# MULTI-CORE ARCHITECTURE CHALLENGES

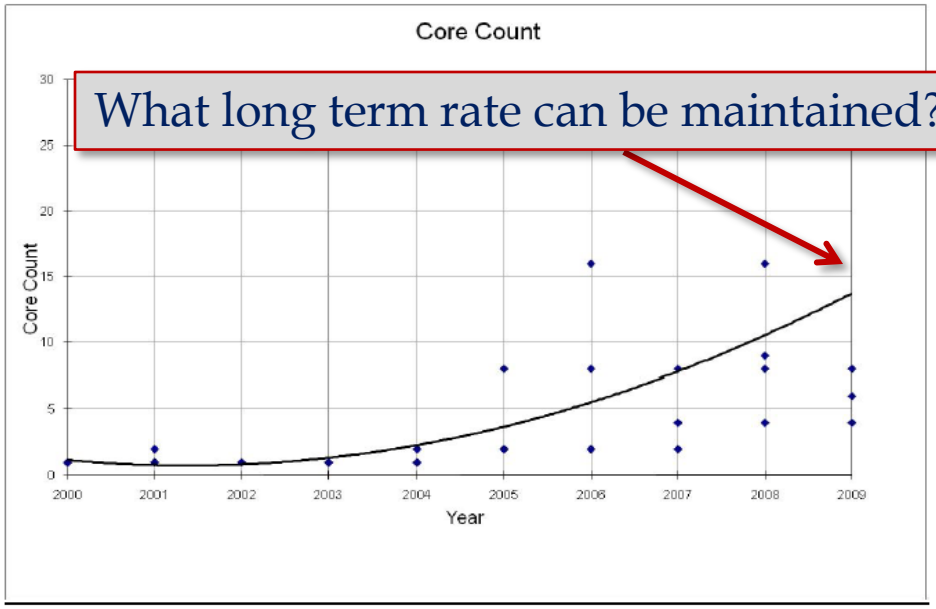
Don Soltis
Sr. Principal Engineer
Intel
Fort Collins

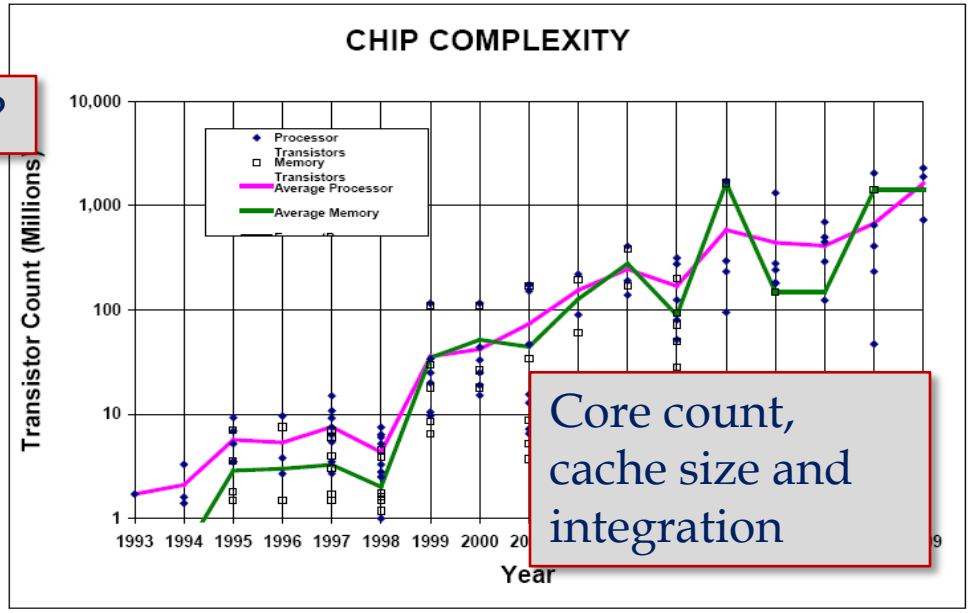
# Intro

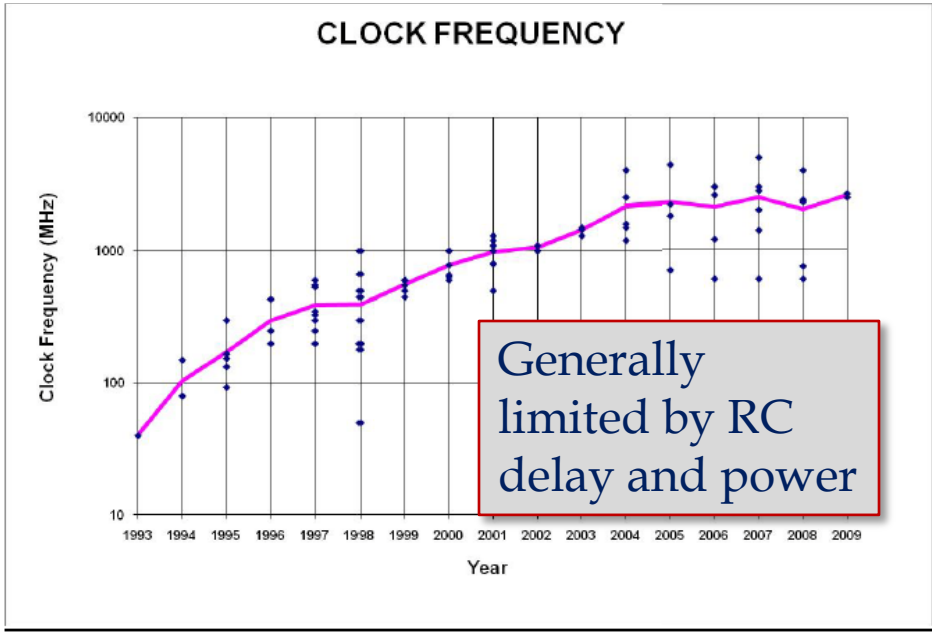
- My background
- Servers perspective
- Start with some current processor trends
- Look in more detail at multi-core and system integration
- Examine hardware multi-core and integration limiters
- Consider hardware and software innovation areas

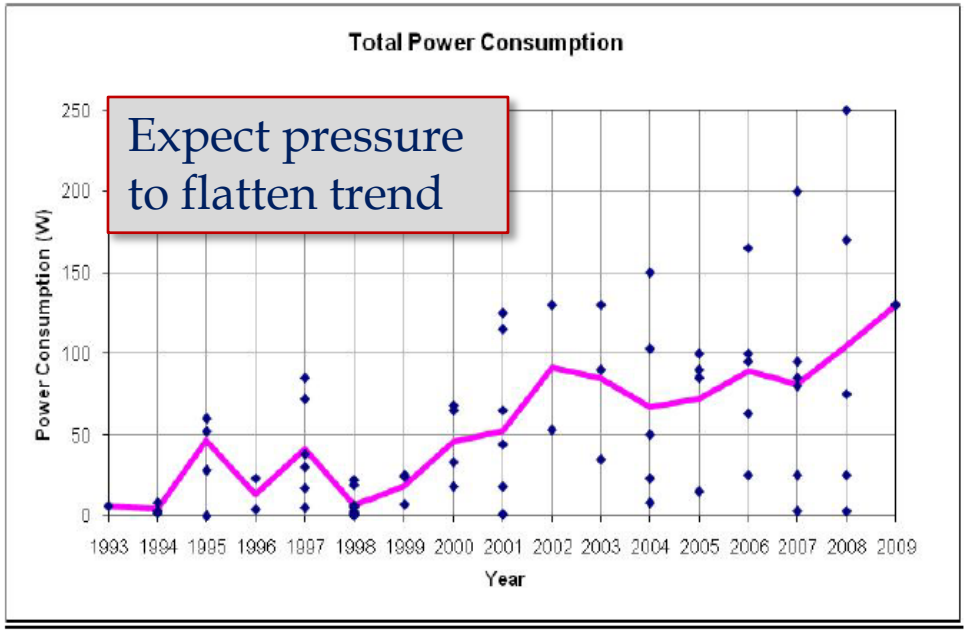
# Microprocessor Trends





# PERCS Hardware (NCSA system)

PERCS – IBM's Programmable Easy-to-use Reliable Computing System

Selected for Phase II of HPCS – DARPA's High Productivity computing system

Power7 technology

| | |
|---|---|
| System Peak | 10.06 PF<br>38,912 8-way 4.04 GHz POWER7 chips; 45 nm technology |
| HPCC HPL | 8.2 PF (estimate) |
| Min/Max Number of OS Images | 4,864 (64 way) to 38,912 (8 way) Linux or AIX OS images |
| FLOPs/Core, FLOPs/Chip, FLOPS/Socket, FLOPS/Supernode | 32.3 GF per core, 258.6 GF per chip, 517.1 GF per socket, 331 TF/supernode |
| Threads/Core | 4-way SMT |
| Total Cache Memory | 1.3 TB |
| Total System Main Memory | 623 TB, IBM Pulsar buffered DIMMS |
| Total Main Memory Available to Users | 556 TB (38,912 SMPs), 574 TB (4,864 SMPs) |
| Total Memory Bandwidth | 5.0 PB/s (B/F=0.5; L1: B/F=6; L3: B/F=3) |
| HPCC STREAM | 3.10 PB/s (estimate) |
| Peak Interconnect Bandwidth | 1.37 PB/s |
| Disk Storage | 26.3 PB raw, 23.3 usable (not including RAID6+ with spares) |
| Archival Storage | Up to 1 EB |
| Total Storage Bandwidth | 4.38 TB/s raw, 2.02 TB/s sustained (disk) + 100 GB/s (tape) |

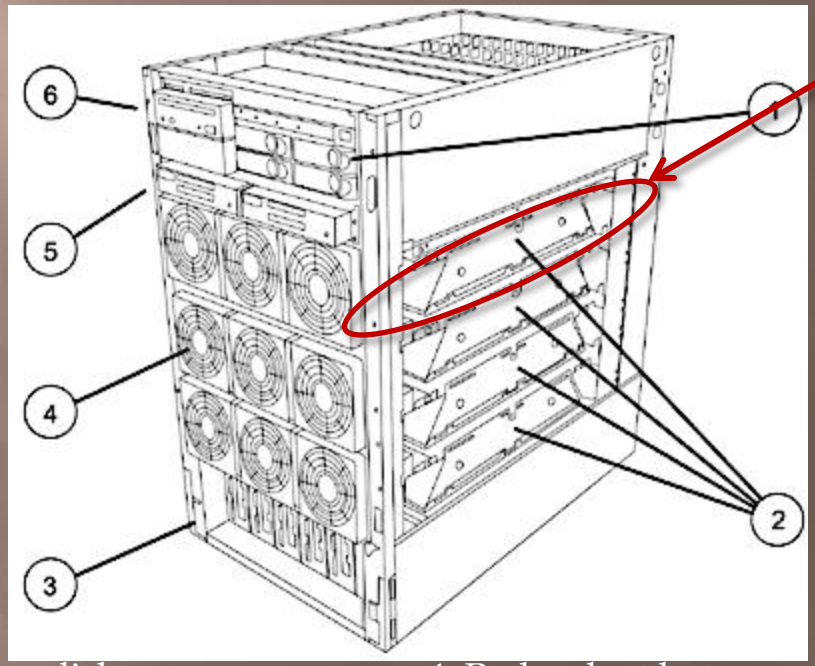
8 cores/chip

32 threads/die

IBM presentation at: http://www.it.utah.edu/leadership/committees/IT_Managers/papers/IBMinEducation.ppt

Content Copyright IBM 2008 extracted from "IBM in Education "2008

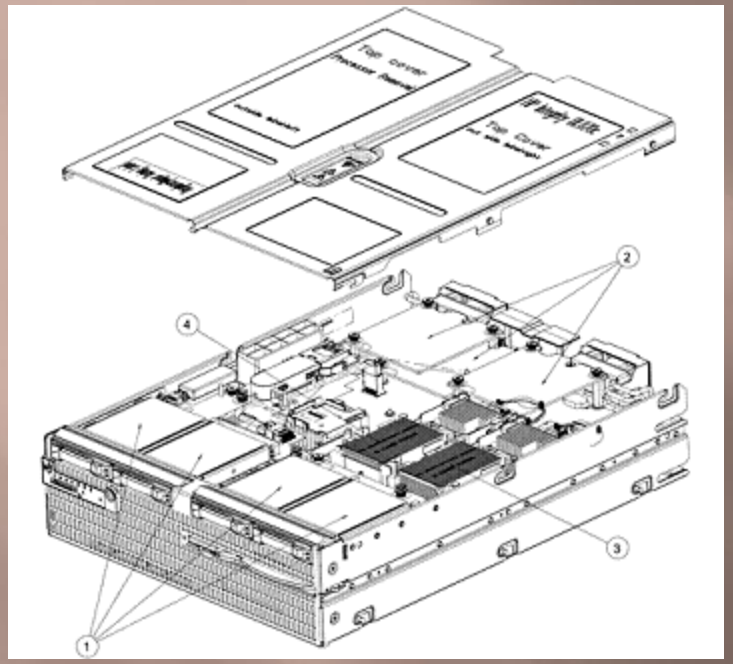
# HP Integrity rx8640 Server System Overview-Front View



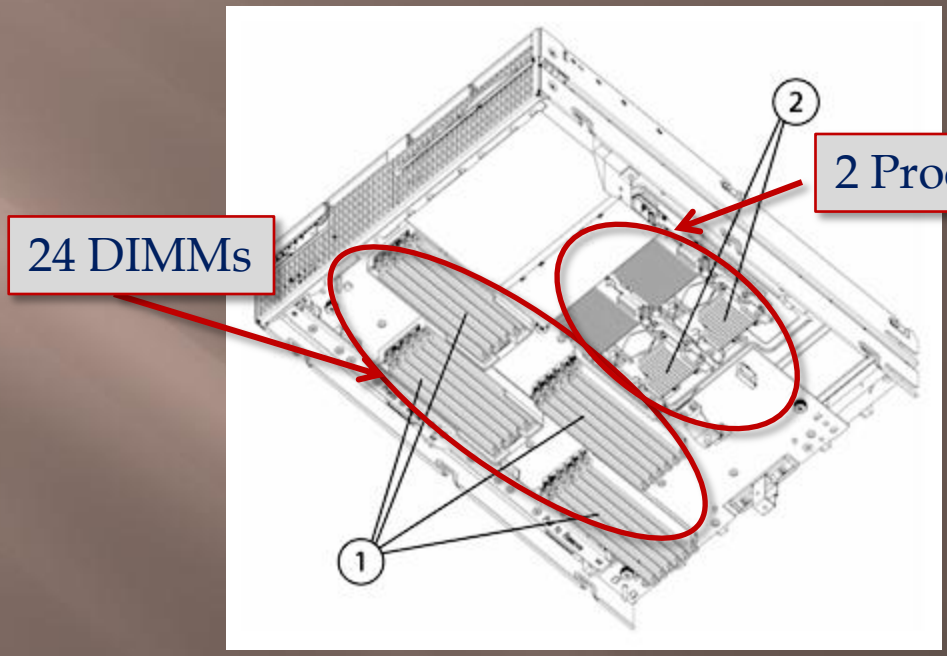

2 Processors and Memory require 1/8 of volume of 8 socket system

1. Hot plug disks
2. Cell boards
3. Redundant hot-swap power
4. Redundant hot-swap fans
5. PCI power supplies
6. Removable media - DVD/DAT

DIMM capacity more than doubled every two years up to 1995. Past ten years growth has been less than double every two years. Source: ITRS Winter conference 2007.



Top View
1. Four small form factor (SFF) hot-plug drive bays
2. Three PCIe Mezzanine slots
3. Processor 0 of up to four Intel® Itanium® 9100 series processors
4. Processor 1 of up to four Intel® Itanium® 9100 series processors



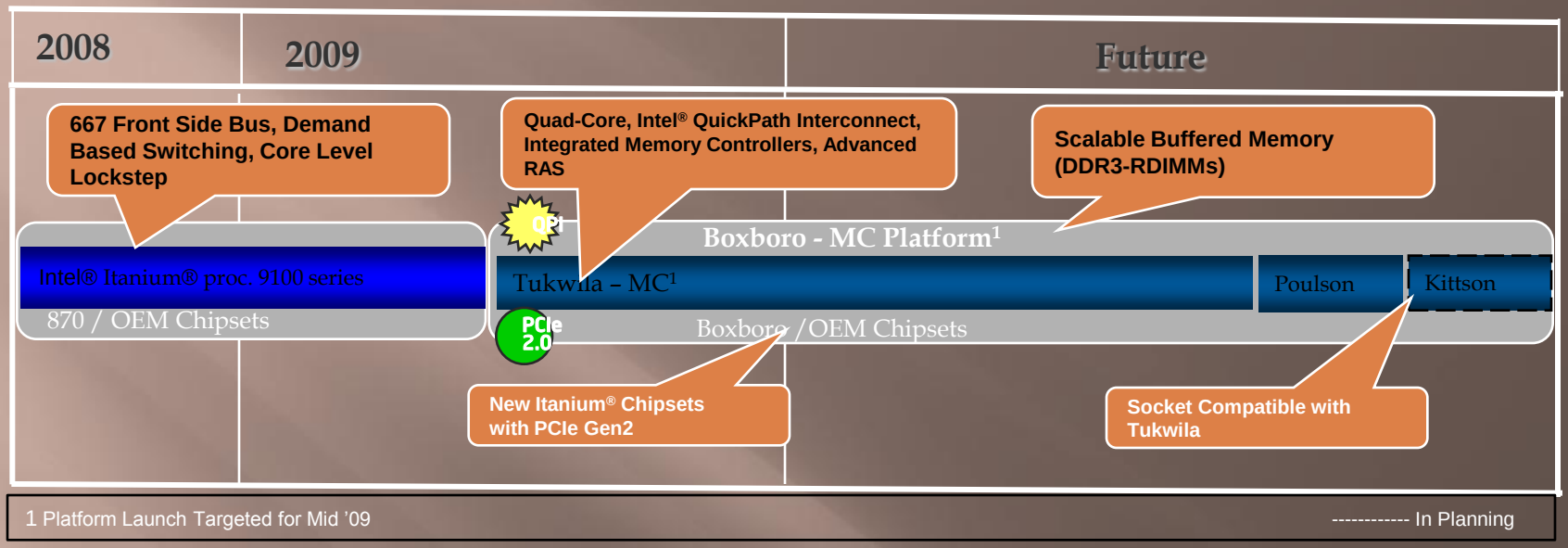

Bottom View
1. Twenty-four (24) DDR2 DIMM Slots
2. Processor 2 and 3 of up to four Intel® Itanium® 9100 series processors

Content Copyright HP 2009, extracted from "HP Integrity BL870c Server Blade QuickSpecs" http://h18004.www1.hp.com/products/quickspecs/12926_na/12926_na.pdf

# Core Count Futures

- Limiters
  - Memory bandwidth, capacity or both
  - Application and OS scaling, applicability or both
- Enablers
  - Memory bandwidth and capacity technology innovations
  - Software innovations
    - Address Amdahl's Argument
      - Scalar vs. throughput performance paradigm

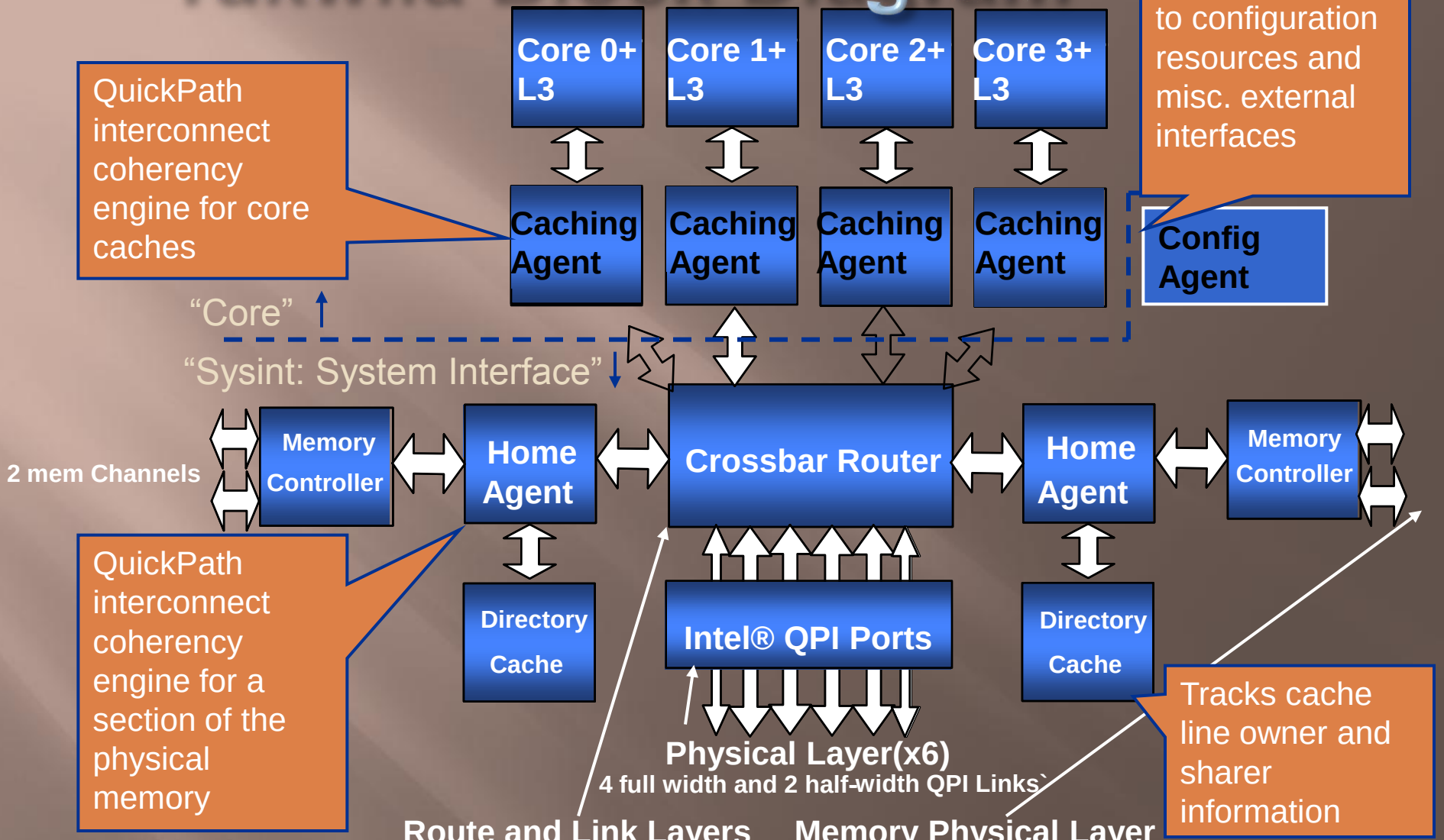
# Mission Critical (MC) Platform Roadmap




- Largest scale Intel based platforms
- Powering systems with RISC/Mainframe-class reliability and data center security
- Greater choice and flexibility than proprietary RISC solutions

**Roadmap focused at Scalability & RAS**

# Tukwila Block Diagram

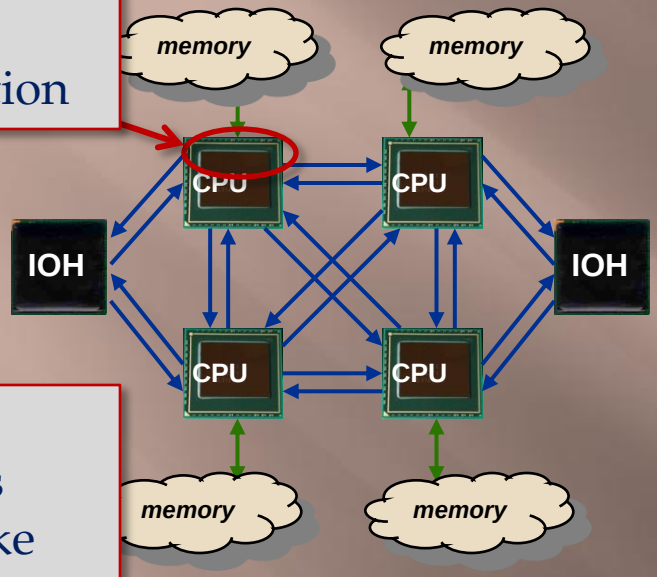



* Intel® QPI = Intel® QuickPath Interconnect

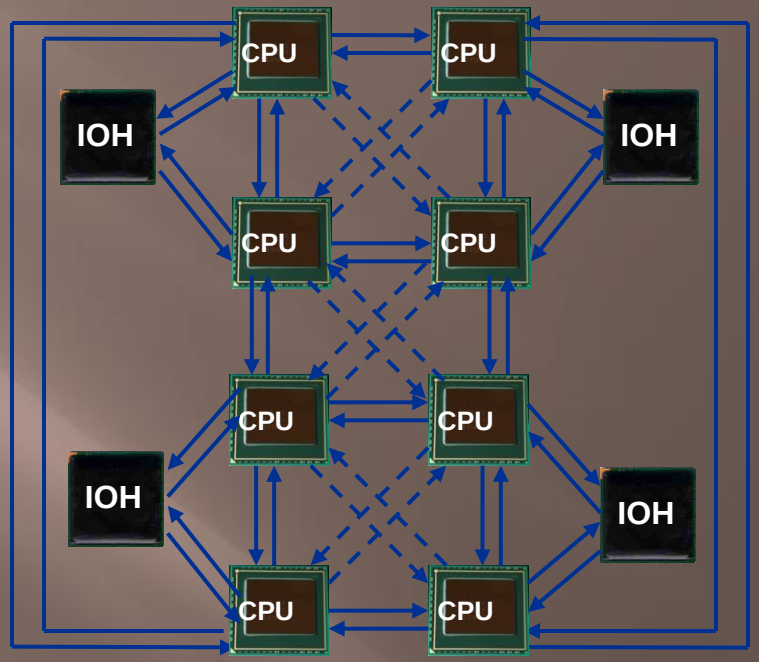
# Example Glueless topologies

Increased system integration

New 4S systems more like past, large, NUMA, scale-up systems




4 CPU topology




8 CPU topology (memory not shown)

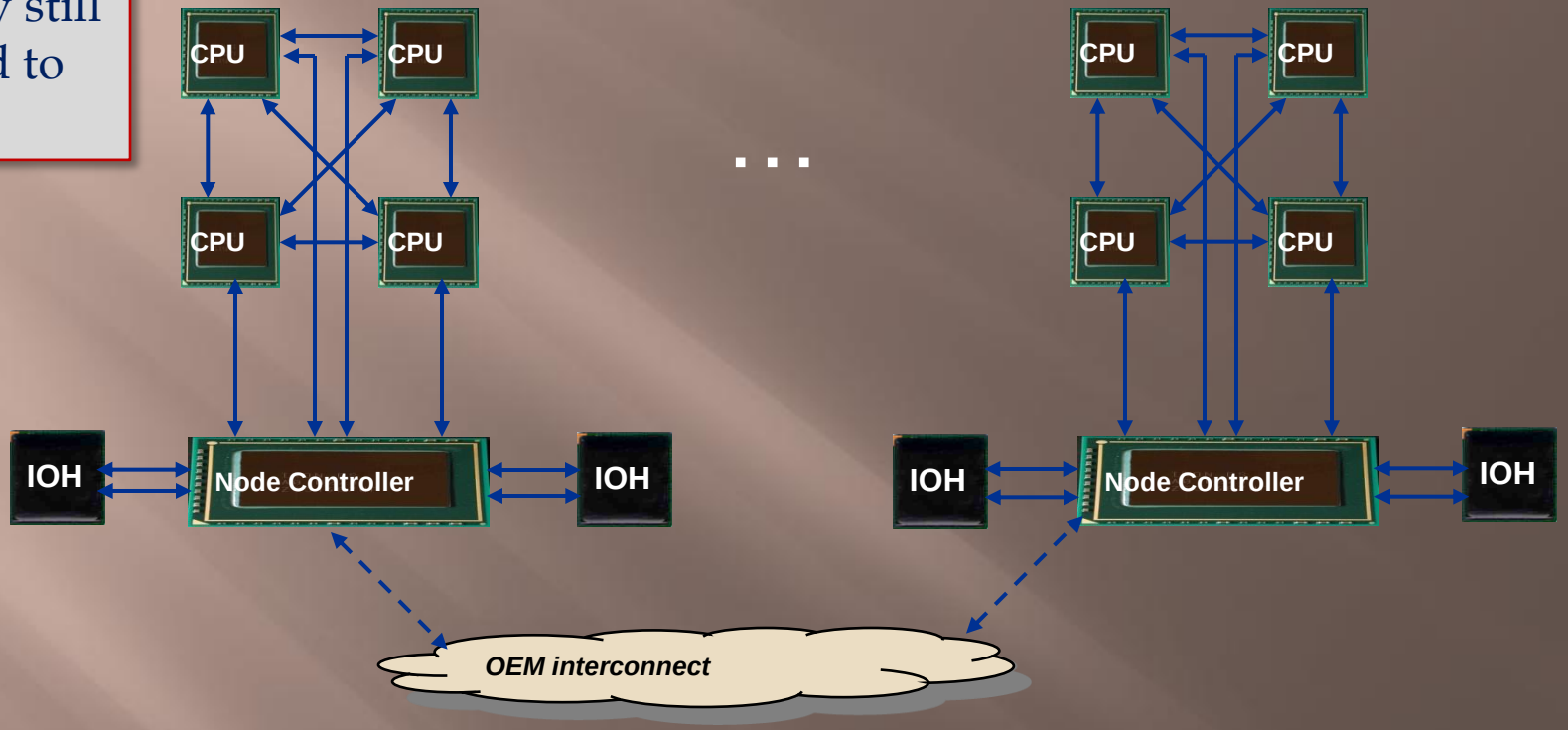
Full Width (20 lane uni-directional) channel (@4.8GT/s)

Half Width (10 lane uni-directional) channel (@4.8GT/s)

IOH: I/O Hub (e.g. bridge to PCIe)

# Example Hierarchical SMP

Memory still attached to CPU




Full width (20 lane) channel per direction(@4.8GT/s)

Number, type, and size are OEM dependent.

# Core - Cache – Memory Topology

- Limiters
  - Cores (applications) may cooperate or compete over cache bandwidth and capacity
  - More memory bandwidth sharing among cores
  - Memory capacity per core/thread is fairly flat
- Enablers
  - Innovations in hardware-software communication to maximize system performance under varied workloads.
  - Performance monitoring may play a part.
  - Fast thread to thread communication and synchronization

# System Component Integration

- Limiters
  - What's left?
  - Socket pin count
- Enablers
  - Increased bandwidth per pin
  - Flexibility for different system designs

# Summary

- Increase in applications competing for compute resources in horizontally scaled environments
  - Cloud computing, Virtual machines, Consolidated machines, Utility computing are examples of this environment
  - Security becomes even more important
- Increase in applications cooperating for compute resources in vertically scaled environments
  - HPC, Medical, Energy and Simulation applications are a good example of this environment