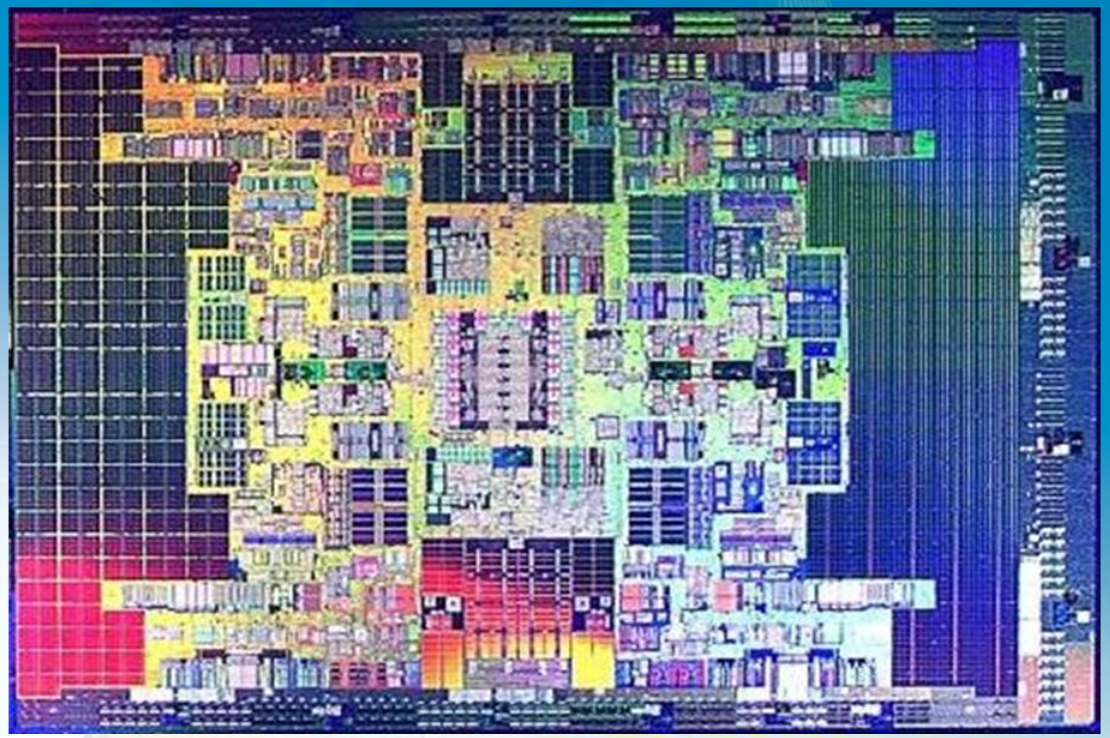
Intel® Itanium® Quad-Core Architecture for the Enterprise

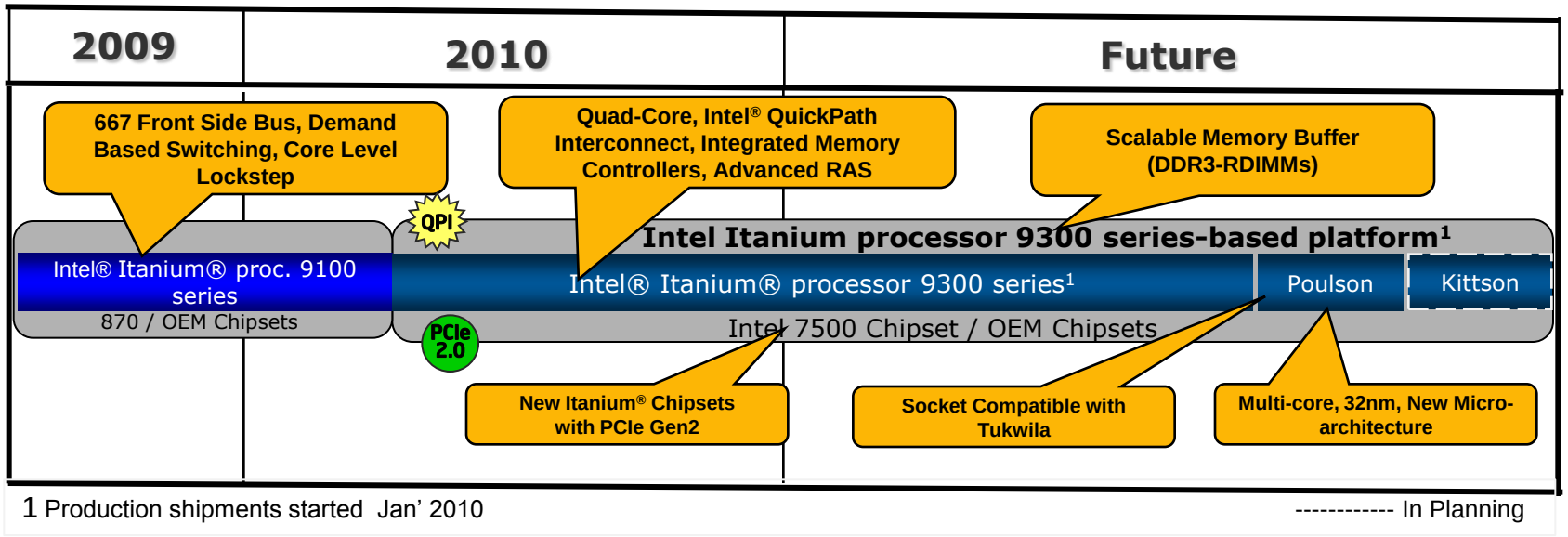
Lambert Schaelicke
Eric DeLano



Agenda

- Introduction
  - Intel® Itanium® Roadmap
  - Intel® Itanium® Processor 9300 Series Overview
- Key Features
  - Pipeline Overview
  - Multithreading
  - Integrated System Interface
  - Performance Monitoring Unit
  - Power and Frequency Management
  - RAS and Manageability
- Questions & Answers

Intel® Itanium® Platform Roadmap



- Multiple product generations on the roadmap
- Powering systems with leading-edge scalability, world-class resiliency, high-end virtualization and intelligent energy efficiency
- Greater choice and flexibility than proprietary RISC solutions

Tukwila Overview

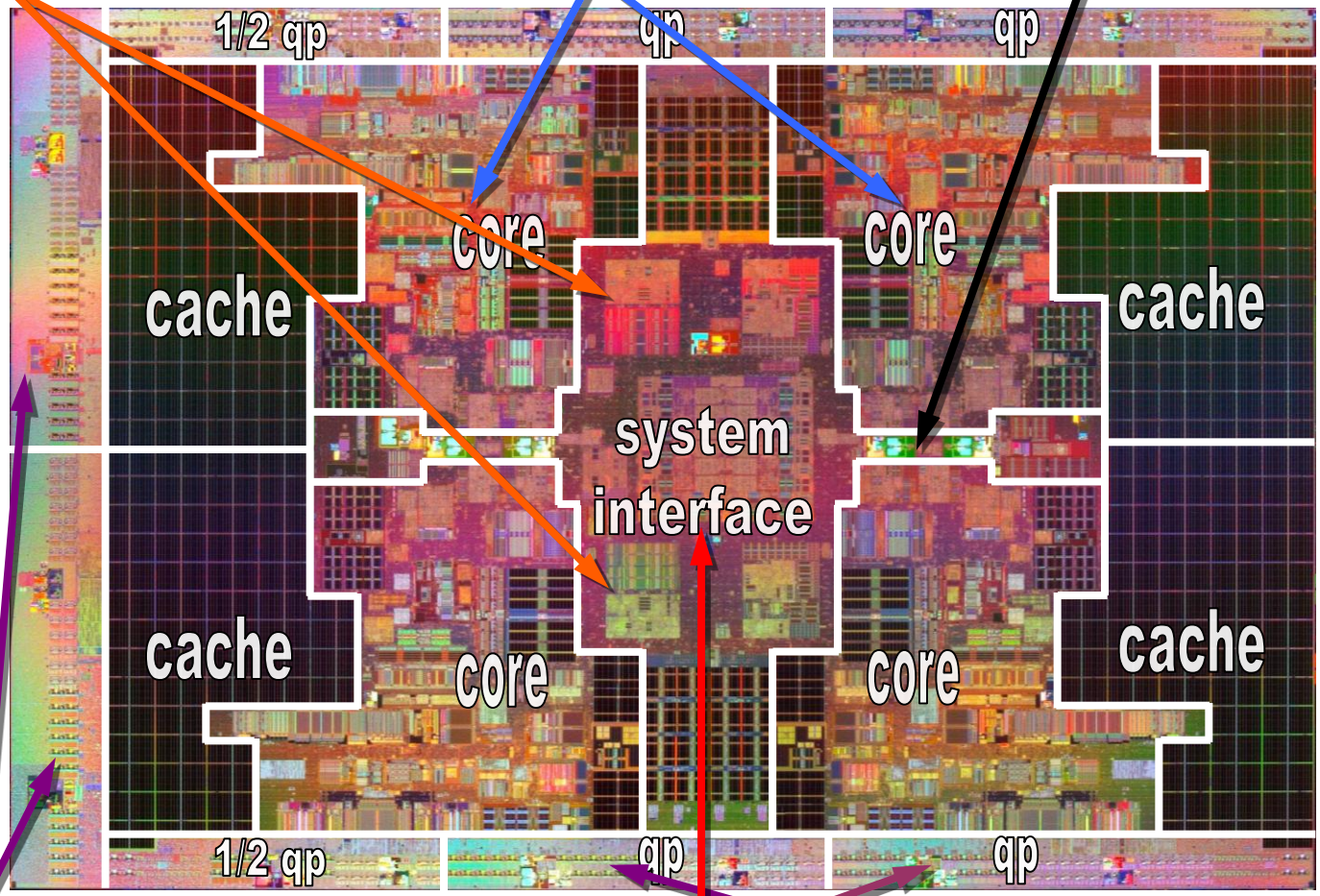
- Performance
  - 4 cores with 2 threads/core
  - High memory and interconnect bandwidth
  - Instruction-level and thread-level parallelism improvements
- System integration and new system interfaces
  - Integrated Memory Controllers and Router
  - Intel QuickPath System Interconnect
  - Intel Scalable Memory Interconnect
- Scalability
  - Directory coherency with ~2MB of directory cache
  - Up to 8 socket glueless (without a Node Controller chip)
  - Higher scalability with OEM chipsets
- Power efficiency and performance balance
  - Dynamic Voltage/Frequency management
- Reliability, Availability, Serviceability, Manageability

Die Photo

Dual Integrated Memory Controllers

Quad Multi-Threaded Cores

Power/Thermal/Freq Management

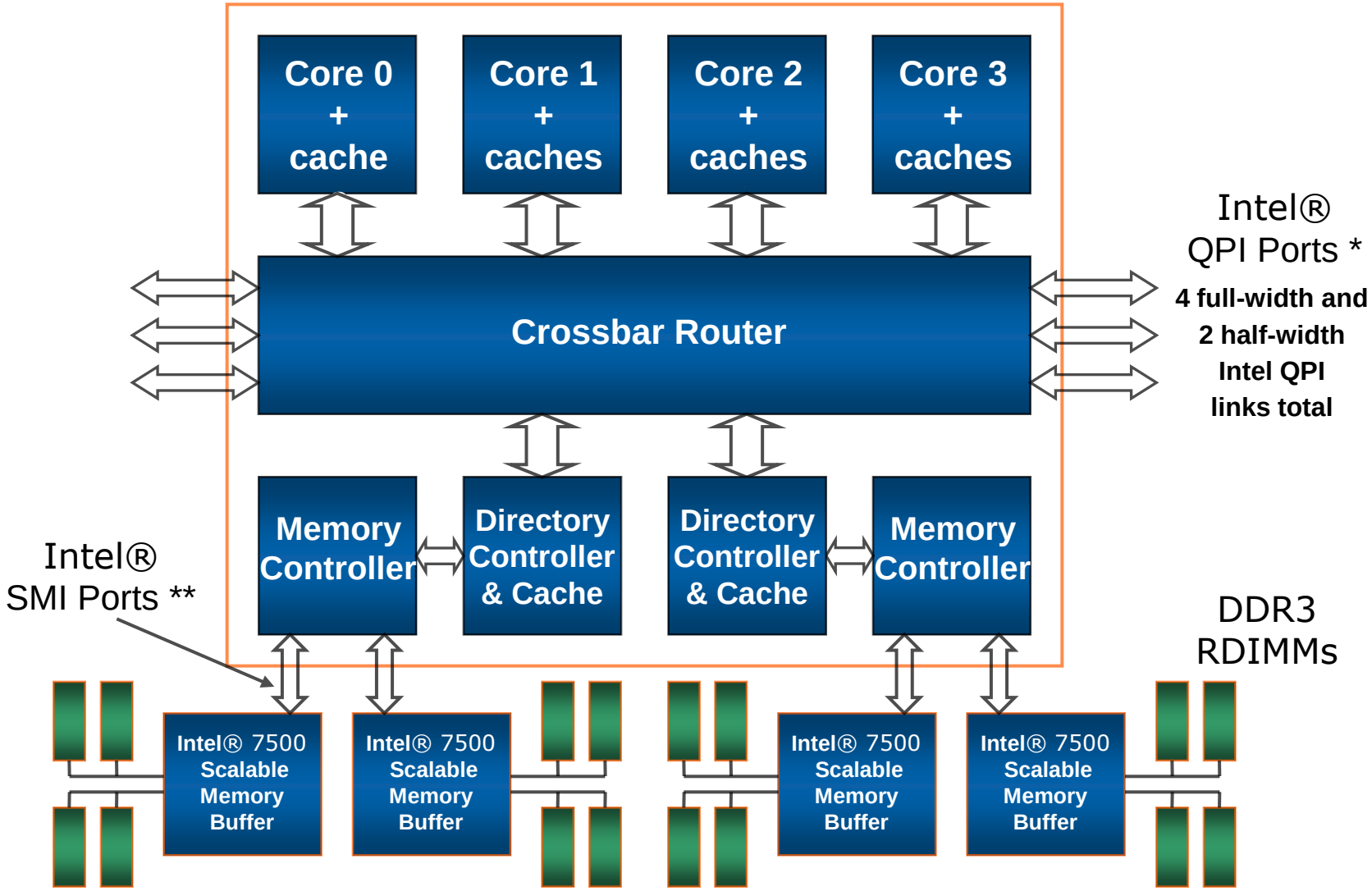


Memory Interface

CrossBar Router

Intel QuickPath Interconnect

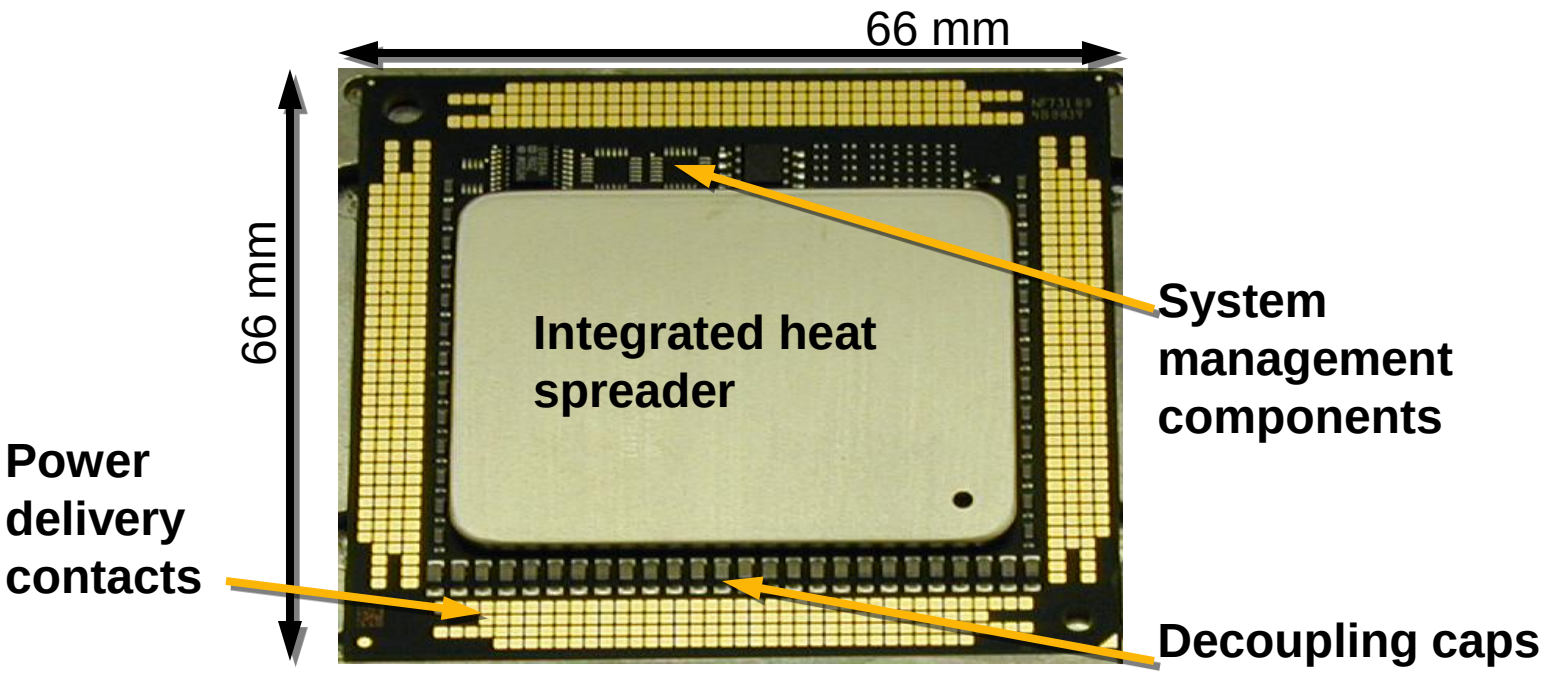
Block Diagram



* Intel® QPI = Intel® QuickPath Interconnect

** Intel® SMI = Intel® Scalable Memory Interconnect

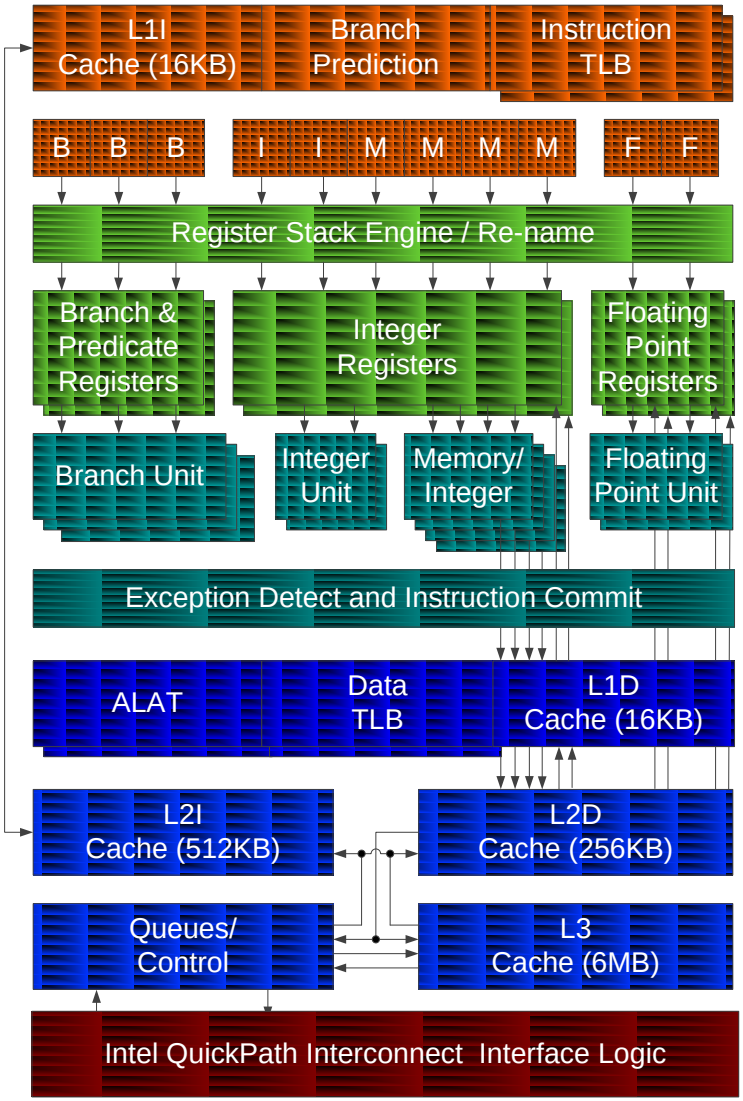
Tukwila Die and Package details



- 65nm bulk CMOS, 8 layer Cu interconnect
- 2.05 billion transistors
- 32.5mm X 21.5mm = 700 mm2 die size
- 811 signal package pins in a 1248 pin LGA socket
- 5 major voltage and frequency domains

Execution Cores

- Same high ILP* core as on the 9000/9100 series processor
  - 6 wide instruction fetch, issue & execution
    - 6 integer units, 2 FP units, 4 memory units, 3 branch units
    - 1 cycle L1 data cache
    - Separate L1I, L1D, L2I, and L2D caches
  - Short 8 stage pipeline – high performance and energy efficiency
  - Predication
  - data prefetching and speculation
- Increased core frequency
- Improved memory hierarchy
  - Increased memory bandwidth
  - 4k/8k/16k page size support in first level data TLB



* ILP = instruction-level parallelism

Cache Hierarchy

	L1 Inst Cache per core	L1 Data Cache per core	L2 Inst Cache per core	L2 Data Cache per core	L3 Cache per core
Size	16 KB	16 KB	512 KB	256 KB	6 MB
Line Size	64-byte	64-byte	128-byte	128-byte	128-byte
Ways	4-way	4-way	8-way	8-way	12-way
Write Policy	-	WT	-	WB	WB
Latency	1 clk	1 clk	7 clks	5,7,9 clks[1]	15+ clks[1]
Protection	SEC[2] via Parity	SEC[2] via Parity	SEC[2] via Parity	SEC[2] via ECC	SEC[2] via ECC

[1] **Add 1 clk for FP loads**

[2] **SEC: Single Bit Error Correction**

Note: Higher level caches do not force inclusion

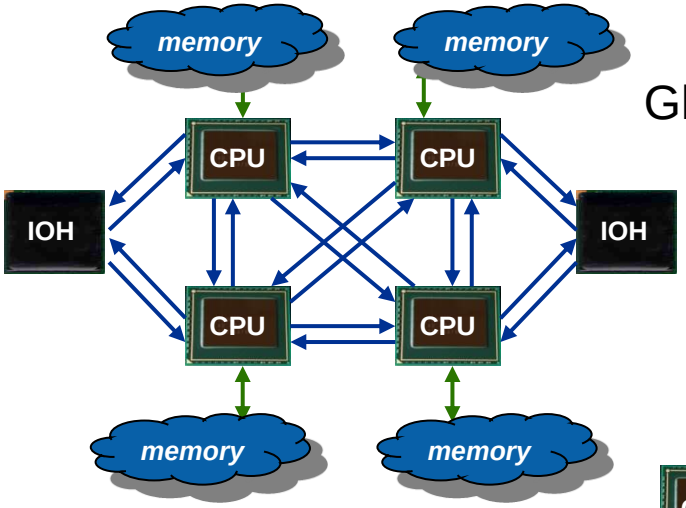
Switch-on-Event Multithreading

- Switch to alternate thread on specific events
  - Last-level cache demand load miss
  - Data return to last-level cache for inactive thread
  - Explicit thread switch hint instruction
  - Timeout - ensure fairness between thread
- New switch events
  - Semaphore release
  - Extended stalls of execution pipeline
    - Secondary load miss after data prefetch
    - Allows aggressive use of software data prefetching while retaining multithreading benefit
- Thread Switching Controlled by Urgency Logic
  - Tracks relative importance of both threads
  - Switch events modify urgency values

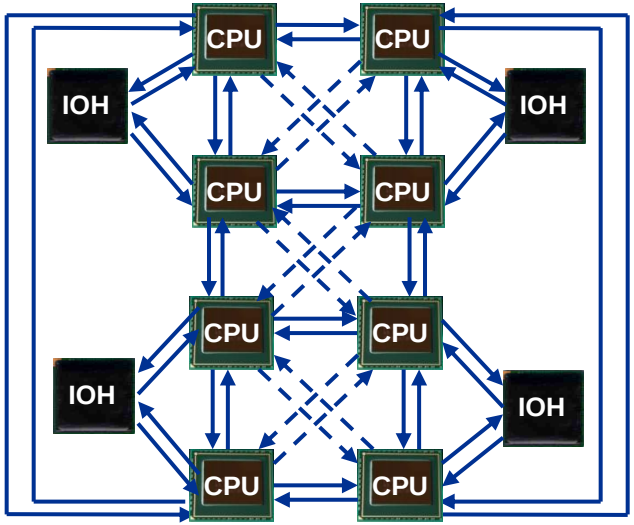
Integrated System Interface

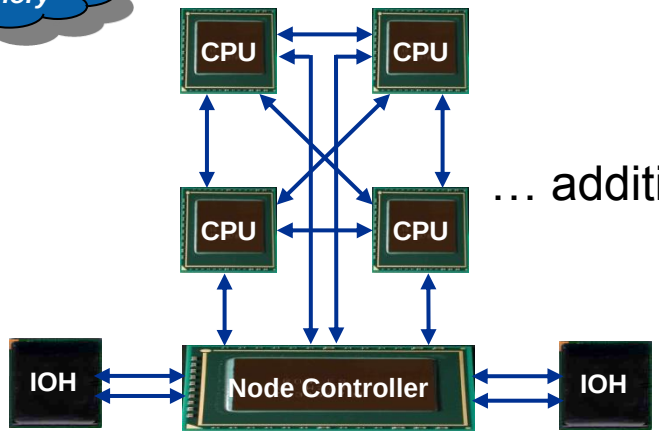
- On-chip Router
  - 12 ports
    - 4 cores, 2 home agents, 4 full-width and 2 half-width external links
- 6 Point-to-point Intel QPI links
    - Higher bandwidth and better scaling than shared buses
- Supports glueless 2 to 8 socket systems
  - One-hop remote memory latency only 40% higher than local memory latency
  - Larger system supported via OEM-provided node controllers
- Directory coherency for efficient scaling
  - Less snoop traffic than broadcast-snoop protocols
- Integrated memory controller
  - Drives external memory buffers to support large memory with high bandwidth

Example Topologies



Glueless systems



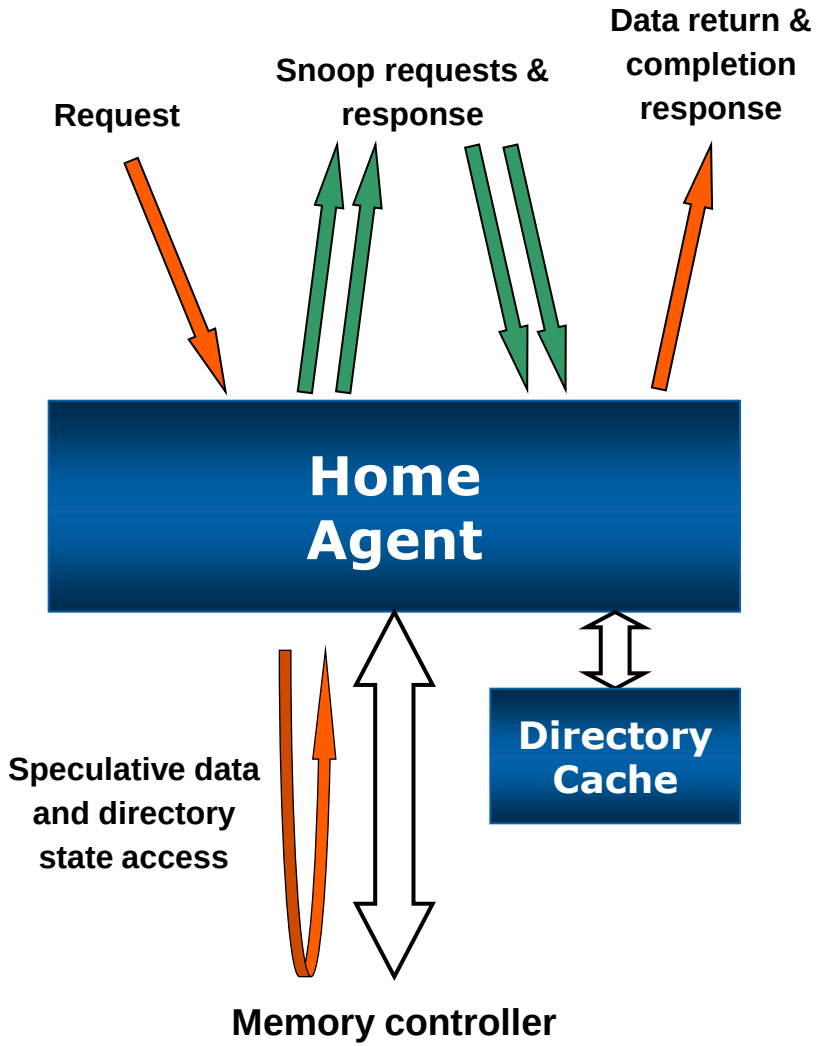


… additional cells

Node controller system

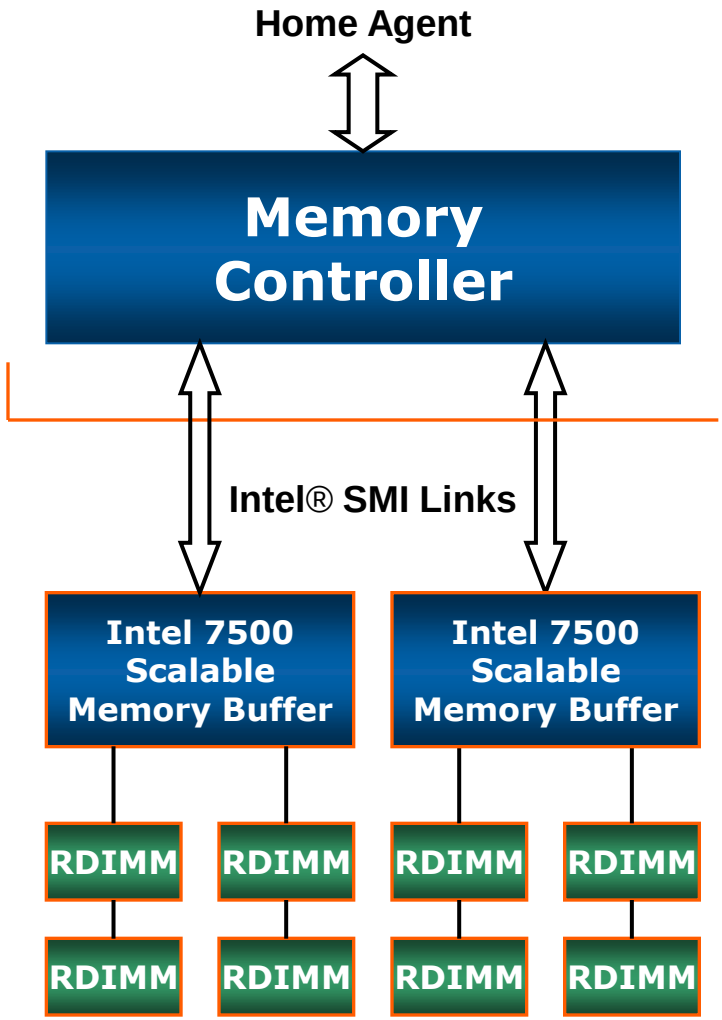
Directory Coherency

- Improved Scalability
  - Snoop only cache line owners and sharers
  - Snoop traffic is dependent only on workload characteristics
    - Does not scale with the square of system size as in broadcast-snoopy coherency
    - Serial workloads execute with minimal snoop overhead even on large systems
- Home Agent
  - Two home agents per processor
  - Receive all requests & issue snoops based on directory state
  - Receive snoop responses & generate final directory and MESI state
  - Data return from home agent or other caches
- Directory Cache
  - Caches directory state to reduce snoop latency
  - Snoops overlapped with memory access rather than in series



Integrated Memory Controller

- Memory Controller
  - Associated with home agent
    - Two memory controllers per processor
  - Each drives two Intel Scalable Memory Interface (Intel SMI) channels in lock-step
    - Higher bandwidth and better error correction capability than single channel
- External Memory Buffer
  - Converts point-to-point SMI channel into two independent DDR-3 busses
  - Improved fan-out for greater memory capacity
    - RDIMMs can be located further away from processor for easier board layout and routing
  - Maintains high bandwidth for large memory capacities
    - Electrical load does not slow down high-speed channels



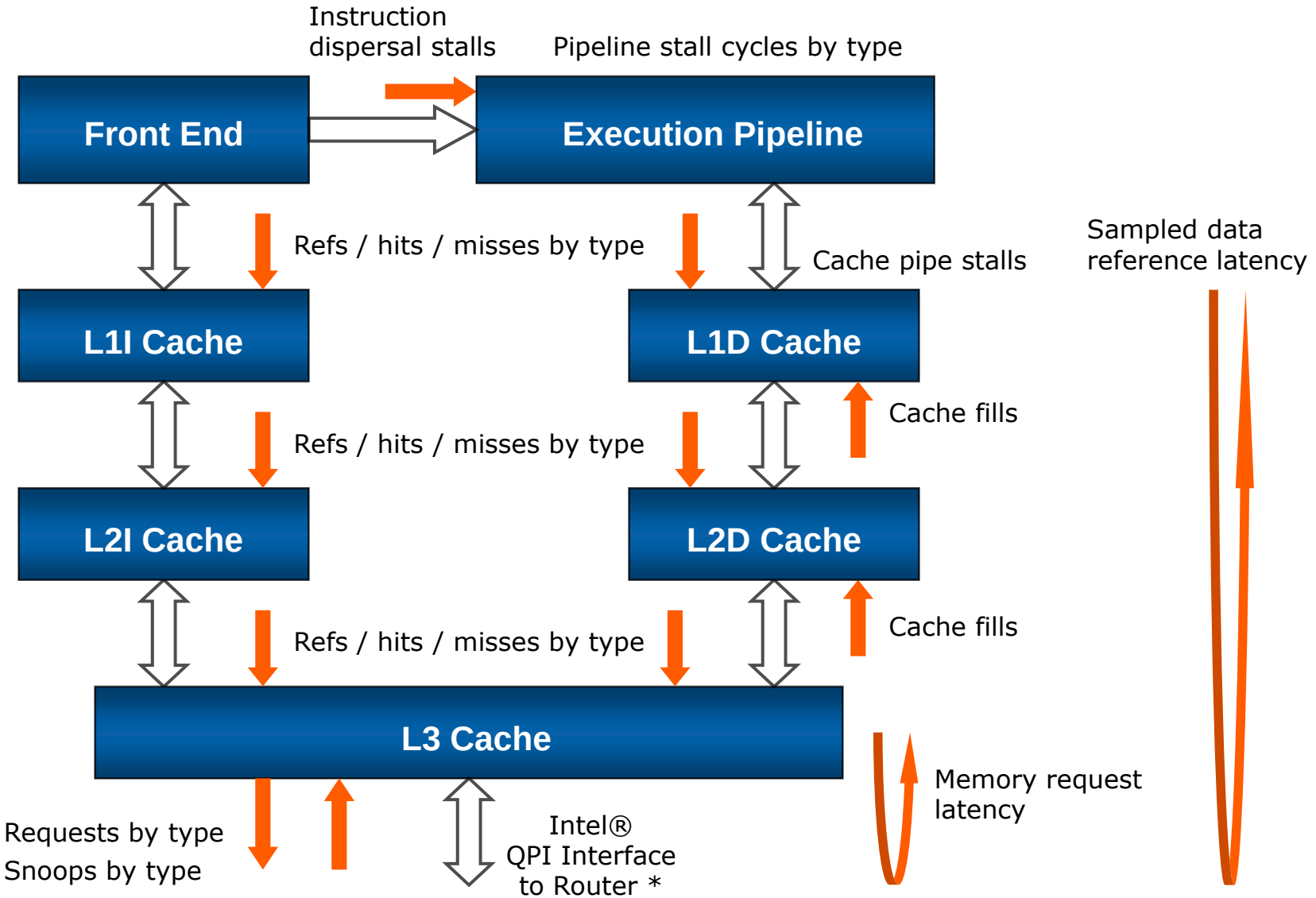
Performance Monitoring Unit

- Core PMU
  - Leveraged from previous Itanium designs
    - Pipeline metrics: instruction issue, stalls, thread switches etc.
    - Cache & TLB* metrics: hits, misses, various conflicts
    - Intel QPI Interface: request breakdown, latency
  - Reports CPI** breakdown & memory hierarchy behavior
- System Interface PMU
  - Metrics for each major block
    - Router, directory controller, memory controller
    - Transaction breakdown
    - Transaction latency
    - Queuing delay
  - Allows constructing detailed latency breakdown

•TLB = translation look-aside buffer

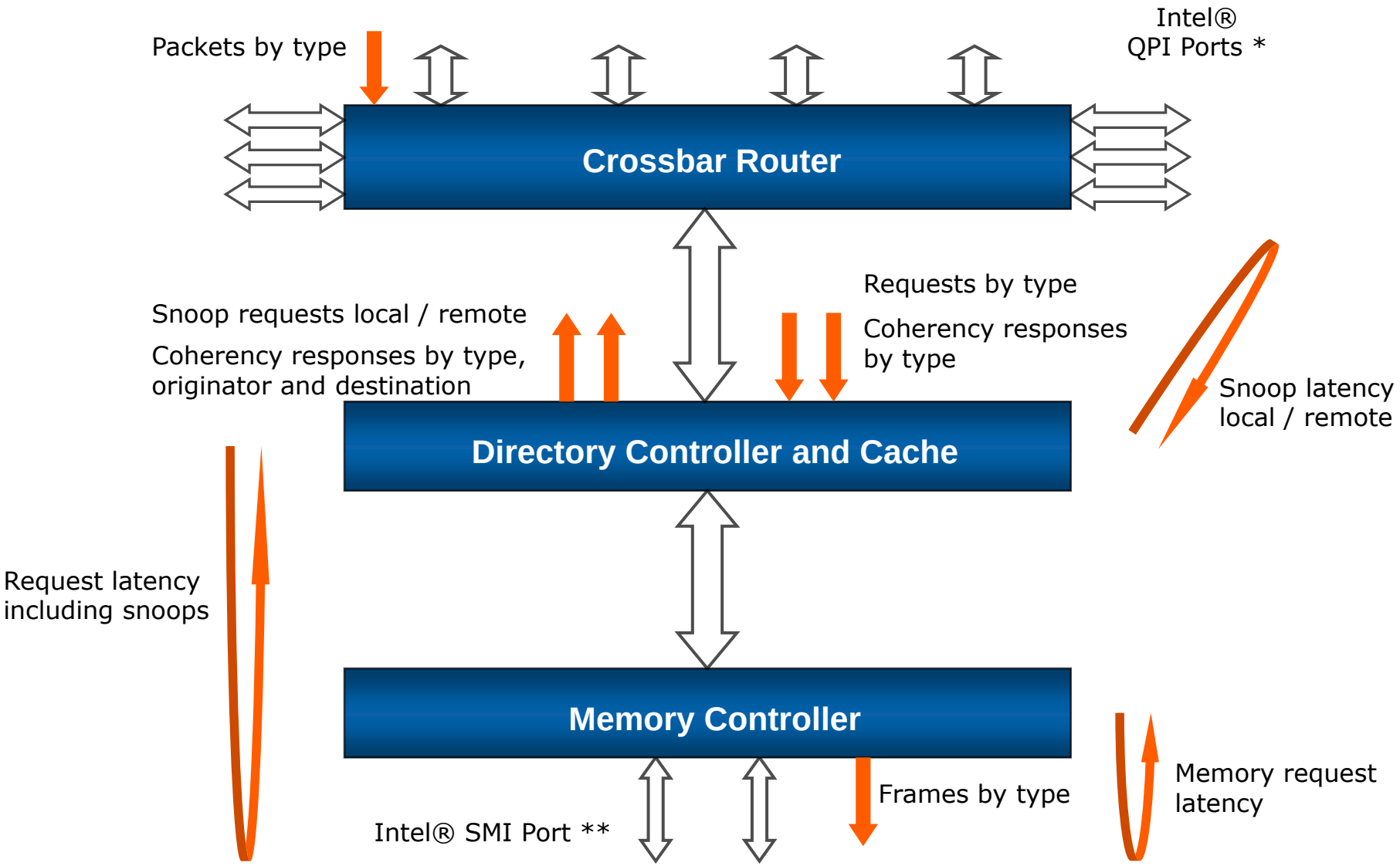
** CPI = cycles per instruction

Core PMU



* Intel® QPI = Intel® QuickPath Interconnect

System Interface PMU



* Intel® QPI = Intel® QuickPath Interconnect

** Intel® SMI = Intel® Scalable Memory Interconnect

Dynamic Power Management

- Dynamically adjusts voltage and frequency
  - Maximize performance for a given power and thermal envelope
  - Different applications may run at different frequencies
  - Adjustments are transparent to the OS
  - Digitally estimated power leads to deterministic voltage/frequency changes
- Balances power across all cores

Processor RAS

- All major structures are protected
  - Extensive ECC and parity protection on core and system interface structures and datapaths
- Extensive use of Soft Error hardened (SE-hardened) latches and registers
  - SE-hardened latches/registers are 100x/80x (respectively) less susceptible to soft errors due to alpha particles
- Intel® Cache Safe Technology added to L2I, L2D, and Directory Cache (in addition to L3)
  - Cache lines that show too many errors in the field are mapped out by the processor

Memory RAS Features

- Memory Channel Protection
  - Retry, channel reset & retraining, Intel SMI lane failover
- Patrol and demand scrubbing
  - Prevent uncorrectable errors by writing good data/ECC when encountering a correctable error
    - In background (patrol), or due to a request (demand)
- DRAM Protection – Double Device Data Correction (DDDC)
  - Tukwila can fix both single and double memory hard-errors and still correct an additional single bit error
- Memory channel Hot-Plug Support
  - Replace faulty component without bringing down the system
- Memory Migration and DIMM sparing

System Interconnect RAS

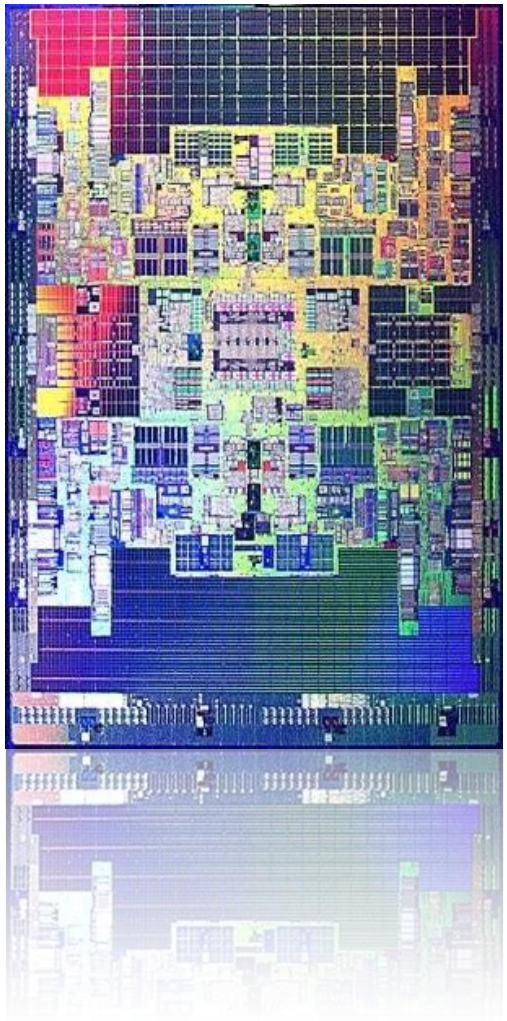
- Intel QuickPath Interconnect protection
  - Retry, reset, retraining
  - Intel QPI lane fail-over
- Hot-Plug Support
- Timeout mechanisms
  - Aids detection of faulty components
- Error Containment mechanisms
  - Error signaling indicates scope of error

Summary

- Improved performance
  - Better thread-level and instruction-level parallelism
  - Higher memory bandwidth
- Improved scalability and platform flexibility
  - Integrated system interface
  - Directory-based coherence protocol
- Power and frequency management capabilities
- Mission Critical RAS: Processor – memory – interconnect

- Launched February 8, 2010
- Multiple OEMs working on Tukwila-based systems
- Performance benchmark results becoming available now

Q&A



Legal Disclaimer

- 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. INTEL PRODUCTS ARE NOT INTENDED FOR USE IN MEDICAL, LIFE SAVING, OR LIFE SUSTAINING APPLICATIONS.
- Intel may make changes to specifications and product descriptions at any time, without notice.
- All products, dates, and figures specified are preliminary based on current expectations, and are subject to change without notice.
- Intel, processors, chipsets, and desktop boards 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.
- Tukwila, Poulson, Kittson and other code names featured are used internally within Intel to identify products that are in development and not yet publicly announced for release. Customers, licensees and other third parties are not authorized by Intel to use code names in advertising, promotion or marketing of any product or services and any such use of Intel's internal code names is at the sole risk of the user
- Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.
- Intel, Intel Inside, Intel Itanium, Intel QuickPath Interconnect, Intel Cache Safe, IA-64, FBD2, VT-i, SDDC, DDDC and the Intel logo are trademarks of Intel Corporation in the United States and other countries.
- *Other names and brands may be claimed as the property of others.
-


intel
Leap ahead™