CS 423

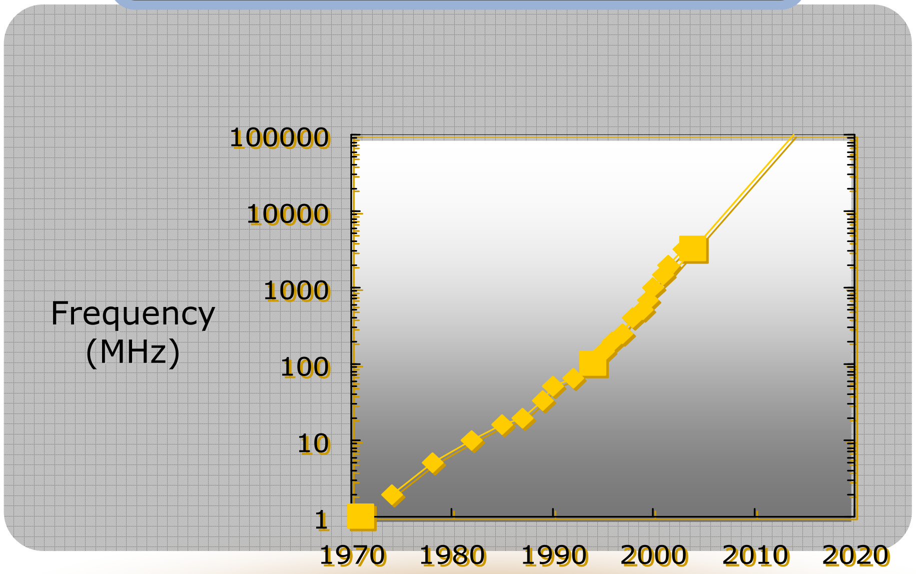
CMP Architecture - I

Design choices

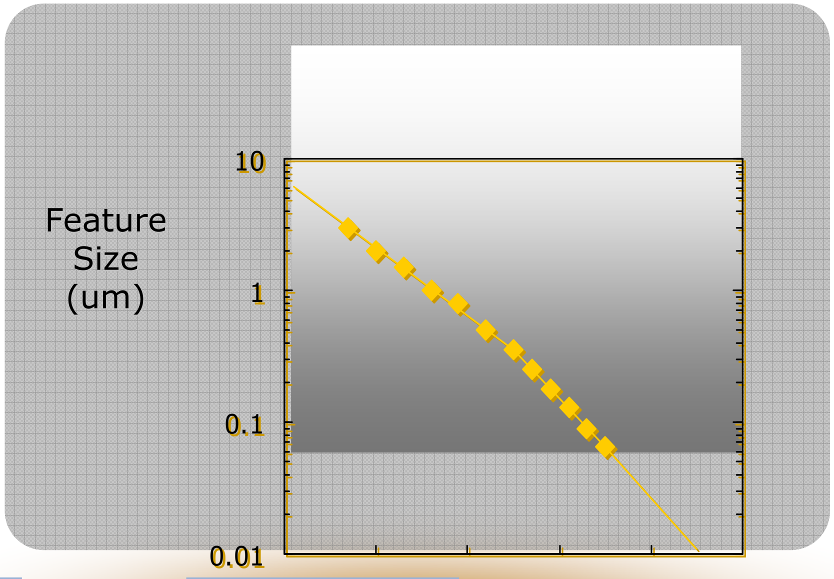
- Processing elements
  - Number
  - Type
  - Homogeneous or heterogeneous
- Memory
  - Size
  - Hierarchy
  - Private vs shared
  - Software managed vs hardware managed
- Interconnection networks
  - Topology
  - Protocol

Historical Driving Forces

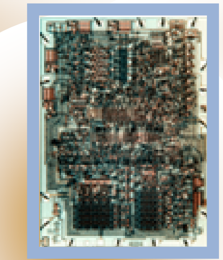
Increased Performance via Increased Frequency



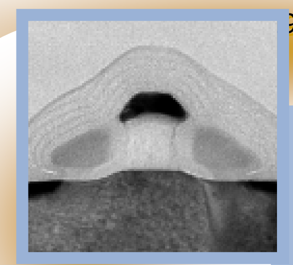
Shrinking Geometry



1946
20 Numbers
in Main Memory



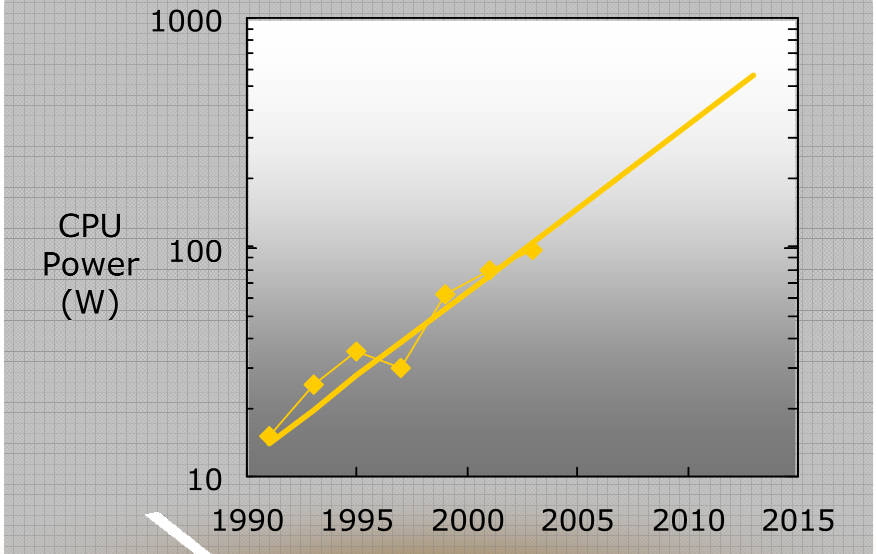
1971
I4004 Processor
2300 Transistors



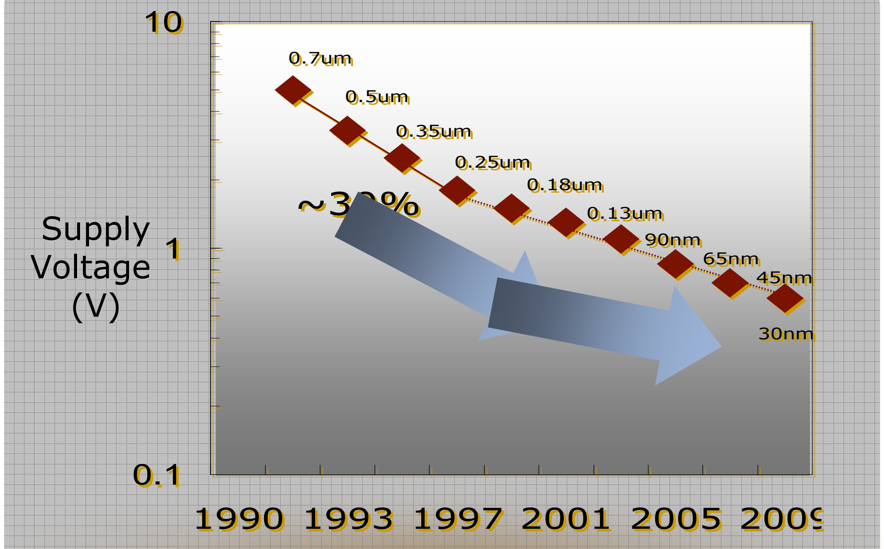
2005
65nm
1B+ Transistors

The Challenges

Power Limitations



Diminishing Voltage Scaling



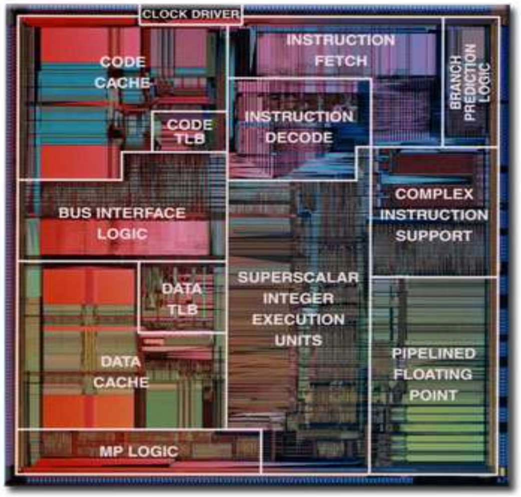
Power = Capacitance x Voltage^2 x Frequency
also
Power ~ Voltage^3

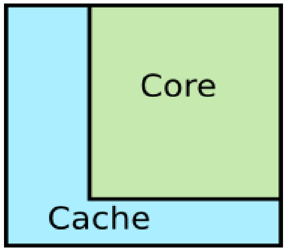
History (List of Intel microprocessors)

- **The 4-bit processors**
  4004, 4040
- **The 8-bit processors**
  8008, 8080, 8085
- **The 16-bit processors: Origin of x86**
  8086, 8088, 80186, 80188, 80286
- **The 32-bit processors: Non x86**
  iAPX 432, 80960, 80860, XScale
- **The 32-bit processors: The 80386 Range**
  80386DX, 80386SX, 80376, 80386SL, 80386EX
- **The 32-bit processors: The 80486 Range**
  80486DX, 80486SX, 80486DX2, 80486SL, 80486DX4
- **The 32-bit processors: The Pentium ("I")**
  Pentium, Pentium MMX
- **The 32-bit processors: P6/Pentium M**
  Pentium Pro, Pentium II, Celeron, Pentium III, PII and III Xeon
  Celeron(PIII), Pentium M, Celeron M, Intel Core, Dual Core Xeon LV
- **The 32-bit processors: NetBurst microarchitecture**
  Pentium 4, Xeon, Pentium 4 EE
- **The 64-bit processors: IA-64**
  Itanium, Itanium 2
- **The 64-bit processors: EM64T-NetBurst**
  Pentium D, Pentium Extreme Edition, Xeon
- **The 64-bit processors: EM64T- Core microarchitecture**
  Xeon, Intel Core 2

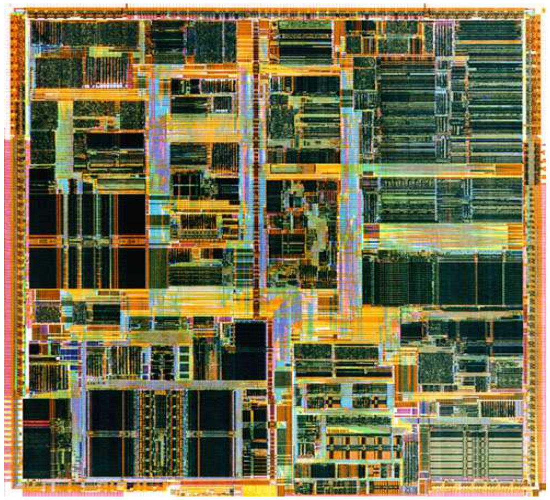
Pentium Processors

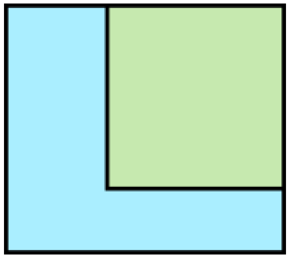
Pentium I



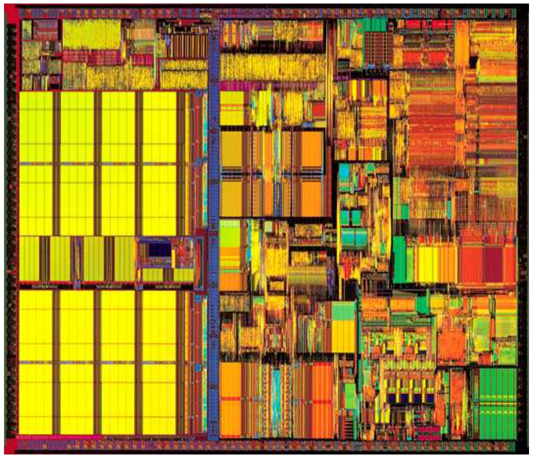


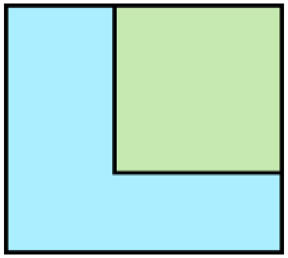
Pentium II



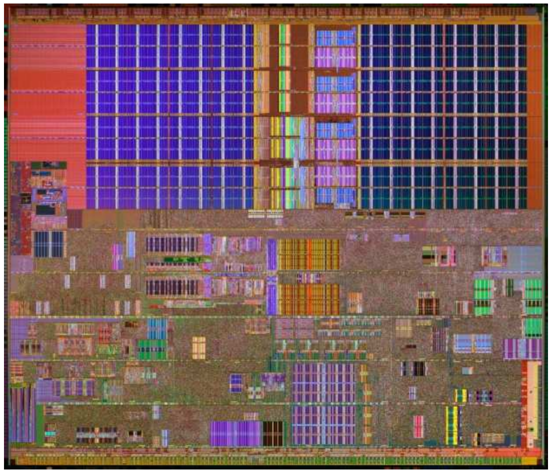


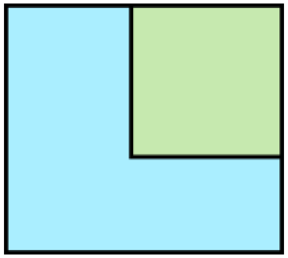
Pentium III





Pentium IV





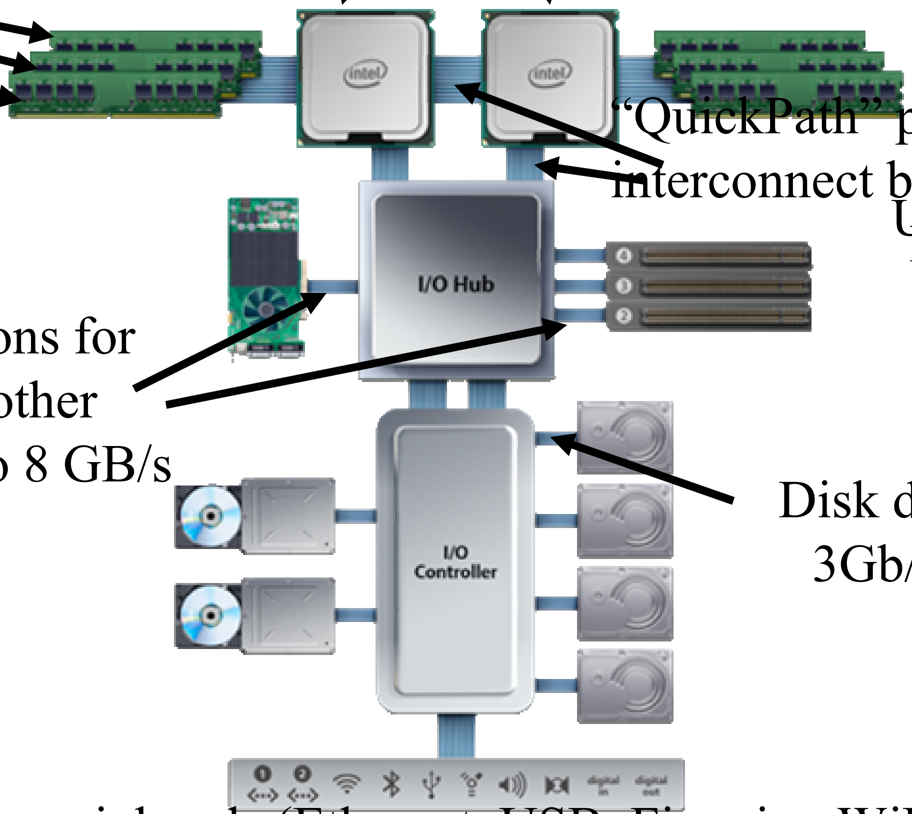
What is happening at Intel?

Intel has more than 15 multi-core related projects underway and plans to increase its software and solutions enabling product lines, tools, investment and programs to further spur design and validation

Nehalem System Example: Apple Mac Pro

Each chip has three DRAM channels attached, each 8 bytes wide at 1.066Gb/s (3*8.5GB/s).

Two Nehalem Chips (“Sockets”), each containing four processors (“cores”) running at up to 2.93GHz



Can have up to two DIMMs on each channel (up to 4GB/DIMM)

“QuickPath” point-point system interconnect between CPUs and I/O.

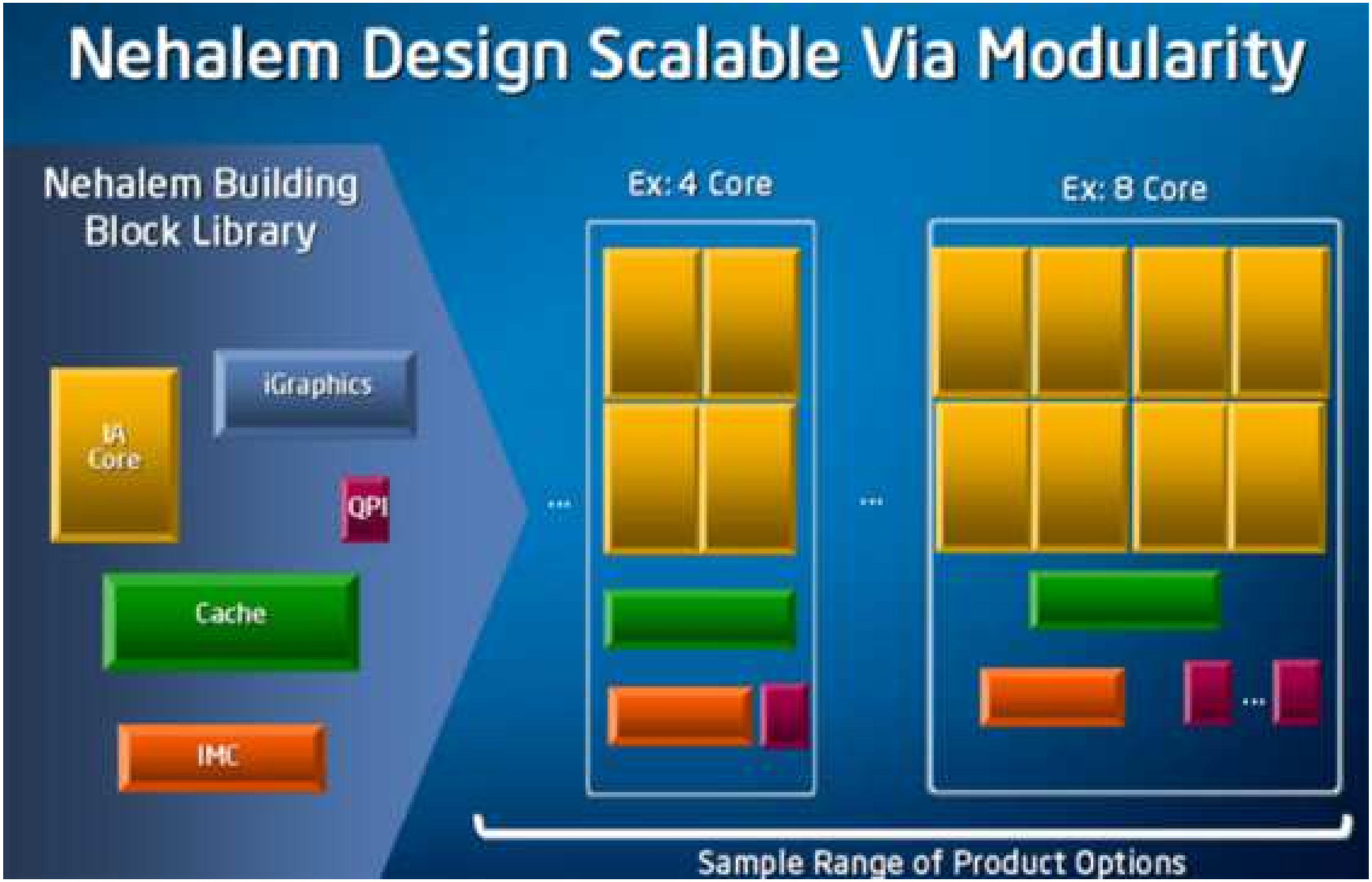
Up to 25.6 GB/s per link.

PCI Express connections for Graphics cards and other extension boards. Up to 8 GB/s per slot.

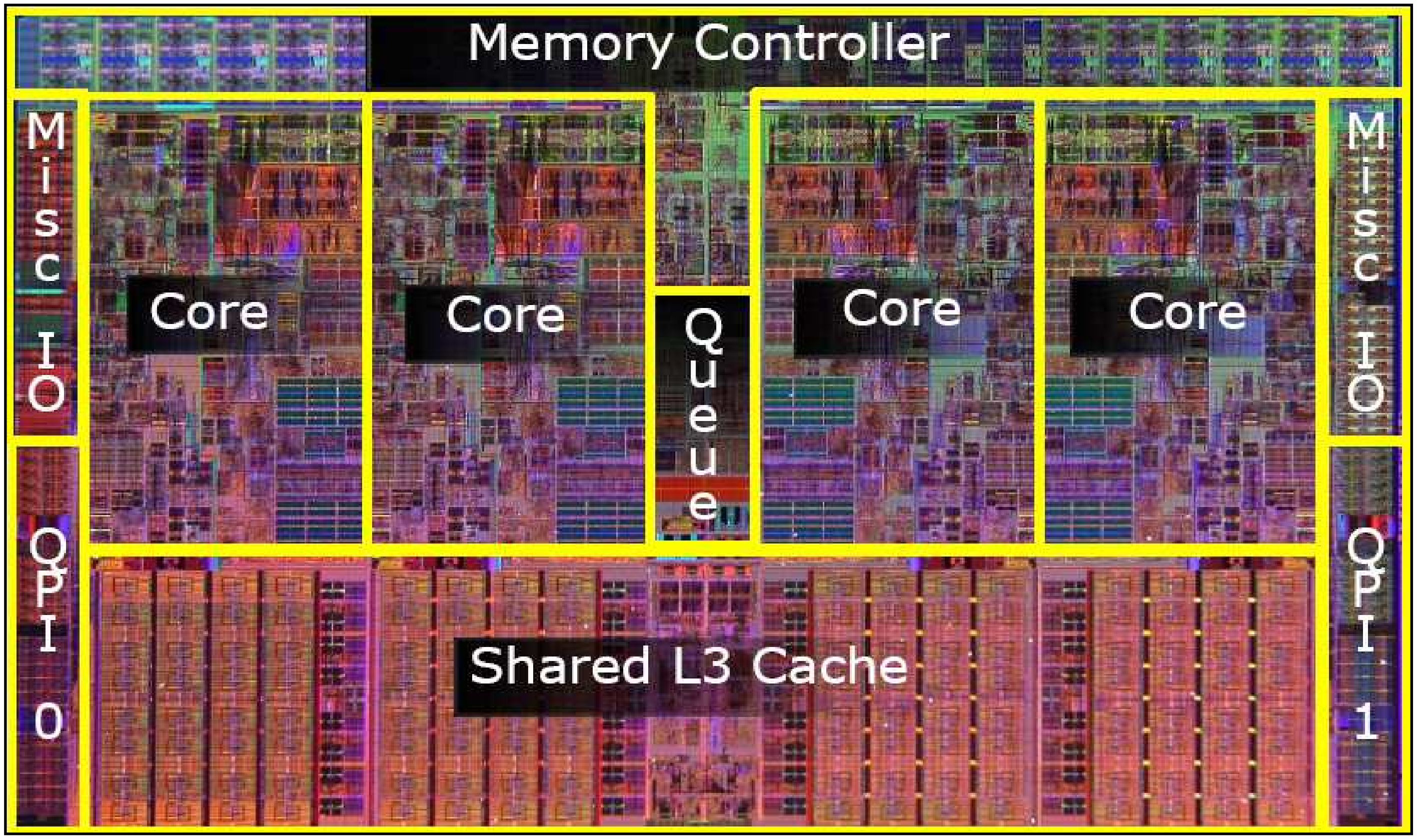
Disk drives attached with 3Gb/s serial ATA link

Slower peripherals (Ethernet, USB, Firewire, WiFi, Bluetooth, Audio)

Building Blocks

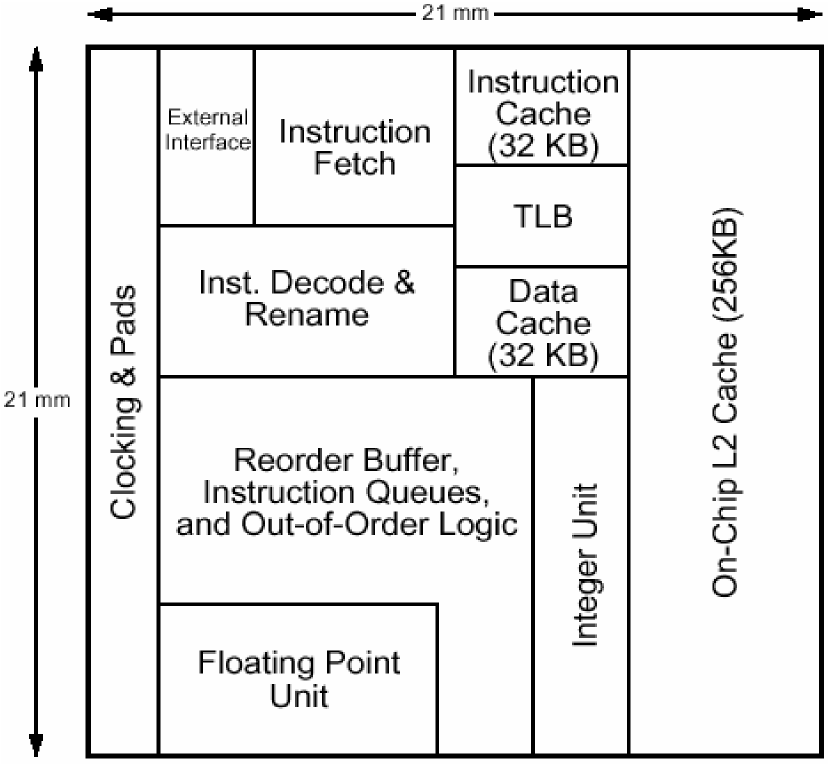

Nehalem Design Scalable Via Modularity
Nehalem Building Block Library
IA Core
iGraphics
QPI
Cache
IMC
...
Ex: 4 Core
...
Ex: 8 Core
...
Sample Range of Product Options

Example of modern core: Nehalem

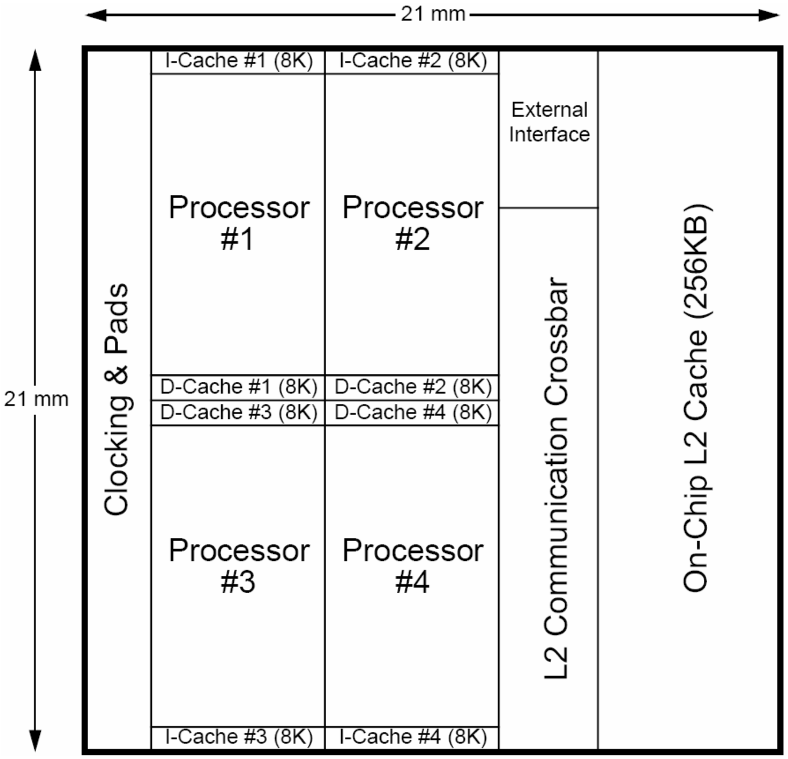


- ON-chip cache resources:
  - For each core: L1: 32K instruction and 32K data cache, L2: 1MB
  - L3: 8MB shared among all 4 cores
- Integrated, on-chip memory controller (DDR3)

Superscalar (SS) – Multicore (MP)



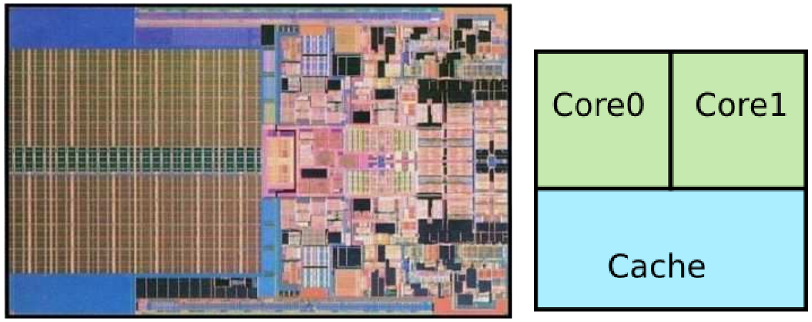
6-way superscalar (SS) microarchitecture



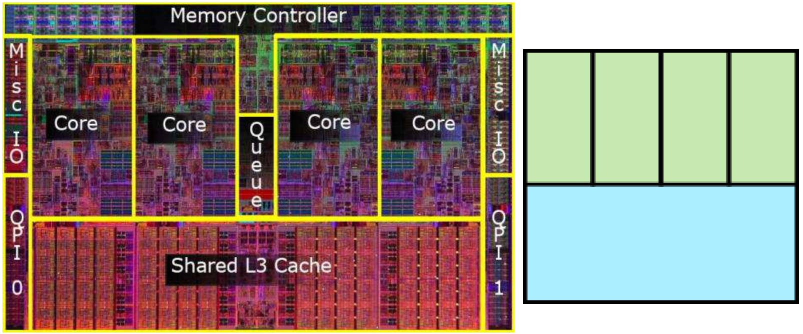
Multiprocessor (MP) microarchitecture
(4 identical 2-way superscalar processors)

Multicore Processors

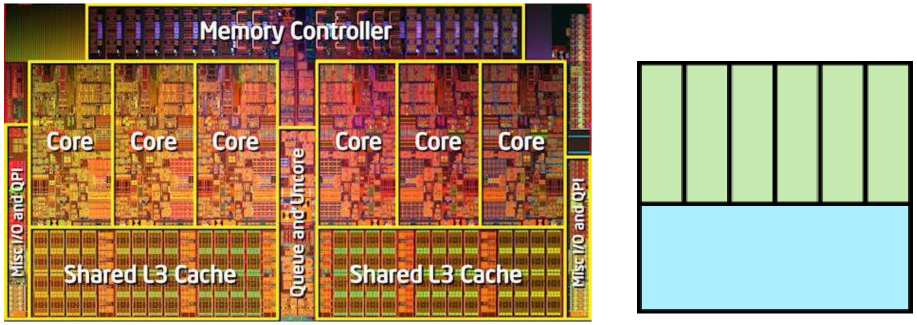
Penryn



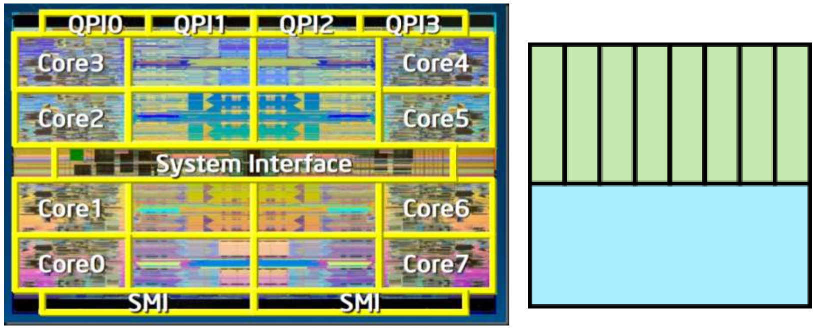
Bloomfield



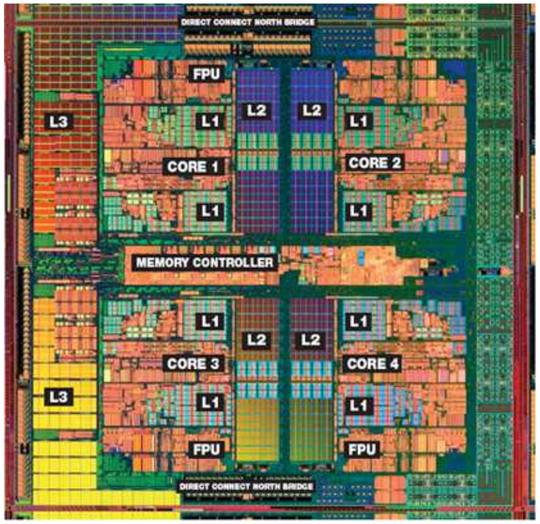
Gulftown

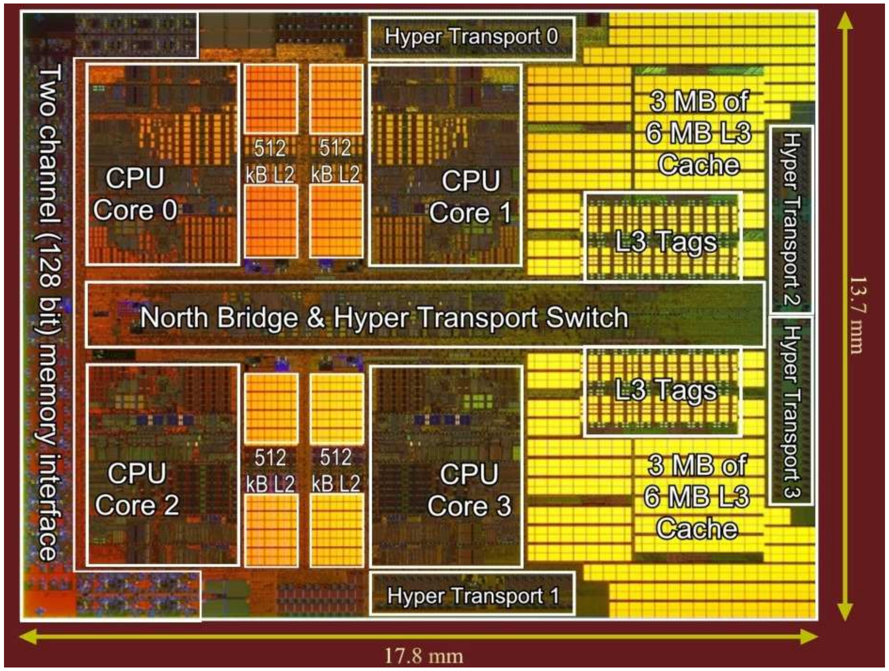


Beckton



AMD Phenom II Quad-Core)





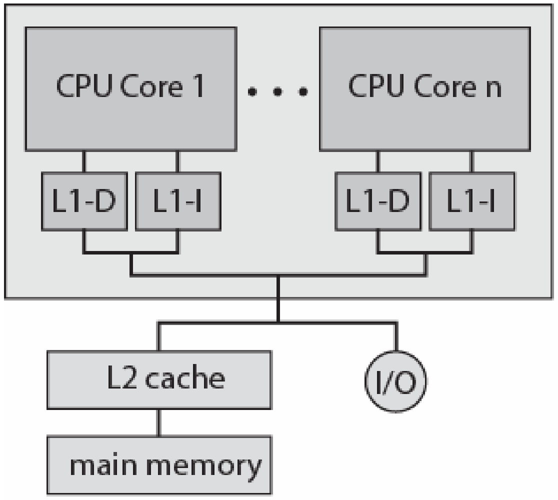
- AMD K10 (Barcelona)
- Code name “Deneb”
- 45nm process
- 4 cores, private 512KB L2
- Shared 6MB L3 (2MB in Phenom)
- Integrated Northbridge
  - Up to 4 DIMMs
- Sideband Stack optimizer (SSO)
  - Parallelize many POPs and PUSHs (which were dependent on each other)
    - Convert them into pure loads/store instructions
  - No uops in FUs for stack pointer adjustment

ARM11 MPCore

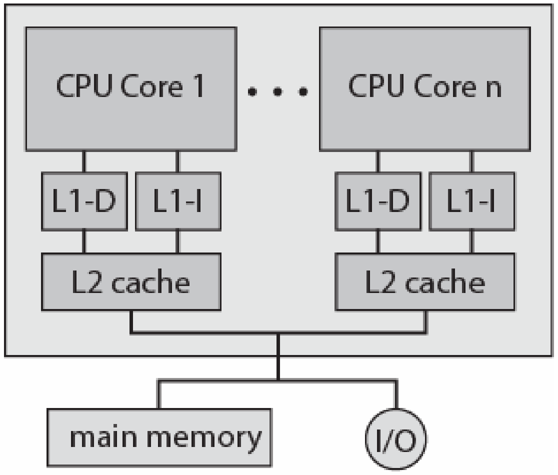
- Up to 4 processors each with own L1 instruction and data cache
- Distributed Interrupt Controller (DIC)
  - Recall the APIC from Intel’s core architecture
- Timer per CPU
- CPU interface
  - Interrupt acknowledgement, masking and completion acknowledgement
- CPU
  - Single ARM11 called MP11
- Vector floating-point unit (VFP)
  - FP co-processor
- L1 cache
- Snoop control unit
  - L1 cache coherency

Multicore Organization Alternatives

ARM11 MPCore



(a) Dedicated L1 cache

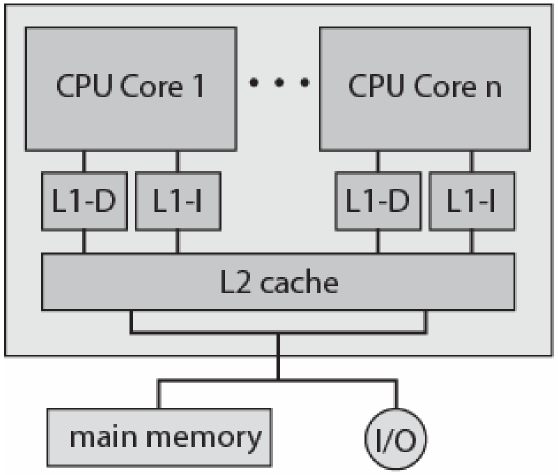


(b) Dedicated L2 cache

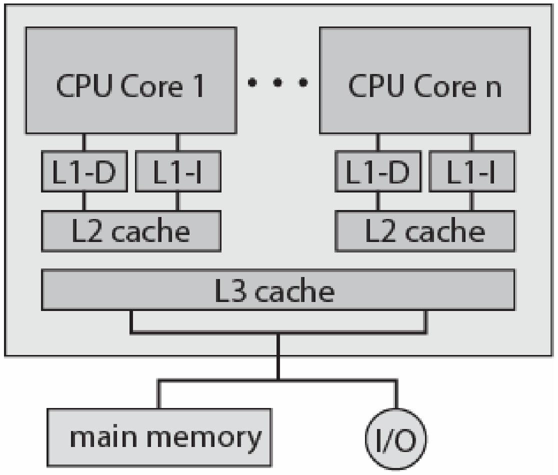
AMD Opteron

No shared

Intel Core Duo



(c) Shared L2 cache



(d) Shared L3 cache

Intel Core i7

Shared

Cell History

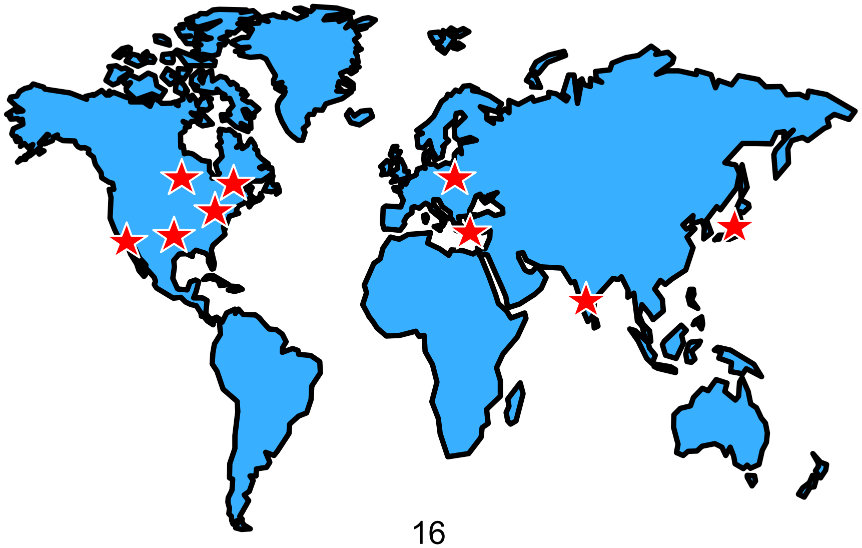
- IBM, SCEI/Sony, Toshiba Alliance formed in 2000
- Design Center opened in March 2001
  - Based in Austin, Texas
- Single Cell BE operational Spring 2004
- 2-way SMP operational Summer 2004
- February 7, 2005: First technical disclosures
- October 6, 2005: Mercury Announces Cell Blade
- November 9, 2005: Open Source SDK & Simulator Published
- November 14, 2005: Mercury Announces Turismo Cell Offering
- February 8, 2006 IBM Announced Cell Blade

SONY COMPUTER ENTERTAINMENT

SONY

TOSHIBA

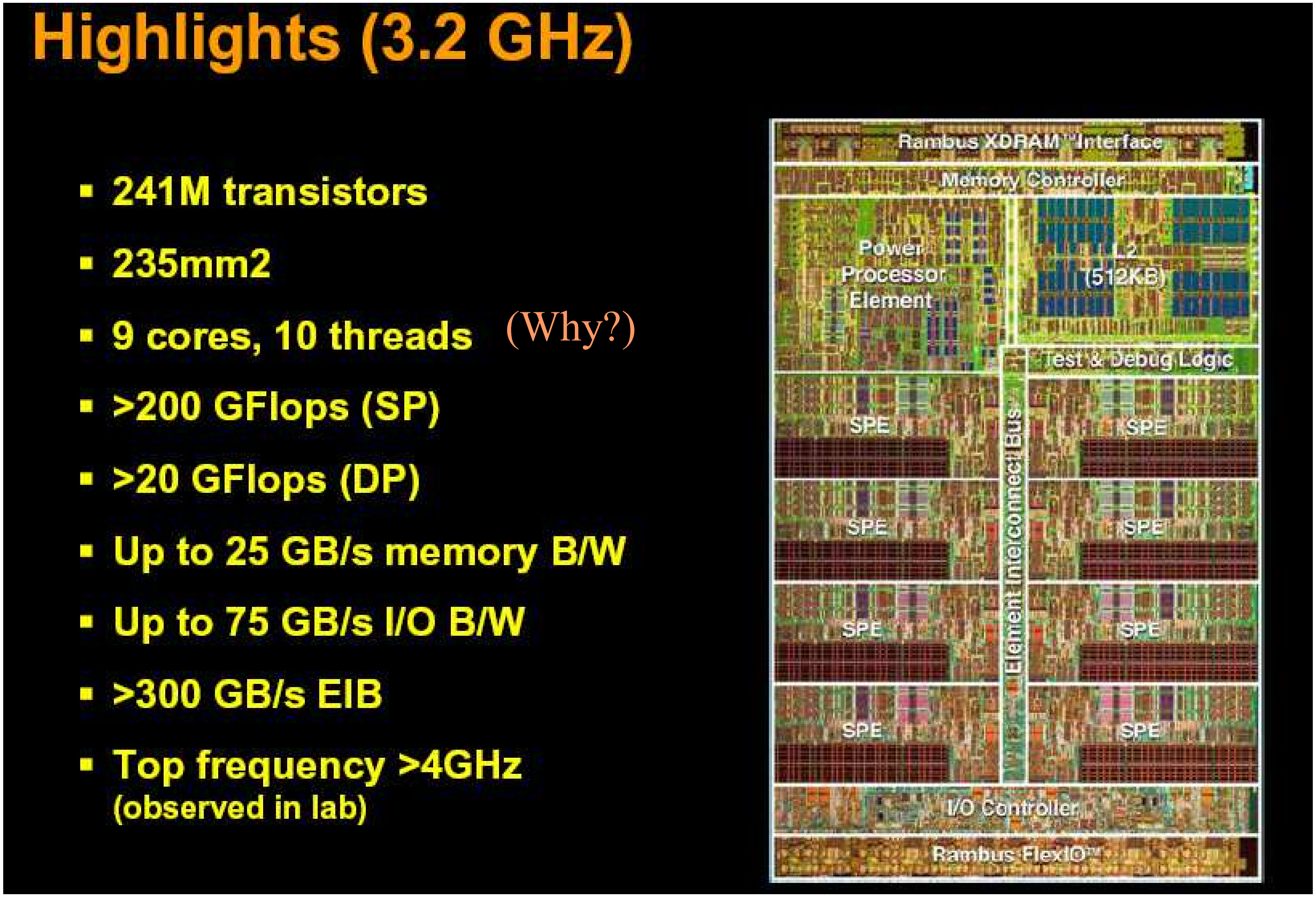
IBM



Cell Synergy

- Cell is not a collection of different processors, but a synergistic whole
  - Operation paradigms, data formats and semantics consistent
  - Share address translation and memory protection model
- PPE for operating systems and program control
- SPE optimized for efficient data processing
  - SPEs share Cell system functions provided by Power Architecture
  - MFC implements interface to memory
    - Copy in/copy out to local storage
- PowerPC provides system functions
  - Virtualization
  - Address translation and protection
  - External exception handling
- EIB integrates system as data transport hub

Cell Chip



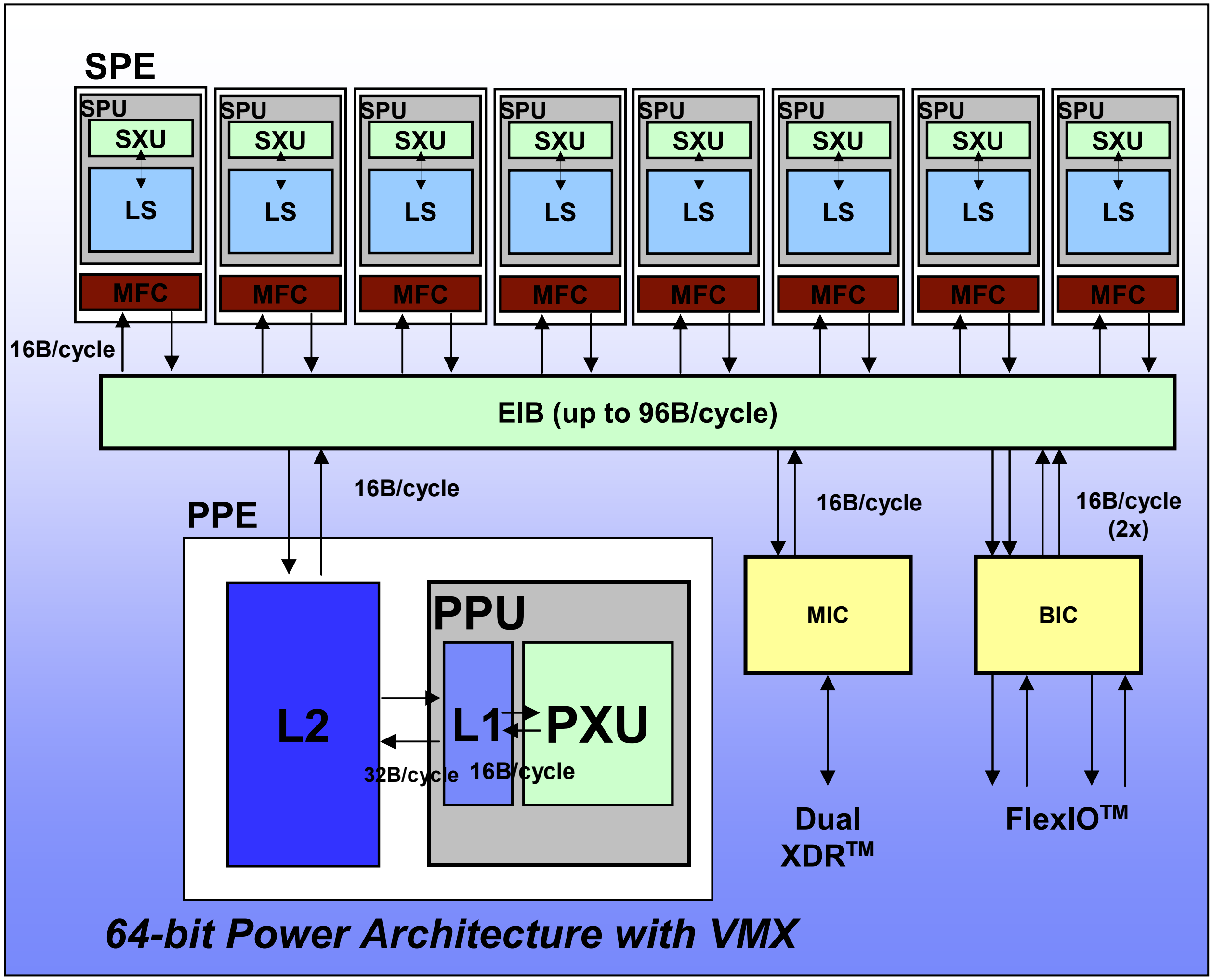
Highlights (3.2 GHz)

- 241M transistors
- 235mm2
- 9 cores, 10 threads (Why?)
- >200 GFlops (SP)
- >20 GFlops (DP)
- Up to 25 GB/s memory B/W
- Up to 75 GB/s I/O B/W
- >300 GB/s EIB
- Top frequency >4GHz (observed in lab)

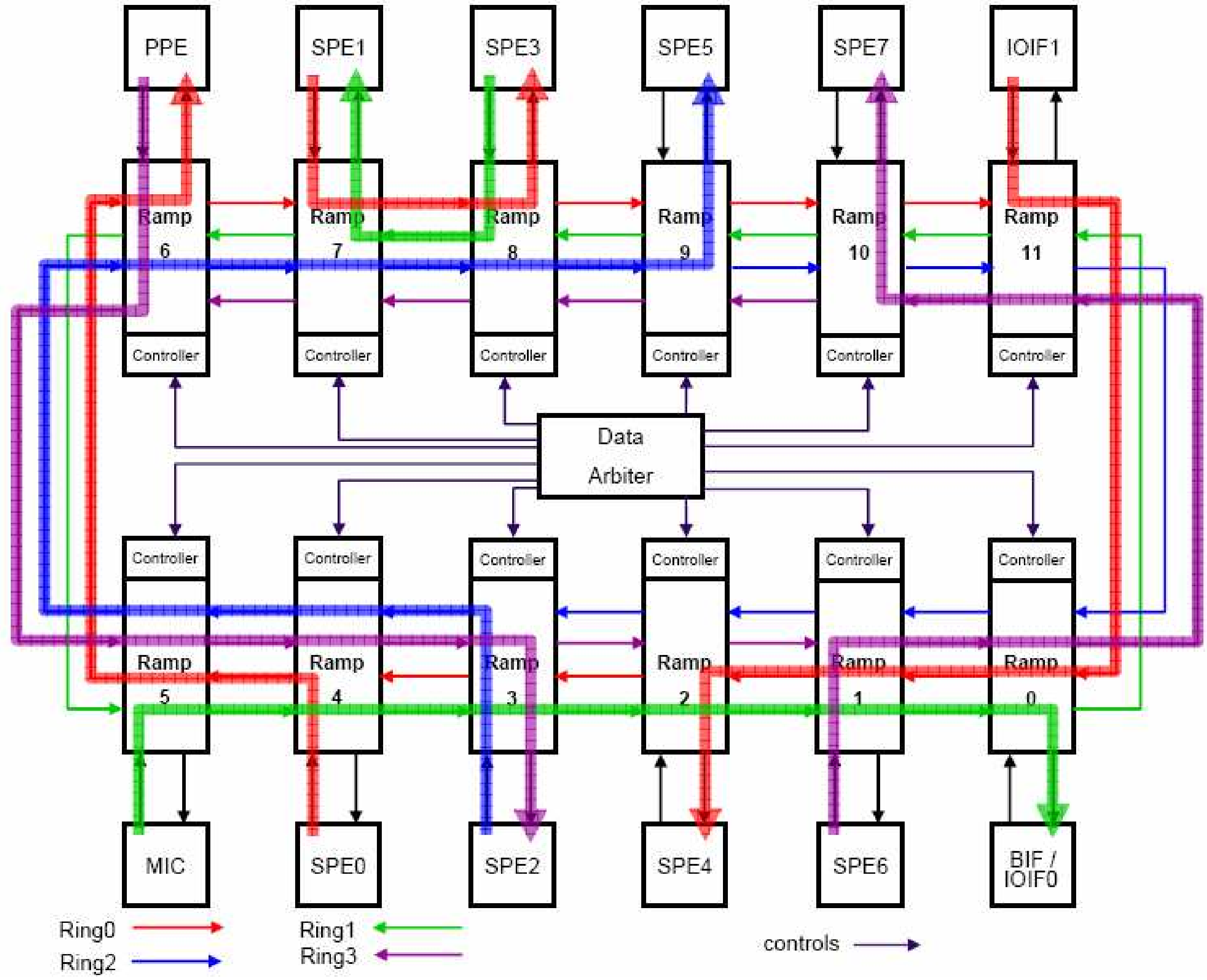
Rambus XDRAM™ Interface
Memory Controller
Power Processor Element
L2 (512KB)
Test & Debug Logic
Element Interconnect Bus
SPE
SPE
SPE
SPE
SPE
SPE
SPE
SPE
I/O Controller
Rambus FlexIO™

Cell Features

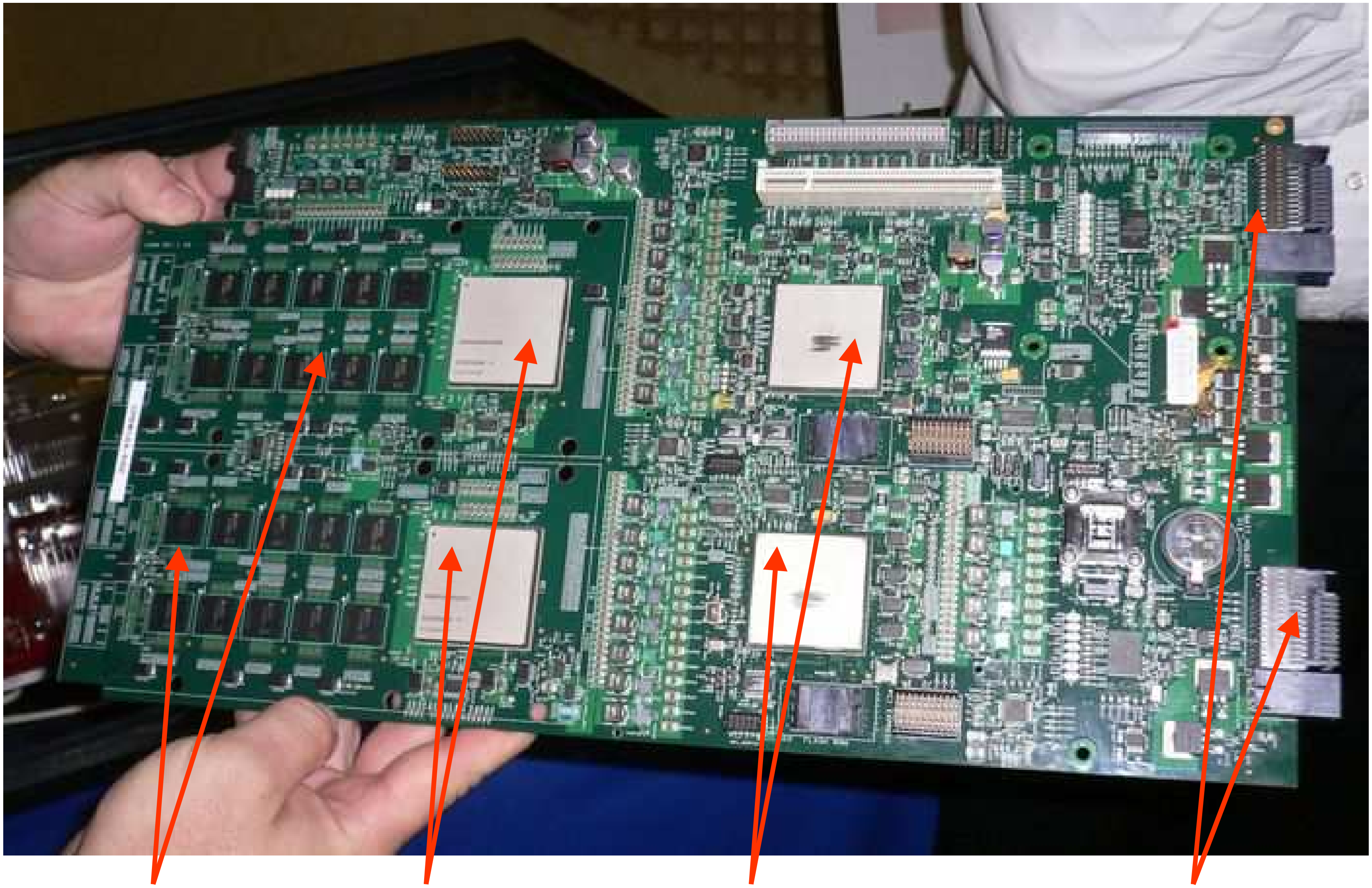
- Heterogeneous multicore system architecture
  - Power Processor Element for control tasks
  - Synergistic Processor Elements for data-intensive processing
- Synergistic Processor Element (SPE) consists of
  - Synergistic Processor Unit (SPU)
  - Synergistic Memory Flow Control (MFC)
    - Data movement and synchronization
    - Interface to high-performance Element Interconnect Bus



Eight Concurrent Transactions



The First Generation Cell Blade



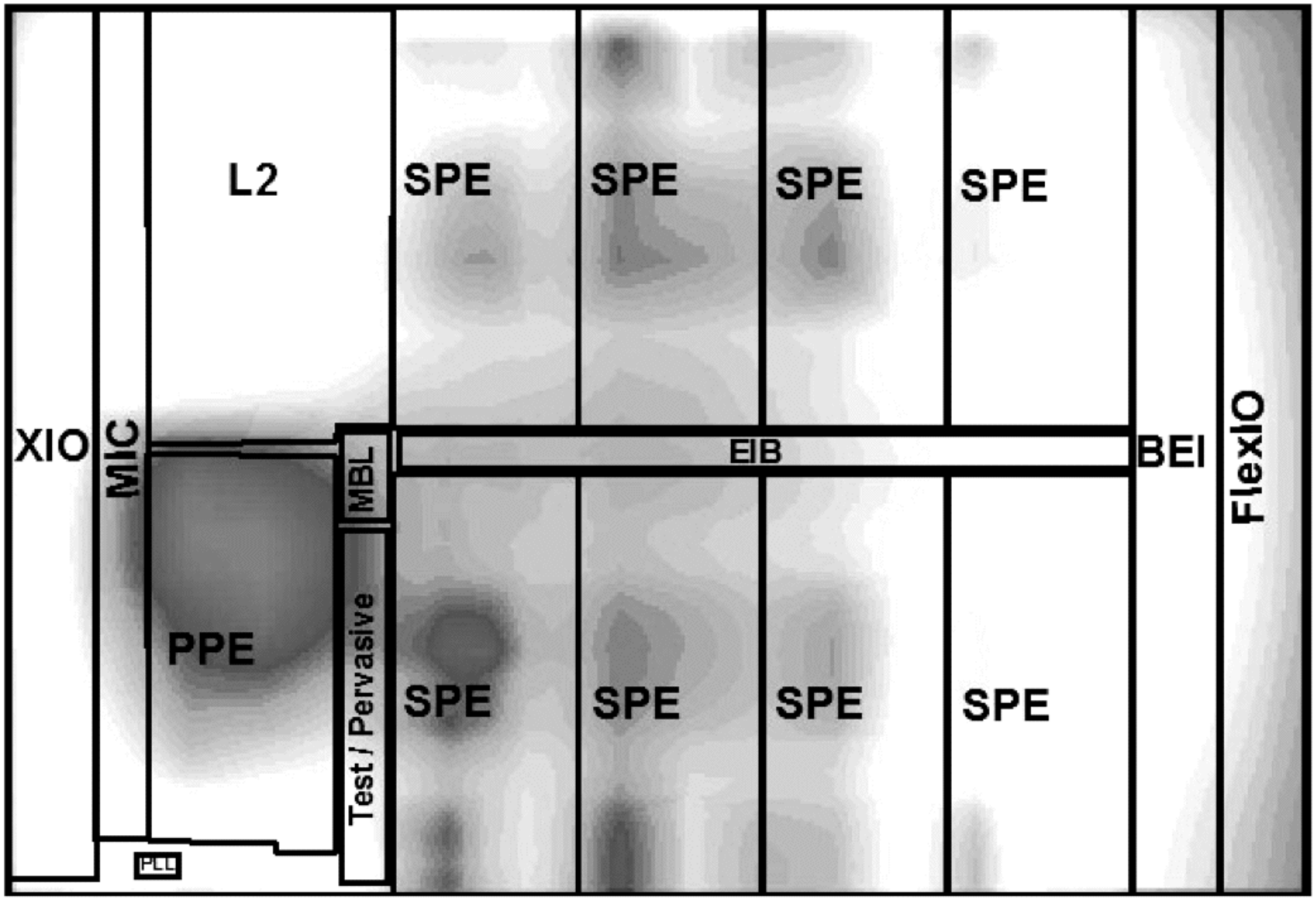
1GB XDR Memory

Cell Processors

IO Controllers

IBM Blade Center interface

Cell Temperature Graph



- Power and heat are key constrains
  - Cell is ~80 watts at 4+ Ghz
  - Cell has 10 temperature sensors
  - Prediction: PS3 will be more like 3 Ghz

Source: IEEE ISSCC, 2005

Code Sample

- PPE code:

```
#include <stdio.h>
#include <libspe.h>
extern spe_program_handle_t hello_spu;
int main(void)
{
        speid_t speid;
        int status;
        speid = spe_create_thread (0, &hello_spu, NULL, NULL, -1, 0);
        spe_wait(speid, &status, 1);
        return 0;
```

- SPE code:

```
#include <stdio.h>
#include <cbe_mfc.h>
#include <spu_mfcio.h>
int main(speid_t speid, unsigned long long argp, unsigned long long envp)
{
        printf("Hello world!\n");
        return 0;
}
```

“PPE-Centric” & “SPE-Centric” Models

- “PPE-Centric”:
  - an offload model
  - main line application code runs in PPC core
  - individual functions extracted and offloaded to SPEs
  - SPUs wait to be given work by the PPC core
- “SPE-Centric”:
  - most of the application code distributed among SPEs
  - PPC core runs little more than a resource manager for the SPEs (e.g., maintaining in main memory control blocks with work lists for the SPEs)
  - SPE fetches next work item (what function to execute, pointer to data, etc.) from main memory (or its own memory) when it completes current work item