

A 32 nm, 3.1 billion transistor, 12 wide issue Itanium® Processor for Mission-Critical Servers

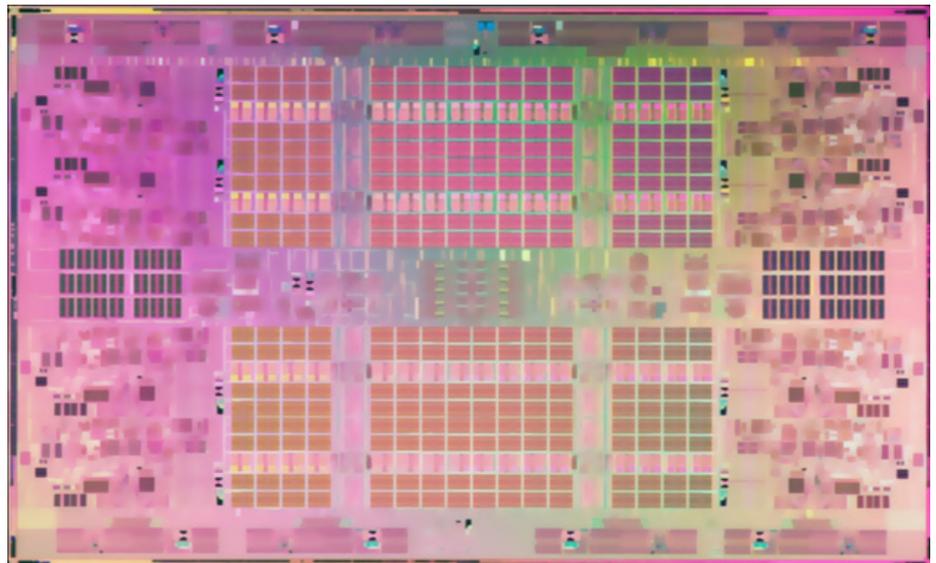


Figure 1: Poulson Die Photo

Authors:

Reid Riedlinger
Rohit Bhatia
Larry Biro
Bill Bowhill
Eric Fetzer
Paul Gronowski
Tom Grutkowski
Intel Corporation

ABSTRACT

An Itanium® processor implemented in 32 nm CMOS with 9 layers of Cu contains 3.1 billion transistors. The die measures 18.2 mm by 29.9 mm. The processor has 8 multi-threaded cores, a ring based system interface and combined cache on the die is 50 MB. High speed links allow for peak processor-to-processor bandwidth of up to 128 GB/s and memory bandwidth of up to 45 GB/s. The Poulson processor is currently in Intel and customer labs for validation and is on track with targeted availability in 2012.

July 2011
Revision 1.1

Poulson High Level Summary

The next generation in the Intel® Itanium® processor family, code named Poulson, has eight multi-threaded 64 bit cores. Poulson is socket compatible with the current Intel® Itanium® Processor 9300 series (Tukwila) [1]. The new design integrates a ring based system interface derived from portions of previous Xeon® and Itanium® processors, and includes 32 MBs of Last Level Cache (LLC). The processor is designed in Intel's 32 nm CMOS technol-

while lowering the thermal design power (TDP) by 15 watts to 170 and increases the top frequency of the IO and memory interfaces by 50% to 6.4GT/s.

The design introduces a new core micro-architecture and floor plan that significantly improves frequency, and power efficiency. The core implements an 11-stage in-order, decoupled frontend and backend pipeline which employs replay and flush mechanisms versus the previous global stall micro-architecture. The decoupled

96 entry distributed instruction buffer, replicated for each thread, decouples the frontend and backend pipelines while storing replay information. To further reduce data access penalties the core implements new features including: a hardware data prefetcher, data access hints and improved concurrent TLB accesses. The core improves performance through fine grain multi-threading, replicated instruction buffers and D-side TLBs, a 4 cycle integer multiplier, and an additional 32 entries to the integer register file. A three level cache hierarchy supports these parallel execution resources with the first level single cycle 16K Instruction (I) and Data (D) cache that is backed by two second level caches a nine cycle 512K I cache and an eight cycle 256K D cache.

The core floor plan was optimized around the performance sensitive single cycle integer execution and first level data cache (Figure 3). Timely delivery of virtual addresses to the first level TLB required placing it within the IEU data path. Way muxing and data rotation circuitry was positioned directly below the integer vertical center line to speed data return from the FLD to the IEU. The floor plan optimizations helped enable a 25% reduction in cycle time for the pre-silicon design target over the previous generation core.

Technical Details

To further improve Vmin operation the pre-silicon frequency analysis was primarily done at low voltage with heuristic runs at high voltage. Vmin sensitive circuits, including pulse latches, dynamic logic and NFET-only latches were avoided to enable robust low voltage operation and lower power. Register files (RF) used fully interrupted feedback cells for low voltage writes and static global bit lines where possible. (Figure 4) A full-scan methodology supports an ATPG scan-based testing flow which enables excellent manufacturing test coverage. All RFs contain a local direct access testing port. A fine-grain

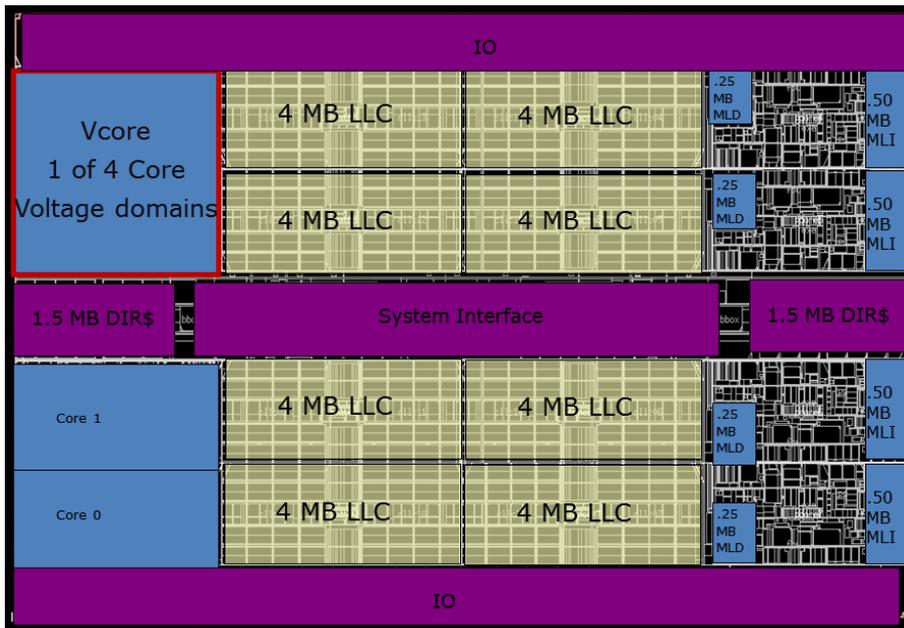


Figure 2. Poulson Processor Block Diagram

ogy utilizing high-K dielectric metal gate transistors [2] combined with nine layers of copper interconnect. The 18.2 mm by 29.9 mm die contains 3.1 billion transistors, with 720 million allocated to the eight cores (Fig 2). A total of 54 MB of on die cache is distributed throughout the core and system interface. Poulson implements twice as many cores as Tukwila

pipelines enable an increase in resource utilization and throughput. The frontend pipeline fetches six instructions per cycle while the backend executes and retires up to twelve instructions per cycle. The backend execution resources include six ALUs, two integer units, two floating-point units, two memory units and three branch units distributed across twelve ports. A

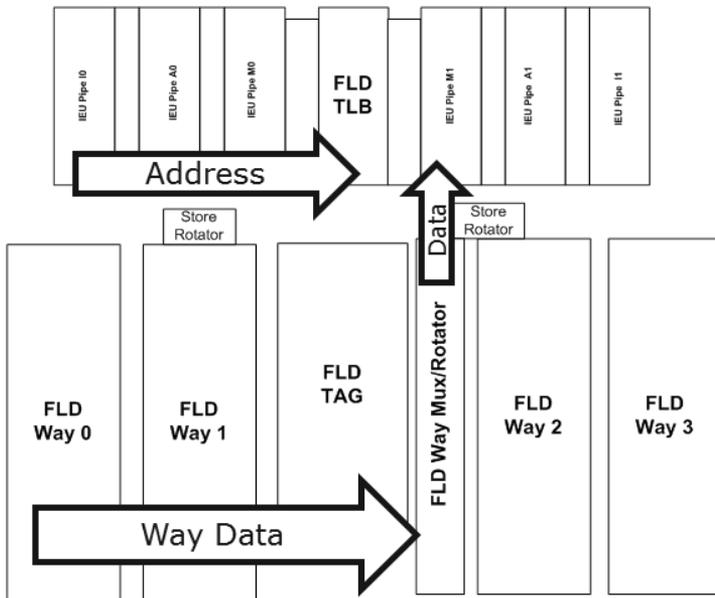


Figure 3. EXE/FLD datapath integration

and data events to predict core power consumption. To monitor the power supply in real time, the design features an on-die droop measurement [4] with the ability to introduce controlled droop events to improve debug. The power supplies are divided into four individually regulated core pairs with additional supplies for the system interface/large caches and the I/O subsystem. The power on Poulson is distributed as follows: ~55% across the 8 cores, ~35% in the uncore and ~10% consumed in the IO. Core pairs can be turned off for core defeated configurations and for power savings. Individual regulation of the core pairs allows for optimization of frequency by compensating for within die technology variations with voltage adjustments unique to each core pair. Core power supplies are bypassed by an embedded array capacitance in the central layers of the package to improve di/dt-

clock vernier system facilitates speed path debug through the insertion of 45000 clock skew adjustment points. The clock vernier [3] circuit has separate edge controls enabling both duty cycle adjustment and insertion delay modification to improve frequency or robustness without doing a silicon stepping. (Figure 5)

Techniques such as eliminating domino logic in large data-paths, replacing architectural stalls with replays, eliminating glitches, and increasing dynamic clock gating efficiency to over 85% effectively reduced dynamic power in the core by an additional 60% beyond the technology scaling (Fig 5). Post-timing analysis circuit downsizing maximizes the use of lower-leakage devices in the core (>81%) and uncore (76%), reducing overall leakage by 30%.

An improved digital activity sensor includes the ability to monitor the >30% dynamic power consumed in data patterns and reacts to power events in under 1µs. This system monitors 1834 architectural

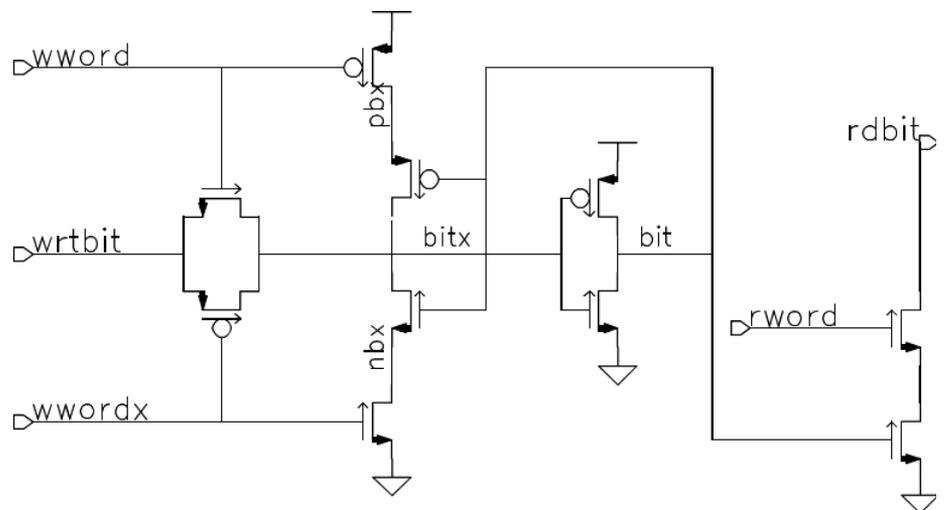


Figure 4. Fully interrupted RF bit cell

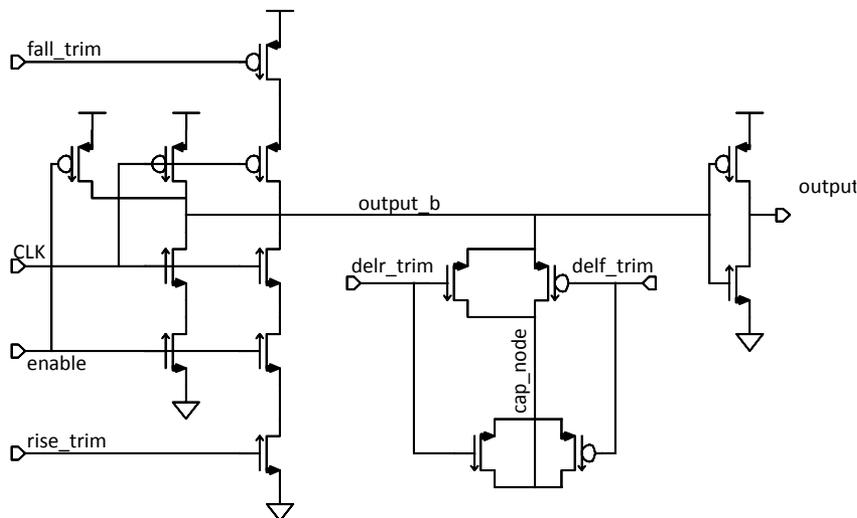


Figure 5. Clock Vernier circuit

several improvements were made to the design. These improvements include:

1. Large cache arrays covered by ECC including the large L3 utilizing DECTED and protecting the MLI/MLD with inline SECTED.
2. Extensive parity protection and parity interleaving on nearly all RFs.
3. End-to-end parity protection with recovery-support on all critical internal buses and data paths including the ring.
4. Residue protection on key logic and execution data paths.
5. The adoption of radiation-hardened (RAD) sequential latching elements for vulnerable architectural and state.

The adoption of these techniques required design methodology and automation enhancements to ensure the RAS goals were achieved without dramatically increasing induced noise.

System Interface Details

The system inter-connect is provided by an integrated 10-port router which connects to external IO and processors via four full-width and two half-wide point-to-point 6.4 GT/s Quickpath™ (QPI) link interfaces. The ring based system interface provides a theoretical peak bandwidth of 700GB/sec (Figure 7). The chip includes two integrated memory controllers each supporting two Scalable Memory Interconnect (SMI) links operating in lockstep. These four SMI ports provide a 6.4 GT/s connection to up to 512 GBs memory per socket. The SMI and QPI links provide Reliability, Availability, and Scalability (RAS) support for features such as clock and lane failover.

In order to enable Mission Critical servers and the associated high RAS requirements

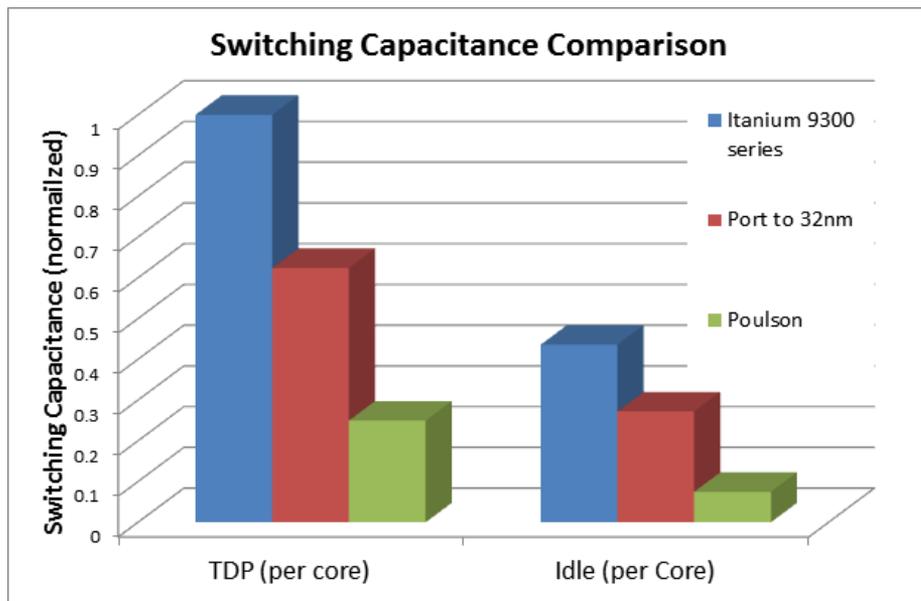


Figure 6. Power reductions in the Poulson Core

die area and power. The random logic synthesis (RLS) flows optimally select hardened sequential usage based upon attack vulnerability and timing criticality. Additionally, a new error micro-architecture combines uniform logging and configuration with key hardware/firmware hooks enabling increased availability and serviceability. These techniques enabled 2X the cores and 1.5X the transistors with a lower overall susceptibility to radiation induced error events as compared to the previous generations.

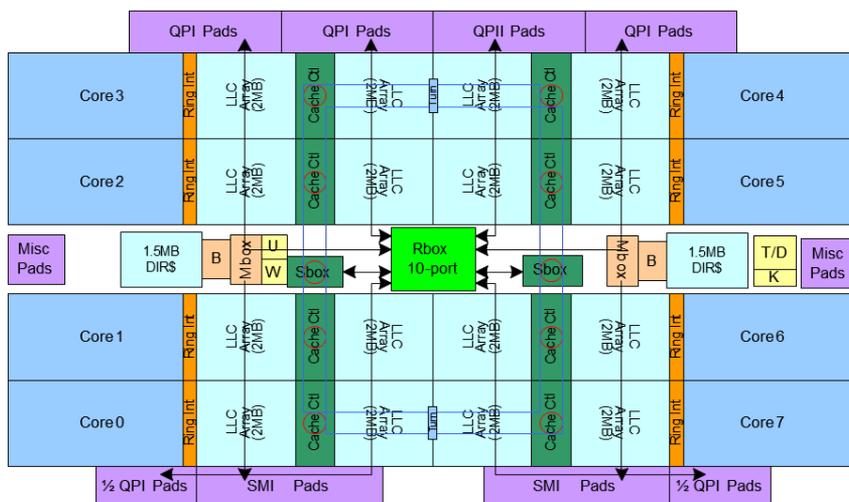


Figure 7. Uncore block diagram

The next generation in the Intel® Itanium® processor family code named Poulson is implemented in 32 nm CMOS, contains 3.1 billion transistors, adds new features to deliver increased performance, has 2X the cores of the previous generation, and adds new features to deliver increased performance. For more information on the Intel® Itanium® processor, visit www.intel.com/go/itanium.

References

- Stackhouse, B.; et al.; "A 65 nm 2-Billion Transistor Quad-Core Itanium® Processor," JSSCC 2009.
- Packan, P; et al.; "High Performance 32nm Logic Technology Featuring 2nd Generation High-k + Metal Gate Transistors," IEDM 2009.
- Mahoney, P; Fetzer, E.; et al., "Clock distribution on a dual-core, multi-threaded Itanium®-family processor," Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International
- Petersen, R.; et al.; "Voltage transient detection and induction for debug and test," Test Conference, 2009. ITC 2009. International , vol., no., pp.1-10, 1-6 Nov. 2009

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Configurations: [describe config + what test used + who did testing]. For more information go to <http://www.intel.com/performance>. Intel does not control or audit the design or implementation of third party benchmarks or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmarks are reported and confirm whether the referenced benchmarks are accurate and reflect performance of systems available for purchase. Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families. See www.intel.com/products/processor_number for details.

Intel Turbo Boost Technology requires a system with Intel® Turbo Boost Technology capability. Consult your PC manufacturer. Performance varies depending on hardware, software and system configuration. For more information, visit <http://www.intel.com/technology/turboboost>. Intel Virtualization Technology requires a computer system with a processor, chipset, BIOS, virtual machine monitor (VMM) and applications enabled for virtualization technology. Functionality, performance or other virtualization technology benefits will vary depending on hardware and software configurations. Virtualization technology-enabled BIOS and VMM applications are currently in development. 64-bit computing on Intel architecture requires a computer system with a processor, chipset, BIOS, operating system, device drivers and applications enabled for Intel® 64 architecture. Performance will vary depending on your hardware and software configurations. Consult with your system vendor for more information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web site at www.intel.com.

Copyright © 2011 Intel Corporation. All rights reserved. Intel, the Intel logo, and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

