# Table of contents

Abstract .................................................................................................................................................................................. 2
Introduction ............................................................................................................................................................................... 2
  Application level RAS ........................................................................................................................................................ 3
  File level RAS ................................................................................................................................................................... 3
  OS level RAS .................................................................................................................................................................... 3
  Virtualization .................................................................................................................................................................... 3
  Hardware RAS .................................................................................................................................................................... 4
  Integrity server RAS differentiators .................................................................................................................................. 5
  Processor RAS .................................................................................................................................................................... 5
  Memory RAS ..................................................................................................................................................................... 6
  General RAS ..................................................................................................................................................................... 7
  HP rx2800 i4 Server RAS differentiators ................................................................................................................................ 7
  HP Integrity i4 server blade RAS differentiators .................................................................................................................. 8
  Integrity Superdome 2 RAS differentiators ................................................................................................................................ 9
    Fault-tolerant fabric ........................................................................................................................................................ 9
    HP nPartitions ............................................................................................................................................................... 9
    Clock redundancy .......................................................................................................................................................... 10
    Isolated I/O paths .......................................................................................................................................................... 11
    Advanced PCI error handling with automatic error recovery .......................................................................................... 11
    PCIe hot-swap ............................................................................................................................................................... 11
    Serviceability ............................................................................................................................................................... 11
    Analysis Engine ............................................................................................................................................................. 12
    Capacity on demand capabilities ................................................................................................................................... 13
    Competitive comparisons ............................................................................................................................................ 14
    Summary of RAS features ............................................................................................................................................... 16
    Conclusion ..................................................................................................................................................................... 19
Abstract

Reliability, availability, and serviceability—collectively known as “RAS,” is one of the top reasons customers deploy mission-critical workloads on HP Integrity servers with HP-UX. This paper provides an overview of the RAS features of the HP Integrity i4 servers running HP-UX 11i v3. This document helps IT Managers, system administrators, and CTOs or CIOs to better understand the RAS capabilities of the Integrity portfolio and objectively compare them with x86 systems running Linux.

Introduction

Today’s global economy operates 24/7 and consumers expect access to products and services anytime and anywhere. HP Integrity servers and HP-UX are designed for high availability, ensuring that your mission-critical workloads are always on and secure without compromise. The latest generation of HP Integrity i4 servers consists of Integrity rx2800 i4 Servers, Integrity BL860c i4, BL870c i4, BL890c i4 Server Blades, and Integrity Superdome 2 (Figure 1).

Figure 1. HP Integrity servers

HP Integrity i4 servers, based on the Intel® Itanium® processor 9500 series, feature an architecture that increases reliability and performance for data-intensive applications. The HP Integrity rx2800 i4 Server is a reliable and secure 2U two-socket UNIX system that is a great fit for branch offices and data centers using rack-mount servers supporting a wide range of mission-critical workloads. With up to 3x performance, 21% less energy consumption, and a 33% lower total cost of ownership (TCO) over rx2800 i2 Servers, the rx2800 i4 Server does more with less. HP Integrity i4 server blades offer an easy upgrade by simply adding them to an existing HP BladeSystem environment. Integrity i4 Server Blades build on a common base two-socket (2s) blade design to include 4s and 8s systems for easy scalability to meet application needs. The Integrity Superdome 2 offers a complete solution for mission-critical infrastructure needs. Combining a common-bladed architecture with the sx3000 enterprise chipset and cross-bar (Crossbar) architecture, Integrity Superdome 2 extends reliability and availability to new levels.
HP Integrity servers offer RAS features in key hardware subsystems—processor, memory, and I/O—and provide the ideal foundation for the HP-UX 11i v3 operating environment. HP-UX 11i v3 reflects the growing emphasis on availability by ensuring that your business is always on and providing availability through a layered approach that offers application, file system, and operating system protection. The HP-UX 11i v3 operating environment is available in four operating environments (OEs) to best fit your application:

- **HP-UX 11i v3 Base OE (BOE)** – for basic installations or single-system management applications
- **HP-UX 11i v3 High Availability OE (HA-OE)** – for large, business-critical applications requiring clustering capability
- **HP-UX 11i v3 Virtual Server OE (VSE-OE)** – for applications requiring high resource utilization for workloads that can be virtualized and consolidated
- **HP-UX 11i v3 Data Center OE (DC-OE)** – offers a superset of all HP-UX 11i v3 versions

HP-UX 11i v3 and HP Integrity servers provide a comprehensive RAS strategy that covers all layers—from application data to hardware components.

**Application level RAS**
HP Serviceguard, available with the HP-UX 11i v3 HA-OE edition, monitors the availability and accessibility of your critical IT services including databases, standard applications, and custom applications. Those applications—and everything they rely upon to do their job—are meticulously monitored for any fault in hardware, software, operating system, virtualization, storage, or network. When a failure or threshold violation is detected, HP Serviceguard automatically and transparently resumes your normal operations in mere seconds by restarting the service in the right way and in the right place to enable improved performance. HP Serviceguard Storage Management Suite (SMS) also allows you to use a clustered file system to achieve the highest levels of availability and outstanding performance for applications benefitting from improved manageability and scalability. Furthermore, you can extend the comprehensive protection of HP Serviceguard beyond the walls of your data center. Metrocluster and Continental Clusters offer robust recovery mechanisms for geographically dispersed clusters and enable your business to remain online even after a catastrophic event with disaster recovery solutions.

**File level RAS**
Mirrordisk/UX or VxVM mirroring offers data replication on storage media for cost-effective recovery solutions and maximizes application availability. HP-UX also offers the Encrypted Volume and File System (EVFS) service. EVFS is a software-based data encryption service that fulfills the need to store files in a way that they cannot be read by unauthorized parties who may have obtained physical access to your storage. EVFS operates transparently—application-level changes are not required.

**OS level RAS**
RAS-related functions such as mapping out bad memory and identifying and bypassing a processor prevent inoperative components from shutting down the system. The interrupt strobing function of HP-UX can be adjusted to minimize the impact that CPU interrupt handling will have on performance and availability.

**Virtualization**
Virtualization offers greater utilization of resources, resulting in higher availability of services. The HP-UX 11i v3 Virtual Server OE (VSE-OE) edition allows systems to host virtual partitions and virtual machines. HP-UX 11i v3 offers a virtualization layer allowing the migration of virtual machines (VMs) from one server to another—without any downtime—thereby enhancing availability.
HP-UX Virtual Partitions
HP-UX Virtual Partitions (vPars) is a partitioning technology delivered as part of the HP-UX 11i v3 VSE-OE. HP-UX vPars allow you to increase server utilization and flexibility by dividing a single electrically isolated partition called an nPar or a whole server into several smaller servers called vPars. vPars allow finer granularity than nPars partitioning but still have the performance of a dedicated processor core and memory. Each vPar hosts its own HP-UX11i v3 operating environment along with its own subset of dedicated processor, memory, and shared or dedicated I/O resources. Any application or OS failure will affect only the vPar in which it is executing—other vPars in the same system will not be affected. vPars enable stronger application and operating system fault isolation, with near bare metal performance, while increasing the overall system availability.

HP Integrity Virtual Machine
Integrity Virtual Machines (VMs) share CPU, memory, and I/O by virtualizing these resources for a dynamic resource-optimized environment. This results in maximum workload consolidation that features online mobility of workload for maximum application availability.

VMs allow you to simultaneously execute multiple virtual servers, sharing all of the available resources (cores, memory, and I/O) for maximum server utilization. This software virtualization technology enables the movement of a VM to a completely different system, eliminating the need for planned downtime for server upgrades and maintenance, with online migration of VM workloads.

These RAS capabilities strengthen the reputation of HP-UX 11i v3 for being a rock-solid operating environment, ideally suited to those mission-critical workloads that are vital to your enterprise environment.

HP-UX Containers
HP-UX Containers allow shared OS virtualization to consolidate multiple workloads within a single instance of the HP-UX 11i operating environment. HP-UX Containers provide built-in dynamic resourcing and the sharing of processor cores, memory, and I/O. Security is enhanced since a process running in one container cannot communicate with a process running in another container. Availability is also enhanced since the granular resource control of HP-UX Containers permit applications to use just the resources necessary, allowing more applications to run.

Hardware RAS
Integrity servers fully realize our design philosophy for systems handling mission-critical workloads, which is to implement, when applicable, a four-stage RAS strategy of detection, logging, analyzing, and repair (Figure 2).

Figure 2. HP Hardware RAS strategy
**Integrity server RAS differentiators**

While features such as hot-swap n+1 power supplies and single-/multi-bit memory error correction have become common in the industry, a number of RAS differentiators set Integrity servers apart from industry-standard servers. Integrity i4 servers offer several types of RAS differentiators:

- Processor RAS
- Memory RAS
- General RAS

**Processor RAS**

HP Integrity servers use the Intel® Itanium® 9500 series processors. The Itanium 9500 series features 8-core technology and includes extensive capabilities for detecting, correcting, and reporting hard and soft errors. Processor RAS features include Intel Cache Safe Technology®, error hardened latches, register store engine, memory protection keys, and double device data correction. In addition, Itanium’s Advanced Machine Check Architecture (MCA) and MCA recovery allow the HP-UX 11i v3 operating system to recover from errors that would cause crashes on other systems.

**Dynamic Processor Resiliency**

The flagship processor RAS feature for Integrity servers is the HP Dynamic Processor Resiliency (DPR). DPR is a set of error monitors that will flag a processor as being degraded when it has experienced a certain number of correctable errors over a specific time period. These thresholds help identify processor modules that are likely to cause uncorrectable errors that will bring the system or partition down. DPR effectively “idles” these suspect processors, and marks them as unavailable (de-allocated) on the next reboot cycle. Dynamic Processor Resiliency has been 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.

**Intel Cache Safe technology**

The majority of processor errors are bit flips in the processor’s local (cache) memory. These cache errors are similar to DRAM DIMM errors, which can be transient or persistent. Itanium processors implement Error Correction Code (ECC) in most layers of the cache and parity checking in layers that have copies of data from ECC protected layers. Therefore, all levels of the cache are protected from any single-bit error, and some caches have double bit error correction and triple error detection.

To improve on ECC, Itanium uses 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. If the error is persistent, the data is moved out of that cache location to a spare location. The bad location is permanently removed from use and operation continues. With the combination of ECC and Intel Cache Safe technology, the cache is protected from most multi-bit errors. Transient errors are handled as described in the following section.

**Soft error-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 are 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 9500 series includes new circuit topologies that dramatically reduce the susceptibility of the core latches and registers to such particle strikes.
Advanced 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, or an OS instance. 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.

Memory RAS
Main memory failures have been a significant cause of hardware downtime. Integrity servers use several technologies for enhancing the reliability of memory: proactive memory scrubbing, double-chip memory sparing, and address/control signal sparing. HP DIMMs are qualified to provide both performance and quality.

Proactive memory scrubbing
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. This 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. Accumulated memory DIMM errors can result in multi-bit errors that cannot be corrected and can result in data corruption. Proactive memory scrubbing is a hardware function included in Integrity i4 servers that finds memory errors before they accumulate. Corrected data is rewritten back to the appropriate memory location.

Double-chip memory sparing
The industry standard for memory protection is single error detecting and double error detecting (SECDED) 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 a failed DRAM and a single-bit error. Though detected, these will cause a system to crash.

Combined with memory scrubbing, ECC prevents multiple-bit, transient errors from accumulating. However, persistent errors can put the memory 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 is included in the Itanium 9300 processor and 9500 series 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).

Double-Chip Sparing drastically improves system uptime, as fewer failed DIMMs need to be replaced. This technology delivers up to a 17x improvement in the number of DIMM replacements versus those systems that use only Single-Chip Sparing technologies. Furthermore, Double-Chip Sparing in Integrity servers significantly reduces the chances of memory related crashes compared to systems that only have Single-Chip Sparing capabilities.

Address/control signal sparing
Integrity servers perform address/control parity for DIMM sets. If a fatal error is detected, the associated DIMM set is de-allocated on the next reboot.
General RAS

Integrity servers include RAS functions that ensure reliable intra-processor communication, memory access, and I/O operation.

QuickPath Interconnect and Scalable Memory Interconnect RAS

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 self-healing mechanisms that allow for continued operation through a hard failure such as a failed link.

With QPI self-healing, full-width QPI links will automatically be reduced to half-width when persistent errors are recognized on the QPI bus (Figure 3). Similarly, half-width ports will be reduced to quarter-width. Though there is a loss in bandwidth, overall operation can continue until repairs can be made.

Figure 3. QPI self-healing.

With SMI lane fail-over, the SMI uses a spare lane to replace a lane experiencing persistent errors. For uninterrupted operation without loss in performance, the processor and the memory controllers perform the fail over automatically. SMI lane fail-over and QPI self-healing thus prevent persistent errors from eventually crashing the system. In some cases, continually correcting persistent errors can affect performance more than self-healing techniques that reduce the bandwidth.

HP rx2800 i4 Server RAS differentiators

The HP Integrity rx2800 i4 Server is a reliable and secure 2U two-socket UNIX server. In addition to the above RAS features, the rx2800 i4 also allows workload consolidation and software licensing cost savings through HP-UX Virtual Partitions and Integrity Virtual Machines. HP Insight software and Integrated Lights-Out 4 increase staff productivity by enabling efficient and remote management of servers. HP-UX 11i v3 and HP Serviceguard support provides always-on resiliency and disaster recovery capability.
**HP Integrity i4 server blade RAS differentiators**

HP Integrity i4 server blades offer specific RAS capabilities suited for the modular, high-density environment of blade servers.

**Electrical isolation/nPARs**

HP nPartitions (nPARs) capability is provided as a default feature for the HP Integrity BL870c i4 and BL890c i4 Server Blades. This feature, a key component of the virtualization solution set, enables electrically isolated hard partitions at the hardware blade level. HP nPartitions provide improved scalability, availability, and fault isolation, for an ultimately richer set of virtualization options within the Blade Link Domain. The nPartitions capability is seamlessly integrated into the converged infrastructure environment, including Virtual Connect and HP Insight management tools such as HP Systems Insight Manager (SIM). Moreover, the system administrator will find nPartitions configuration management quick and easy to use.

An HP Integrity blade nPartition behaves in the same manner as an individual server, but with the added enhanced capability to programatically scale (shrink or grow) its physical resources to match the required workload or application compute resource needs. At a physical hardware level, the HP Integrity BL870c i4 and BL890c i4 Server Blades are composed of multiple blades plus a Blade Link to create a Blade Link Domain (Figure 4). The blades within the Blade Link Domain are electrically connected through a processor-to-processor fabric via the Blade Link. nPartitions technology enables the processors to be electrically tied together (conjoined) at blade-level granularity, while being electrically isolated from any other processors outside of the given defined nPartition.

**Figure 4. HP Integrity four-, two-, and one-blade configurations**

Note: Blade link and blades comprise a BL domain.
DMP: Domain management processor
MMP: Monarch management processor
AMP: Auxiliary management processor

**HP Virtual Connect FlexFabric**

HP Integrity i4 server blades offer HP Virtual Connect FlexFabric connectivity, which increases network scalability and configuration flexibility while reducing infrastructure costs by converging LAN and SAN traffic on the same connection. Each Integrity i4 server blade has two, dual-port HP Virtual Connect FlexFabric connections for connectivity to both network and storage, reducing the need for additional I/O mezzanine cards. With wire-once connectivity, IT administrators can manage all subsequent “rewiring” virtually, significantly reducing cabling. HP Virtual Connect FlexFabric results in fewer required components effectively increasing HP Integrity blade RAS.
### Integrity Superdome 2 RAS differentiators

Integrity Superdome 2 implements the HP Mission-Critical infrastructure strategy of attacking IT sprawl with standards-based modular architecture. The heart of Integrity Superdome 2 architecture is the fault-tolerant HP Crossbar Fabric consisting of passive midplanes with end-to-retry and link failover functionality. Integrity Superdome 2 also uses HP’s Dynamic Processor Resiliency and the innovative sx3000 enterprise system chipset that includes extensive self-healing, error-detection, and error correction capabilities.

### Fault-tolerant fabric

Superdome 2 with the sx3000 chipset sets the industry standard for fault-tolerant fabric resiliency. 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 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 the 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 is 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 (Figure 5). Thus, end-to-end retry guarantees reliable communication for any disruption or failure in the communication path including bad cables and chips.

![Figure 5. Fault-tolerant fabric.](image)

The system crossbar provides unprecedented containment between partitions. High reliability for single partition systems is accomplished by offering high-grade parts for the crossbar chipset, and fault-tolerant communication paths between 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.

### HP nPartitions

Resiliency is a prerequisite for true hard partitions. HP nPartitions (nPars) is a hard partition technology providing electrical isolation, enabling you to configure a single blade-based server as one large server or as multiple, smaller, independent servers. Each nPartition 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 manipulate the hardware physically.

Many systems use a shared backplane, where all blades are competing for the same electrical bus (Figure 6A). This raises the potential for a number of shared failure modes. For example, high queuing delays and saturation of the shared backplane limit performance scaling. On the HP Superdome 2 system (Figure 6B), the crossbar fabric logically separates the physical partitions, providing performance and isolation for a more reliable and scalable system.
Clock redundancy

The fully redundant clock distribution circuit contains the clock source and continues through the distribution to the blade itself. Unlike legacy Superdomes that have the crossbar switches integrated onto the midplane, all Superdome 2 midplanes are completely passive.

The system clocks are powered by two 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. If only one HSO is installed then its output is used (assuming it is of valid amplitude). If both HSOs are plugged in and both outputs are valid, then one of the two is selected by the clock switch logic on the blade. If one of the HSO outputs fails to have the correct amplitude, the clock switch logic will use the valid HSO as the source of clocks and send an alarm to the system indicating which HSO failed. A green LED will be lit on the good HSO and a yellow LED will be lit on the failed HSO. This clock source can then be repaired through a hot-plug operation.

Figure 7. RAS clock failover.
Isolated I/O paths

Isolated I/O paths allow 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 ends can be eliminated. The operating system 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 constitute a substantial majority of all I/O error sources.

Advanced PCI error handling with automatic error recovery

The Advanced 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, the PCIe slots operate in hard-fail mode. If a PCIe error occurs when a slot is in hard-fail mode, an MCA will occur and the system will crash.

With automatic error recovery, the PCIe cards operate 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. If automatic error handling fails to detect an error and isolate the slot, the Online Replace Add Delete (OLRAD) command and the Attention Button can be used for online recovery to restore the slot, card, and driver to a usable state. Coupled with Multi-pathing, automatic error recovery can remove a substantial amount of I/O errors that would otherwise result in system downtime.

PCIe hot-swap

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 high availability since the system can remain active while an I/O adapter is installed. You can then configure the added card online and make it available to the operating environment and applications. PCI OLA is an easy-to-use feature enhanced by doorbells and latches on HP intelligent rack systems.

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. These features significantly enhance the availability of the I/O subsystem. When Superdome 2 PCI I/O is replaced, you can dynamically replace I/O in the virtualization pool with HP-UX vPars and Integrity VM v6.3. vPars and Integrity VM v6.3 can recognize the new I/O and add it to the virtualization pool where the Virtualization Services Platform provides the ability to assign the new resource to either a VM or a vPar. This PCI OLAR support in a VMS/vPar environment enables physical Host Bus Adapters and Network Interface Controllers to be replaced while they are in use by VMS/vPars for non-critical resources.

Serviceability

Superdome 2 has been designed to be highly serviceable. The overall design leverages the HP ProLiant c-Class, with components designed to the same service standards. Service repairs can be done quickly and efficiently—usually without tools.

Figure 8 shows Superdome 2 key components, many of which are redundant and hot-swappable:

- Power supplies and cooling fans
- Global Partition Services Modules (GPSMs)
- I/O switches
- Onboard administrators
Analysis Engine

The Superdome 2 Analysis Engine is the latest generation in the evolutionary development of methods to create administrative alerts. In the Superdome classic design, the Integrated Lights Out (iLO) management processor signaled core platform OS agents when it detected a problem that needed administrator attention. These server health agents alerted the administrator through protocols such as IPMI, SNMP, or WEBM. This worked well for small servers and those without partitions, such as the Integrity BL860c i2 Server Blade and Integrity rx2800 i2 Server.

Later Superdome legacy systems used a set of management processors for monitoring the shared system hardware. Separate components were involved in monitoring the partition-specific hardware. Since these servers contain multiple OS partitions, every OS partition was notified when a management processor detected a problem in shared hardware. For example, if a power supply failed, every OS partition was notified. Consequently, every OS partition sent an alert and the administrator was inundated with redundant error messages. Conversely, problems found only on a single partition’s hardware were not shared with monitoring components in other partitions or with the main management processor. Thus, administrators needed to check multiple, separate health logs for complete system information.

In Superdome 2, the core platform OS agents have been 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 or, 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 iLO 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. The Superdome toolbox is always available regardless of the state of the system (up, down, rebooting, or OS not yet loaded).

**Capacity on demand capabilities**

To bring in capacity when it is needed (from an availability perspective), HP Instant Capacity (iCAP) and Global Workload Manager (gWLM) are features that can reduce the cost of high availability.

iCAP supports the ability to reallocate active cores among partitions while they are running, and gWLM can be used to automate this process. gWLM monitors the load in each of the partitions, automatically deactivating idle cores in one partition while activating iCAP cores in another if workloads in the other partition can use them. gWLM deactivates the core as soon as it becomes idle, thereby minimizing the amount of time the cores are active. gWLM can also active additional iCAP cores (to increase the CPU count) if reallocation is not possible.

The policy-based, automated, in-server load-balancing capability provided by the combination of iCAP and gWLM can be extended to provide automated load balancing between all the servers that are part of a GiCAP group. Along with this functionality, GiCAP and gWLM can be used to automate the activation of pooled temporary instant capacity available to all the servers in a GiCAP group.
Competitive comparisons

While other architectures are catching up to Integrity’s prior generation of RAS features, HP and Intel (through Integrity and Itanium) continue to raise the bar for mission critical servers. Furthermore, HP continues to make advances in its HP-UX 11i v3 mission-critical operating systems to fully exploit the RAS features of the hardware.

Within the Intel processor line, the Itanium processor 9500 series is the processor of choice for mission critical systems. Though the Xeon processor 7500 series has made significant improvements in its RAS features, there remain gaps in comparison to the Itanium processor 9500.

Only the Itanium processor 9300/9500 series has Double-Chip Sparing (DDDC) and soft error hardened latches. Though Intel Cache Safe Technology is available on both Itanium processor 9300/9500 and Xeon processor 7500 series, there is deeper coverage with the Itanium processors. 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. While a new feature for the Xeon, MCA recovery has existed for many generations on the Itanium. Not only is MCA more mature on Itanium, there are more error conditions that can be recovered. In addition, HP-UX 11i v3 implements the maximum amount of coverage for MCA recovery.

Tables 1 thru 3 show the RAS comparisons across processor types.

**Table 1. Processor RAS comparison across processor types.**

<table>
<thead>
<tr>
<th>Processor RAS</th>
<th>Intel Itanium</th>
<th>Intel Xeon</th>
<th>AMD Opteron</th>
<th>Oracle M5/M6</th>
<th>IBM Power8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cache parity/ECC</td>
<td>Yes</td>
<td>Limited</td>
<td>Limited</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Data bus error CRC or ECC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dynamic processor resiliency</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Instruction retry [1]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Built-in sensors &amp; thermal control</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Processor lockstep</td>
<td>[3]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bad data containment (data poisoning)</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Internal logic soft error checking</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cache line deletion</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Intel virtualization</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Intel active management technology</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes:

[1] Requires OS support.
[2] Xeon-EX.
[3] Not used on HP Integrity servers. Replaced with Integrity NonStop function for better protection.
### Table 2. Memory RAS comparison across processor types.

<table>
<thead>
<tr>
<th>Memory RAS</th>
<th>Intel Itanium</th>
<th>Intel Xeon</th>
<th>AMD Opteron</th>
<th>Oracle M5/M6</th>
<th>IBM Power8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data bus CRC or ECC protection</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Scrubbing</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Chip spare/advanced ECC/Chip kill/SDDC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Double-Chip spare/DDDC</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Address/control bus parity protection</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Mirror/RAID/DIMM spare</td>
<td>(see note)</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: Not used on Integrity servers—replaced by better and more economical Integrity memory features such as DDDC.

### Table 3. Infrastructure and I/O RAS comparison across processor types.

<table>
<thead>
<tr>
<th>RAS</th>
<th>HP Integrity Superdome 2</th>
<th>Intel Xeon</th>
<th>AMD Opteron</th>
<th>Oracle M5/M6</th>
<th>IBM Power8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot-swap, dual-grid (N+N) power</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hot-swap redundant fans</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Tool-less field-replaceable units</td>
<td>Yes</td>
<td>System-dependent</td>
<td>System-dependent</td>
<td>System-dependent</td>
<td>System-dependent</td>
</tr>
<tr>
<td>Electrically isolated partitions</td>
<td>Yes [1]</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Redundant I/O paths</td>
<td>Yes</td>
<td>System-dependent</td>
<td>System-dependent</td>
<td>Yes</td>
<td>System-dependent</td>
</tr>
<tr>
<td>Hot-swap PCIe cards</td>
<td>Yes [2]</td>
<td>System-dependent</td>
<td>System-dependent</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fault-Tolerant I/O fabric</td>
<td>Yes [2]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes:

[1] Integrity blades and Superdome 2
Summary of RAS features

HP Integrity systems deliver best-in-class RAS support in a mission-critical platform with features that result in tangible benefits such as the ones described in Table 4.

Table 4. Integrity RAS features and benefits.

<table>
<thead>
<tr>
<th>RAS type</th>
<th>Features</th>
<th>rx2800 i4</th>
<th>Integrity i4 blades</th>
<th>Superdome 2</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>• Cache error correction and detection</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>• Extreme levels of availability with no compromise of system performance or any added hardware cost</td>
</tr>
<tr>
<td></td>
<td>• Self-healing</td>
<td></td>
<td></td>
<td></td>
<td>• Itanium reliability is &gt;2x that of industry volume processors</td>
</tr>
<tr>
<td></td>
<td>• Soft error hardened latches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Core logic ECC and parity protection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Core deactivation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Advanced MCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• MCA recovery with HP-UX assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• QPI path retry, detection, correction, lane failover and link width reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Poison data error containment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Dynamic Processor Resiliency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Core/socket de-activation and de-configuration</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Partitioning</td>
<td>• Electrical isolation/nPartitions—hardware and software isolation between partitions</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>• Integrity BL890c i4 Server Blades can support up to four hard partitions. A 32S system supports up to 16 nPartitions or 32 vPars. More partitioning is available with HPVM.</td>
</tr>
<tr>
<td></td>
<td>• Redundant and hot-swap clock</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>• Superdome 2 supports up to 16 hard partitions enabling greater consolidation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Superdome 2 enables true server consolidation. 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.</td>
</tr>
</tbody>
</table>
### Table 4. Integrity RAS features and benefits. (Continued)

<table>
<thead>
<tr>
<th>RAS type</th>
<th>Features</th>
<th>rx2800 i4</th>
<th>Integrity i4 blades</th>
<th>Superdome 2</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>System infrastructure</td>
<td>• Redundant packet-based management fabric with automatic failover</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>• Reliable system management</td>
</tr>
<tr>
<td></td>
<td>• Ease of service—hardware can be repaired without bringing down multiple partitions</td>
<td></td>
<td></td>
<td></td>
<td>• No downtime required for replacing power and cooling components</td>
</tr>
<tr>
<td></td>
<td>• 2N Power &amp; Power grid redundancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Redundant Fans</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Passive mid-planes</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Analysis Engine</td>
<td>• Error analysis</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>• Interprets events to improve diagnosis and reduce repair times.</td>
</tr>
<tr>
<td></td>
<td>• Automatic failover &amp; hot-swap of OA</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td>• DRAM protection (ECC, SDDC, DDDC, address and control parity)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>• 17x fewer DIMM replacements and 3x fewer memory related crashes than with traditional chip-spare</td>
</tr>
<tr>
<td></td>
<td>• Scrubbing</td>
<td></td>
<td></td>
<td></td>
<td>• 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</td>
</tr>
<tr>
<td></td>
<td>• Channel protection</td>
<td></td>
<td></td>
<td></td>
<td>• CPUs: Enterprise-class reliability for Enterprise customers</td>
</tr>
<tr>
<td></td>
<td>• SMI link CRC error detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blade, I/O, fabric links</td>
<td>• Link level entry and link width reduction</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>• 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</td>
</tr>
<tr>
<td></td>
<td>• End-to-end retry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• IDX attached to XFMs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossbar/syste m Fabric</td>
<td>• Redundant links to Superdome 2 server blades</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
<td>• A key enabler of HP’s leadership partitioning strategy.</td>
</tr>
<tr>
<td></td>
<td>• Explicit support for hard partitioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAS type</td>
<td>Features</td>
<td>rx2800 i4</td>
<td>Integrity i4 blades</td>
<td>Superdome 2</td>
<td>Benefits</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------</td>
<td>----------</td>
<td>--------------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>I/O slots</td>
<td>• Error detection and correction</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>• Moves I/O errors from one of the major contributors of system, improving system uptime 20x to 25x.</td>
</tr>
<tr>
<td></td>
<td>• PCI failure isolation to a single slot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Enhanced I/O error recovery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Multi-pathing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• PCI card OLRAD</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>• Online repair reduces downtime and enhances the fault avoidance capabilities.</td>
</tr>
<tr>
<td>sx3000 chipset</td>
<td>• Internal data path error detection</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
<td>• HP’s value added chip-set puts performance and availability above all else.</td>
</tr>
<tr>
<td></td>
<td>• Hardened latches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• L4 cache line sparing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Infrastructure includes the enclosure’s power distribution, cooling, and passive midplanes.
Conclusion

HP Integrity i4 servers continue to build on the long Integrity tradition of combining high performance with resiliency and security. The robust architectures and built-in RAS capabilities of HP Integrity i4 servers practically eliminate single-point-of-failure possibilities and set the foundation for a reliable infrastructure you can trust to be flexible, scalable, and always available.

HP Integrity i4 servers running HP-UX 11i v3 provide a comprehensive platform that offers multi-layered RAS capabilities to enable your business to be always-on and extend your mission-critical application availability.

Resources, contacts, or additional links

HP Integrity rx2800 i4 white paper

HP Integrity server blade white paper

HP Integrity Superdome 2 white paper

HP Serviceguard

HP Instant Capacity (iCAP) white paper

HP-UX 11i v3 Operating Environments Data Sheet

Learn more at
hp.com/go/integrity