



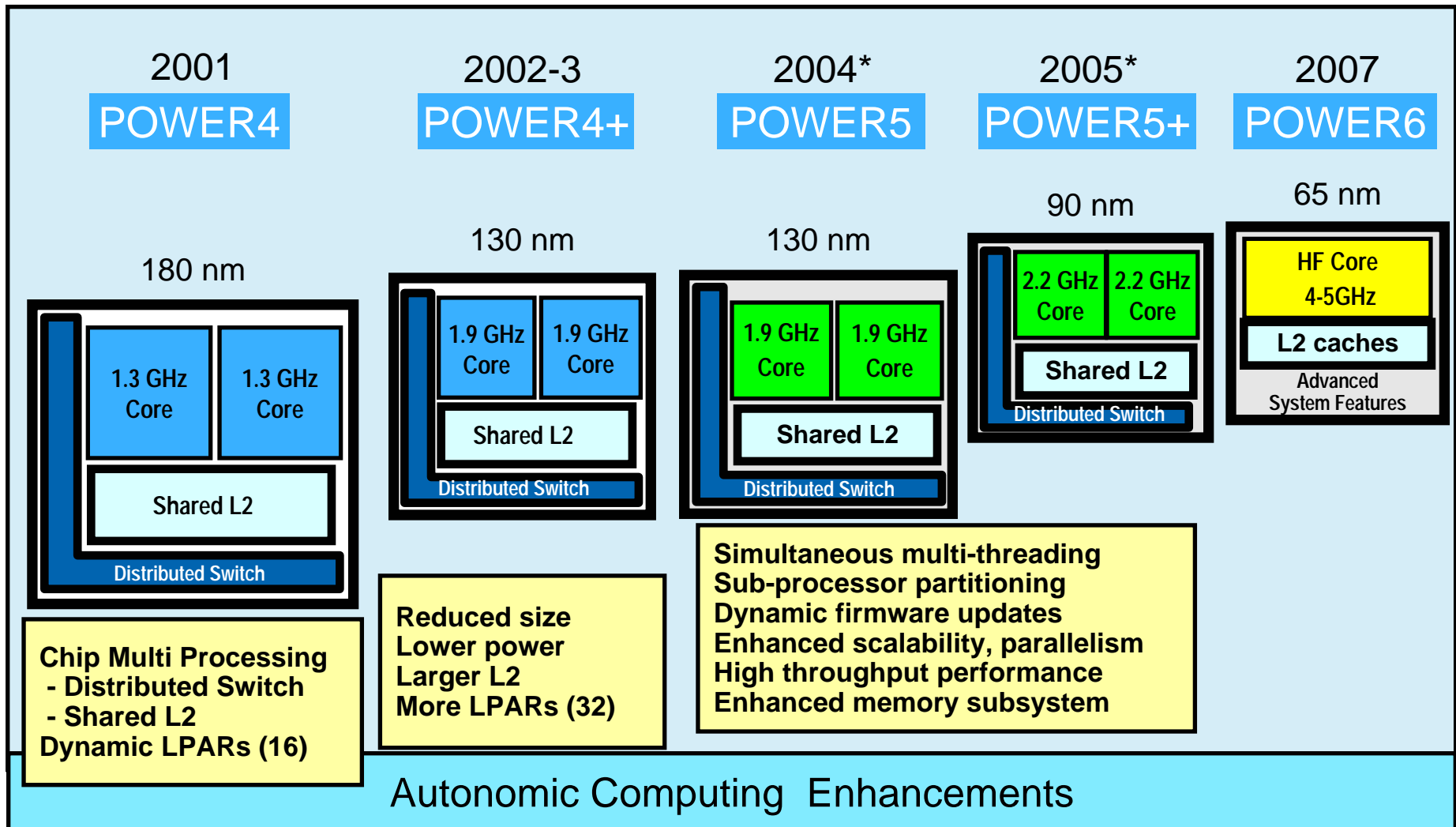
IBM's Micro Processor Design and Methodology

Ron Kalla
IBM Systems and Technology Group

Outline

- POWER5
- POWER6
- Design Process
- Power Aware Design

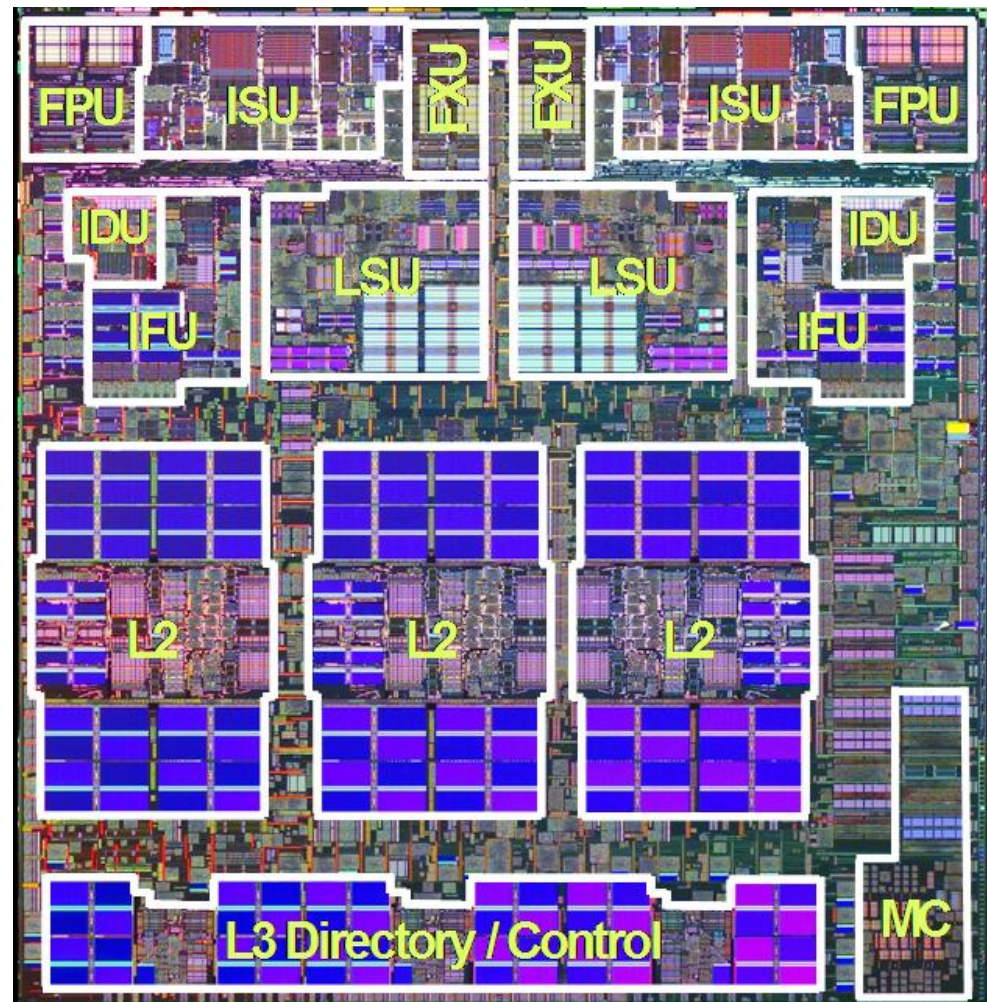
POWER Server Roadmap



*Planned to be offered by IBM. All statements about IBM's future direction and intent are subject to change or withdrawal without notice and represent goals and objectives only.

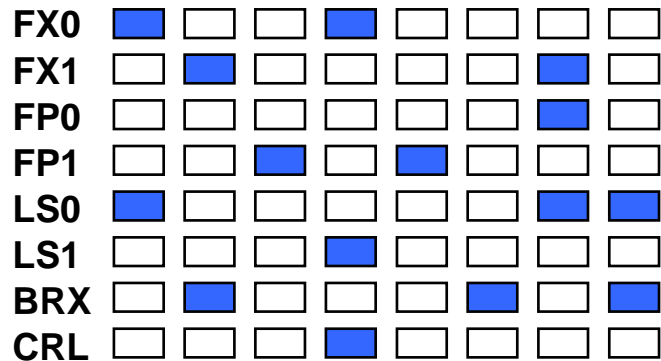
POWER5

- Technology: 90nm lithography, Cu, SOI
- 245mm² 300M Transistors
- Dual processor core
- 8-way superscalar
- Simultaneous multithreaded (SMT) core
 - ▶ Up to 2 virtual processors per real processor

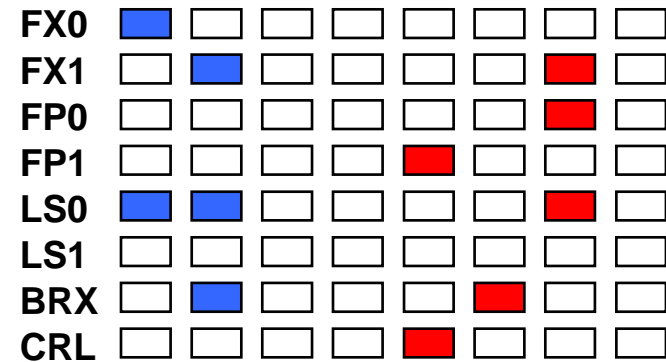


Multi-threading Evolution

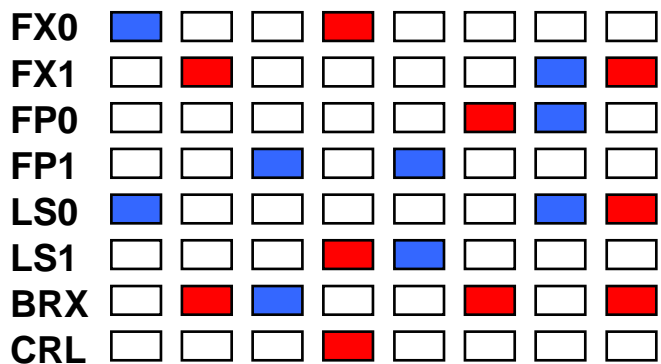
Single Thread



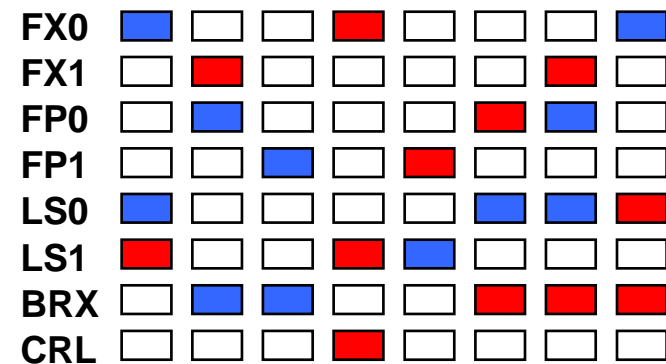
Coarse Grain Threading



Fine Grain Threading



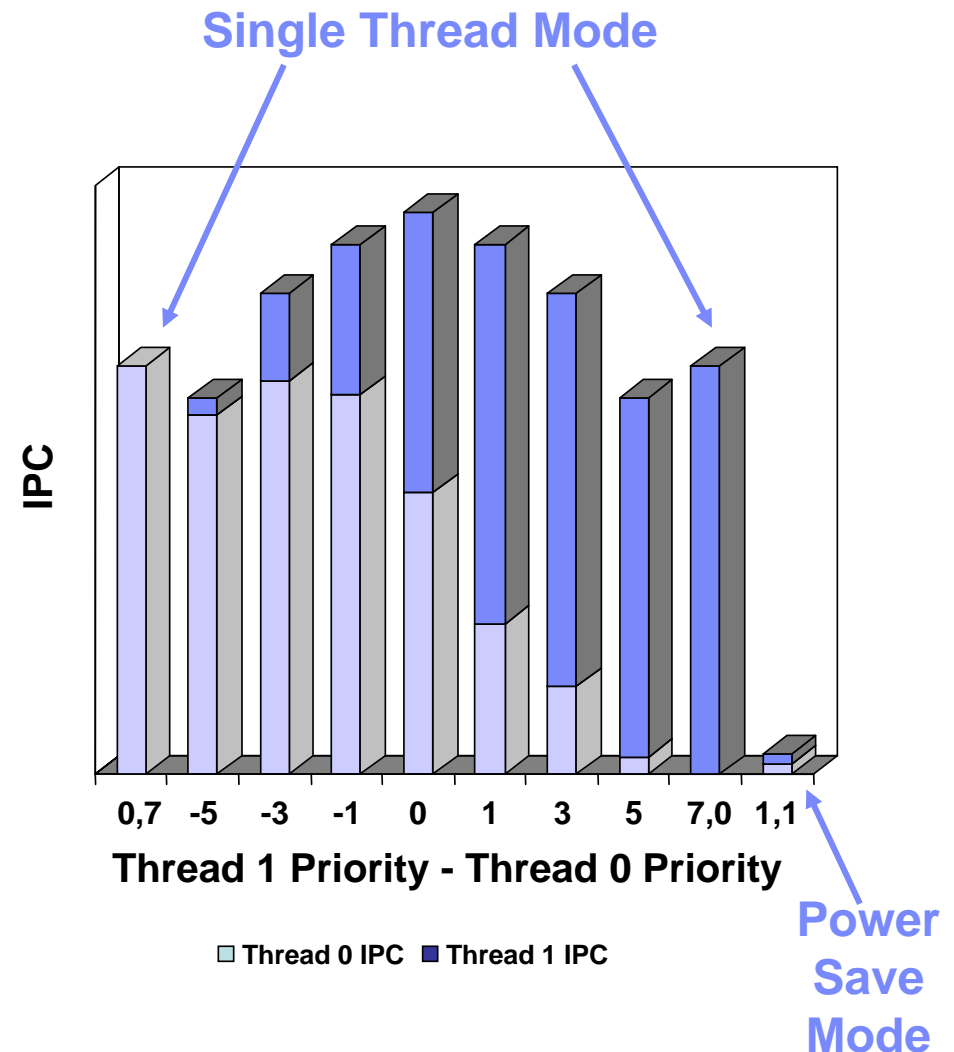
Simultaneous Multi-Threading



█ Thread 0 Executing
 █ Thread 1 Executing
 □ No Thread Executing

Thread Priority

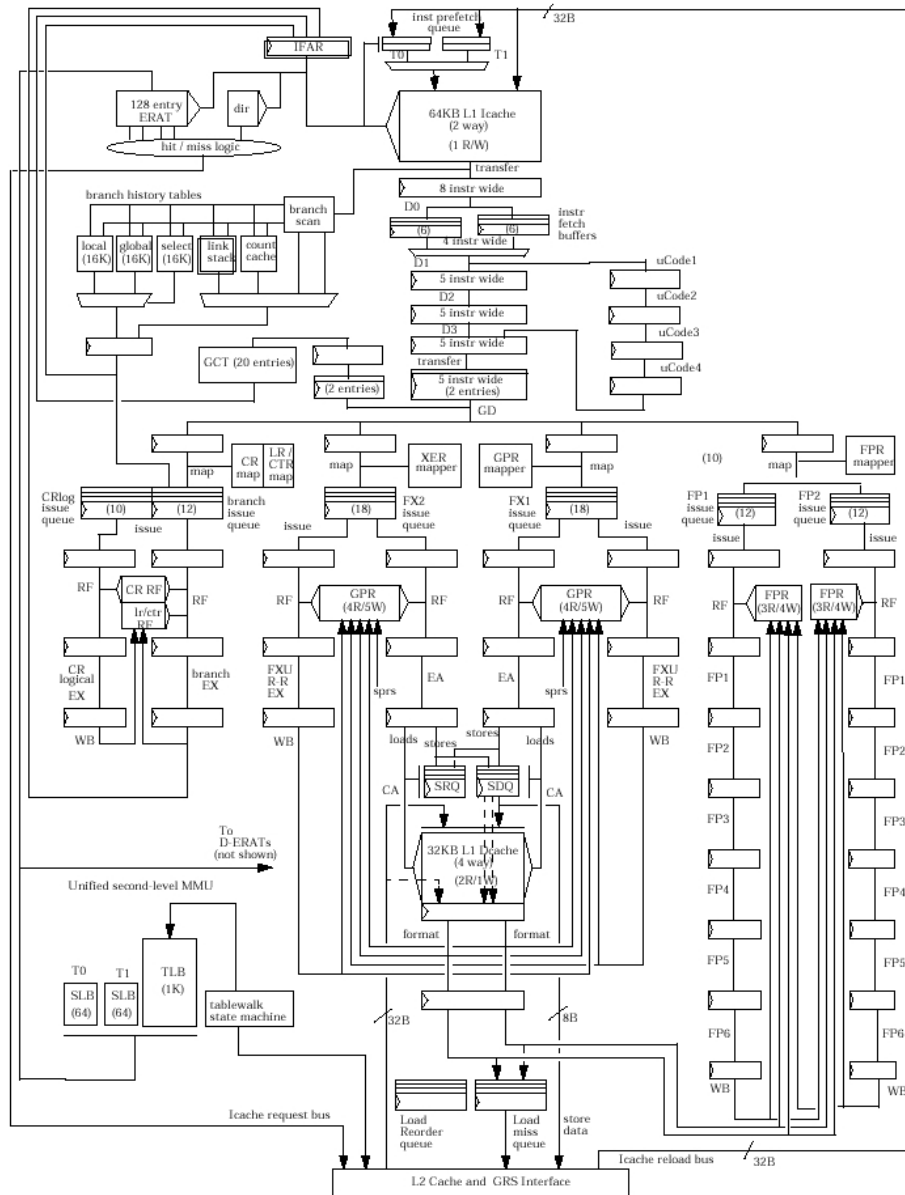
- Instances when unbalanced execution desirable
 - ▶ No work for opposite thread
 - ▶ Thread waiting on lock
 - ▶ Software determined non uniform balance
 - ▶ Power management
 - ▶ ...
- Solution: Control instruction decode rate
 - ▶ Software/hardware controls 8 priority levels for each thread



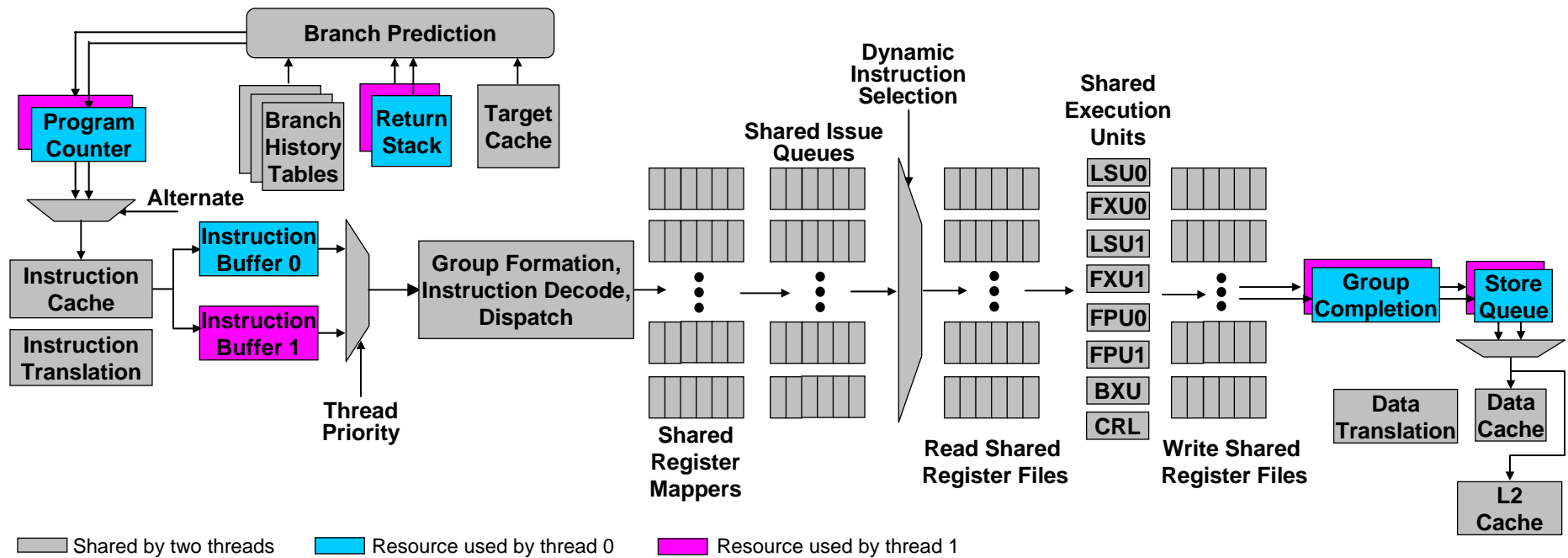
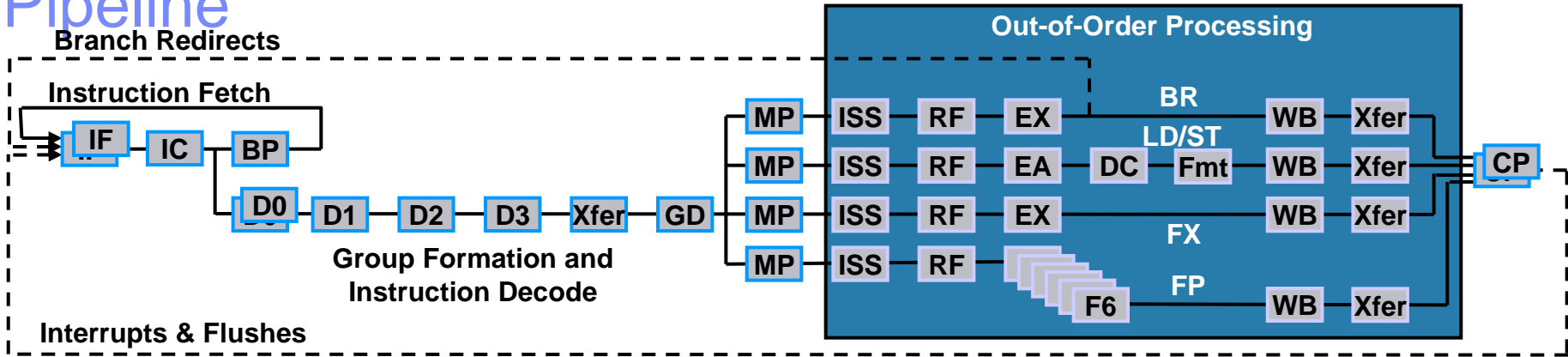
Terminology

- PowerPC Addresses
 - ▶ Virtualization drives more levels
 - ▶ Effective>(SLB)>Virtual>(Page Table)>Real>(LPAR)>Physical
- Instruction Execution
 - ▶ I-fetch
 - ▶ Decode
 - ▶ Dispatch
 - ▶ Issue
 - ▶ Finish
 - ▶ Complete

IBM's Micro Processor Design and Methodology

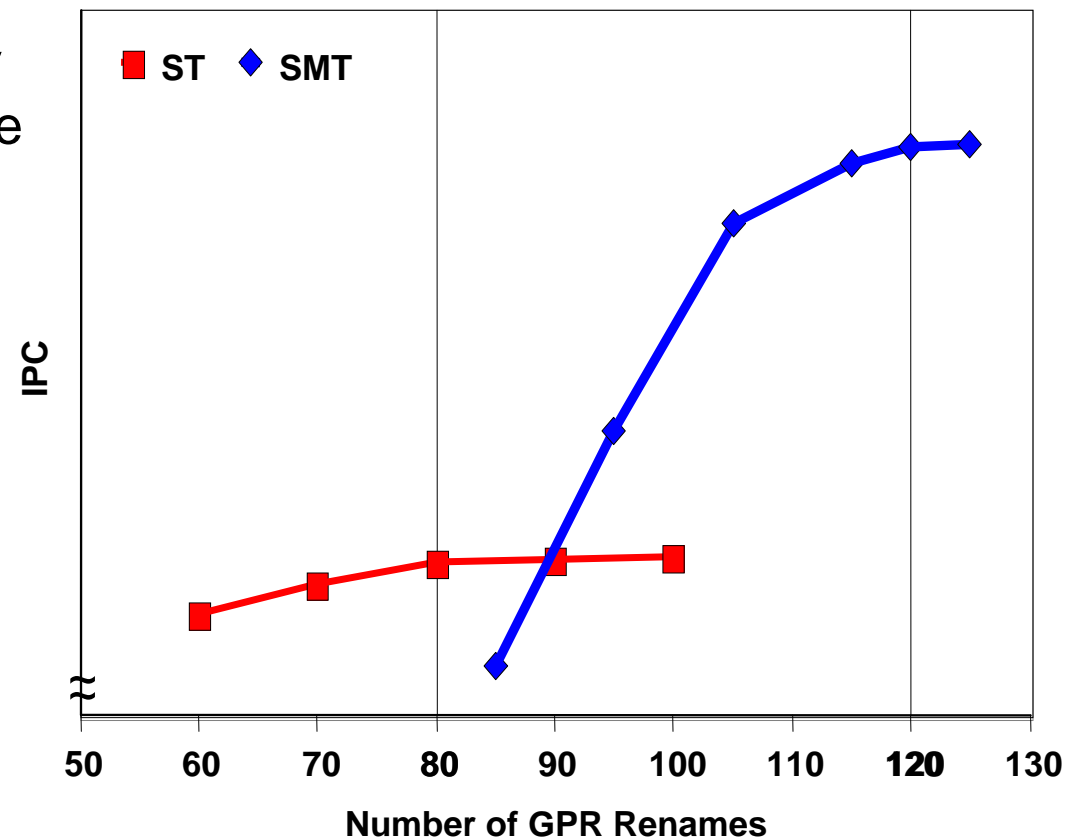


Multithreaded Instruction Flow in Processor Pipeline



Resource Sizes

- Analysis done to optimize every micro-architectural resource size
 - ▶ GPR/FPR rename pool size
 - ▶ I-fetch buffers
 - ▶ Reservation Station
 - ▶ SLB/TLB/ERAT
 - ▶ I-cache/D-cache
- Many Workloads examined
- Associativity also examined

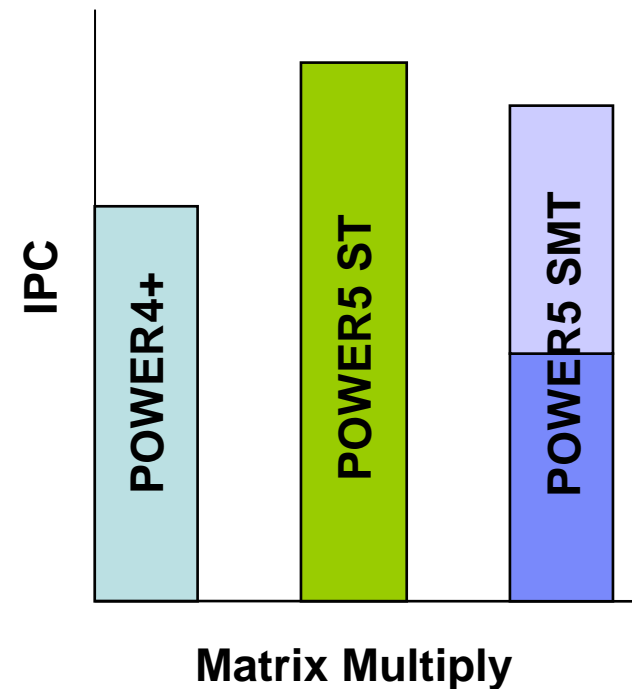


Results based on simulation of an online transaction processing application

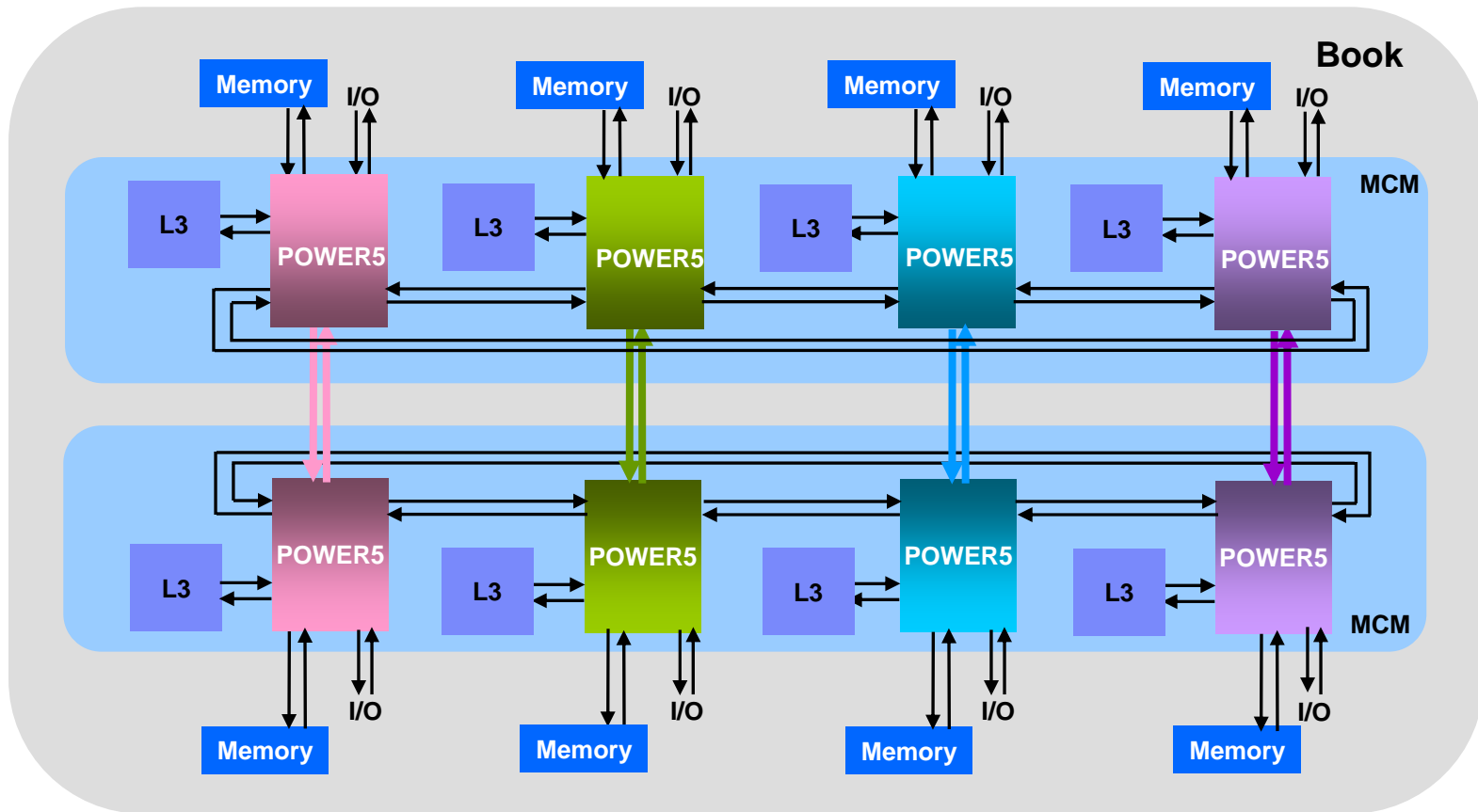
Vertical axis does not originate at 0

Single Thread Operation

- Advantageous for execution unit limited applications
 - ▶ Floating or fixed point intensive workloads
- Execution unit limited applications provide minimal performance leverage for SMT
 - ▶ Extra resources necessary for SMT provide higher performance benefit when dedicated to single thread
- Determined dynamically on a per processor basis

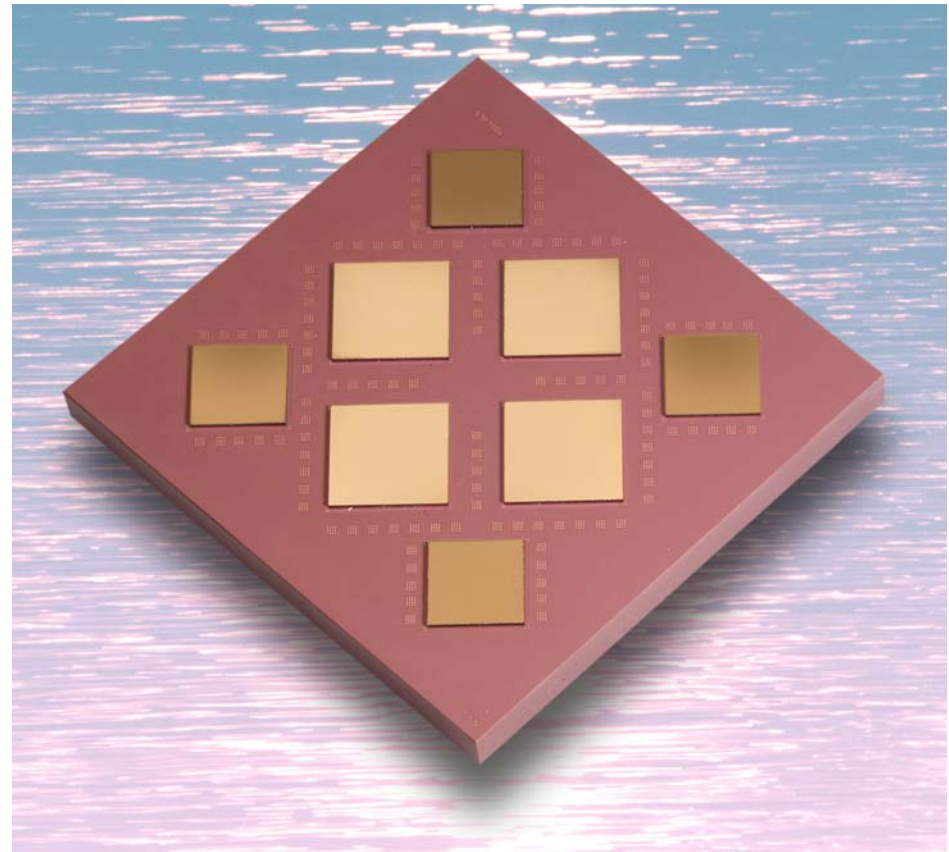


16-way Building Block



POWER5 Multi-chip Module

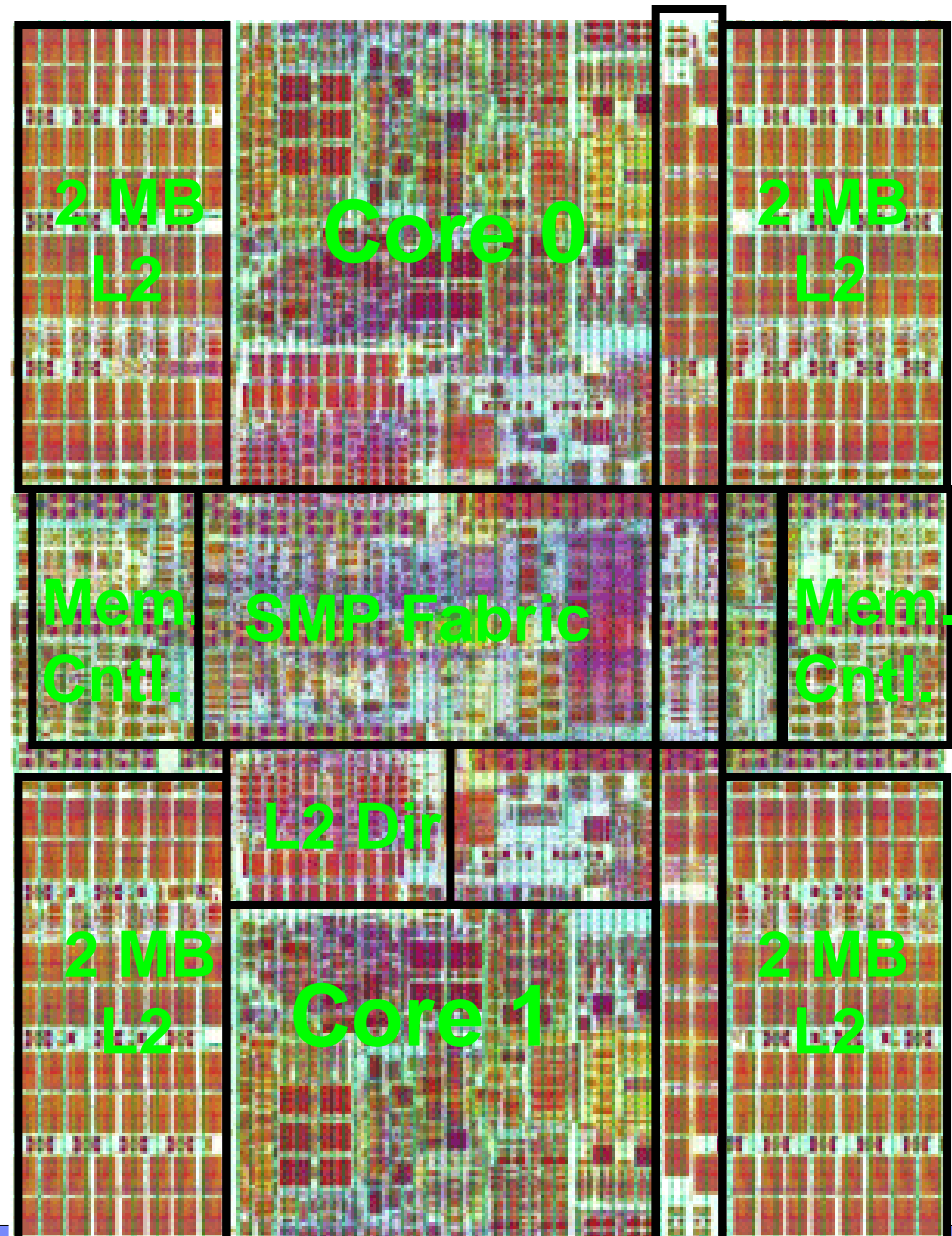
- 95mm × 95mm
- Four POWER5 chips
- Four cache chips
- 4,491 signal I/Os
- 89 layers of metal



POWER6

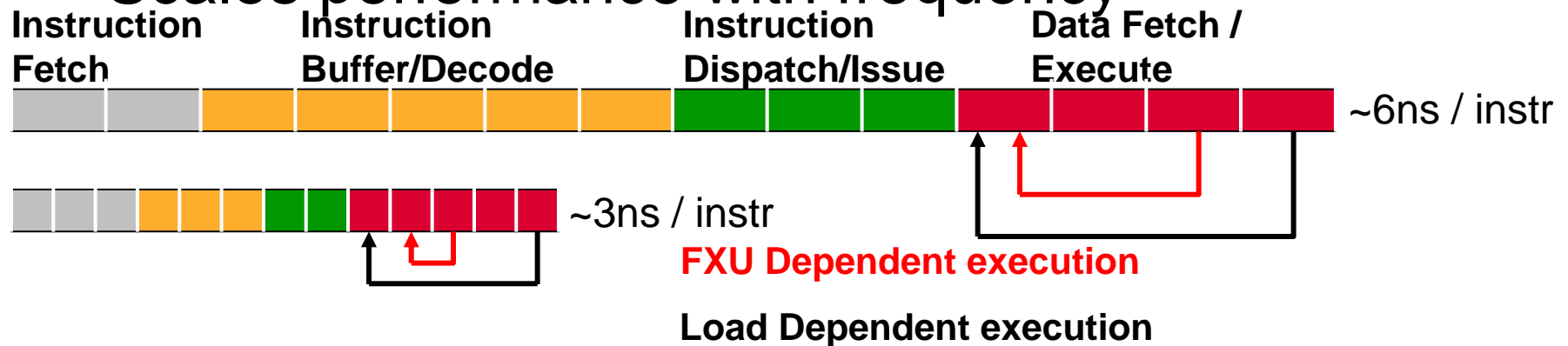
POWER6 Physical Overview

- 5+ GHz operation
- >790M transistors
- 341mm² die
- 65nm SOI process with 10 levels of Cu interconnect and low-k dielectric on 1st 8 levels
- 2 superscalar, SMT cores
- 8 MB Level-2 cache
- Support for 32MB L3
- 2 memory controllers
- Two-tier SMP Fabric



POWER6 Core

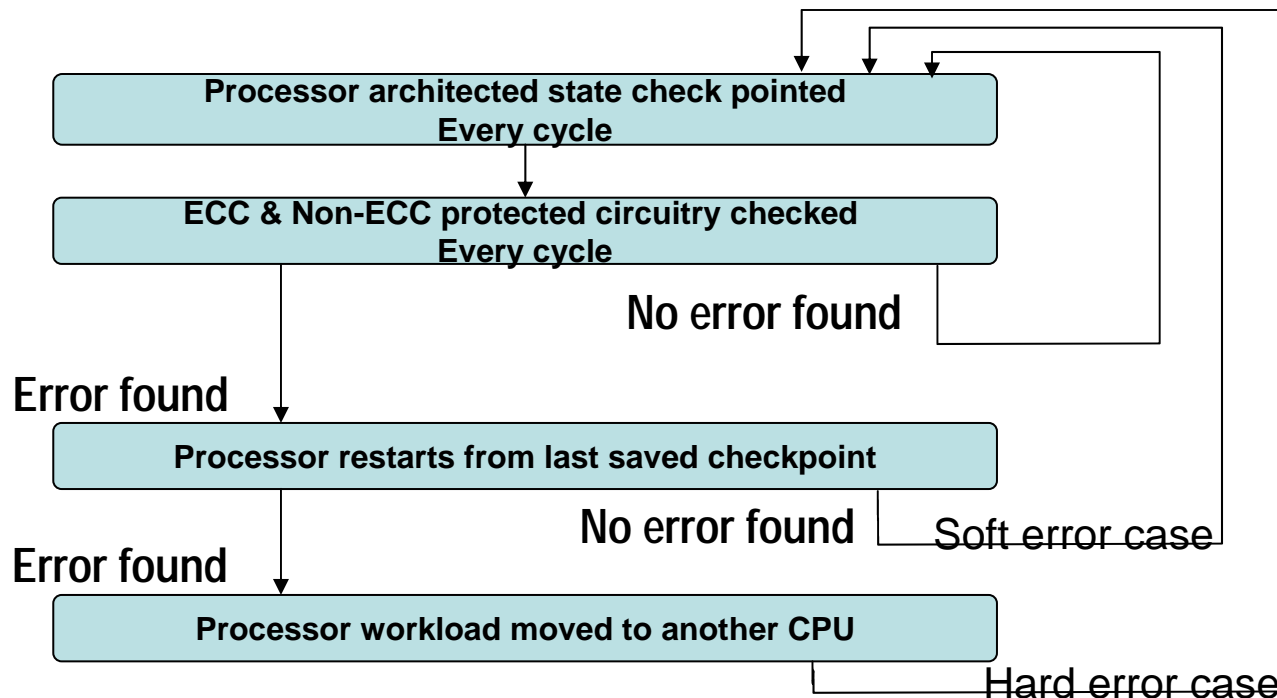
- POWER6 offers ~2X the frequency of POWER5 (4 to 5+ GHz).
- POWER6 maintains POWER5's instruction pipeline depth
 - Achieves same power envelope
 - Scales performance with frequency



- POWER6 extends functionality of POWER5 Core
 - Enhanced 2-way SMT with 7 instruction dispatch
 - 64K, 4-way I Cache; 64K, 8-way D Cache
 - Out of order floating point
 - Speculative load look-ahead and enhanced data prefetch
 - 2 FXU, 2 FPU, 2 LSU, 1 Branch Unit
 - VMX Unit
 - Decimal Floating Point Unit

Bullet-Proof Computing

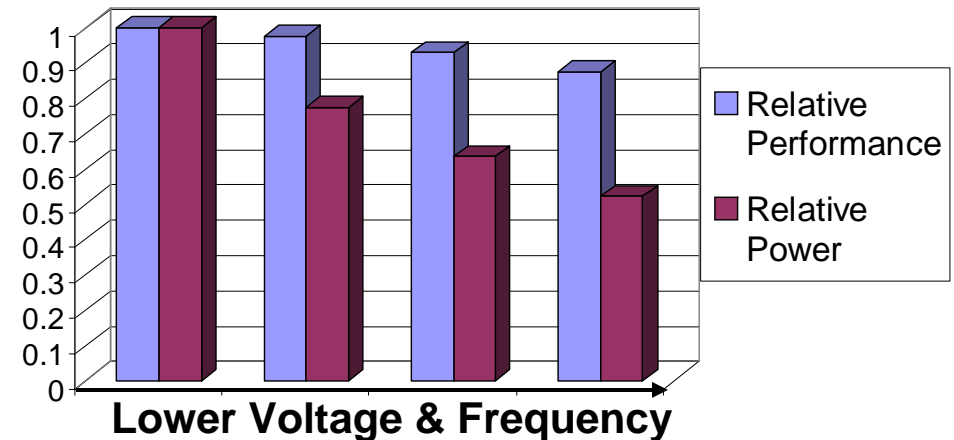
- Error Detection
 - 100% ECC protection for large caches, interfaces, and architected state
 - >99% of small SRAMs and Register files parity protected
 - Dataflow & control protected by parity and logical consistency checkers
 - Experiments indicate ~3400 random soft errors needed to cause 1 undetected data corruption
- Error Recovery



POWER6 Enables Energy Efficiency

- Supports a variety of energy policies
 - Power capping
 - Energy reduction
 - Acoustic optimization
 - Performance optimization
- Extensive hardware controls
 - Wide voltage / frequency range
 - Architected idle state (Nap) for increased clock gating
 - Memory request throttling
 - Power down of memory ranks
 - Programmable fetch / dispatch throttling

Benefits of Voltage Frequency Slewing



Impact of Nap Mode on Power

