and double-bit correct error correction coding with single chip kill for memory are examples of industry-wide standard server RAS features. Though the new Integrity systems have these too, this white paper will instead focus on the differentiating RAS features that set them above the industry standard servers.

New RAS Features for the New Decade

HP uses Intel's most reliable processor, the Itanium processor 9300 series (Tukwila-MC) to drive its mission critical line of Integrity servers. The Itanium processor 9300 series is packed with over one billion transistors in the most reliable, four-core processor technology. Some of the RAS features of the Itanium processor 9300 series include multiple levels of error correction code (ECC), parity checking, and Intel Cache Safe Technology.

With Blade Link technology, the Integrity server blades provide mission-critical level RAS in a cost-effective, scalable blade form factor, from one to eight processor sockets.

For higher processor socket counts, larger memory footprints, greater I/O configurations, or mainframe-level RAS; Superdome 2 is the platform of preference. To provide the tremendous scaling and performance; Superdome 2 uses the new HP sx3000 custom chip set that ultimately gives 4.5x better fabric availability than legacy Superdome servers.

At the platform level, HP further enhances the hardware with more RAS features including HP's mission-critical operating systems: HP-UX and OpenVMS. Because HP has full control of the entire RAS solution stack, these operating systems integrate tightly with the server hardware to provide proactive system health monitoring, self-healing, and recovery beyond what the hardware alone can do. But what about non-proprietary operating systems such as Windows? Can HP still provide better serviceability features when using an industry standard operating system? The Superdome 2 does with its newly introduced Superdome 2 Analysis Engine.

The Analysis Engine takes the proactive RAS monitoring and predictive abilities of the HP mission-critical operating systems and packages it in firmware that runs within the hardware, no operating system required! And because it is always on, all the time, the Analysis Engine is comprehensively more accurate at diagnosing problems in the making.

Top RAS Differentiators

## Integrity server blades (BL860c i2) and rx2800 i2 RAS

The Integrity server blades have up to two times the reliability of comparably configured industry standard volume systems. Integrity servers provide mission-critical resiliency and accelerate business performance. Higher RAS is provided in all key areas of the architecture.

The key differentiating RAS features on the Integrity server blades and the rx2800 i2 are:

- Double-Chip Memory Sparing
- Dynamic Processor Resiliency
- Enhanced MCA Recovery
- Processor Hardened latches
- Intel Cache Safe Technology
- QPI & SMI error recovery features (point-to-point & self-healing)
- Passive backplane

Figure 1 shows the basic two-processor socket, building block of the rx2800 i2 rack-mount and the BL860c i2 server blade architectures. The two-socket blade building blocks are conjoined in pairs to make the four-socket BL870c i2; and are conjoined in multiples of four to make the eight-socket BL890c i2.

**Figure 1:** BL860c i2 and rx2800 i2 Platform RAS

3
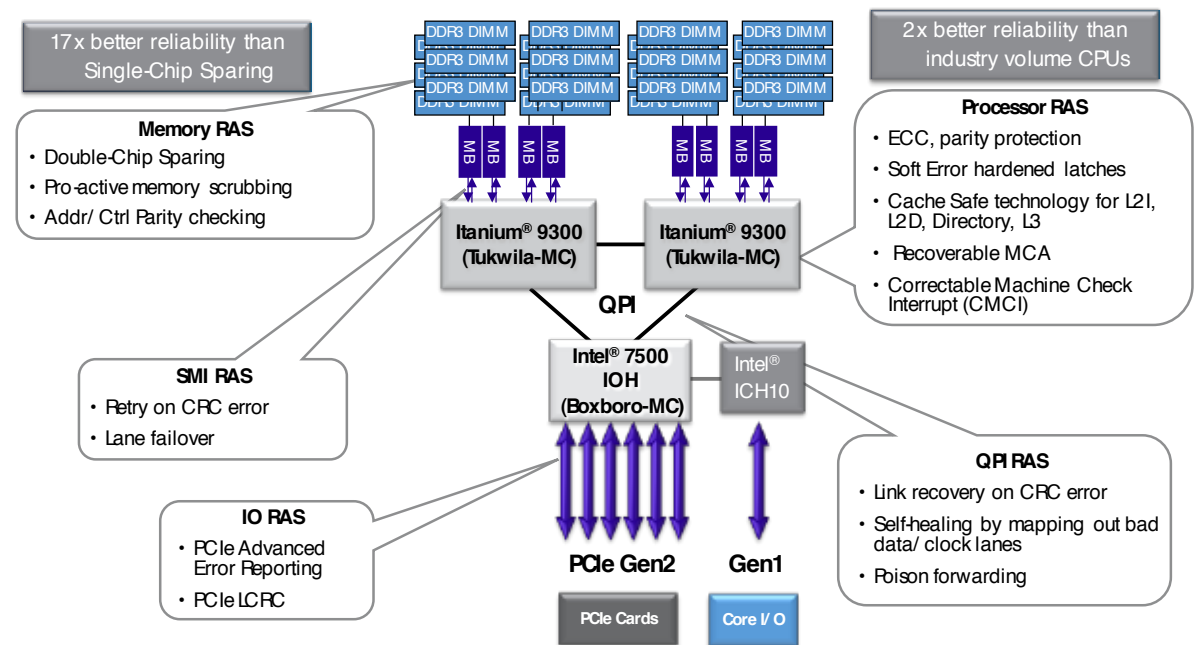
Double-Chip Memory Sparing

The industry standard for memory protection is 1-bit correction and 2-bit detection (ECC) of data errors. Furthermore, virtually all servers on the market provide Single-Chip Sparing also known as Advanced ECC and Chip-kill. These protect the system from any single-bit data errors within a memory word; whether they originate from a transient event such as a radiation strike, or from persistent errors such as a bad dynamic random access memory (DRAM) device. However, Single-Chip Sparing will generally not protect the system from double-bit faults. Though detected, these will cause a system to crash.

In order to better protect memory, many systems including Integrity servers implement a memory scrubber. The memory scrubber actively parses through memory looking for errors. When an error is discovered, the scrubber rewrites the correct data back into memory. Thus scrubbing combined with ECC prevents multiple-bit, transient errors from accumulating. However, if the error is persistent then the memory is still at risk for multiple-bit errors.

Double-Chip Sparing in Integrity servers addresses this problem. First implemented in the HP zx2 and sx2000 custom chipsets of the prior generation of Integrity servers, this capability has been moved into the Itanium 9300 processors along with the memory controllers. Double-Chip Sparing (or Double Device Data Correction (DDDC)) technology determines when the first DRAM in a rank has failed, corrects the data and maps that DRAM out of use by moving its data to spare bits in the rank. Once this is done, Single-Chip Sparing is still available for the corrected rank. Thus, a total of two entire DRAMs in a rank of dual in-line memory modules (DIMMs) can fail and the memory is still protected with ECC. Due to the size of ranks in Integrity, this amounts to the system essentially being tolerant of a DRAM failure on every DIMM and still maintaining ECC protection. Note that Double-Chip Sparing requires x4 DIMMs to be used (currently all DIMMs 4 GB or larger). If a x8 DIMM is used (currently 2 GB only), the systems default to Single-Chip Sparing mode.

Figure 2: Improvement to Reliability with Double-Chip Sparing
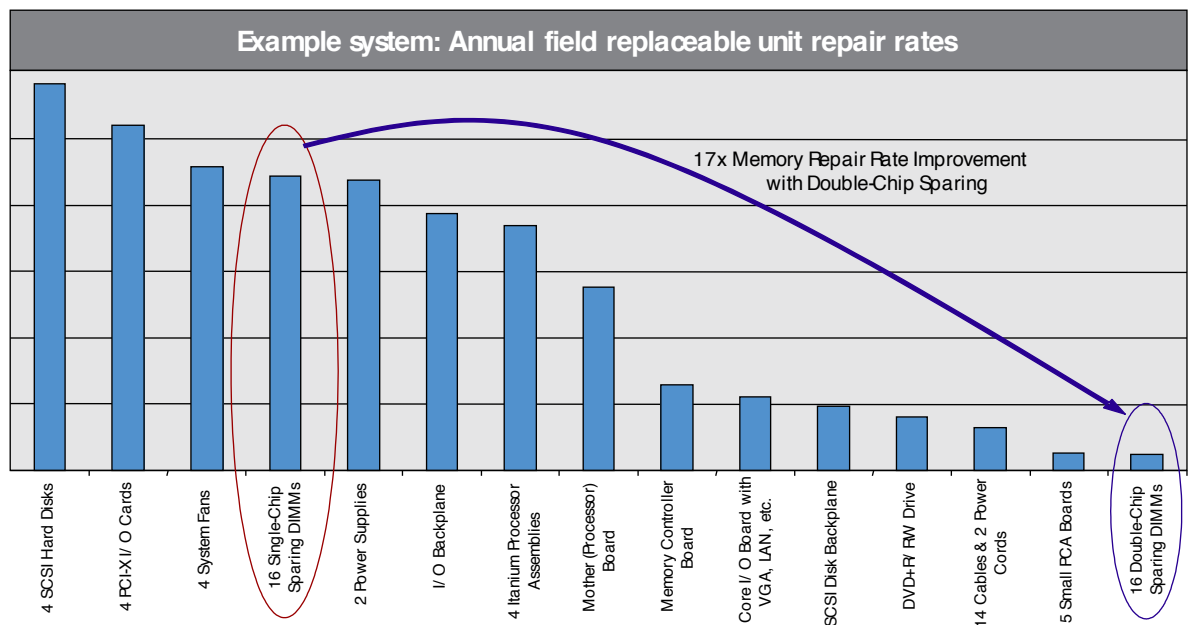
This maximizes the coverage of HP's unique protection mechanism without degrading performance. Unlike memory mirroring, DIMM spares, and RAID memory; no extra DIMMs are required for Double-Chip Sparing. It more efficiently uses the same DRAMs used for Single-Chip Sparing.

Double-Chip Sparing drastically improves system uptime, as fewer failed DIMMs need to be replaced. As shown in Figure 2, this technology will deliver about a 17x improvement in the number of DIMM replacements versus those systems that use only Single-Chip Sparing technologies. Furthermore, with repair rates on DIMMs lowered to be on the order of cables, the top four failing field replaceable units (FRUs) are all hot-swappable (Disks, IO cards, Fans, and power supplies) in a four processor entry-level Integrity system.

Furthermore, Double-Chip Sparing in Integrity servers reduces memory related crashes to be one third that of systems that have Single-Chip Sparing capabilities only.

Intel Cache Safe Technology

The majority of processor errors are bit flips in the processor local memory, known as the cache. These cache errors are similar to DRAM DIMM errors in that they can be either transient or persistent. Itanium processors implement ECC in most layers of the cache, and parity checking in layers that have copies of data from ECC protected layers. Thus, all levels of the cache are protected from any single-bit problem.

To improve on ECC, Itanium implements a more advanced feature known as Intel Cache Safe Technology. To keep a multi-bit error from occurring, the system determines if the failed cache bit is persistent or transient. When the error is persistent, the data is moved out of that cache location and put in a spare. The bad cache location is permanently removed from use and the system continues on. Thus with the combination of ECC and Intel Cache Safe Technology, the cache is protected from most multi-bit errors.

Processor Soft Error (SE) Hardened Latches and Registers

One of the most common sources of naturally occurring transient computer errors is high-energy particles striking the nuclei in electrical circuits. There are two common sources for high-energy particles. The first comes in the form of alpha particles released from radioactive contamination of materials used in circuits. The second is high-energy neutrons that get launched when solar radiation ionizes molecules in the upper atmosphere. Alpha particles can be minimized through stringent manufacturing processes, but high energy neutron strikes cannot be.

Such particle strikes can cause the logic to switch states. DRAMs on DIMMs and the memory caches of the processor have ECC algorithms and correction mechanisms to recover from such random occurrences. But what is done to protect the core of the CPU?

The Intel Itanium processor 9300 series includes new circuit topologies that dramatically reduce the susceptibility of the core latches and registers to such particle strikes. Estimates show that the soft error hardened latches reduce soft errors by about 100x, and the soft error hardened registers reduce soft errors by about 80x.

Dynamic Processor Resiliency

The flagship processor RAS feature for Integrity servers is HP's Dynamic Processor Resiliency (DPR). DPR is a set of error monitors that will flag a processor as degraded when it has experienced a certain number of correctable errors over a specific time period. These thresholds help identify processor modules which are likely to cause uncorrectable errors in the future that can panic the OS and bring the system or partition down. DPR effectively "idles" these suspect CPUs (deallocate), and marks those CPUs to not be used (deconfigured) on the next reboot cycle.

Dynamic Processor Resiliency has continued to be enhanced with each generation of Integrity server to deal with new recoverable error sources, such as register parity errors. These further differentiate Integrity servers from the competition.

The new Itanium processor 9300 series provides 2x better reliability than industry volume processors by utilizing features such as: Intel Cache Safe Technology®, error hardened latches, register store engine, memory protection keys, double device data correction, and CPU sparing and migration. In addition, Itanium's Enhanced Machine Check Architecture (MCA) and MCA recovery allow the HP-UX operating system to recover from errors that would cause crashes on other systems.

Enhanced MCA Recovery

Enhanced MCA Recovery is a technology that is a combination of processor, firmware, and operating system features. The technology allows for errors that can't be corrected within the hardware alone to be optionally recovered from by the operating system. Without MCA recovery, the system would be forced into a crash. However, with MCA recovery the operating system examines the error, determines if it is contained to an application, a thread, an OS instance or not. The OS then determines how it wants to react to that error.

When certain uncorrectable errors are detected, the processor interrupts the OS or virtual machine and passes the address of the error to it. The OS resets the error condition and marks the defective location as bad so it will not be used again and continues operation.

QPI & SMI error recovery features (point-to-point & self-healing)

Both QuickPath Interconnect (QPI) and Scalable Memory Interconnect (SMI) have extensive cyclic redundancy checks (CRCs) to correct data communication errors on the respective busses. They also have mechanisms that allow for continued operation through a hard failure such as a failed lane or clock.

With SMI lane fail-over, the SMI uses a spare lane to fail-over a bad one due to persistent errors. This is done automatically by the processor and the memory controllers for uninterrupted operation with no loss in performance.
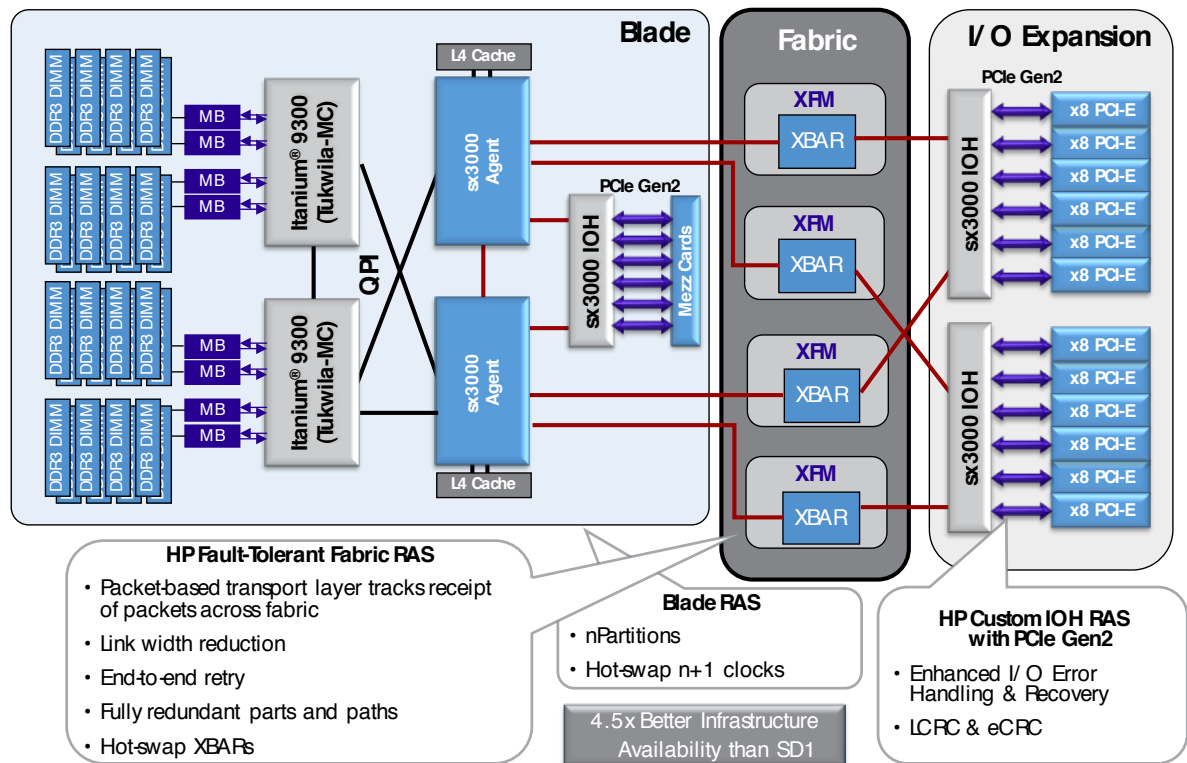
With QPI self-healing, full-width QPI lanes will automatically be reduced to half-width when persistent errors are recognized on the QPI bus. Similarly half-width ports will be reduced to quarter-width. Though there is a loss in bandwidth, overall operation can continue. SMI lane fail-over and QPI self-healing thus prevent persistent errors from eventually crashing the system. Furthermore in some cases, continually correcting persistent errors can affect performance more than self-healing techniques that reduce the band-width.

Superdome 2 RAS

In addition to the RAS features found in the Integrity server blades, Superdome 2 has numerous innovative self-healing, error detection, and error correction features to provide the highest levels of reliability and availability. New hardware features contribute to the increased reliability of the Superdome 2 infrastructure. The key differentiating Superdome 2 RAS features are:

- Electrically isolated hard partitions (nPartitions)
- Fault Tolerant Fabric
- Superdome 2 Analysis Engine
- Fully redundant, hot-swap system clock tree
- Custom PCIe Gen2 IOH with Advanced I/O RAS features
- C7000 enclosure-like serviceability of major components and FRUs

HP Fault-Tolerant Fabric RAS
- Packet-based transport layer tracks receipt of packets across fabric
- Link width reduction
- End-to-end retry
- Fully redundant parts and paths
- Hot-swap XBARs

Blade RAS
- nPartitions
- Hot-swap n+1 clocks

4.5x Better Infrastructure Availability than SD1

HP Custom IOH RAS with PCIe Gen2
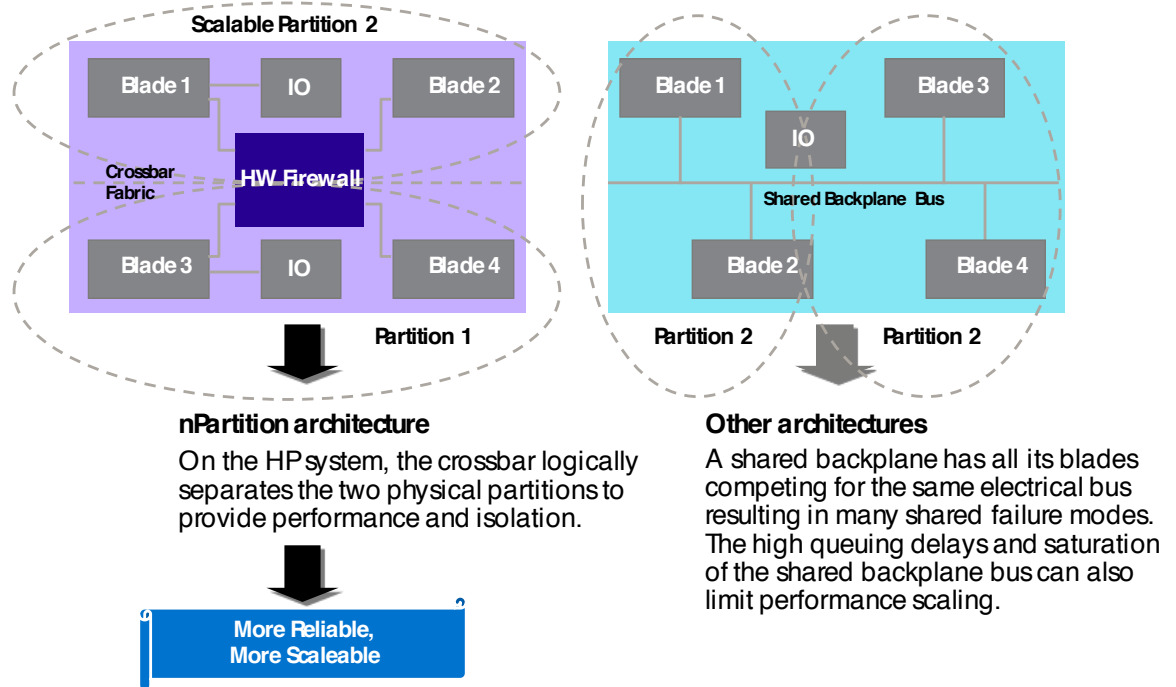- Enhanced I/O Error Handling & Recovery
- LCRC & eCRC

Superdome 2 Availability Overview

Availability is the ability of a system to do useful work at an adequate level despite failures, patches, and other events. One method of obtaining high availability is to provide methods to add, replace or remove components while the system is running. All of the components that can be serviced live in the c-Class systems can also be physically removed and replaced while partitions continue to run in Superdome 2. In addition, the unique Superdome 2 components including the Crossbar Fabric Modules (XFMs) and Global Partitions Service Modules (GPSMs) are on-line serviceable. The sx3000 chipset, firmware, and OS provide the capability to add, replace or delete blades, XFMs, and I/O cards in the IOX while the application is running. Blade OLRAD supports workload balancing, capacity on demand, power management, and field service.

Electrically Isolated Hard Partitions

Resiliency is a prerequisite for true hard partitions. Each hard partition has its own independent CPUs, memory, and I/O resources consisting of resources of the blades that make up the partition. Resources may be removed from one partition and added to another by using commands that are part of the System Management interface, without having to physically manipulate the hardware. With a future release of HP-UX 11i, using the related capabilities of dynamic reconfiguration (e.g. on-line addition, on-line removal), new resources may be added to a partition and failed modules may be removed and replaced while the partition continues in operation.

Figure 4: Hard partition error containment

Scalable Partition 2

Blade 1   IO   Blade 2

Crossbar Fabric   HW Firewall

Blade 3   IO   Blade 4

Partition 1

Blade 1   Blade 3

IO

Shared Backplane Bus

Blade 2   Blade 4

Partition 2   Partition 2

nPartition architecture

On the HP system, the crossbar logically separates the two physical partitions to provide performance and isolation.

More Reliable, More Scaleable

Other architectures

A shared backplane has all its blades competing for the same electrical bus resulting in many shared failure modes. The high queuing delays and saturation of the shared backplane bus can also limit performance scaling.

Crossbar (XBAR) Fault Resiliency

The system crossbar provides unprecedented containment between partitions, along with high reliability for single partition systems. This is done by providing high-grade parts for the crossbar chipset, hot-swap redundant crossbar replaceable units (XFMs), and fault-tolerant communication paths between HP Integrity Superdome 2 blades and I/O. Furthermore, unlike other systems with partitioning, HP provides specific hardware dedicated to guarding partitions from errant transactions generated on failing partitions.

Fault-Tolerant Fabric

Superdome 2 with the sx3000 chipset implements the industry's, best-in-class, fault tolerant fabric. The basics of the fabric are redundant links and a packet-based transport layer that guarantees delivery of packets through the fabric.

The physical links themselves contain availability features such as link width reduction that essentially allow individual wires or I/O pads on devices to fail and the links are reconfigured to eliminate that bad wire. Strong CRCs are used to guarantee data integrity.

Beyond the reliability of the links themselves, the next stage of defense is end-to-end retry. When a packet gets transported, the receiver of the packet is required to send acknowledgement back to the transmitter. If there is no acknowledgement, the packet is retransmitted over a different path to the receiver. Thus end-to-end retry guarantees reliable communication for any disruption or failure in the communication path including bad cables and chips.

When combined with the hot-swappable crossbars, end-to-end retry and link width reduction results in a Superdome 2 fabric that is 4.5x more reliable than the already impressive legacy Superdome fabric.

Clock Redundancy

The fully redundant clock distribution circuit contains the clock source and continues through the distribution to the blade itself. All mid-planes of Superdome 2 are completely passive, unlike legacy Superdome where the crossbar switches are integrated onto the mid-plane.

The system clocks are powered by 2 fully redundant and hot-pluggable Hardware Reference Oscillators (HSOs) which support automatic, "glitch-free" fail-over/ reconfiguration and are hot-pluggable under all system operating conditions.

During normal operation, the system selects one of the two HSOs as the source of clocks for the platform. Which one gets selected depends whether the oscillator is plugged into the backplane and on whether it has a valid output level. If only one HSO is plugged in and its output is of valid amplitude then it gets selected. If both HSOs are plugged in their output amplitudes that are valid then one of the two is selected as the clock source by logic on the MHW FPGA.

Figure 5: Clock redundancy
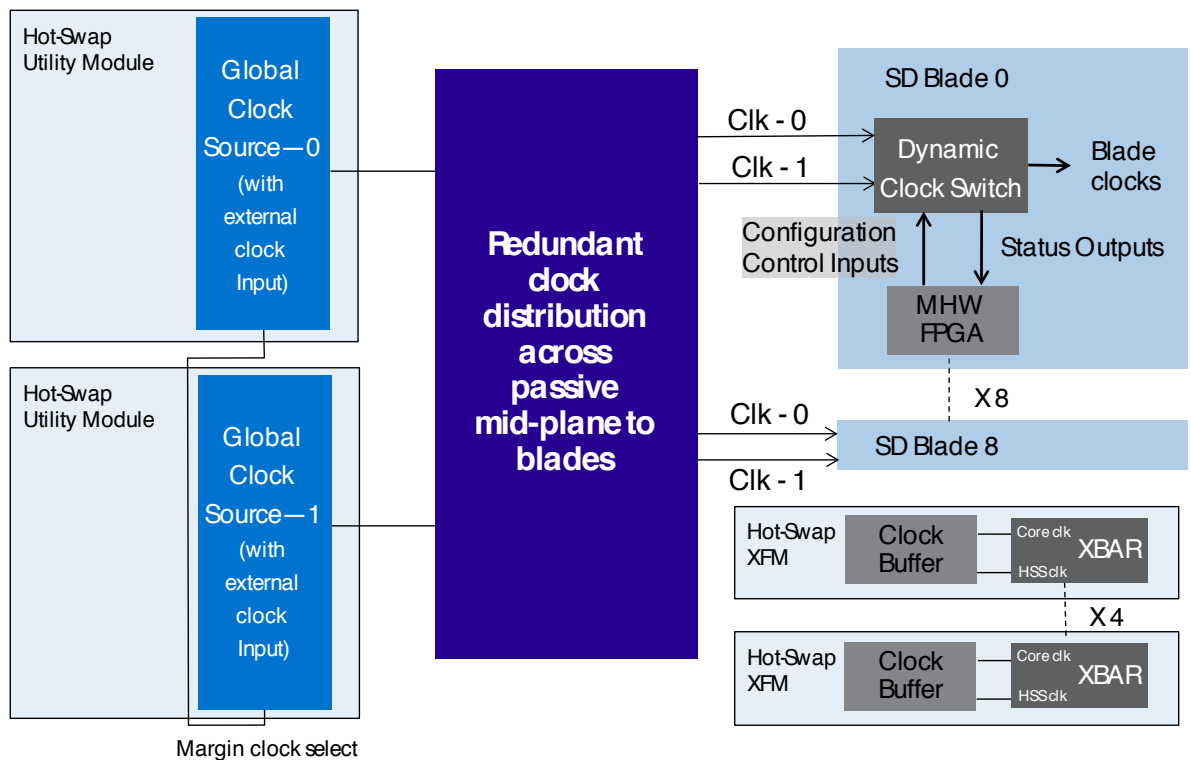
If one of the HSOs outputs fails to have the correct amplitude then the RCS will use the good one as the source of clocks and send and alarm signal to the system indicating which HSO failed. The green LED will be lit on the good HSO and the yellow LED will be lit on the failed HSO. This clock source can then be repaired through a hot-plug operation.

Isolated I/O Paths

This feature allows accessibility to a storage-device/networking-end-node through multiple paths. The access can be simultaneous (in an active-active configuration) or streamlined (in an active-passive configuration). With this feature, points of failure between two end points can be eliminated. The system software can automatically detect network/storage link failures and can failover (online) to a standby link. This feature makes the system fault tolerant to any I/O cable, crossbars (XBAR), and device-side I/O card errors, which are estimated to be at least 90% of all I/O error sources.

Advanced I/O Error Handling and Recovery

The PCI Error Handling feature allows an HP-UX system to avoid a Machine Check Abort (MCA) or a High Priority Machine Check (HPMC), if a PCIe error occurs (for example, a parity error). Without the PCI Error Handling feature installed, the PCIe slots are set in hard-fail mode. If a PCIe error occurs when a slot is in hard-fail mode, an MCA will occur, and then the system will crash.

With Advanced I/O Error Handling, the PCIe cards will be set in soft-fail mode. If a PCIe error occurs when a slot is in soft-fail mode, the slot will be isolated from further I/O, the corresponding device driver will report the error, and the driver will be suspended. The OLRAD command and the Attention Button can be used to online recover, restoring the slot, card, and driver to a usable state. PCI advanced error handling, coupled with Multi-pathing, is expected to remove upwards of 90% of I/O error causes from system downtime.

PCIe Hot-Swap or OL* (Online addition, replacement, and deletion)

The system hardware uses per-slot power control combined with operating system support for the PCIe Card online addition (OLA) feature to allow the addition of a new card without affecting other components or requiring a reboot. This feature enhances the overall high availability solution for customers since the system can remain active while an I/O adapter is being added. All HP supported PCIe cards (Gigabit Ethernet, Fiber Channel, SCSI, Term I/O, etc.) and the corresponding drivers have this feature. The new card added can be configured online and quickly made available to the operating environment and applications. PCI OL* is an easy to use feature in HP products, enhanced by the inclusion of doorbells and latches.
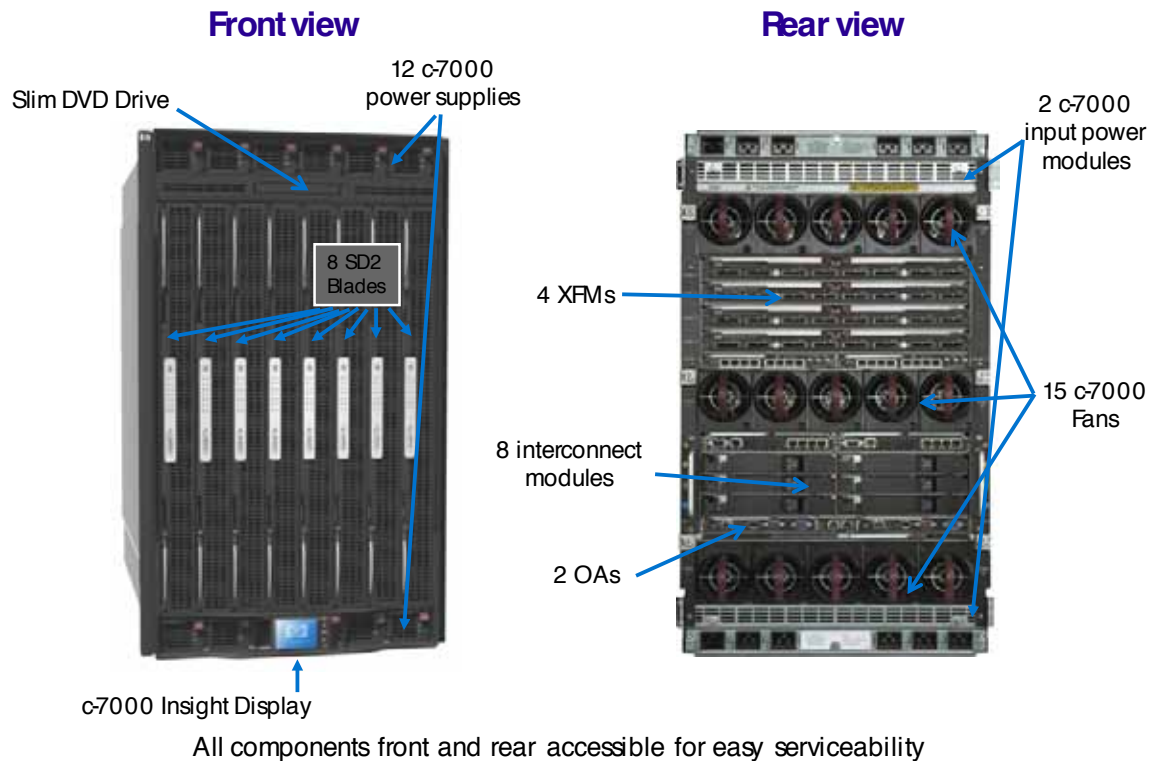
Furthermore, I/O cards can fail over time, resulting in an automatic failover to the secondary path, or a loss of a connection to a non-critical device (For those devices that do not warrant dual-path I/O). PCI online replacement (OLR) allows a user to repair a failed I/O card, online, restoring the system to its initial state without incurring any customer visible downtime.

The Advanced PCIe I/O RAS features are unique to Integrity systems and are enhanced by HP's mission-critical operating systems like HP-UX and OpenVMS. The combination of these features result in 20x to 25x better availability of the I/O subsystem.

Superdome 2 Serviceability

Superdome 2 has been designed to be highly serviceable. Many components have been leveraged from the c-Class and the new ones have been designed to the same service standards. Service repairs can be done quickly and efficiently usually with no tools.

Figure 6: Superdome 2 Components

Front view

Slim DVD Drive

12 c-7000 power supplies

8 SD2 Blades

c-7000 Insight Display

Rear view

2 c-7000 input power modules

4 XFMs

15 c-7000 Fans

8 interconnect modules

2 OAs

All components front and rear accessible for easy serviceability

Superdome 2 Analysis Engine

The Classic Health Monitoring Model

HP Server platforms use management processors such as Integrated Lights-Out 3 (iLO 3) to monitor fundamental hardware health such as voltage, temperature, power, and fans. In this classic design, the management processor signals software agents running on the OS when it detects a problem that needs an administrator's attention. These server health agents then alert the administrator through protocols such as IPMI, SNMP, or WEBM. This classic picture works very well for small servers and those without partitions, such as the Integrity BL860c i2 and rx2800 i2.
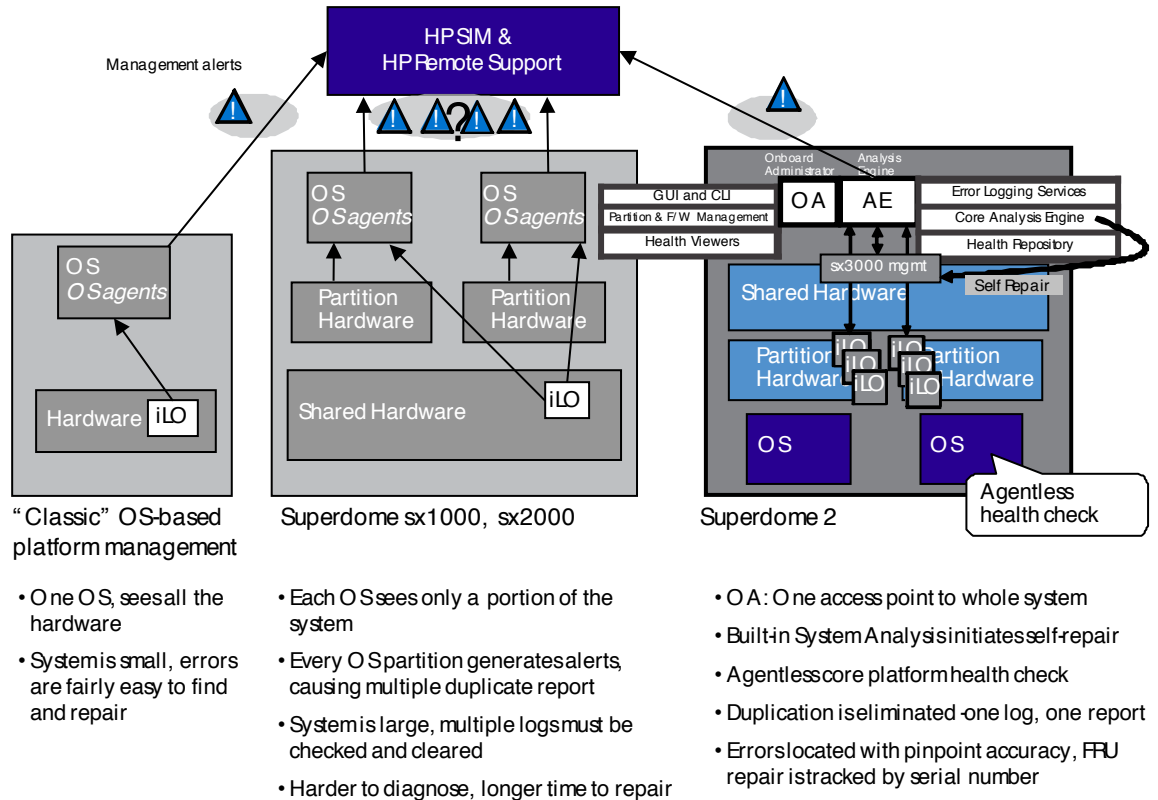
The Legacy Superdome Health Monitoring Model

HP extended the classic model and applied it to the previous generation of sx1000- and sx2000-based Superdome servers. In these systems, there is a set of management processors, monitoring the shared system hardware. In addition, there are separate components monitoring the partition-specific hardware.

Because these servers contain multiple OS-partitions, every OS-partition is notified when a management processor detects a problem in shared hardware. For example, if a power supply fails, every OS-partition is notified. Consequently, every OS-partition sends an alert and the administrator is inundated with redundant error messages.

Conversely, problems found only on a single partition's hardware are not shared with monitoring components in other partitions or with the main management processor. Thus administrators must check multiple, separate health logs for complete system information.

Figure 7: Superdome 2 Analysis Engine

"Classic" OS-based platform management

- One OS, sees all the hardware
- System is small, errors are fairly easy to find and repair

Superdome sx1000, sx2000

- Each OS sees only a portion of the system
- Every OS partition generates alerts, causing multiple duplicate report
- System is large, multiple logs must be checked and cleared
- Harder to diagnose, longer time to repair

Superdome 2

- OA: One access point to whole system
- Built-in System Analysis initiates self-repair
- Agentless core platform health check
- Duplication is eliminated - one log, one report
- Errors located with pinpoint accuracy, FRU repair is tracked by serial number

The Superdome 2 Health Monitoring Model — The Analysis Engine

In Superdome 2, the core platform OS agents have been removed and replaced with analysis tools that run in the management processor subsystem. Administrative alerts come directly from the Superdome 2 Analysis Engine, not from each OS partition, thus eliminating duplicate reports.

The Analysis Engine does much more than just generate alerts. It centrally collects and correlates all health data into one report. It then analyzes the data and can automatically initiate a self-repair without any operator intervention.

Since the Analysis Engine is a part of the firmware, error handling rules are updated only in one location. It is available with or without on-line OS-diagnostics and errors can be analyzed even if a partition cannot boot. The Superdome 2 Analysis Engine has a single command line interface for reporting the health of the server, including the replacement history of parts. When a fault is detected, the Analysis Engine automatically attempts to resolve the problem and reports any problems that require the system to be serviced. It can report directly to customers; and for systems under warranty, to HP Customer Support via Remote Support Pack (RSP) or HP Insight Remote Support (Insight RS).

Every Superdome 2 blade, and thus every partition has an (iLO 3) built into it. The entire enclosure and all of the iLOs are managed through the Superdome 2 Onboard Administrator (OA). The server health and configuration is managed through the Superdome 2 Onboard Administrator (OA), eliminating the need for an external management station.

In addition, the Superdome 2 OA contains a full Superdome toolbox to manage an OS partition. This new Superdome toolbox is always available regardless of the state of the system (up, down, rebooting, or OS not even yet loaded).

Competitive Comparisons

Other architectures are catching up to Integrity's prior generation of RAS features. Meanwhile, HP and Intel, through Integrity and Itanium, continue to raise the bar for mission critical servers. Furthermore, HP continues to make advances in its mission-critical operating systems such as HP-UX and OpenView to best exploit the RAS features of the hardware.

Within the Intel processor line, the Itanium processor 9300 series (Tukwila-MC) is still the processor of choice for mission critical systems. Though the Xeon processor 7500 series (Nehalem-EX) has made significant improvements in its RAS features, there still are gaps in comparison to the Itanium processor 9300.

Only the Itanium processor 9300 series has Double-Chip Sparing (DDDC) and SE hardened latches. Though Intel Cache Safe Technology is available on both Itanium processor 9300 and Xeon processor 7500 series, there is deeper coverage with the Itanium processor 9300. Not all levels of the cache are covered in the Xeon processor 7500 series.

Lastly, MCA recovery requires complex integration between the processors, the firmware, and the operating system on systems. It is a new feature for the Xeon while it has existed for many generations on the Itanium. Not only is it more mature on Itanium, there are more error conditions that can be recovered. Plus, HP's mission-critical operating HP-UX, has implemented the most coverage of MCA recovery in the industry.

These are the most notable differences on a processor basis. A detailed list of comparisons are in the following tables.

Table 1: Processor RAS Features

Processor RAS	Intel® Itanium® (Tukwila)	Intel® Xeon® (Nehalem-EX)	AMD Opteron®	Sun Ultra-Sparc T2®	IBM Power7®
Cache parity/ ECC		Limited	Limited		
Data bus error CRC or ECC Protection					
Enhanced MCA error handling & error logging[1]		Limited	Limited	Limited	Proprietary
Dynamic Processor Resiliency[1]				Limited	
Instruction Retry[1]				No[2]	
Built-in sensors & thermal control					
Processor Lockstep	[3]				
Bad data containment (Data Poisoning & Signaling for Error Recovery)					
Internal logic soft error checking					
Cache-line deletion (Intel Cache Safe Technology or equiv)					
Intel virtualization technology					
Intel active management technology					

[1] Requires OS support
[2] Sun Ultra-Sparc line does not support this, but the NEC SPARC64 does
[3] This Itanium feature is not implemented in HP Integrity servers because resources are used in an expensive way. A better level of protection is obtained with Integrity NonStop.

Table 2: Memory RAS Features

Memory RAS	Intel Itanium® (Tukwila)	Intel® Xeon® (Nehalem-EX)	AMD Opteron®	Sun Ultra-Sparc T2®	IBM Power7®
Data Bus CRC or ECC protection					
Scrubbing					
Chip Spare/ Advanced ECC/ Chip Kill™/ SDDC					
Double-Chip Spare/ DDDC					
Address/ Control Bus Parity protection					
Mirror/ RAID/ DIMM Spare	*				

* This Itanium feature is not implemented in HP Integrity servers because resources are used in an expensive way.
  The same level of protection is obtained with the suite of other Integrity memory features.

Table 3: Infrastructure and I/ O RAS Features

Infrastructure RAS	HP Integrity	Intel® Xeon® (Nehalem-EX)	AMD Opteron®	Sun Ultra-Sparc T2 Systems®	IBM Power7 Systems®
Hot-swap, Dual-Grid (N+N) Power					
Hot-swap, Redundant Fans					
Tool-less Field Replaceable Units					
Electrically isolated partitions	+				

I/ O RAS	HP Integrity	Intel® Xeon® (Nehalem-EX)	AMD Opteron®	Sun Ultra-Sparc T2 Systems®	IBM Power7 Systems®
Redundant I/ O Paths					
Hot-Swap PCIe Cards	+				
Fault-Tolerant I/ O Fabric	+				

+ Available only on Superdome 2

  System-dependent feature

Note that most vendors are discontinuing hot-swap I/ O.

Summary of Integrity Reliability and Availability Features

Integrity systems continue to incorporate new RAS features that have real benefits to real mission critical deployments. Many of the top differentiating features were discussed in this paper. We conclude with Table 4 that summarizes the top RAS features found in the BL860c i2, BL870c i2, BL890c i2, rx2800 i2, and Superdome 2 systems.

Table 4: System reliability and availability features

Location	Features	Customer experience
Memory system	• DRAM Protection (ECC, SDDC, DDDC) Double device data correction in memory • Memory Scrub (Patrol & Demand) • Memory Channel Protection (Retry, Reset, Lane Failover) • Can distinguish SMI link CRC error from Memory ECC error	• 17x fewer DIMM replacements and 3x fewer memory related crashes than with traditional chip-spare • Extreme levels of availability with no compromise of system performance or any added hardware cost • Risk of memory data corruption is drastically reduced to near zero with HP's DIMM enhancements
Processor	• Cache error detection/correction • Self healing L2, L3, and directory caches • Soft error (SE) hardened latches • Core logic ECC & parity protection • Dynamic Processor Resiliency • Core deconfiguration • Advanced Machine Check Architecture with new CMCI support • MCA Error Recovery with assistance from HP-UX • QPI Interconnect path detection/correction (CRC, Retry, Reset, Lane Failover)	• Covers all cache errors and the majority (70%) of the CPU core errors resulting in much better error coverage and data integrity that can be expected with x86 CPUs: Enterprise-class reliability for Enterprise customers. • Itanium reliability is >2x that of industry volume processors.
Blade, I/O, and fabric links	• Link level retry Link width reduction • End-to-end Retry • IOX attached to XFMs • Online replaceable XFMs	• Fault resilient links means partitions that stay up. This feature eliminates errors due to environmental glitches and latent manufacturing imperfections, common causes of field server failures. Able to service links without bringing the system down.
Crossbar/System fabric	• Redundant links to Superdome 2 blades • Explicit support for hard partitioning	• A key enabler of HP's leadership partitioning strategy.
I/O slots	• Error detection/correction • PCI failure isolation to a single slot • Enhanced I/O error recovery • Multi-pathing • PCI card OLARD	• Moves I/O errors from one of the major contributors of system, improving system uptime 20x to 25x. The ability to online repair further enhances the fault avoidance capabilities.
SX3000 Chipset	• Internal data path error detection/correction • Hardened latches • L4 cache line sparing	• HP's value added chip-set puts performance and availability above all else.

Location	Features	Customer experience
Partitioning/ System Infrastructure	• nPartitions—Hardware and Software isolation between partitions • Blade OLARD • Redundant and hot-swap clock • Automatic failover & hot-swap manageability models (OA & GPSM) • Redundant packet based management fabric with automatic failover • Ease of service—hardware can be repaired without bringing down multiple partitions • 2N Power & Power grid redundancy • Redundant Fans • Passive mid-planes • Superdome 2 Analysis Engine	• Superdome 2 enables true server consolation. With a measured infrastructure[#] MTBF of greater than 300 years, combined with HP's two generations of hard partitioning experience, a customer can be ASSURED that a Superdome 2 server broken up into hard partitions is an excellent approximation of an array of smaller boxes, but without all the system management, reliability, and cost of ownership headaches.

[#] Infrastructure includes the enclosure power distribution, cooling, and passive midplances.

For more information

To learn more about the new generation of Integrity Servers, visit www.hp.com/ go/ Integrity and www.hp.com/ go/ CI