# The Intel® Itanium® Processor 9300 Series

## A Technical Overview for IT Decision-Makers

White Paper

**The Intel® Itanium® Processor 9300 Series**


## Executive Summary

The Intel® Itanium® processor family provides the foundation for mainframe-class computing systems that offers greater flexibility and better cost models than traditional mainframe and high-end RISC systems. The new Intel® Itanium® processor 9300 series delivers a major step forward in this processor family. It provides more than double the performance[1] of the Intel® Itanium® processor 9100 series, along with greater scalability and enhanced reliability, availability and serviceability (RAS) features. Systems based on this new processor are ideal for running demanding enterprise applications, such as database, business intelligence and enterprise resource planning (ERP).

The most significant improvements in the Intel Itanium processor 9300 series include:

- **Double the number of processing cores:** The new processor has four high-performance cores, which will substantially boost performance for many applications and provide greater capacity for consolidating more applications and heavier workloads per server.
- **Many-fold increases in bandwidth:** With two integrated memory controllers and new Intel® QuickPath Interconnect Technology, memory interconnect bandwidth has been improved by up to a factor of six and system interconnect bandwidth by up to a factor of nine. These increases help to sustain high utilization rates for all four processing cores, enabling substantial gains in performance and throughput for demanding workloads. They also provide a resilient and high-performance foundation for building large multiprocessor systems.
- **Greater physical memory capacity:** The memory subsystem with its new Intel® Scalable Memory Interconnect (Intel® SMI) supports up to eight times more memory than the previous-generation processor using standard DDR3 memory components. Server manufacturers can take advantage of this enhancement to design systems with more memory capacity for memory-intensive workloads and customers can leverage the volume economics of DDR3 memory.
- **Next-generation RAS and manageability features:** The Intel Itanium processor 9300 series builds on the mainframe-class RAS features of its predecessor, extending and enhancing the ability to detect, correct and manage errors; add and replace components without downtime; and dynamically allocate resources among multiple running partitions.

Systems based on the Intel Itanium processor 9300 series are binary compatible with existing software and are designed to provide major performance improvements without the need for additional software optimization. New Intel Itanium processors in development today are being designed for both socket and binary compatibility with today's Intel Itanium processor 9300 series-based systems and software. This will enable organizations to scale performance and capacity through simple component upgrades and without software recompilation, so they can continue to expand and adapt their mission-critical computing systems.

# Table of Contents



**Authors:**

Rajesh Agny, Itanium Platform Marketing Manager, Data Center Group, Intel Corporation

Eric DeLano, Senior Principal Engineer, Intel Architecture Group, Intel Corporation

Mohan Kumar, Senior Principal Engineer, Intel Architecture Group, Intel Corporation

Murugasamy (Sammy) Nachimuthu, Lead Server Firmware Architect, Intel Architecture Group, Intel Corporation

Robert Shiveley, Mission Critical Server Product Manager, Intel Corporation

## Overview of Key Architectural Advancements

The new Intel Itanium processor 9300 series is the world's first processor with more than two billion transistors, and Intel engineers have used this abundance of resources to integrate a variety of fundamental architectural improvements (Table 1). The most obvious advancement is the addition of two more processing cores. Just as important, however, are improvements that enable data and instructions to be delivered to these cores much faster than on previous-generation processors, which help to sustain high levels of utilization for all four cores. Based on internal tests by Intel engineers, the Intel Itanium processor 9300 series can be expected to deliver more than double the performance of its predecessor for many applications.[1]

Since Itanium®-based servers are often deployed in mission-critical computing environments, data integrity and high availability are critical. A great deal of the design effort went into extending and enhancing the mainframe-class reliability, availability and serviceability (RAS) features that were implemented in the previous-generation processor. Improvements are implemented across many levels, from silicon-level advancements that make individual logic gates more resistant to errors, to system-level advancements that help organizations add, remove and allocate resources more effectively among running partitions. These and other architectural improvements are discussed in more detail in the following sections.

## Higher Performance through Enhanced Thread-Level Parallelism (TLP)

With four multi-threaded cores per processor, the Intel Itanium processor 9300 series can handle twice as many simultaneous software threads as its predecessor. This can substantially improve performance and scalability for heavily threaded software code, such as database and decision support applications. The benefits can be equally compelling for consolidated environments in which multiple applications and operating systems are hosted on a single system.

Thread management has also been improved. In the previous generation, a processor core would switch to another thread whenever the active thread was stalled due to a high latency event, such as waiting

Table 1. Architectural Enhancements in the Intel® Itanium® processor 9300 series

| Key Characteristics | Intel® Itanium® processor 9100 series | Intel® Itanium® processor 9300 series |
|---|---|---|
| Cores | 2 | 4 |
| Total On-Die Cache[a] | 27.5 MB | 30 MB |
| Software Threads per Core | 2 | 2 (with enhanced thread management) |
| System Interconnect (bandwidth per processor for a 3-load or 2-socket system) | Front Side Bus<br>▪ Peak bandwidth per processor: 5 GB/s | Intel® QuickPath Interconnect Technology<br>▪ Peak bandwidth: 48 GB/s (up to 9x improvement)<br>▪ Enhanced RAS<br>▪ Enables common Input/Output Hubs (IOHs) with next-generation Intel® Xeon® processors |
| Memory Interconnect (bandwidth per processor for a 3-load or 2-socket system) | Front Side Bus<br>▪ Peak bandwidth per processor: 5 GB/s | Dual Integrated Memory Controllers<br>▪ Peak bandwidth 34 GB/s (up to 6x improvement) |
| Memory Capacity (4-socket system) | 128-384 GB | 1TB (using 16 GB RDIMMs) – up to 8x improvement |
| Partitioning and Virtualization | Intel® Virtualization Technology (Intel® VT-i) | Intel® VT-i2 (includes extensions to improve performance in both processing and I/O efficiency) |
| RAS | Mainframe Class | Enhanced<br>▪ Advanced Machine Check Architecture error recovery for increased uptime<br>▪ Error detection, correction and avoidance extended and/or improved across all key components (processor, memory, interconnect)<br>▪ Improved support for partitioning, virtualization and resource management (including component hot add/remove/replace if supported by OS) |
| Energy Efficiency | Demand Based Switching (DBS) | ▪ Enhanced DBS (voltage modulation in addition to frequency)<br>▪ Intel® Turbo Boost Technology<br>▪ Advanced CPU and Memory Thermal Management |
| SMP Scalability | ▪ 64-bit Virtual Addressability<br>▪ 50-bit Physical Addressability<br>▪ Home snoop coherency | ▪ 64-bit Virtual Addressability<br>▪ 50-bit Physical Addressability<br>▪ Directory coherency for better performance in large SMP configurations<br>▪ Up to 8-socket glueless systems (higher scalability with OEM chipsets) |

[a]Size includes all on-die cache arrays which are comprised of the cache tag and data arrays for the three-level cache hierarchy (9100 and 9300 series) as well as directory cache arrays (9300 series only).

for data from main memory. The newer processor supports switching for medium latency events and for spin-lock loops. These and other enhancements to thread-switching logic help to boost core utilization, which can improve both application response times and overall system throughput.

The large and fast cache structures of the Intel Itanium processor 9300 series also help to sustain high levels of core utilization, by delivering data to the cores at or near clock speed. The processor is available with up to 30 MB of on-die cache versus 27.5 MB on its predecessor. Each core has its own dedicated L1, L2 and L3 cache. Although the dedicated cache architecture may increase cache miss rates somewhat compared with a shared cache solution, it provides lower latency, higher bandwidth and better quality of service (QoS). This can be especially important when running multiple mission-critical applications per system. It ensures that each core has dedicated cache resources to provide more predictable performance and throughput.

## Higher Performance through Enhanced Instruction-Level Parallelism (ILP)

ILP refers to a processor's ability to simultaneously process multiple instructions on each software thread. A high degree of ILP increases throughput and decreases response times for operations and transactions that must be performed sequentially, which are common in transactional applications and business intelligence queries. It can also help to reduce latencies for individual software threads in multi-threaded workloads.

The Intel Itanium architecture is based on the Explicitly Parallel Instruction Computing (EPIC) model, which was specifically designed to enable very high ILP. The processor employs an exceptionally wide and short pipeline (six instructions wide and only eight stages deep) and disperses instructions among 11 functional units. It also supports zero cycle load-use penalties and zero-cycle branch re-steers, and implements extensive bypasses so a thread is less likely to stall the pipeline.

Since the basic core structure for the Intel Itanium processor family has been tuned and optimized over many generations, no major changes were made to the cores in the Intel Itanium processor 9300 series. Instead, the focus for improving ILP was to dramatically increase the processor's ability to feed its multiple, high-performance cores with data and instructions (see next section).

However, one significant enhancement to the core architecture was implemented. The first-level Data Translation Lookaside Buffer (DTLB), which translates virtual addresses to physical addresses, now supports larger pages (8 K and 16 K). The larger page sizes help to reduce the number of misses (and resulting stall cycles), which can improve performance for many applications.

## Improvements in Scalability and Headroom

Communication channels within the processor have received a major overhaul to help keep the cores operating at high levels of utilization and to provide even better support for memory- and I/O-intensive applications.

### New Intel® QuickPath Interconnect Technology

The Intel Itanium processor 9300 series marks the first implementation of Intel QuickPath Interconnect Technology in the Intel Itanium processor family. Intel QuickPath Interconnect Technology replaces the Front Side Bus of previous processors with a point-to-point architecture that is more scalable and resilient. With four full-width Intel QuickPath Interconnect links and two half-width links, the Intel Itanium processor 9300 series provides peak processor-to-processor and processor-to-I/O communications up to 96 GB/s, or up to nine times that of its predecessor. It also provides a solution that scales automatically in larger system designs. As more processors are added, system bandwidth increases accordingly.

Intel QuickPath Interconnect Technology also provides a strong foundation for scaling system bandwidth in future processor generations. More links can be added as needed to support additional cores. Since the links are point-to-point and unidirectional, higher bandwidth per pin is technically feasible, which furnishes another avenue for future scaling. Intel QuickPath Interconnect Technology is also being integrated into Intel® Xeon® processors, so server manufacturers will be able to employ a common chipset for the two processor families. This will provide significant economies of scale and help fuel faster innovation across both architectures.

### Two Integrated Memory Controllers and an Intel® Scalable Memory Interconnect

Communication channels between the processor cores and main memory have been dramatically improved. Each processor has two integrated memory controllers, which provide peak memory bandwidth up to 34 GB/s, which is up to six times the bandwidth of the previous-generation processor. The new processor also includes an Intel® Scalable Memory Interconnect (Intel® SMI), which connects to the Intel® 7500 Scalable Memory Buffer to support larger physical memory configurations using industry-standard DDR3 RDIMMs. The combination of larger physical memory and increased bandwidth communications will deliver higher performance in many environments and will make it easier and more cost-effective for organizations to add memory as workloads grow.

### Directory-based Cache Coherency

Cache coherency ensures that data in memory and cache remain synchronized, so the most current data is used in every transaction. The previous-generation Intel Itanium processor used snoopy-based coherency mechanisms in which every processor on the Front Side Bus had to be "snooped" for every memory transaction. With this approach, coherency traffic is roughly proportional to the square of the number of processors in the system and can become a bottleneck in large multiprocessor servers unless auxiliary coherency (node) controllers are employed.

The Intel Itanium processor 9300 series implements true directory-based coherency. The home agent for each memory controller keeps track of all owners and sharers for each line of cache using information stored in main memory. Since the number of owners and sharers for a given cache line is typically much less than the number of caching agents in the system, coherency traffic tends to scale sub-linearly with increasing system size. This improves cache efficiency and reduces coherency-related traffic that would otherwise contend for available bandwidth. It also helps to sustain low-latency data access times as workloads increase in large multiprocessor systems.

### Glueless System Designs Up to Eight Sockets

Systems based on the previous-generation Intel Itanium processor were limited to four processor loads per Front Side Bus due to the electrical challenges of the shared bus design. The Intel Itanium processor 9300 series supports up to eight-socket glueless systems, with "glueless" meaning that no additional node controller is required (Figure 1). This enables simpler system designs with lower chip counts and smaller board footprints.[2] It also provides exceptionally fast processor-to-processor communications.

Larger systems can be built using one or more node controllers, which are designed independently by the various server vendors. These node controllers are used to aggregate multiple, multi-socket Intel QuickPath Interconnect "nodes" into larger symmetric multiprocessing (SMP) systems.[3]

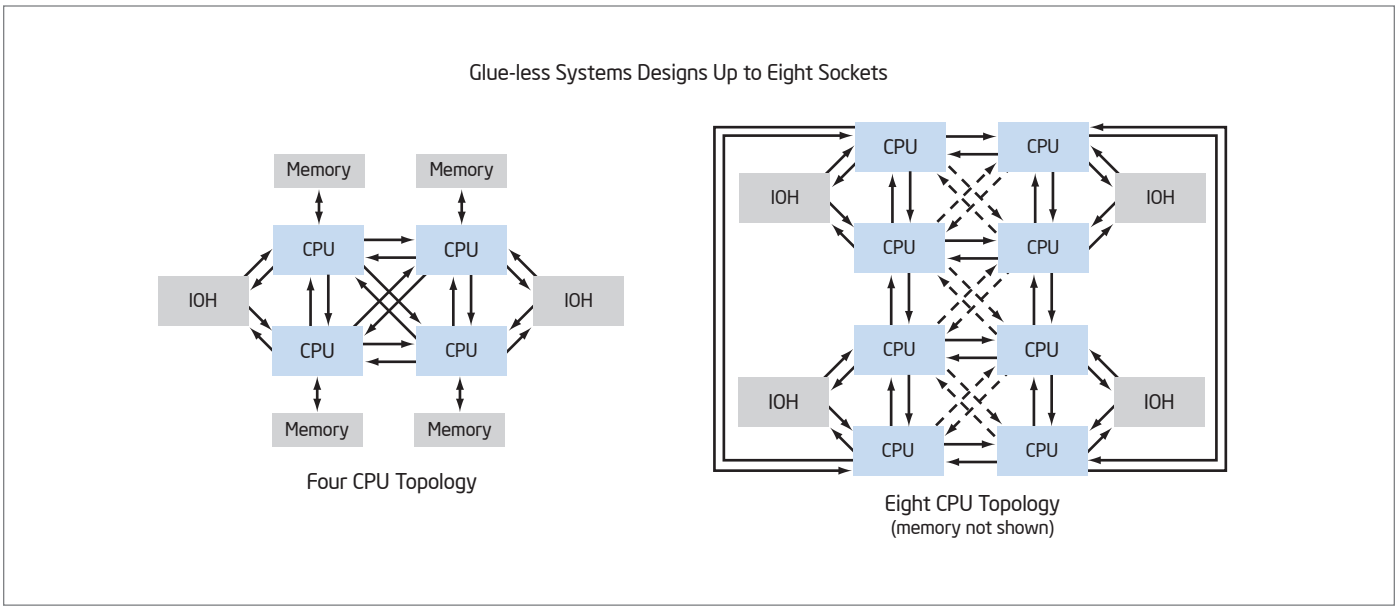



**Figure 1.** The quad-core Intel® Itanium® processor supports up to eight-socket glueless system designs. Larger symmetric multiprocessing (SMP) systems are built using node controllers designed by the individual system vendors.

## Better Energy Efficiency through Advanced Power and Thermal Management

Power consumption is increasingly important in today's data centers, and considerable effort was devoted to optimizing performance per Watt in the Intel Itanium processor 9300 series. The previous-generation processor supported Demand Based Switching (DBS), in which the OS monitors activity and throttles frequency when workloads are light to improve energy efficiency.

The Intel Itanium processor 9300 series provides enhanced power and thermal management. This includes an enhanced form of DBS, in which both voltage and frequency are modulated to enable even better energy efficiency. It also includes Intel® Turbo Boost Technology,[4] which can automatically increase frequency and voltage settings when workloads are high to provide the best possible performance without exceeding the processor's thermal design power (TDP) envelope. Performance is maximized by monitoring 120 events in each of the four cores and adjusting the core voltage and frequency every 6µs. This function can be used in conjunction with DBS to fine-tune performance versus power consumption and to meet different needs at different times (e.g., peak performance for end-of-month financial closes and peak energy efficiency when workloads are relatively light).

## Improvements in Reliability, Availability and Serviceability (RAS)

With the ongoing move toward real-time business models, data integrity and uninterrupted operation are becoming more important than ever in mission-critical computing environments. Data errors can be propagated quickly across multiple applications and downtime can have a major impact on revenue, customer satisfaction, vendor relationships and brand perception.

Consolidation of multiple applications per system further magnifies the potential impact of system downtime. A fatal error in a virtualized server could potentially bring down all hosted applications. Systems must therefore be highly resilient in order to maintain service levels while capturing the full benefits of high consolidation ratios (lower costs, simpler management and less drain on data center resources). To address growing needs, the Intel Itanium processor 9300 series extends the mainframe-class Reliability, Availability and Serviceability (RAS) features that were integrated into previous-generations of Intel Itanium processors.

In any computing system, errors in data or processing can impact data reliability, system availability or both. There are basically two kinds of errors:

- **Soft errors** are typically caused by an alpha particle or other micro event that changes the logic state of one or more silicon gates. These are transient errors and can be fixed simply by correcting the logic state.
- **Hard errors** are more persistent. They imply a failure at the hardware level and require a permanent fix to ensure correct and uninterrupted operation.

The Intel Itanium processor 9300 series incorporates extensive capabilities for detecting, correcting and reporting both kinds of errors on the processor die, in attached components and along the pathways that connect those components (Figure 2 on the next page).

### Processor RAS Features

An extensive set of RAS features are supported in the processor die itself.

- **Error Avoidance, Detection and Correction Across All Major Core Structures:** Single bit errors are avoided, detected and corrected throughout each core using a variety of mechanisms, including:
  - **Soft error (SE) hardened latches and registers[5]:** These new circuit topologies were designed by Intel to improve resistance to soft errors (see the sidebar, Reducing Soft Errors in Latches by Up to 100x).
  - **Error correcting code (ECC) or parity:** These widely used algorithms are implemented in hardware to monitor data and identify errors that occur during transmission.
  - **Intel® Cache Safe technology:** Heuristics are used to monitor the number of errors per cache index and map out bad cache lines if error frequencies reach a specified threshold. Cache data is also automatically scrubbed to correct single bit errors that might otherwise accumulate and create uncorrectable problems. Intel® Cache Safe technology has been extended in the Intel Itanium processor 9300 series to cover the second-level and directory cache arrays. In the previous generation, only the third-level cache was covered.

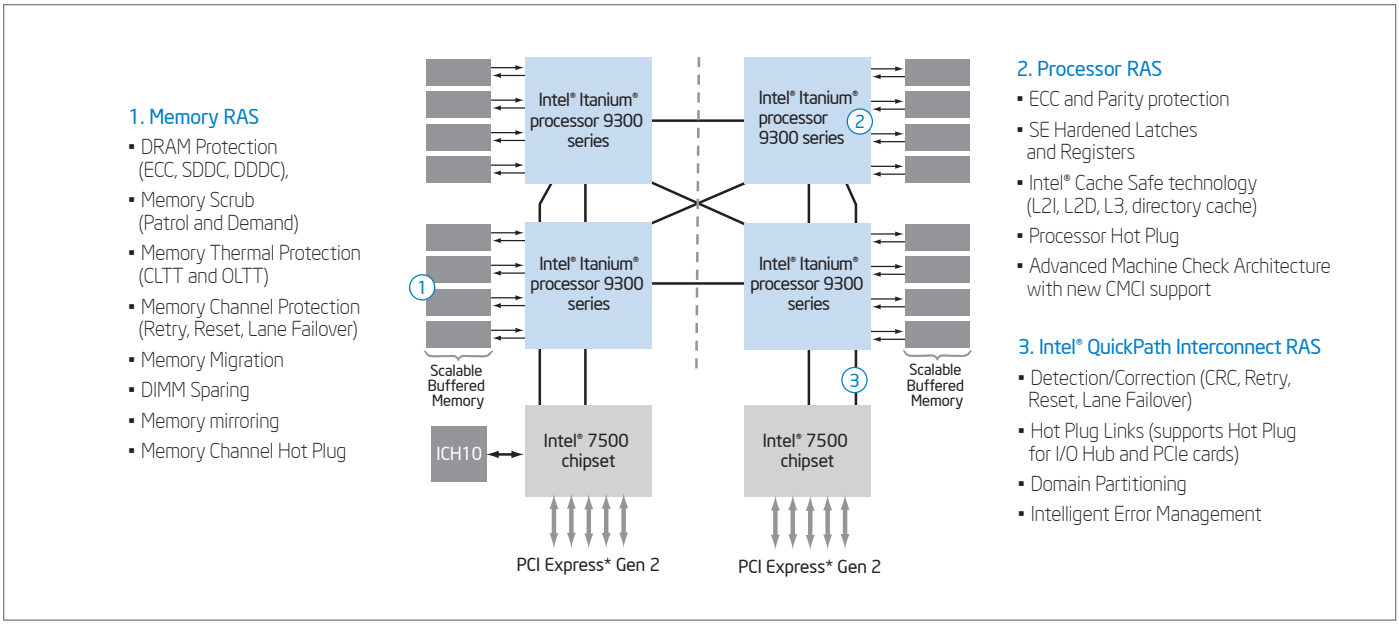



**Figure 2.** The Intel® Itanium® processor 9300 series provides comprehensive support for error avoidance, detection, correction, containment and reporting across all major structures.

- **Processor Hot Plug[6]:** An entire processor or processor board can be physically removed, replaced or added without bringing down the system, either to replace a failed or failing component or to add resources as workloads grow. Processors or cores can also be logically added or removed from running partitions, either to reallocate resources among partitions or to functionally integrate a spare processor into a running partition.
- **Advanced Machine Check Architecture:** The Advanced Machine Check Architecture enables coordinated error handling across the hardware, firmware and operating system (see page 9 and Figure 4 for more information).

### Memory RAS Features

Extensive RAS features are integrated to detect and correct errors throughout the memory subsystem.[7]

- **DRAM Protection:** Error Correcting Code (ECC) mechanisms are implemented to detect and fix errors in attached memory components. Both Single Device Data Correction (SDDC) and Double Device Data Correction (DDDC) are supported. SDDC is strong enough to correct multi-bit errors[8] in a single DRAM device, to map out a failed device, and to continue correcting single-bit errors after a device is mapped out. DDDC is even stronger. It can correct multi-bit errors in two DRAM devices, map out two failed devices, and continue correcting single-bit errors after the devices are mapped out. These mechanisms can improve system uptime and reduce DIMM replacement rates. There is no performance penalty for mapping out the devices.

## Reducing Soft Errors in Latches by Up to 100x

One of the most common non-human sources of server error is caused by naturally occurring high-energy particles striking nuclei in processors, chipsets and memory components. In some cases, the energy from these events can cause a logic gate to switch states, resulting in a "soft" error that can corrupt data and even bring down an entire server.

As discussed in this paper, the Intel® Itanium® processor 9300 series has extensive mechanisms for detecting, correcting and containing these errors. It also includes new circuit topologies that dramatically reduce the frequency of soft errors. Estimates show that the new soft error (SE) hardened latches are up to 100 times less susceptible to soft errors than standard latches, and the new SE hardened registers are up to 80 times less susceptible than standard registers.

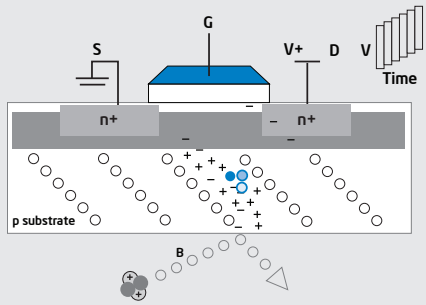



**Figure 3.** A random particle strike can change a logic state, causing a "soft error" that must be detected and corrected.

- **Memory Scrubbing (Patrol and Demand):** The protections described above are activated only when a memory location is read. However, errors can occur to data or instructions in memory locations that are not accessed. If these errors accumulate, they can result in multi-bit errors that cannot be corrected and could result in data corruption and even system failure. Memory scrubbing in the Intel Itanium processor 9300 series employs an integrated hardware engine to find and correct memory errors before they accumulate. Scrubbing is performed periodically and automatically on all populated memory locations (Patrol Scrubbing). In addition, when errors are discovered for data in transit, the corrected data is rewritten back to the appropriate memory location (Demand Scrubbing).

- **Memory Thermal Protection:** Overheating of memory components can cause or accelerate component failure. The Intel Itanium processor 9300 series supports two mechanisms for throttling commands issued to the memory channels to protect against overheating. Closed loop thermal throttling (CLTT) is triggered by a thermal sensor in the DIMM that sends a signal to the memory controller. Open loop thermal throttling (OLTT) is triggered when the rate of memory commands per DIMM exceed a configurable limit for a configurable time window. Alternatively, the firmware can be configured to increase system fan speed in response to these same triggers.

- **Memory Channel Protection:** A Cyclic Redundancy Check (CRC) mechanism is used to detect errors in the memory channels. When an error is detected, a series of progressively stronger corrective actions are triggered: 1) the transaction is retried, several times if necessary, which corrects most soft errors; 2) the memory channel is physically reset (reinitialized), which corrects most persistent errors; 3) if the problem persists, the affected lane on the memory channel is mapped out. This corrects hard errors without degrading performance.

- **Memory Migration and DIMM Sparing:** An algorithm in system firmware continuously monitors memory errors. If it determines that a memory component is failing, a hardware engine can copy the contents of the failing component to another location. This process can be completely transparent to the OS. Two mechanisms are supported. With DIMM sparing, the contents of the failing DIMM are copied to a spare DIMM on the same memory channel. With memory migration, the contents are copied to the memory of any other memory controller on the system. DIMM sparing requires less memory overhead (just a single spare DIMM per memory channel). However, memory migration enables hot-swap capabilities for an entire memory card and can help to support hot-swap functionality for processors.

- **Memory Mirroring:** The Intel Itanium processor 9300 series can be configured to automatically maintain a backup copy of main memory. If a failure is detected, the correct data can be accessed from the backup. Since the probability of simultaneous errors in parallel memory locations on two different memory DIMMs is extremely small, this provides exceptionally strong protection against memory errors. However, it does require that the system be configured with twice the memory capacity to support the backup. If memory capacity is an issue, memory mirroring can be configured only for selected memory controllers.

- **Memory Channel Hot Plug:** Memory channels can be put in an electrically idle state. This enables IT personnel to logically reallocate memory resources among running partitions. It also enables them to physically add or remove memory riser cards without bringing down the system. With these capabilities, memory upgrades, faulty memory card replacements, and resource management can all be performed without downtime.

### Intel® QuickPath Interconnect RAS Features

The Intel QuickPath Interconnect offers extensive, multi-level protections against hard and soft errors to support very high levels of data reliability and system availability for processor-to-processor and processor-to-I/O Hub communications.

- **Error Detection and Correction:** Error protection in the interconnect subsystem works very much like the progressive memory channel protections described above: 1) CRC is used to detect errors; 2) transactions can be retried multiple times; 3) the channel can be physically reset; and 4) bad lanes can be mapped out. Although mapping out lanes may impact performance by reducing a full-width link to half-width or a half-width link to quarter-width, it does enable uninterrupted performance and it does protect against most multi-bit hard errors.
- **Clock Failover:** In the event of a clock failure, clocks can be redirected to one of two failover clock lanes to enable uninterrupted operation.
- **Hot Pluggable Interconnect Links:** Interconnect links can be put in an electrically idle state without bringing down the system, which enables a component on the other side of the link, such as an I/O Hub or another processor, to be physically replaced. This capability can also be used to create a hard partition between connected components and to dynamically reconfigure partition boundaries during runtime to prevent downtime and to utilize available resources more efficiently.
- **Intelligent Error Management:** Intel QuickPath Interconnect Technology goes beyond error detection and correction. It also provides information on the type, scope and source of an error. For example, a single dropped or lost packet often leads to a cascade of lost packets resulting in a large number of transaction timeouts. By sorting and analyzing the dropped packets, Intel QuickPath Interconnect Technology helps to identify the initial source of the error. The firmware and OS can then use this information to recover in the least disruptive manner. Intel QuickPath Interconnect Technology also provides mechanisms to ensure that errors do not corrupt non-volatile storage, such as a hard drive.

### Advanced Machine Check Architecture

Many of the RAS mechanisms discussed above are supported entirely in hardware. Others require support from the firmware or the OS. The Intel® Itanium® microarchitecture implements an Advanced Machine Check Architecture that coordinates error handling across all these levels, using well-defined interfaces that enable server vendors to integrate and extend RAS capabilities in their system designs and management applications. This sophisticated error management greatly reduces the likelihood of data corruption. It also improves the reliability of the system, since it enables hardware to work in conjunction with system firmware and the OS to recover from otherwise uncorrectable errors (Figure 4).

The Intel Itanium processor 9300 series includes a number of enhancements to the Machine Check Architecture. For example, Corrected Machine Check Interrupts (CMCI) reporting capabilities have been added to the memory and interconnect subsystems. CMCI enables error reporting through the local processor interface which helps to localize errors faster and provides a foundation for building predictive failure solutions.

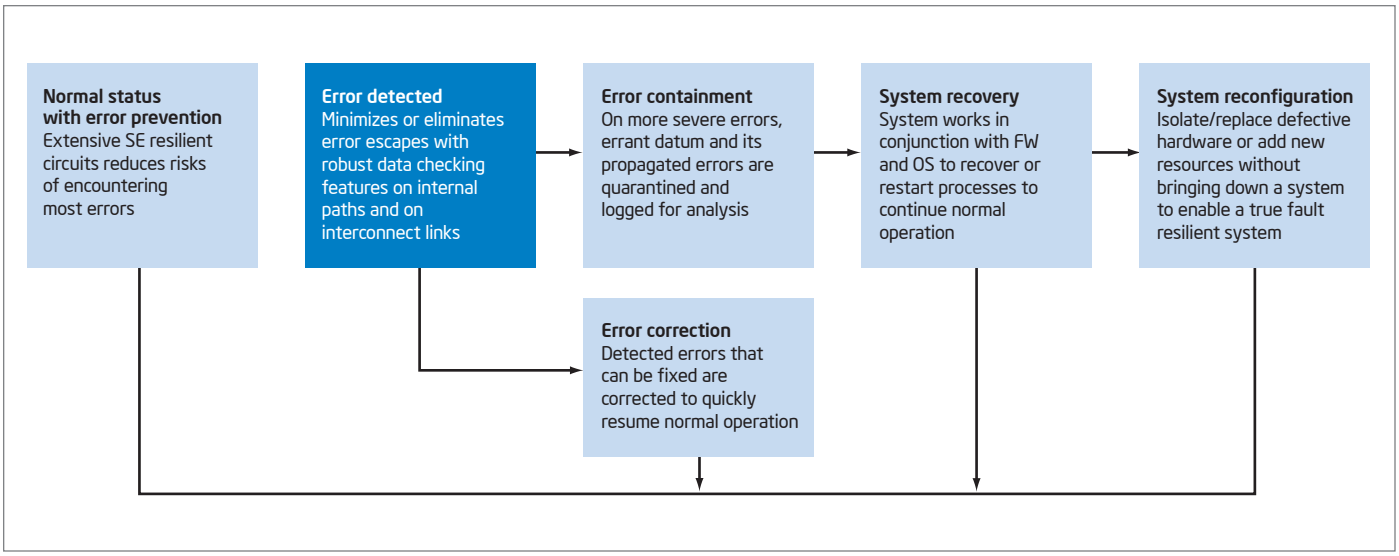



**Figure 4.** The extensive error detection and correction mechanisms in the Intel® Itanium® processor 9300 series, combined with its Advanced Machine Check Architecture, enable comprehensive error management to optimize both system uptime and data integrity.

## Manageability: Dynamic Partitioning and Hot Plug Components

The Intel Itanium processor 9300 series includes a rich set of capabilities for partitioning and managing system resources. These capabilities can help organizations maintain uninterrupted operation in response to error events. They can also help them optimize utilization and power efficiency by dynamically responding to changing workloads. Many of these capabilities were supported in earlier Itanium®-based systems, but were based on proprietary technologies developed and integrated by the various system vendors. They are supported natively in the Intel Itanium processor 9300 series, which simplifies integration and helps to establish more consistent implementations.

These new capabilities are best understood within the context of the three types of partitioning models supported by Itanium-based systems (Figure 5).

**Physical (Hard) Partitions** provide full electrical isolation so hardware and software faults in one partition will not impact any other partition. A number of new features in the Intel Itanium processor 9300 series support flexible and dynamic control of hard partitions. For example, physical partitioning boundaries can be dynamically changed among running partitions and resources can be added, removed or replaced without restarting the system or rebooting the affected operating systems[9] (Table 2). These capabilities are enabled by Intel QuickPath Interconnect Technology, which supports hot pluggable processor, memory and I/O components, as well as dynamic reconfiguration of parameters such as snoop control, interrupt routing, address decoding and packet routing.

**OS Partitions** enable multiple applications to operate under a single OS, while each appears to have a dedicated OS. This capability is provided by special manageability firmware and is therefore independent of the OS.

**Virtual Partitions** are created using Virtual Machine Monitor (VMM) software, which is available from some system vendors and from a number of third-party software vendors. The VMM enables each physical partition to host multiple OS and application images, and provides the most dynamic and granular control of partitioning boundaries and system resources. Most VMM software also provides substantial isolation for software errors, though it does not provide the electrical isolation of a physical partition.

Intel® Virtualization technology[10] (Intel® VT-i) was integrated into the previous-generation processor to provide a better foundation for virtualization. It provides a clean and stable interface for the VMM as well as hardware assists for many core virtualization functions. The Intel Itanium processor 9300 series includes enhanced support for virtualization (Intel® VT-i2). Intel VT-i2 provides additional hardware assistance designed to improve performance by reducing the amount of time spent running the VMM emulation code. It also reduces the number of required transitions between the OS and the VMM.

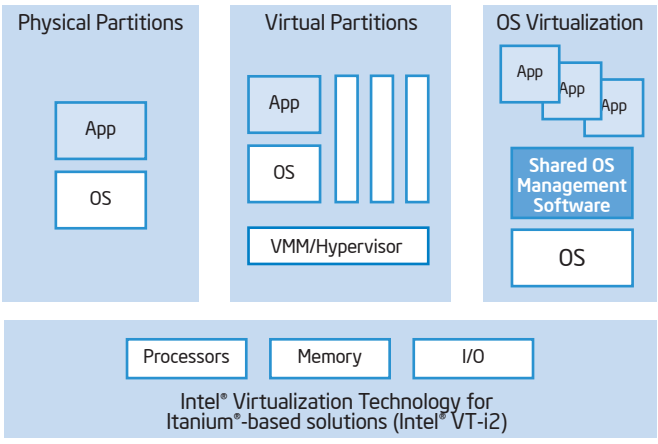



**Figure 5.** Itanium-based servers support multiple partitioning options to enhance IT flexibility and resource utilization. Advancements in the Intel® Itanium® processor 9300 series – such as Intel® VT-i2 and native support for dynamic partitioning and dynamic resource allocation – deliver enhanced support to provide organizations with greater flexibility, better performance and more consistent solutions across vendor implementations.

**Table 2. The Intel® Itanium® processor 9300 series provides mainframe-class support for adding, replacing and managing system resources**

| Function | Component[a] | | |
|---|---|---|---|
| | CPU | Memory | I/O |
| **Onlining/Offlining –** System resources can be logically added to or removed from a running partition (does not imply physical insertion or removal). | Processor socket or core | Channel pair | Hub or PCIe card |
| **Sparing –** A failed or failing component can be taken offline and a spare component can be brought online without downtime. | Processor socket or core | DIMM pair or channel pair | Hub or PCIe card |
| **Physical Hot Swap –** A physical component can be added, removed or replaced without downtime. | Processor socket | Channel pair | Hub or PCIe card |

[a] Many Itanium®-based system vendors add additional redundancy with hot swap and failover capabilities (fans, power supplies, etc.).

## Future Processor Generations

Two more generations of the Intel Itanium processor are currently in development. Both will be socket compatible with the Intel Itanium processor 9300 series and binary compatible with current software code.

- Next-Generation Intel Itanium processor (code-name Poulson): This processor will be based on a new ultra parallel microarchitecture and will be manufactured on Intel's 32 nm process technology. It will include more cores, support more software threads and run at a higher clock frequency. It will also include new RAS features and support a number of instruction-level advancements.
- Future Intel Itanium processor (code-name Kittson): This future Intel Itanium processor is in definition today and can be expected to provide additional capabilities.

## Conclusion

The Intel Itanium processor 9300 series delivers a major leap forward in performance and scalability for Itanium-based servers, along with more and better RAS features and improved support for virtualization and dynamic partitioning. Systems based on this processor provide an exceptional combination of scalable performance and high availability, along with the flexibility to support diverse operating systems, applications and data center environments.

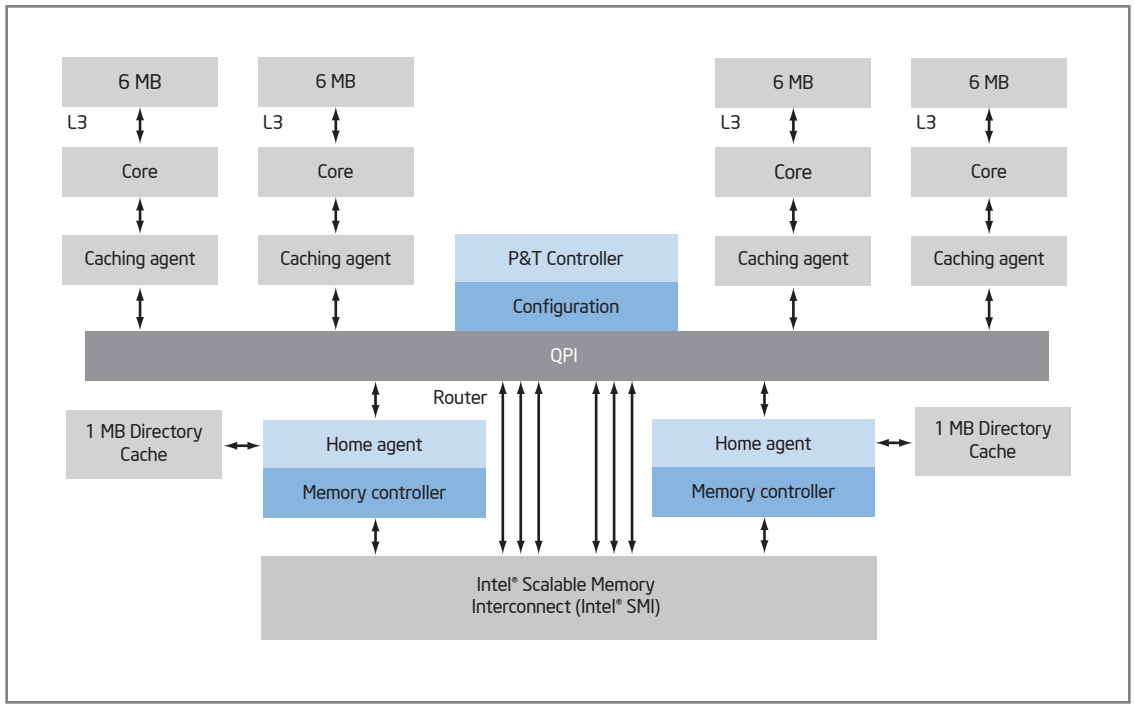
This new processor provides twice as many cores as its predecessor, more on-die cache (30 MB versus 27 MB), up to nine times the interconnect bandwidth and up to six times the memory bandwidth. With these enhancements, it can be expected to deliver more than double the performance (or more) for many applications. It also scales more effectively in large multiprocessor systems. Interconnect and memory bandwidth increase automatically as more processors are added, and memory capacity is up to eight times greater than systems built with the previous-generation processor. System design is also simplified. Glueless designs are supported up to eight sockets and much larger systems can be built by aggregating these four- or eight-socket "nodes" using vendor-supplied chipsets.

New Intel Turbo Boost Technology and an enhanced version of Demand Based Switching (DBS) provide better control of performance versus energy efficiency. DBS can now be used to throttle frequency as well as voltage to further reduce energy conservation during light workloads. Intel Turbo Boost can be used to dynamically increase clock frequency for certain workloads, to maximize performance while staying within the processor's thermal design limit.

The Intel Itanium processor 9300 series also extends the mainframe-class RAS features of its predecessor. Error detection, correction and avoidance are extended and/or improved across all key components and advanced support for partitioning, virtualization and resource management have been integrated directly into the processor to provide simpler and more consistent vendor implementations.

With these advancements, systems based on the Intel Itanium processor 9300 series will provide even more powerful, flexible and robust support for today's most demanding and mission-critical applications. Organizations will also be able to scale the performance and capacity of these systems for many years to come through simple component upgrades, using future Intel Itanium processors that are in development today and are being designed to provide both binary compatibility and socket compatibility.

## Appendix A: Architectural Overview




| | |
|---|---|
| **Processing** | |
| Four cores per processor | Individual cores are very similar to the cores in the previous Intel® Itanium® processor 9000 series, but with significant enhancements to improve performance and virtualization |
| Dedicated cache per core<br>▪ **L1:** Instruction (16 KB), Data (16 KB)<br>▪ **L2:** Instruction (512 KB), Data (256 KB)<br>▪ **L3:** Cache (6 MB) | Similar to the Intel Itanium processor 9000 series. Maintains 128 byte L2 and L3 cache line size, so existing code remains optimized in the new processor. |
| Dedicated Caching Agent per core | Interfaces the core with the Intel® QuickPath Interconnect technology router. Supports 64 concurrent transactions and converts 128-byte cache lines to the 64-byte Intel QuickPath Interconnect technology protocol. |
| Configuration Agent | Manages configuration and interrupt requests and interfaces with side band external interfaces, such as SMBus and flash ROM. |
| Power & Thermal (P&T) Controller | Enables dynamic voltage and frequency control to optimize performance versus energy consumption as desired. Also interfaces with thermal trip points and can trigger throttling to protect against over-heating and preserve data integrity. |
| **Memory** | |
| Two Integrated Memory Controllers | Each memory controller can schedule memory commands at up to 4.8 GT/s and supports: DRAMs running at 800 MHz; up to 48 concurrent DRAM commands; several policies for keeping DRAM pages closed following an access. |
| Dedicated Home Agent per Controller | Maintains cache coherency using true directory-based coherency (1 MB directory per controller), to improve scalability in large SMP configurations. Tracks Exclusive, Shared, Coarse Shared or Invalid status, plus a coarse sharing vector, for each cache line. Supports up to 32 pending transactions. |
| Memory Channel Links (PHY Layers) | Each link supports 4.8 GT/s transfer rates and lane and polarity reversal for flexible board level routing. |
| Intel® Scalable Memory Interconnect | A high performance serial differential interface that can connect with a memory buffer to expand narrow CPU interfaces to wide DDR3 interfaces. |
| **System Interconnect** | |
| Intel® QuickPath Interconnect Technology Router | 12-port crossbar router with dedicated ports for each of the six Intel QuickPath Interconnect links, the two Home Agents and the four Caching Agents. Can route packets from any port to any port. Supports high-speed, high-reliability point-to-point communications among processor cores and between the cores and connected I/O devices. |
| Intel® QuickPath Interconnect Physical Links (PHY Layers) | Each link supports 4.8 GT/s transfer rates and lane and polarity reversal for flexible board level routing. Full width Intel QuickPath Interconnect Technology links are 20 bits wide with differential signaling per direction. |

## Appendix B: Table of RAS Features

| RAS Capability | Intel® Itanium® Processor 9100 Series | Intel® Itanium® Processor 9300 Series |
|---|---|---|
| Processor Core | | |
| Extensive Error Protection/Correction<br>On-core structures and system interface:<br>▪ ECC, parity and/or SER hardened latches and registers are used to avoid, detect and correct errors.<br>On-cache structures:<br>▪ ECC is used to detect and correct errors.<br>▪ Intel® Cache Safe Technology is used to disable failed cache lines for improved availability. | ▪ | Enhanced |
| | L3 only | L2, L3 and Directory |
| Processor Socket | | |
| Advanced Error Protection/Correction on the Processor Links<br>▪ **Dynamic Link Rerouting:** Sustains uninterrupted operation if an Intel® QuickPath Interconnect or Intel® Scalable Memory Interconnect (Intel® SMI) link physically fails. | ▪ | Enhanced |
| **Processor Onlining/Offlining:** A processor can be functionally enabled or disabled without downtime to adjust available resources or to map out a failed component. | OEM-based[c] | Native support |
| **Processor Hot Plug:** A processor can be physically added, removed or replaced without downtime for system upgrades or to replace a problematic component. | OEM-based[c] | Native support |
| **Intel® Virtualization Technology – processor cores:** Hardware-based virtualization support improves ability to implement transparent workload migration to optimize resource utilization and simplify failover. | Intel VT-i | Intel VT-i2 |
| Memory Subsystem | | |
| Memory Error Correction mechanisms include:<br>▪ **Memory ECC Support:** Automatically detects and corrects all single-bit errors and most double bit stored errors (uncorrectable errors are detected and reported). Errors in up to eight consecutive bits can be corrected.<br>▪ **Single Device Data Correction:** Automatically corrects multi-bit errors on a single DRAM device; can map out a failed device and continue correcting single-bit errors.<br>▪ **Dual Device Data Correction:** Automatically corrects multi-bit errors on two DRAM devices; can map out two failed DRAM devices and continue correcting single-bit errors. | OEM-based[c] or Intel chipset required | Native support |
| | OEM-based[c] | Native support |
| **Memory Channel Protection:** Includes three levels of protection, Cyclic Redundancy Check (CRC) to detect and repair transient errors; physical layer reset for persistent errors; and lane failover if the reset fails. | OEM-based[c] | Native support |
| **Memory Scrubbing:** Memory is monitored to correct errors, which protects correctable errors from accumulating and becoming uncorrectable. Performed automatically and periodically (Patrol) and also at the request of the OS (Demand). | OEM-based[c] | Native support |
| **Memory DIMM sparing:** Firmware copies data from a failing DIMM to a spare DIMM on the same memory channel, and maps out the failed component to enable uninterrupted operation. | | ▪ |
| **Memory Migration:** Firmware copies data from a failing DIMM and migrates it to a DIMM on another memory controller of the same or another processor. | OEM-based[c] | Native support |
| **Memory Mirroring:** A backup copy of main memory can be maintained for very high-reliability error correction (if used, requires twice the memory). | OEM-based[c] | Native support |
| **Memory Onlining/Offlining:** One or more DIMMS or memory riser cards can be functionally enabled or disabled without downtime to adjust available resources or to map out a failed component. | OEM-based[c] | Native support |
| **Memory Hot Plug Support:** Memory components (DIMMs) can be physically added, removed or replaced without downtime. Includes OS-visible and OS-transparent capabilities. | OEM-based[c] | Native support[d] |
| Memory Thermal Protection includes:<br>▪ **Closed Loop thermal throttling:** Memory channel activity can be reduced (or fan speed increased) when the temperature of a DIMM exceeds a preset level.<br>▪ **Open Loop Thermal Throttling:** Memory channel activity can be reduced (or fan speed increased) when the number of memory commands per DIMM exceeds a configurable limit over a configurable time interval. | | ▪ |

## Appendix B: Table of RAS Features, continued

| RAS Capability | Intel® Itanium® Processor 9100 Series | Intel® Itanium® Processor 9300 Series |
|---|---|---|
| I/O Hub | | |
| **PCI Express* Hot Plug:** PCIe cards can be physically added, removed or replaced without downtime. | OEM-based[b] | OEM-based[b] or Intel chipset required |
| **I/O Hub Hot Plug:** I/O Hubs and/or associated I/O devices can be physically added, removed or replaced without downtime. | OEM-based[b] | Native Support |
| **I/O Hub Onlining/Offlining:** I/O Hubs and/or associated I/O devices can be functionally enabled or disabled without downtime to adjust available resources or to map out a failed component. | OEM-based[b] | Native Support[c] |
| **Intel® Virtualization Technology – I/O subsystem:** Hardware-based virtualization support improves ability to implement transparent workload migration to optimize resource utilization and simplify failover. | | ▪ |
| Partition | | |
| Static and Dynamic Hard Partitions<br>▪ Node Level | ▪ | ▪ |
| ▪ Processor Level (Enables more granular control) | | ▪[a] |
| Error Handling | | |
| **Bad Data Containment:** Memory locations with corrupted data can be tagged (sometimes called "data poisoning"). This limits the impact to the currently running program and greatly reduces the need to reset the system. | ▪ | ▪ |
| **Advanced Machine Check Architecture:** Standards-based interfaces enable coordinated error handling with appropriate escalations and communications among the hardware, firmware and operating system. | ▪ | Enhanced |

[a]Integrated capability. May be supported by one or more server vendors in future Itanium®-based systems.

[b]Feature was supported by one or more system vendors using proprietary hardware, firmware and/or software components. Vendors may continue to leverage their proprietary solutions or may take advantage of the hardware and firmware support integrated into the Intel® Itanium® processor 9300 series.

[c]Integrated capability, but requires additional support from firmware and/or the operating system. System and OS vendors may or may not take advantage of the integrated support.

The features indicated have been validated by Intel to the extent possible. It is up to server manufacturers to validate and confirm full functionality in a complete platform environment.

## Appendix C: Glossary of Useful RAS Terms

| | |
|---|---|
| RAS | Reliability, Availability and Serviceability |
| Reliability | Assurance that computational results are correct. Errors are detected and corrected when possible and reported if they cannot be corrected. |
| Availability | Assurance that the system is up and running to support an organization's computing needs. |
| Serviceability | Assurance that errors are reported and that faulty components can be identified and replaced. |
| Error Detection/Correction | Ability to detect and correct hard and soft errors to increase reliability and availability. |
| Soft Errors | A transient error that can be corrected by overwriting with the correct data. |
| Hard Errors | A persistent error that cannot be fixed by overwriting with the correct data (e.g., a faulty logic gate). |
| Error Correction Code (ECC), Cyclic Redundancy Check (CRC), Parity | Widely used mechanisms for identifying hard and soft errors. |
| Domain Partitioning | Ability to divide a system into a number of smaller systems, each booting its own OS and operating independently of the others. |
| Physical (Hard) | Each partition is completely isolated, so software and hardware errors in one partition will not impact another. |
| Dynamic | Partitioning can be implemented and partition boundaries modified without shutting down the system or rebooting affected operating systems. |
| Static | Partitioning can be implemented and configured only at boot time or if impacted operating systems are not running. |
| Field Replaceable Unit (FRU) | A system component that can be physically added, removed or replaced in the field, such as a processor board, an I/O Hub, a DIMM or a PCI card. An FRU may or may not be hot pluggable. |
| Hot Plug | A generic term for hot add and hot remove operations. |
| Hot Add/Remove | The physical addition/removal of an FRU without shutting down the system or stopping the OS. Typically requires OS support. |
| Hot Replace | A hot remove followed by a hot add, typically prompted by components that are defective or starting to show signs of defects (e.g. error frequency exceeds a configurable threshold). |
| Hot Swap | A logical replacement of a component with another component that has already been installed in the system (e.g., a spare processor, DIMM or I/O Hub). A hot swap can be OS transparent (accomplished completely via hardware and firmware) or OS assisted. |
| Onlining/Offlining | The logical (non-physical) addition/removal of a component in a running system. An offlined component can remain in the system as a spare, be reallocated to another partition, or removed at a later time. |
| Processor: Socket, Core, Logical Processor | Each Intel® Itanium® processor fits into a single socket on the system motherboard and can contain up to four cores (complete execution units). Each core can be recognized as multiple logical processors by the OS. A logical processor is collection of processing resources capable of running a thread of software code. |

## About the Authors


### Rajesh Agny, Itanium Platform Marketing Manager, Data Center Group, Intel Corporation

Rajesh is responsible for marketing the Intel® Itanium® processor 9300 series and the associated platform ingredients. He has worked in the microprocessor industry for 25 years, first at Harris Semiconductor and then at Intel Corporation. During his time at Intel, he has played key roles in Process Engineering, Quality and Reliability, Strategic Planning, and Itanium Platform Marketing.

### Eric DeLano, Senior Principal Engineer, Intel Architecture Group, Intel Corporation

Eric is chief architect for the Intel Itanium processor 9300 series. He has worked in the microprocessor industry for 25 years, first at Hewlett-Packard and then at Intel corporation. During that time he has played key roles in developing six different HP PA-RISC microprocessors and three Intel Itanium microprocessors, often in the role of lead logic designer or lead architect.

### Mohan Kumar, Senior Principal Engineer, Intel Architecture Group, Intel Corporation

Mohan is responsible for Server Reliability, Availability and Serviceability (RAS) in the Architecture and Planning organization of the Intel Digital Enterprise Group. He is key lead for the engineering working group that developed the RAS flows for the high-end RAS features implemented in the Intel Itanium processor 9300 series and associated Intel components.

### Murugasamy (Sammy) Nachimuthu, Lead Server Firmware Architect, Intel Architecture Group, Intel Corporation

Sammy is responsible for RAS flow development and implementation for the Intel Itanium processor 9300 series. He has been involved in the development of the System Abstraction Layer (SAL) across all Intel Itanium processor product generations.

### Robert Shiveley, Mission Critical Server Product Manager, Intel Corporation

During his 12 years at Intel, Robert has worked in enterprise product and solutions marketing involving enterprise manageability, security, IT infrastructure optimization, IT business value and enterprise regulatory compliance. Prior to Intel, he worked in strategic and product marketing, solution sales and business development at enterprise ISVs.


For more information on the Intel® Itanium® processor 9300 series, visit http://www.intel.com/products/processor/itanium/index.htm

[1] Based on Intel internal measurements.

[2] Since the processor supports a route-through function to forward packets to remote components, it is not necessary for every socket to have a direct connection to every other socket in the system.

[3] Node controllers are an area in which server vendors can add unique value in their system designs (such as additional RAS and manageability features), so some may continue to design and integrate node controllers even in their 4-socket and 8-socket system designs.

[4] Intel® Turbo Boost Technology requires a platform with a processor with Intel Turbo Boost Technology capability. Intel Turbo Boost Technology performance varies depending on hardware, software and overall system configuration. Check with your platform manufacturer on whether your system delivers Intel Turbo Boost Technology. For more information, see http://www.intel.com/technology/turboboost.

[5] More than 99% of the system interface logic latches (including circuits not in the core, such as the router and memory controller) are SE hardened, and 100% of the system interface logic register files are either SE-hardened or have ECC protection.

[6] Operating System support is required to take advantage of the described functionality.

[7] Many of the described features are supported fully in silicon and are performed automatically and transparently. Others require additional support from the firmware, platform or OS and may not be supported in all systems.

[8] Certain multi-bit error patterns cannot be corrected, but are detected and reported.

[9] These capabilities require operating system (OS) support and may not be supported in all operating systems.

[10] Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain platform software enabled for it. Functionality, performance or other benefits will vary depending on hardware and software configurations and may require a BIOS update. Software applications may not be compatible with all operating systems. Please check with your application vendor.




Printed in USA 0210/RS/OCG/XX/PDF Please Recycle 323247-001US