Poulson: An 8 Core 32 nm Next Generation Intel® Itanium® Processor

Steve Undy

Intel®

This slide MUST be used with any slides removed from this presentation

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.
- 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, 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.
- The code names "Poulson" and "Tukwila" presented in this document are only for use by Intel to identify products, technologies, or services in development, that have not been made commercially available to the public, i.e., announced, launched or shipped. They are not "commercial" names for products or services and are not intended to function as trademarks.
- Intel, Intel Itanium, Itanium, Intel Xeon, Intel Core microarchitecture, and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Agenda

Design Space

Overview

Parallelism

Data Integrity

Power

Conclusion

Design Space: Mission Critical

Performance
- Both single-thread speed and total throughput are important
- Goal: > 2x generational performance improvement

Scalability
- Glueless up to 8 sockets
- Support larger topologies via node controllers from system vendors

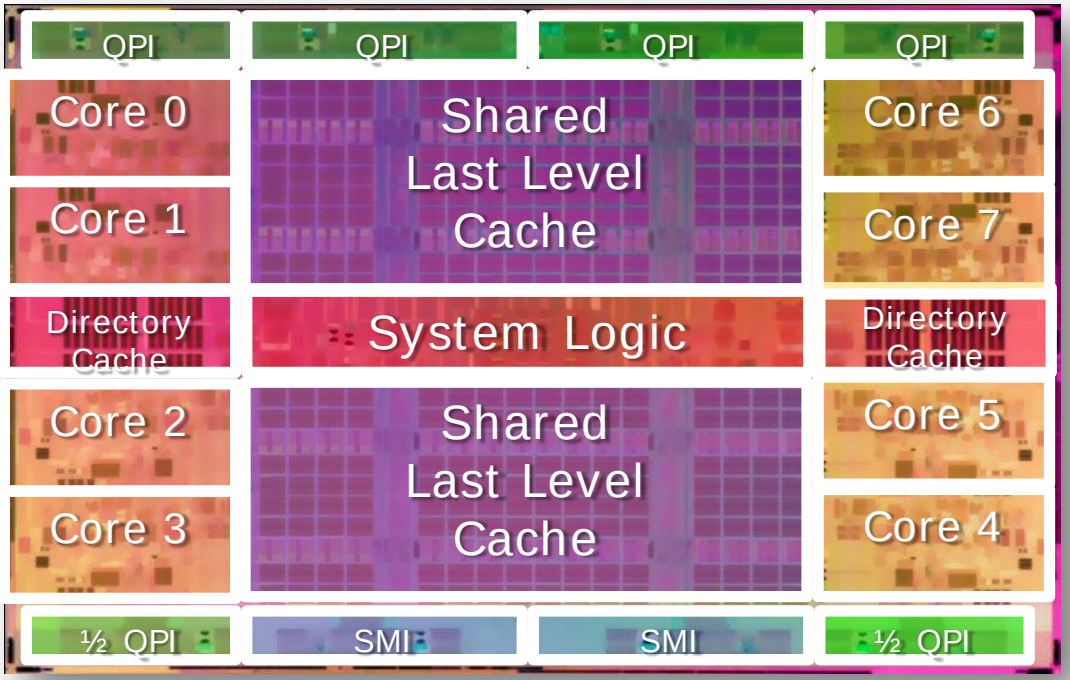
Reliability
- Redundancy and self-correction
- Error detection and recovery via hardware and firmware

Compatibility
- Socket compatible with Itanium® 9300 Processor Series (Tukwila)
- Common platform ingredients with Xeon® E7 Processor Series

Poulson Overview

	Tukwila	Poulson
Process	65nm	32nm
Devices	2.05 Billion	3.1 Billion
Area	700 mm^2	544 mm^2
Power (max TDP)	185W	170W
Itanium® Cores	4	8
Last Level Cache Size	24MB	32MB
Intel® QuickPath Interconnect Links	4 full + 2 half	4 full + 2 half
Intel® QPI Link Speed	4.8GT/s	6.4GT/s
Intel® Scalable Memory Interface Links	4	4
Intel® SMI Link Speed	4.8GT/s	6.4GT/s

Poulson Overview



8 Itanium® Cores

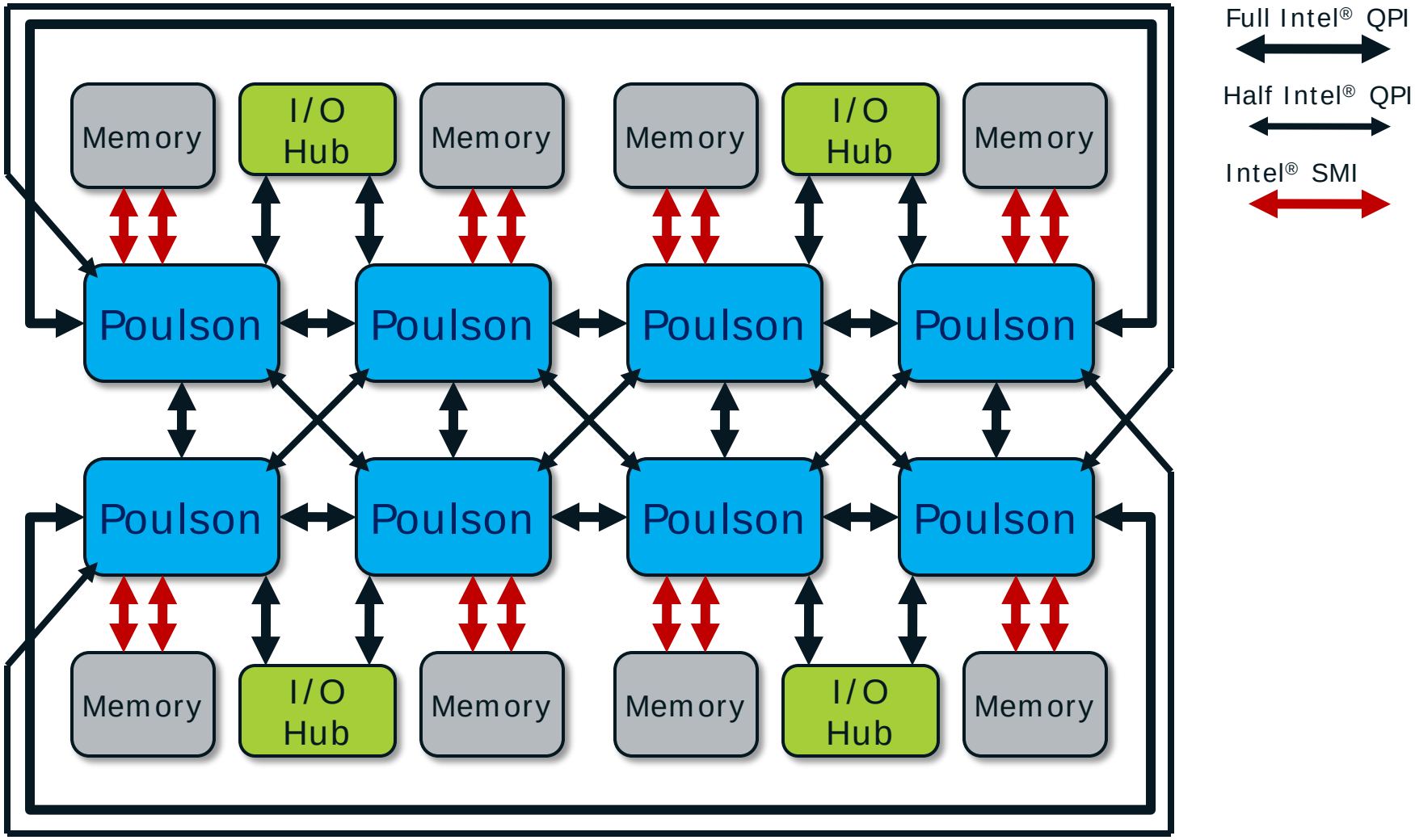
32MB Last Level Cache

4 full-width and 2 half-width Intel® QPI links

2 Directory caches

2 Integrated memory controllers with 2 Intel® SMI links each

Example 8-Socket Topology



Poulson Core

	Tukwila	Poulson
Devices	108 million	89 million
Area	69 mm^2	20 mm^2
Process scaled power @ TDP	1.0	0.4
Process scaled frequency	1.0	>1.2
Threads	2	2+2
Instruction Q size	48	96x2
Max instruction issue/cycle	6	12
Pipeline stages	2 + 6	4 + 7
Pipeline hazard resolution	Interlock+Stall	Replay
Integer RF size	144	185
Integer RF ports	12R 8W	12R 12W
RF protection	Parity	SECDED

All-new Itanium® core

Itanium® New Instructions

Integer operations

- mpy4
- mpyshl4
- clz

Thread Control

- hint@priority

Expanded Data Access Hints

- mov dahr

Expanded Software Prefetch

- lfetch.count

Exploiting Parallelism on All Levels

Instruction Level Parallelism

Memory Parallelism

Thread Parallelism

Multi-core Parallelism

Multi-level parallelism exposed to software control

Instruction Parallelism

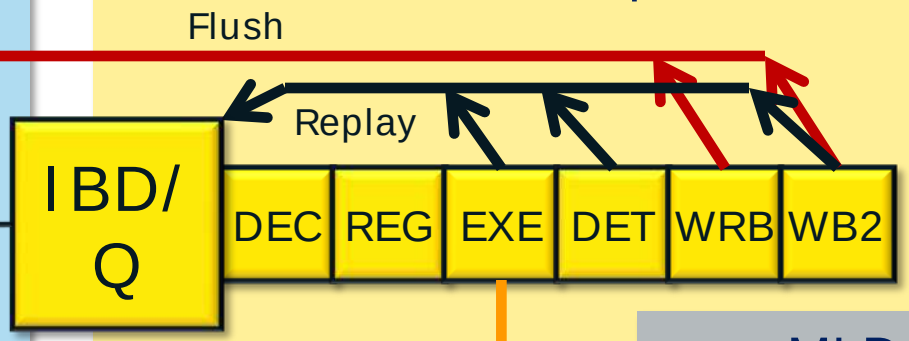
Front-end Pipeline

IPG: Address Gen
FET: Array Access
FDC: Early Decode
REN: Register Rename

Fetch width: 32B/cycle
Multi-level branch Prediction
Renaming: 128 logical to 185 physical registers

Independent thread domain

Main Pipeline

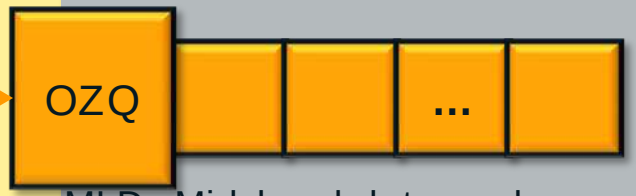


IBD: Instr Buffer/Dispersal
DEC: Instr Decode
REG: Register Read
EXE: Execute
DET: Exception Detect
WRB: Write-back
WB2: Commit

Q holds 96 instructions
In-order issue

Independent thread domain

MLD Pipeline



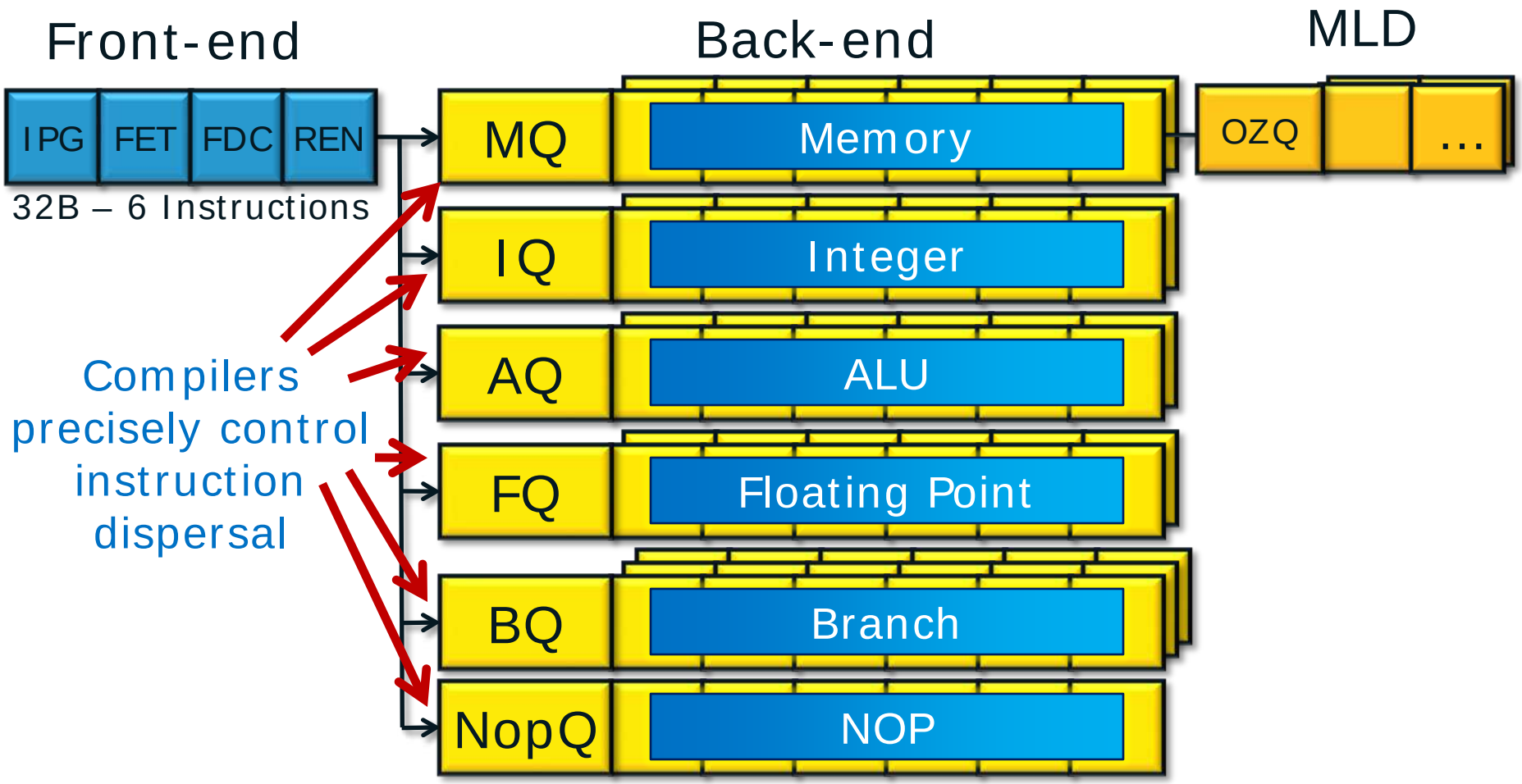
MLD: Mid-level data cache
OZQ: Architectural memory ordering queue

OZQ holds 16 memory ops
OOO issue

Independent thread domain

Concurrency via 3 decoupled pipelines

Instruction Parallelism



12-wide issue under software control

Memory Parallelism – Avoiding Hazards

Expanded Software Hints

- Provides control over allocation, speculation and prefetch policies
- Multi-line software prefetch

Reduced Speculation Costs

- Spontaneous Deferrals on speculative loads
- Deferred load can transform into a prefetch to cache and/or TLB

New Hardware Prefetcher

- Into FLD and/or MLD
- Adaptive based on FLD/MLD miss patterns

Reduced Data Hazards

- Expanded store to consumer bypassing in FLD and MLD
- Single-cycle MLD to FLD line transfers

Pipeline bottlenecks removed

Memory Parallelism – Increasing Throughput

More pending memory operations in each core

- Increased from 48 to 64 operations in MLD
- Increased from 44 to 80 operations in Ring interface

Ring Cache

- 8 Independent LLC pipelines and controllers per socket
- 2 caching agents per socket to generate Intel® QPI transactions

Increased Intel® QPI link bandwidth for multi-socket throughput

Memory Subsystem Improvements

- Home agent in-flight transaction tracker increased by 50%
- MC scheduler algorithm changes focused on performance and power efficiency
- Intel® SMI link speed increased

Improved queuing and B/ W

Thread Parallelism

Intel® Hyper-Threading Technology with Dual Domain Multithreading support

- Front-end and main pipelines are independently threaded
- Pipeline-specific thread switch mechanisms

More per thread resources

- Instruction queues
- Data TLBs
- Hardware page walker handles concurrent walks to both threads

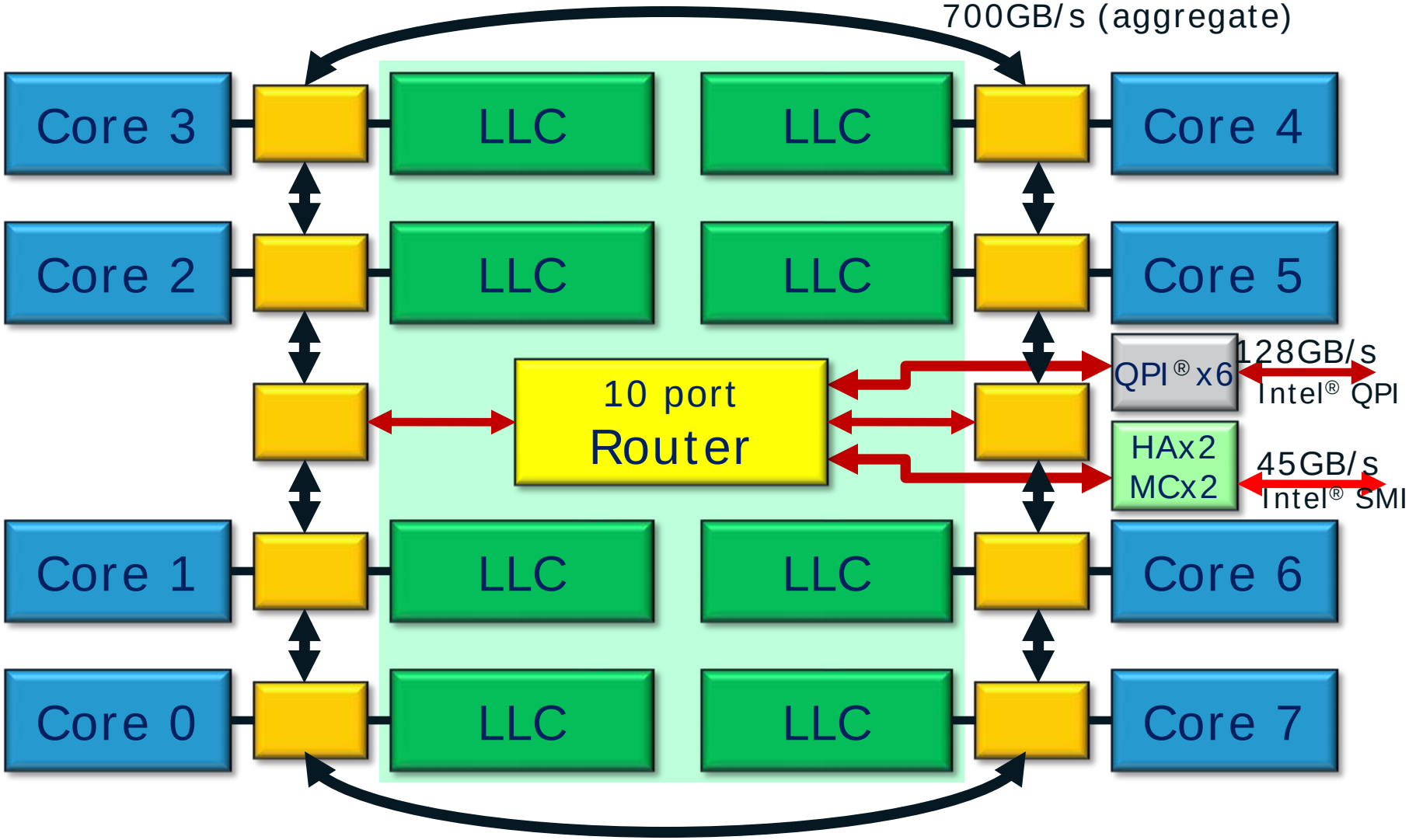
Instructions provide software hints to thread switch hardware

Attention to Thread Fairness

- Hardware to identify unfair behavior
  - Measures instruction retirement and data cache resource usage
- Firmware configurable responses to unfairness
  - By biasing towards victim thread

Increased throughput via improved threading

Multi-core and Multi-socket Parallelism



Reliability, Fault Tolerance and Recoverability

Intel® Instruction Replay Technology

- Main pipeline redesigned for error handling
- Enabled hardware and firmware recovery of correctable errors

Increased Error Detection and Correction

- Arrays
  - MLI, MLD, LLC tag and Directory cache – SECDED
  - LLC data – DECTED
  - Register files (GR and FR) – SECDED
- Logic path error detection in FP ALU via residues
- Key paths are protected “end to end” with invalidation and replay

Reworked Error Detection, Response and Reporting Framework

- Large increase in error detection and response capabilities
- Increased hardware responses and recovery
- Intel® Cache Safe Technology: lockout of failing cache locations

Power Aware Design

Replay-based Pipeline

- Enabled fully gated clocks
- Eliminated slow and power-wasting global pipeline stalls

Aggressive Clock Gating

- Substantial reductions in both idle, typical and max power

Removed Dynamic Logic from FPU

Low-Leakage FETs

Digital Power Prediction

- Measures, weights and averages key core state signals
- Beyond architectural state – uses data patterns
- Estimation error decreased from 35% to 2%

60% TDP, 70% idle C_{dyn} reduction over process scaled Tukwila

Poulson Status

Well into post-Si validation

Booted and being tested on multiple operating systems

Running in many topologies

On track for 2012 shipments

Conclusion

All new core design

- 12-wide issue
- Intel® Hyper-Threading Technology with Dual Domain Multithreading
- Itanium® New Instructions and policies for fine-grain software control of hardware parallelism
- Intel® Instruction Replay Technology for frequency, power and fault tolerance
- Core power efficiency: > 3x better than Tukwila[1]

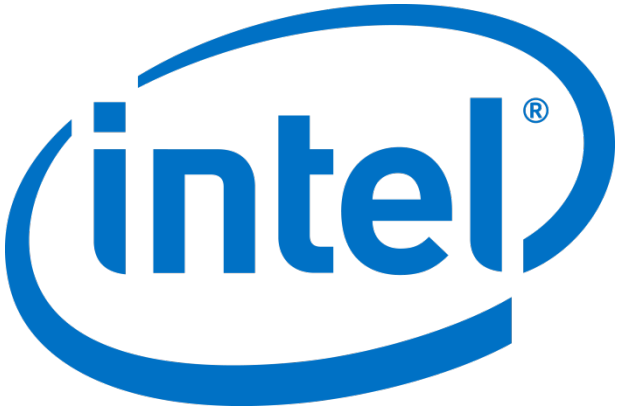
Ring-based design

- Large last level caches
- High bandwidth system interface

Outstanding data reliability

- Enhanced error detection
- Improved error recovery
- Better RAS integration

1 – Internal lab measurement


intel

Glossary

DECTED	Double Error Correction, Triple Error Detection
DTB	Second Level Data TLB
FIT	Failure in Thousands
FLB	First Level Branch Cache
FLD	First Level Data Cache
FLI	First Level Instruction Cache
HA	Home Agent: Intel QPI agent responsible for managing memory
ILP	Instruction Level Parallelism
IPF	Itanium® Processor Family
LLC	Last Level Cache
MC	Memory Controller

MLD	Mid Level Data Cache
MLI	Mid Level Instruction Cache
OZQ	Ordering cZar Queue: architectural memory ordering
QPI	Intel® QuickPath Interface
RAS	Reliability, Availability and Serviceability
RF	Register File
SECDED	Single Error Correction, Double Error Detection
SMI	Intel® Scalable Memory Interconnect
TDP	Thermal Design Power
TLB	Translation Look-aside Buffer