



# The Cell Processor

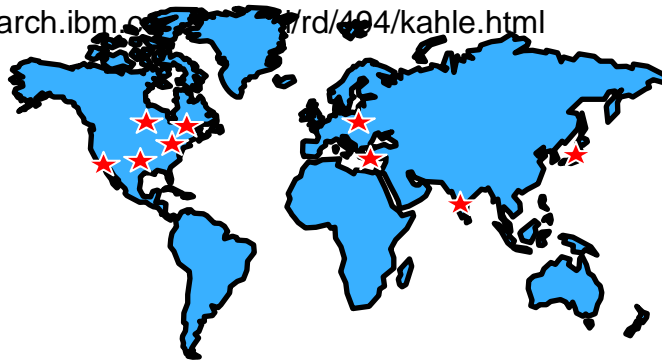
Dr. Jens Leenstra

[leenstra@de.ibm.com](mailto:leenstra@de.ibm.com)

IBM STSM, Cell Processor Development

# Cell History

- IBM, SCEI/Sony, Toshiba Alliance formed in 2000
- Design Center opened in March 2001
- Based in Austin, Texas
- February 7, 2005: First technical disclosures
- May 16, 2005: First public demonstrations at E3
- August 25, 2005: Release of technical documentation
- November, 2005: Release of Cell BE SDK
  - Cell Broadband Engine Architecture documentation can be found at:
    - ✓ <http://www.ibm.com/developerworks/power/cell>
  - Additional publications on Cell can be downloaded from:
    - ✓ <http://www.ibm.com/chips/techlib/techlib.nsf/products/Cell>
  - A paper on Cell in the upcoming issue of the IBM Journal of Research and Development can be found at:
    - ✓ <http://www.research.ibm.com/jrd/404/kahle.html>



SONY

TOSHIBA

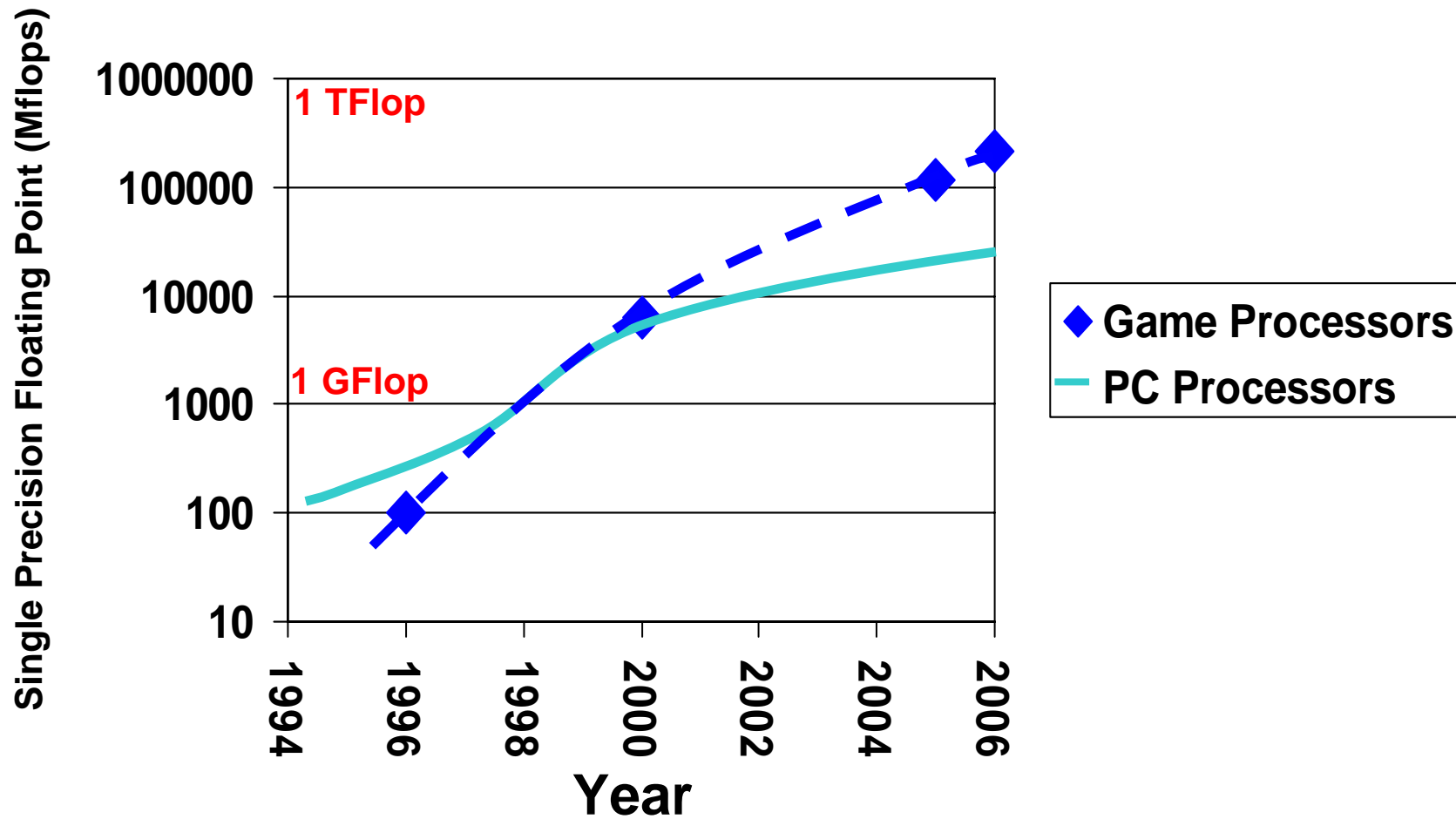
IBM

# Introducing Cell

- **Cell is an accelerator extension to Power**
  - Built on a Power ecosystem
  - Used best know system practices for processor design
- **Sets a new performance standard**
  - Exploits parallelism while achieving high frequency
  - Supercomputer attributes with extreme floating point capabilities
  - Sustains high memory bandwidth with smart DMA controllers
- **Designed for natural human interaction**
  - Photo-realistic effects
  - Predictable real-time response
  - Virtualized resources for concurrent activities
- **Designed for flexibility**
  - Wide variety of application domains
  - Highly abstracted to highly exploitable programming models
  - Reconfigurable I/O interfaces
  - Virtual trusted computing environment for security

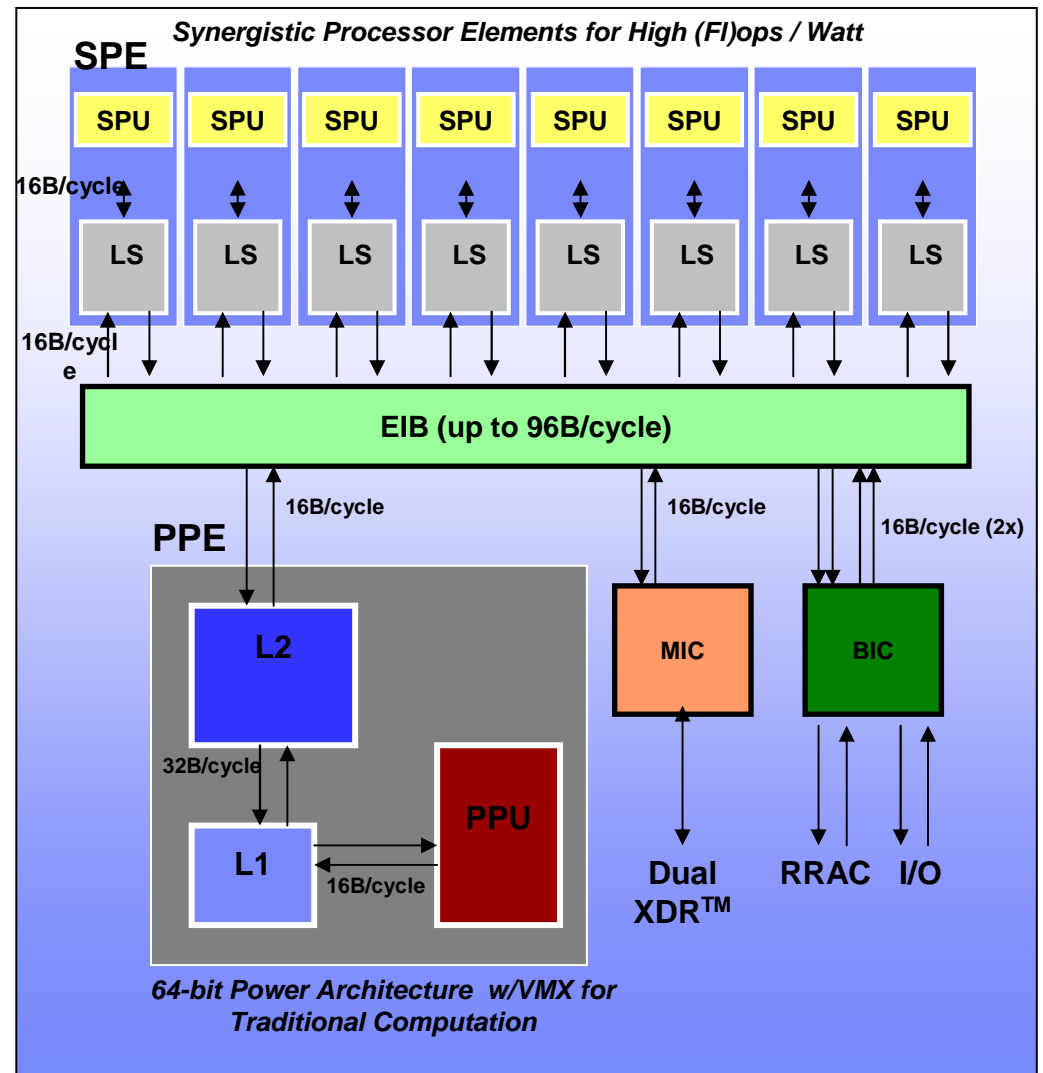
# Performance over Time

Game processors drive media performance

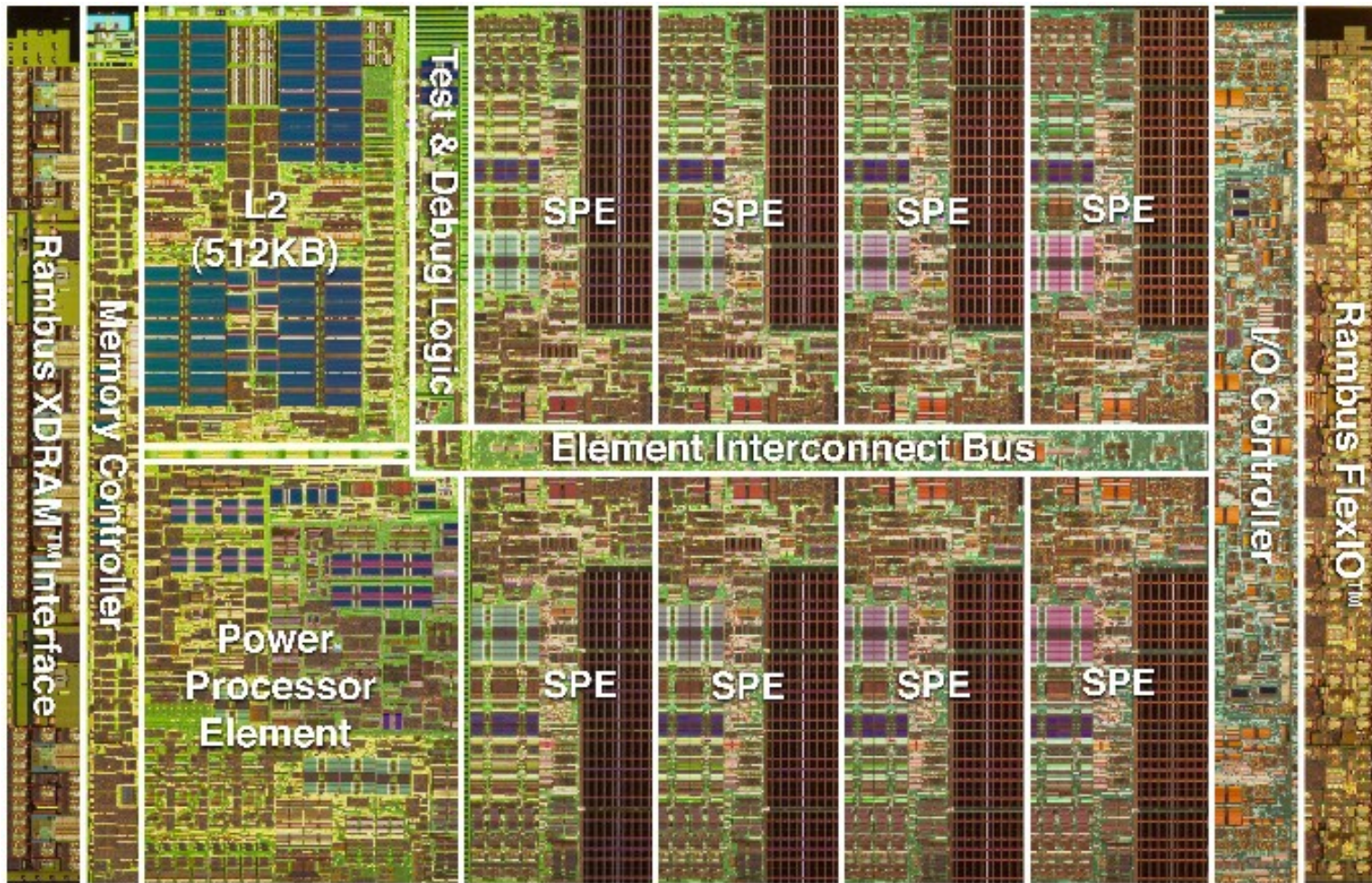


# Key Features

- **The first generation CELL processor consists of:**
  - A Power Processor Element (PPE)
  - 8 Synergistic Processor Elements (SPE)
  - A high bandwidth Element Interconnect Bus (EIB)
  - Two configurable non-coherent IO interfaces (BIC)
  - A Memory Interface Controller (MIC)
  - A Pervasive unit that supports extensive test, monitoring, and debug functions



# Cell Broadband Engine – 235mm<sup>2</sup>



# Cell Processor Components

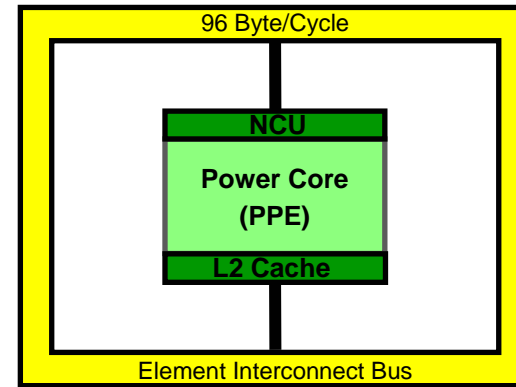
## Power Processor Element (PPE):

- General purpose, 64-bit RISC processor (PowerPC AS 2.0.2)
- 2-Way hardware multithreaded
- L1 : 32KB I ; 32KB D
- L2 : 512KB
- Coherent load / store
- VMX-32
- Realtime Controls
  - Locking L2 Cache & TLB
  - Software / hardware managed TLB
  - Bandwidth / Resource Reservation
  - Mediated Interrupts

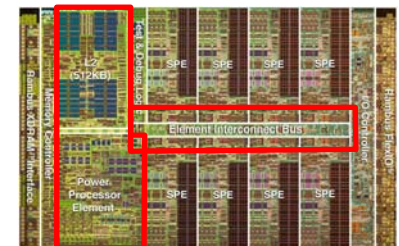
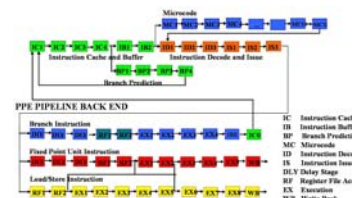
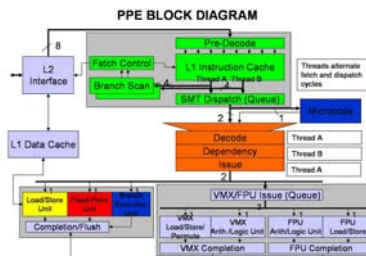
## Element Interconnect Bus (EIB):

- Four 16 byte data rings supporting multiple simultaneous transfers per ring
- 96Bytes/cycle peak bandwidth
- Over 100 outstanding requests

*In the Beginning  
– the solitary Power Processor*



*Custom Designed  
– for high frequency, space,  
and power efficiency*



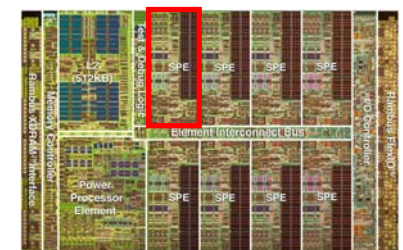
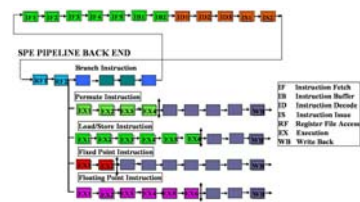
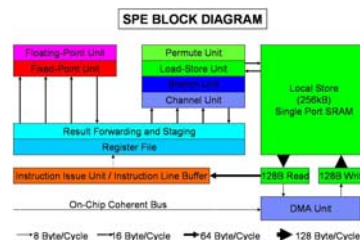
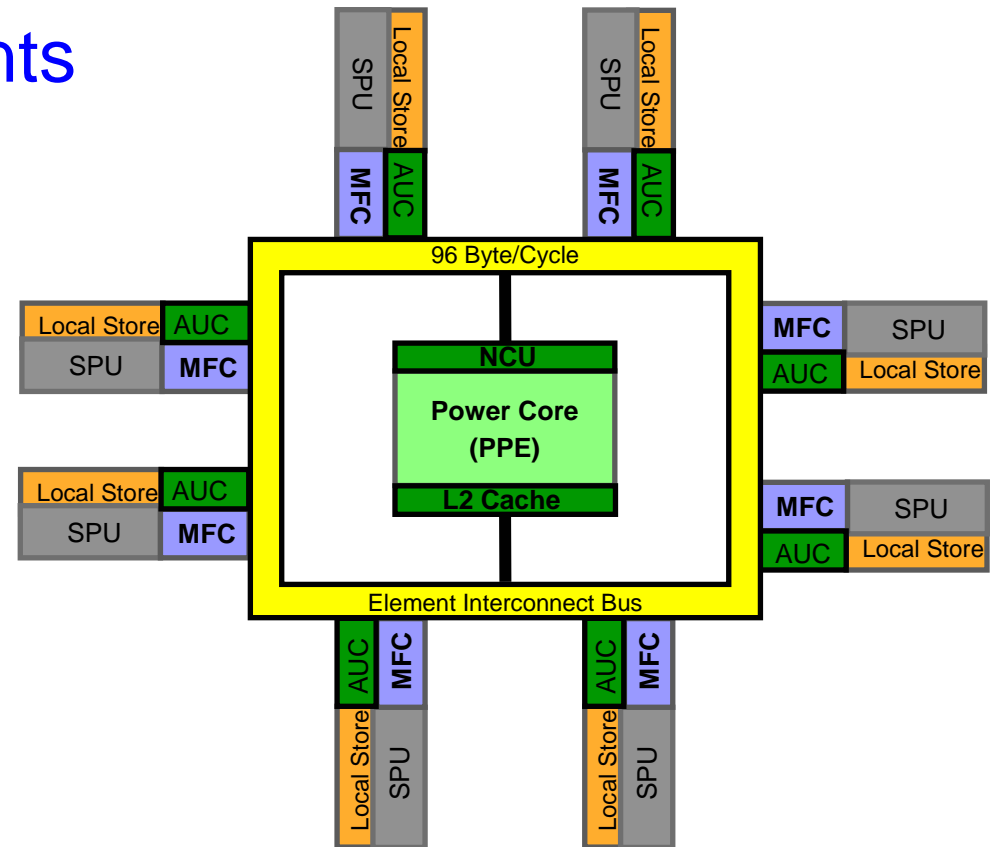
# Cell Processor Components

## Synergistic Processor Element (SPE):

- Provides the computational performance
- Simple RISC User Mode Architecture
  - Dual issue VMX-like
  - Graphics SP-Float
  - IEEE DP-Float
- Dedicated resources: unified 128x128-bit RF, 256KB Local Store
- Dedicated DMA engine: Up to 16 outstanding requests

## Memory Management & Mapping

- SPE Local Store aliased into PPE system memory
- MFC/MMU controls / protects SPE DMA accesses
  - Compatible with PowerPC Virtual Memory Architecture
  - SW controllable using PPE MMIO
- DMA 1,2,4,8,16,128 -> 16Kbyte transfers for I/O access
- Two queues for DMA commands: Proxy & SPU



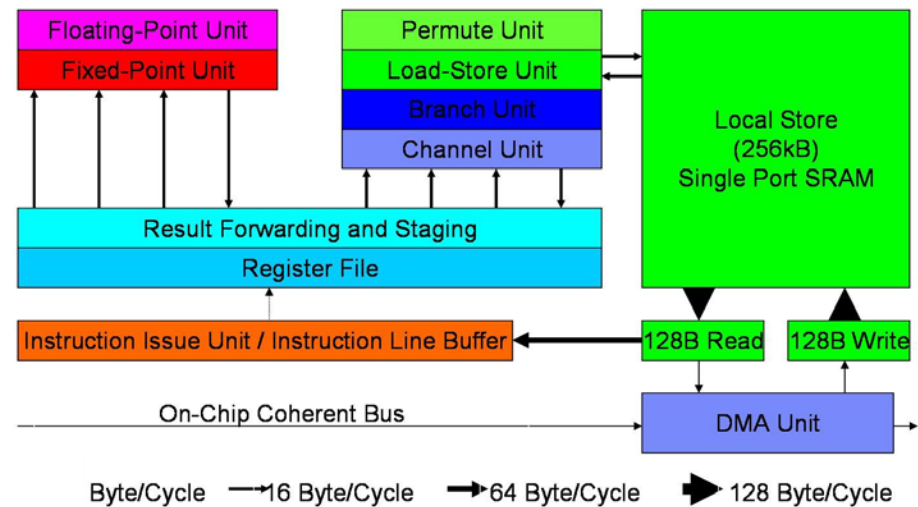


# Cell Synergistic Processor Element (SPE)

## Synergistic Processor Element (SPE):

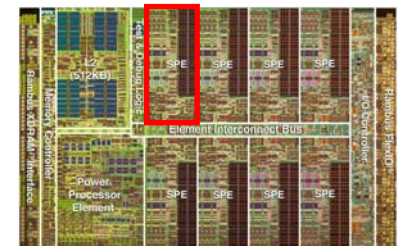
- Provides the computational performance
- Simple RISC User Mode Architecture
  - Dual issue VMX-like
  - Graphics SP-Float
  - IEEE DP-Float
- Dedicated resources: unified 128x128-bit RF, 256KB Local Store
- Dedicated DMA engine: Up to 16 outstanding requests

**SPE BLOCK DIAGRAM**



**SPE PIPELINE BACK END**

IF	Instruction Fetch
IB	Instruction Buffer
ID	Instruction Decode
IS	Instruction Issue
RF	Register File Access
EX	Execution
WB	Write Back



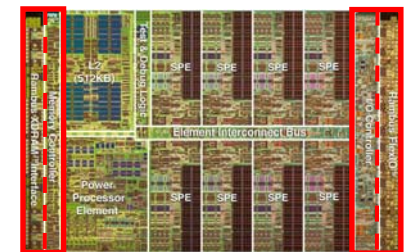
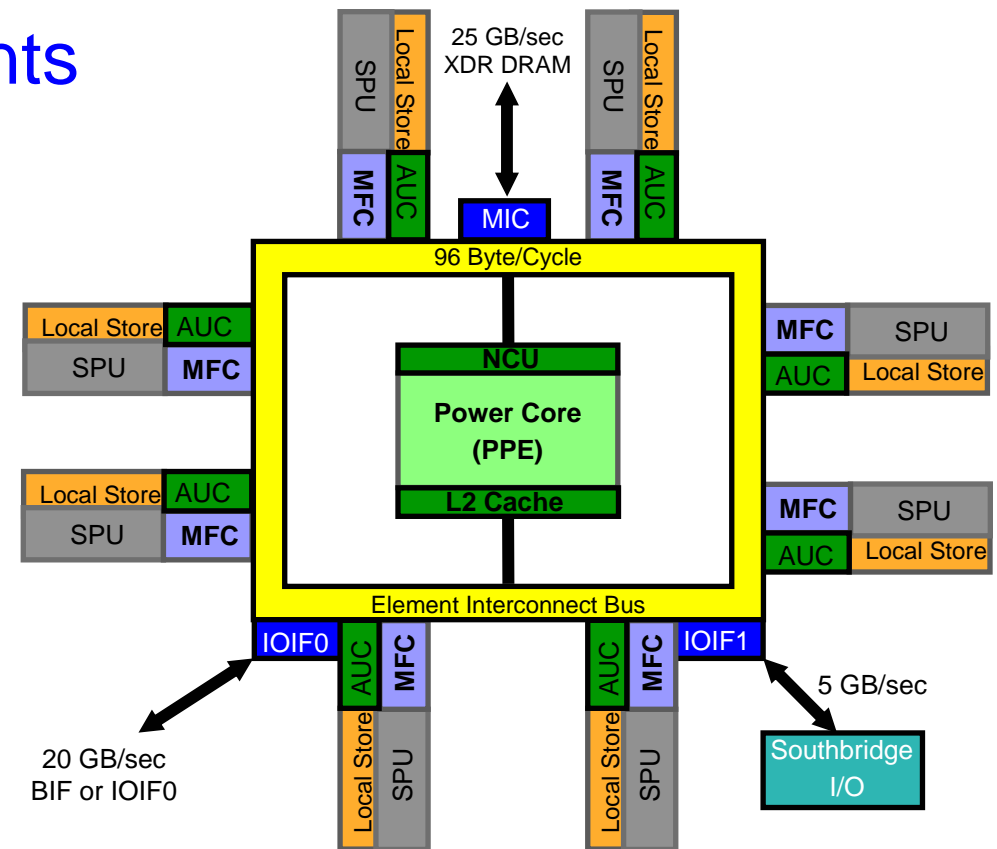
# Cell Processor Components

## Broadband Interface Controller (BIC):

- Provides a wide connection to external devices
- Two configurable interfaces (60GB/s @ 5Gbps)
  - Configurable number of bytes
  - Coherent (BIF) and / or I/O (IOIFx) protocols
- Supports two virtual channels per interface
- Supports multiple system configurations

## Memory Interface Controller (MIC):

- Dual XDR™ controller (25.6GB/s @ 3.2Gbps)
- ECC support
- Suspend to DRAM support



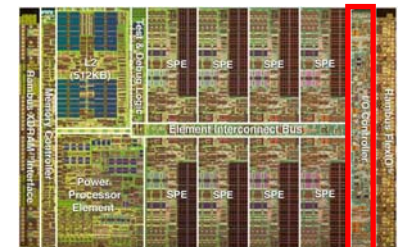
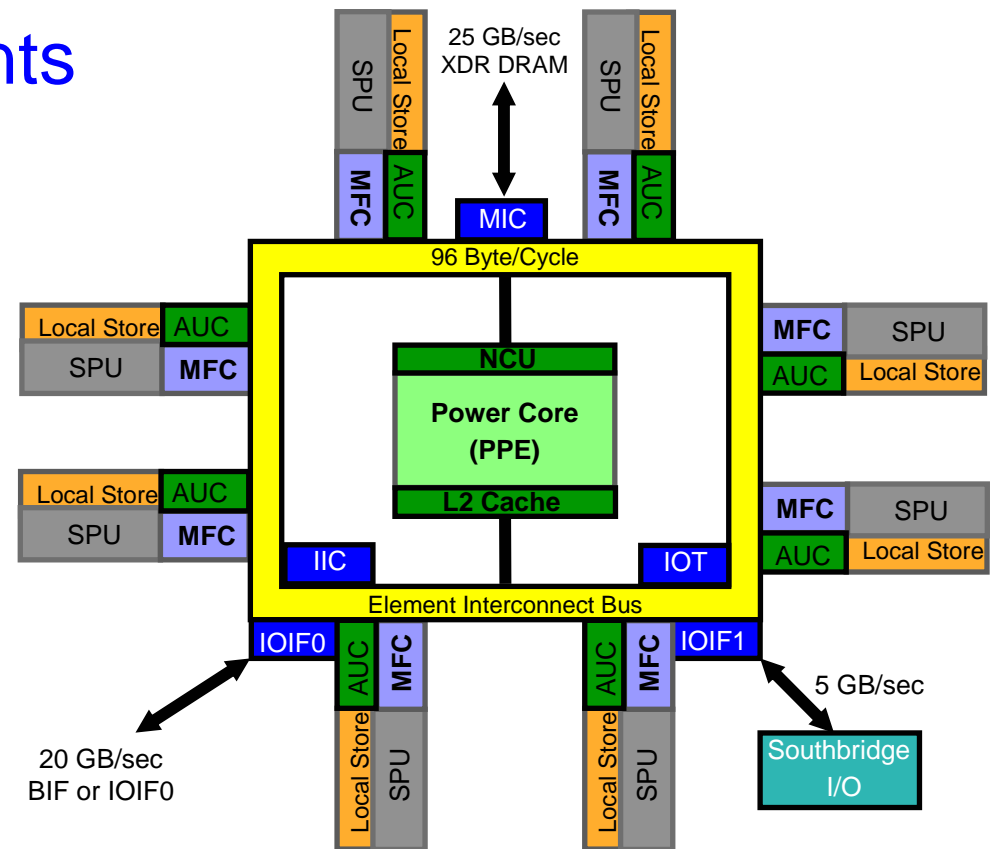
# Cell Processor Components

## Internal Interrupt Controller (IIC)

- Handles SPE Interrupts
- Handles External Interrupts
  - From Coherent Interconnect
  - From IOIF0 or IOIF1
- Interrupt Priority Level Control
- Interrupt Generation Ports for IPI
- Duplicated for each PPE hardware thread

## I/O Bus Master Translation (IOT)

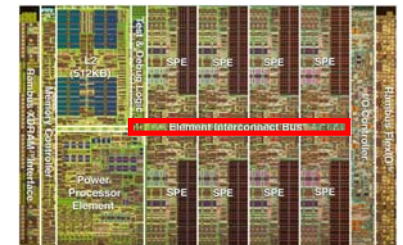
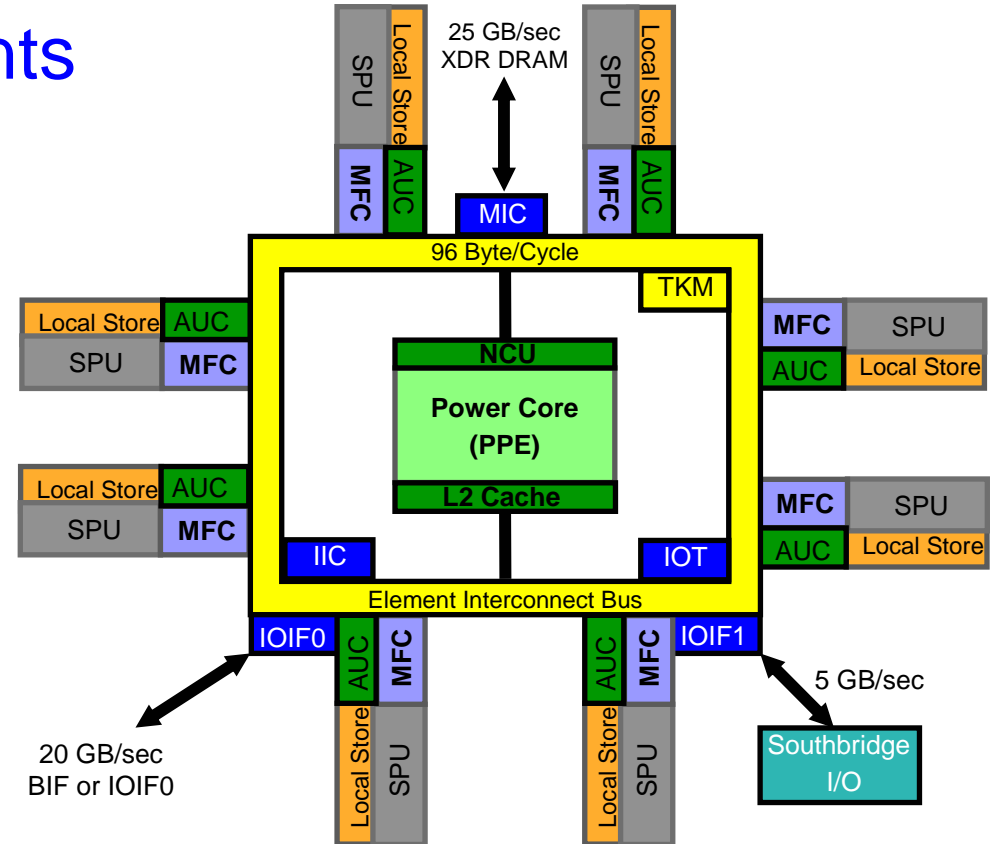
- Translates Bus Addresses to System Real Addresses
- Two Level Translation
  - I/O Segments (256 MB)
  - I/O Pages (4K, 64K, 1M, 16M byte)
- I/O Device Identifier per page for LPAR
- IOST and IOPT Cache – hardware / software managed



# Cell Processor Components

## Token Manager (TKM):

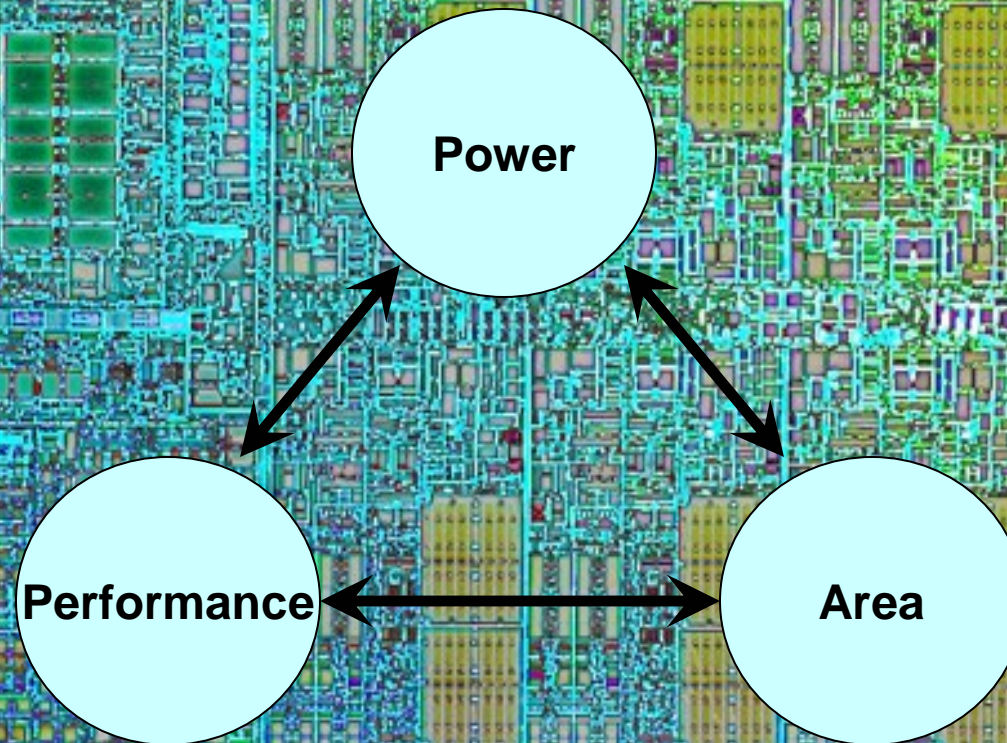
- Bandwidth / Resource Reservation for shared resources
- Optionally enabled for RT tasks or LPAR
- Multiple Resource Allocation Groups (RAGs)
- Generates access tokens at configurable rate for each allocation group
  - 1 per each memory bank (16 total)
  - 2 for each IOIF (4 total)
- Requestors assigned RAG ID by OS / hypervisor
  - Each SPE
  - PPE L2 / NCU
  - IOIF 0 Bus Master
  - IOIF 1 Bus Master
- Priority order for using another RAGs unused tokens
- Resource over committed warning interrupt



# Implementation Challenges

- Technology Scaling
  - Minimize cross chip variations in delay and leakage
  - Array bit cell stability, writability, yield
  - Growing impact of wire RC vs. device speed
- 11FO4 design within air-cooled power envelope
  - Power, Clock, Signal Distribution variation due to hot spots, inductance effects, etc
  - Multi Clock domains
  - Intra-Chip interconnections
  - Global Optimization with “triple constraints”: Frequency, Power, Cost (Die Size and Yield)
- Design for Modularity and Scalability
  - SoC with black box IP
  - Tools & Methodology

# Tripple constrains



# Circuit Design Practices

- Strict design guidelines to minimize design variations
  - Layout topology check and DFM rules for yield
  - Circuit topology and electrical checks
  - Global active clock pulse limiter for dynamic circuits
  - Hold time margin scale with clock path delay
- Reduce design sensitivity to technology leakage
  - Limited dynamic logic circuit usage
  - No Low-Vt devices
- Array yield focus
  - Array redundancy for bit cell stability fails
  - Reduced cell stress during read

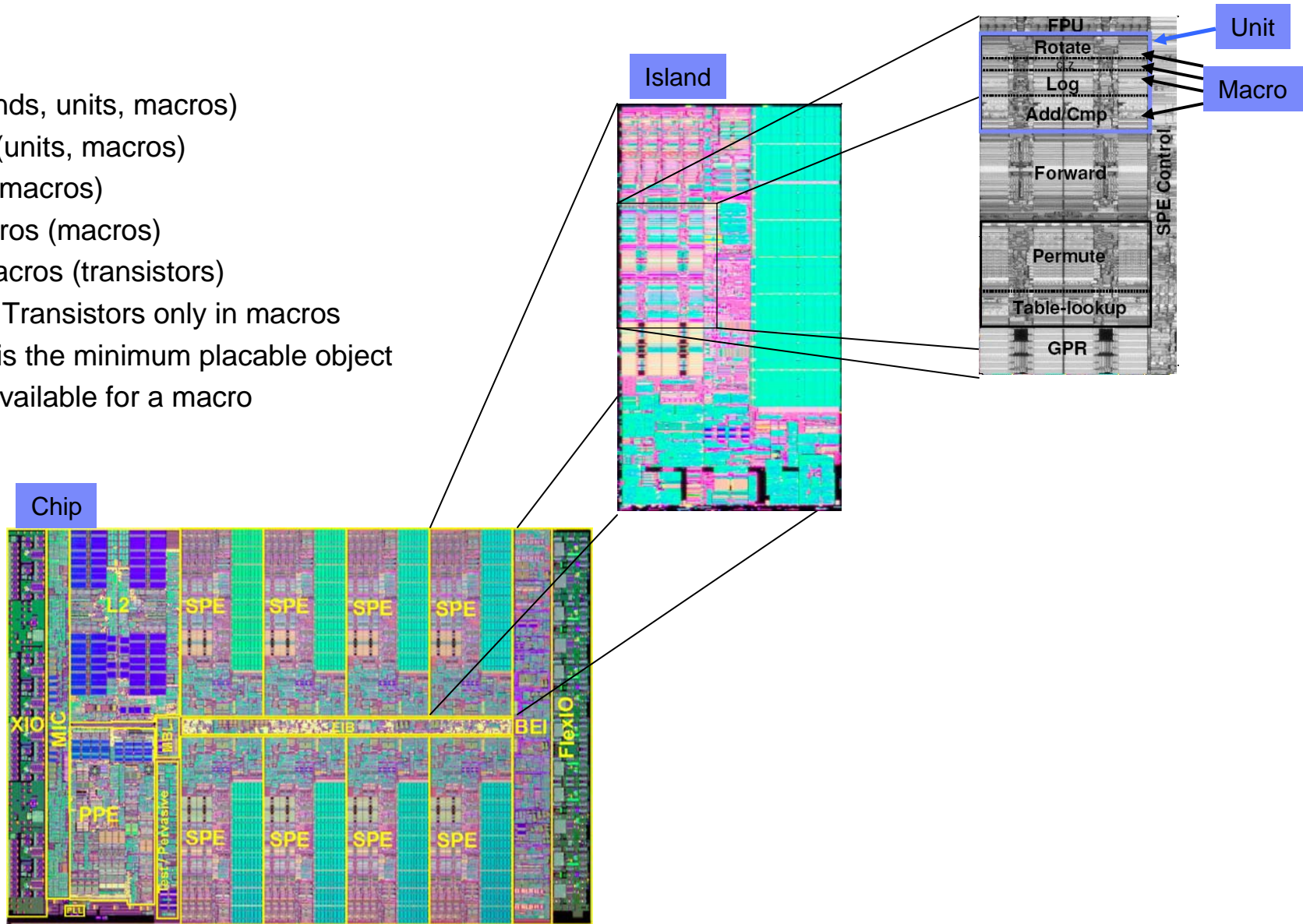
# Illustration of Design Hierarchy

- Chip (islands, units, macros)
- islands (units, macros)
- units (macros)
- macros (macros)
- macros (transistors)

Attention! Transistors only in macros

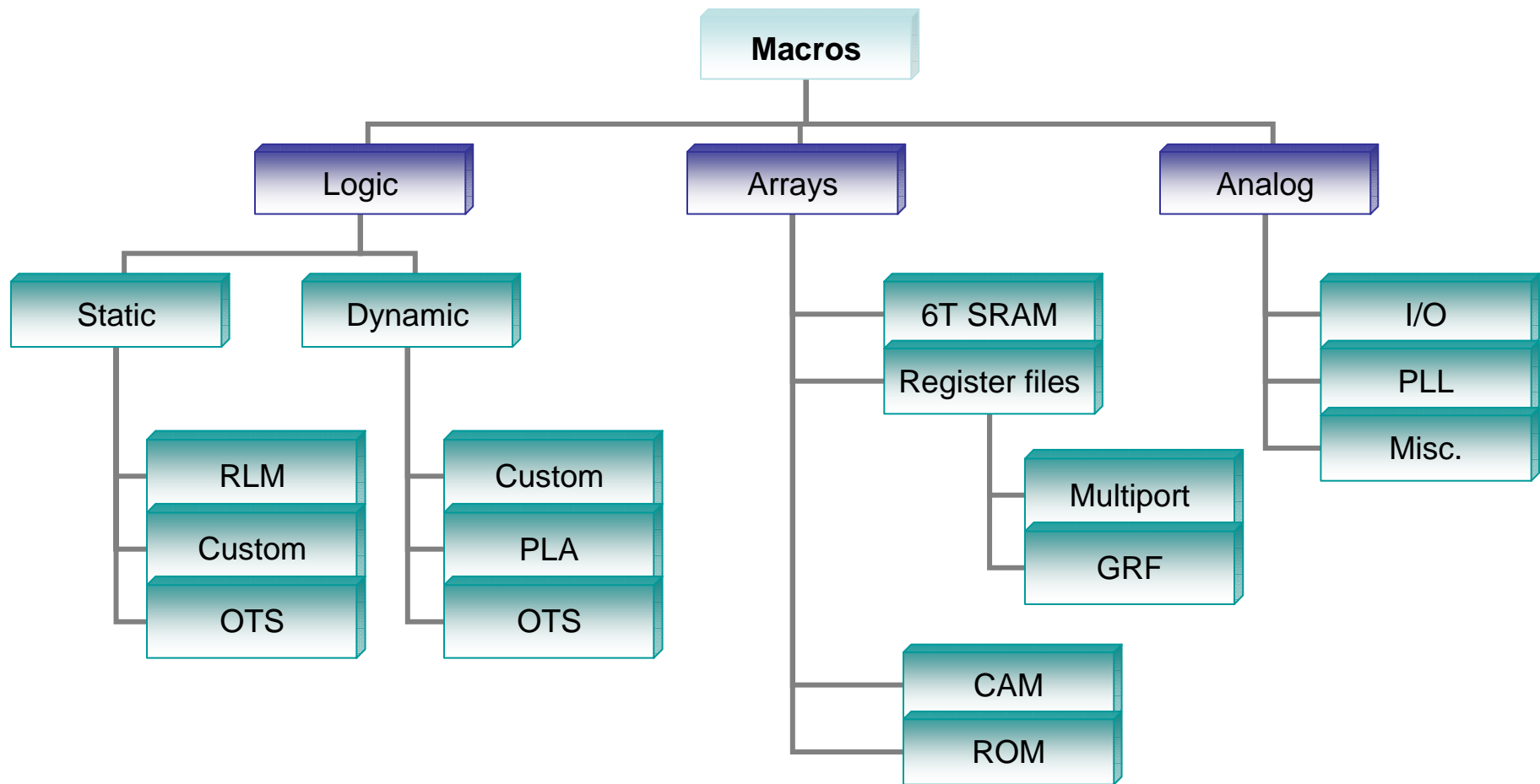
A macro is the minimum placable object

All rules available for a macro





# Macro types



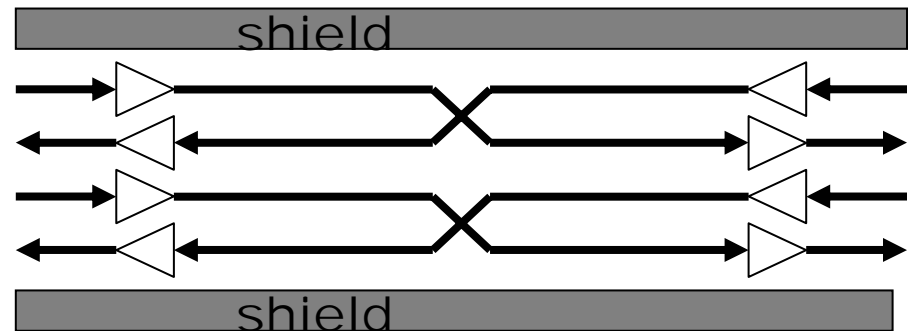
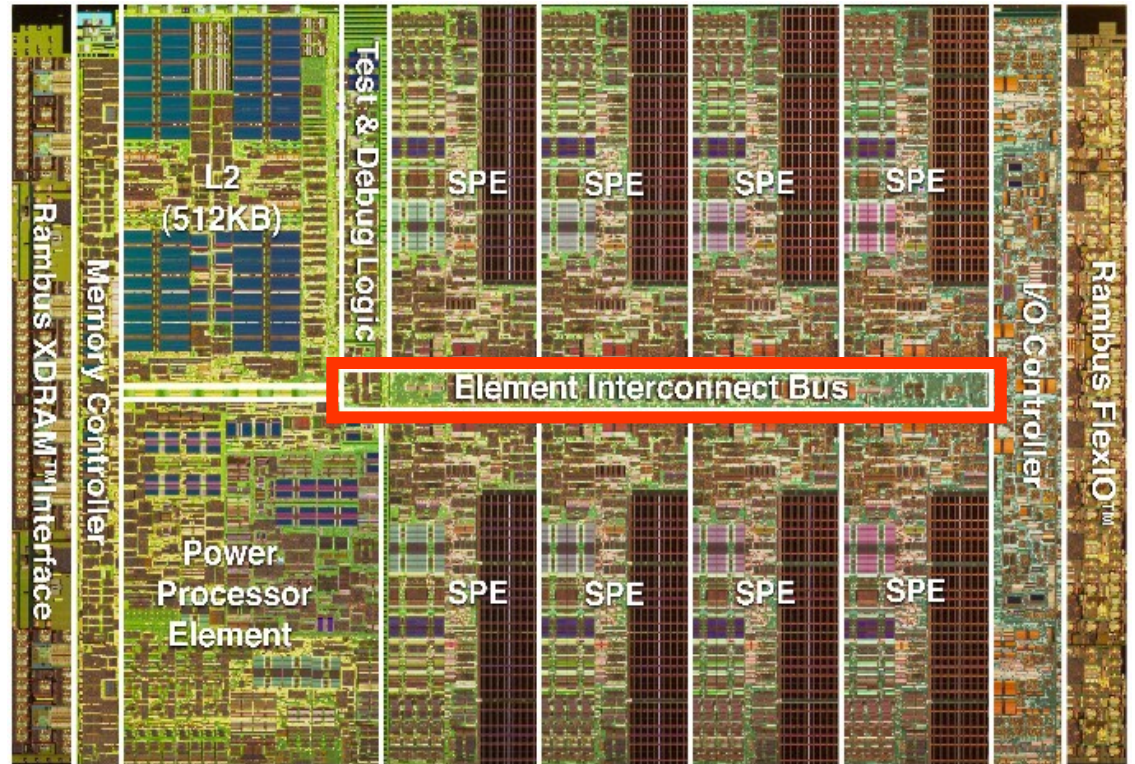
# Static vs. Dynamic Circuits

- **Advantage of using dynamic circuits**
  - Less node capacitance, predominant NFETs in evaluation path (speed)
  - $\mu$ -architectural (combining multiple stages into fewer stages)
  - Tends to be less area
- **Advantage of using static circuits**
  - Design easiness
  - Low switching factor (power)
  - Good tool coverage
  - Technology independent
- **Selective usages of dynamic circuits**
  - Use dynamic circuits only where absolutely necessary
    - Selective data flow logic macros
    - Arrays and register files
    - Dynamic programmable logic array (PLA Control)
    - Multiplexer latch

2/8/2007 Dynamic circuits occupy 19% of non-SRAM area

# Engineering Busses

- EIB design
  - Four 128-bit data rings
  - 64-bit tag
  - Two twisted pairs of wires interleaved with shields.
- Engineering for signal integrity
  - 50% of global nets engineered
  - 32K repeaters added

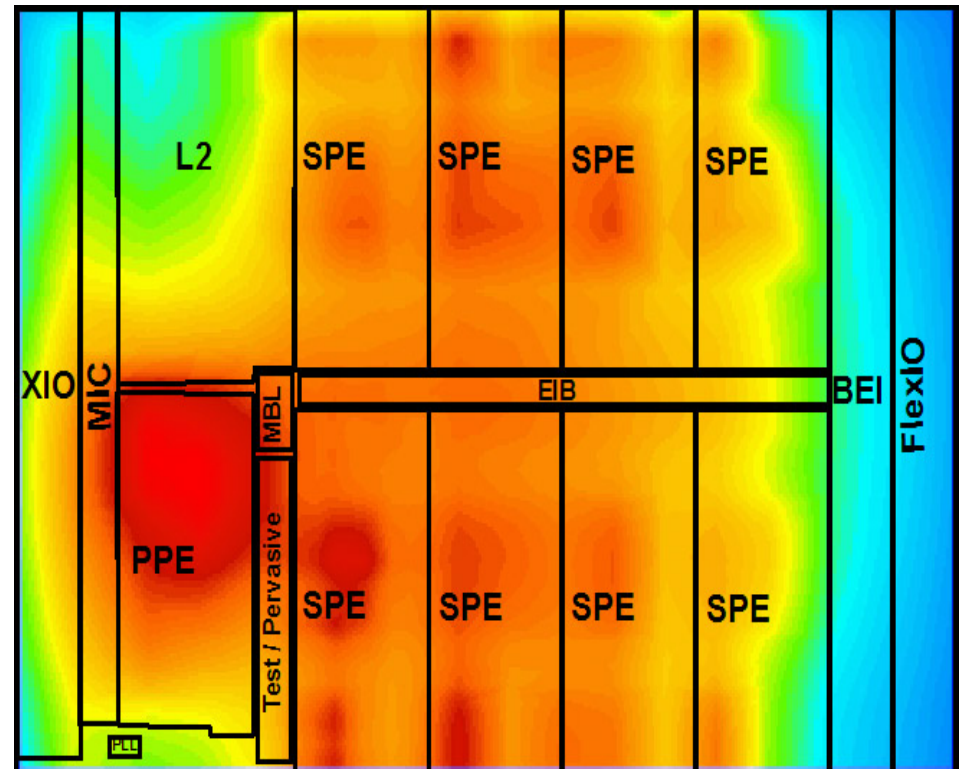


# Power Management Practices

- Dynamic power is controlled by fine-grain clock gating
- Leakage power is managed by adding lower vt devices only where necessary
- Accurate power estimation
  - Macro level uses circuit simulation and generates a power rule (0-50% input switching)
  - Partition/Chip level uses behavior simulation with specific workloads and macro level power rules

# Hot Spot Analysis

- Extensive thermal analysis early in the design cycle
- Power maps created for use with package and heat sink models.
- Steady state and transient thermal behavior simulated
- Analysis feedback to chip floorplan and thermal sensor design

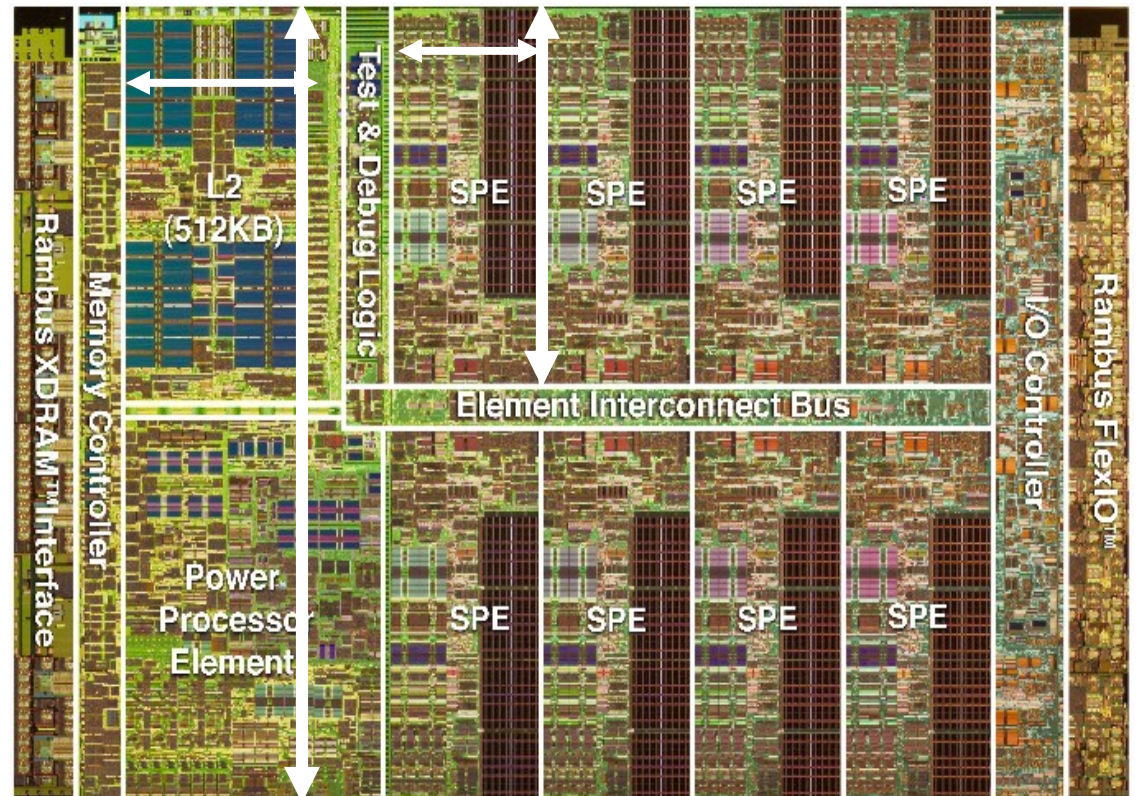


# Test / Pervasive Design Practices

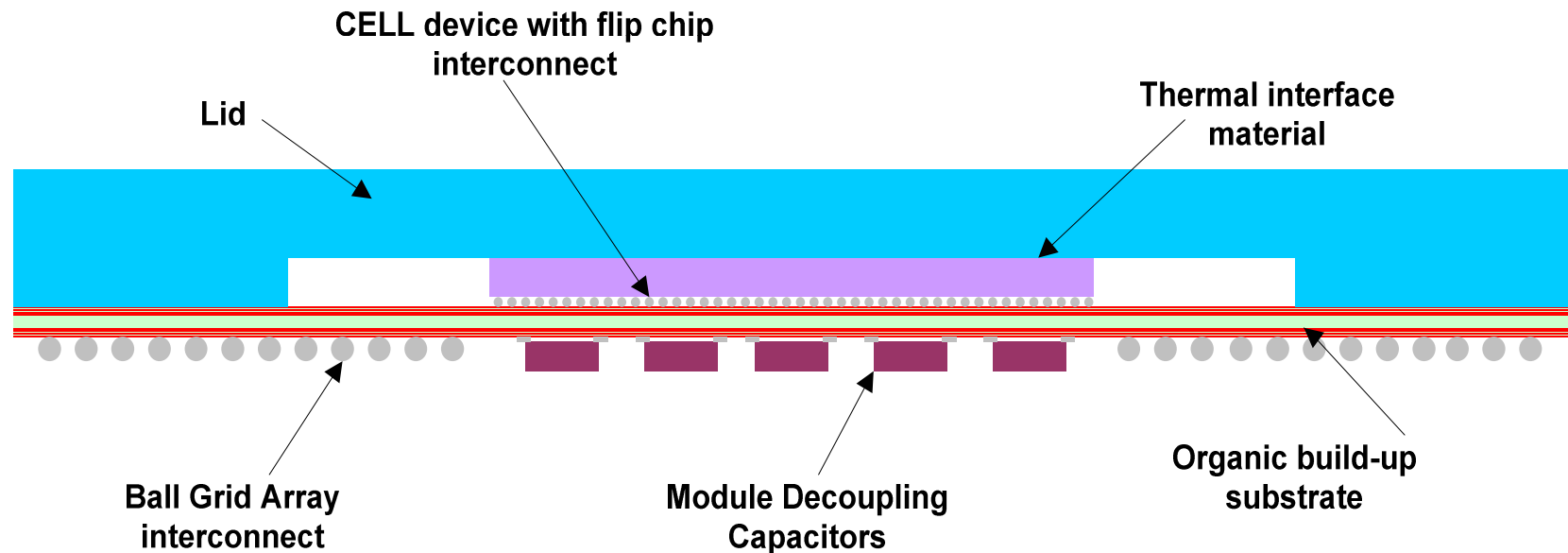
- Modularity at a Global Level
- Distributed test functions
  - LBIST engine for cores
  - ABIST engine for arrays
- Distributed debug features
  - Common debug bus
  - Centralized trace array
- Centralized test and pervasive control
  - Common strategy for logic debug and performance monitoring
  - Monitor some activity externally
- Early focus on design bring up
  - At speed test (internal chip scan, ABIST, programmable LBIST)
  - On chip logic analyzer for debug
  - On chip performance monitor
  - Isolate, start, stop, step controls for lab debug.

# Chip Assembly

- Early design planning for optimum partition aspect ratio
- Modular construction for chip assembly
- 241M transistors
- At top level, 17 major physical partitions, 8912 discrete blocks.
- The chip total, across all levels of hierarchy, 177K blocks, 580K repeaters and 1.55M signal nets



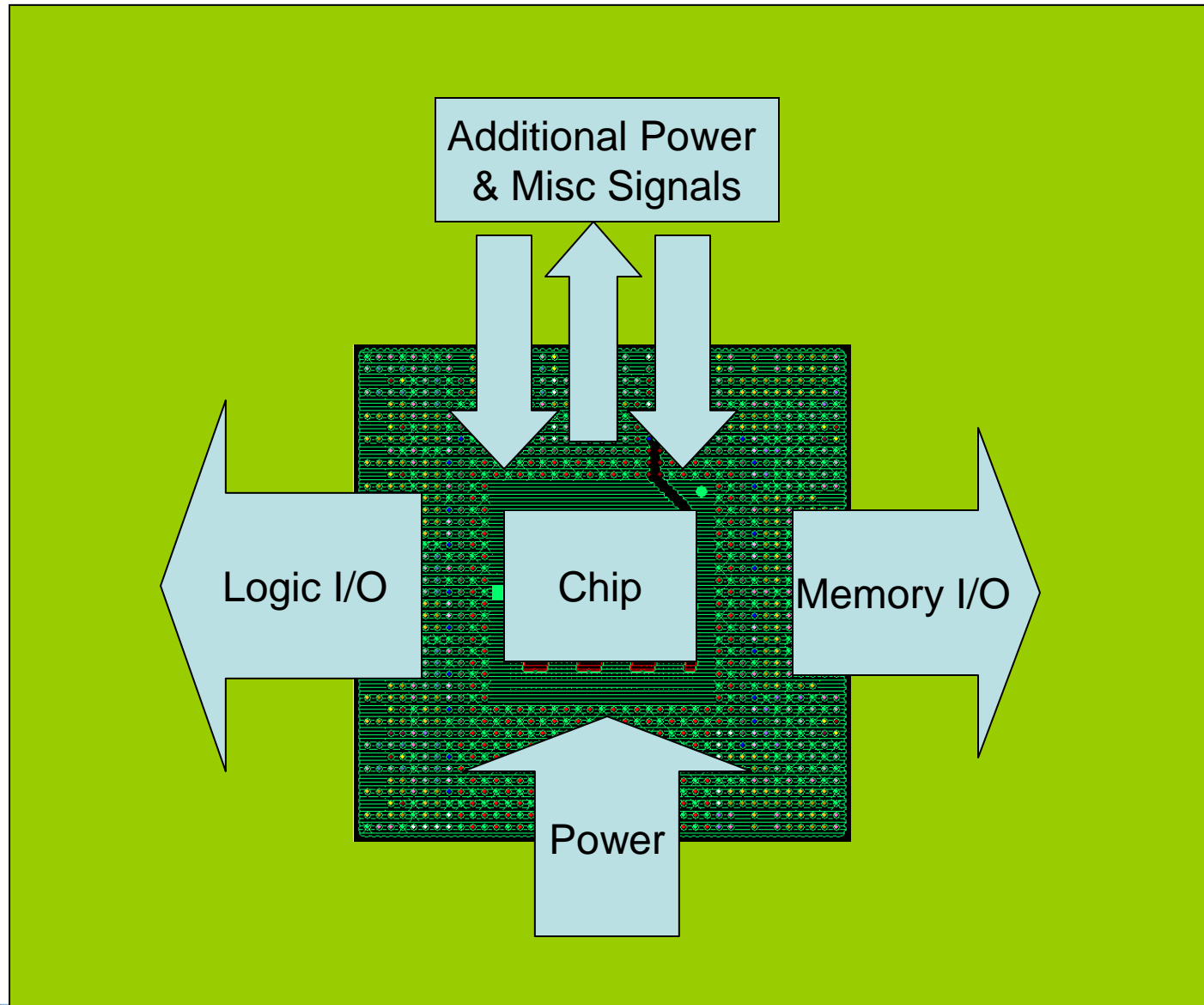
# Packaging



- 3349 C4s, four regions of different pitch
- 3-2-3 Organic 42.5mm x 42.5mm Flip Chip BGA
- High performance thermal lid
- Bottom-side IDC capacitors

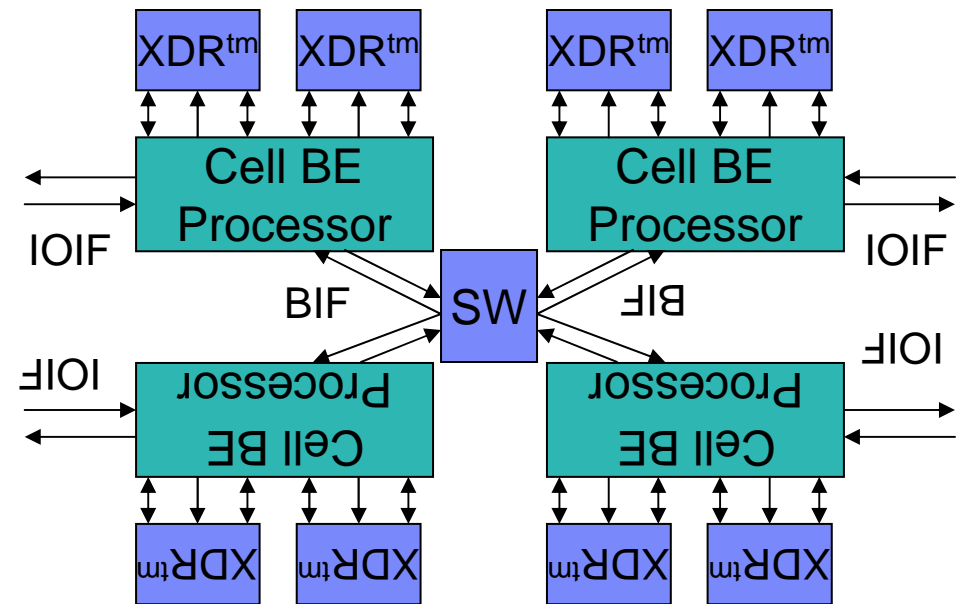
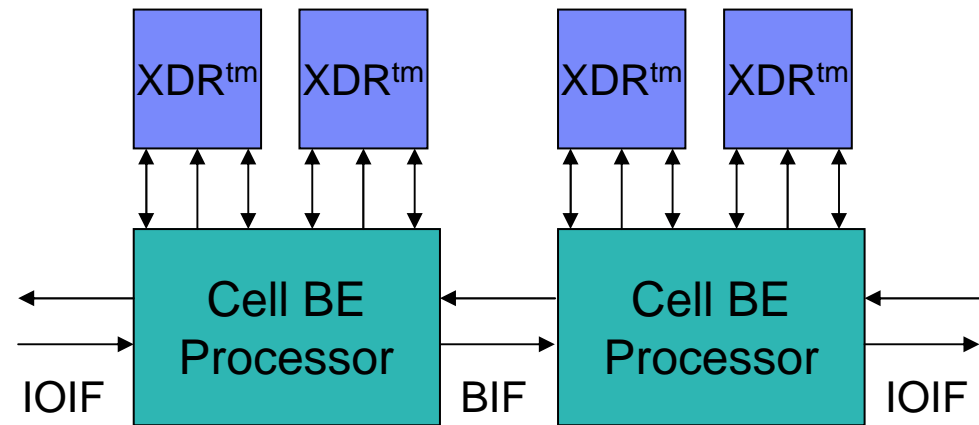
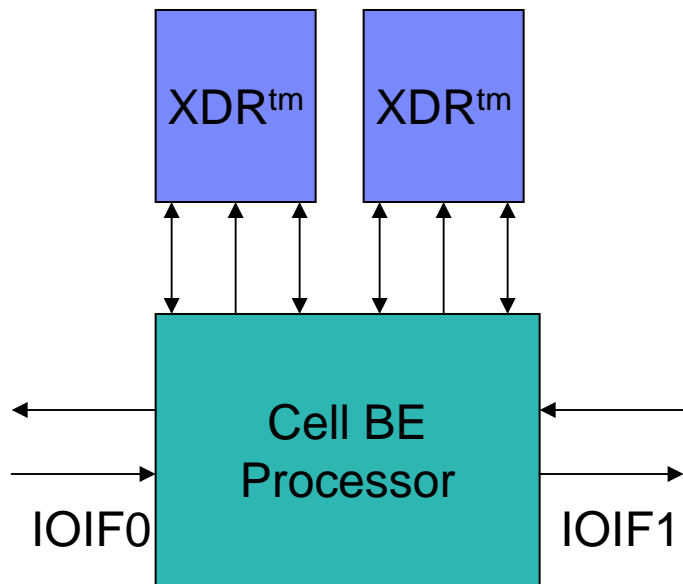


# Board Influences



## Cell BE Processor Can Support Many Systems

- Game console systems
- Blades
- HDTV
- Home media servers
- Supercomputers



## Cell Application Examples

- **TRE – Terrain Rendering Engine**
  - 50x speed up over PowerPC 970 / Apple G5 system
- **FFT – Fast Fourier Transform**
  - 120x speed up over Intel Pentium4 libraries



## TRE Performance

- **2.0 GHz Apple G5**      **0.6 frames/sec**
  - 40% of cycles spent waiting for Memory
- **3.2 GHz Cell**      **30.0 frames/sec**
  - 1% of cycles spent waiting for Memory
- **Cell has 50x advantage**

# Cell Performance Summary

- BE's performance is about an order of magnitude better than traditional GPPs for media and other applications that can take advantage of its SIMD capability

Type	Algorithm	3.2 GHz GPP	3.2 GHz Cell	Cell Perf Advantage
HPC	Matrix Multiplication (S.P.)	24 Gflops (w/SIMD)	200 GFlops* (8SPEs)	8x
	Linpack (S.P.)	16 GFlops (w/SIMD)	156 GFlops* (8SPEs)	9x
	Linpack (D.P.): 1kx1k matrix	7.2 GFlops (IA32/SSE3)	9.67 GFlops* (8SPEs)	1.3x
graphics	Transform-light	170 MVPS (G5/VMX)	256 MVPS** (per SPE)	12x
	TRE	1 fps (G5/VMX)	30 fps* (Cell)	30x
security	AES encryp. 128-bit key	1.03 Gbps	2.06Gbps** (per SPE)	16x
	AES decryp. 128-bit key	1.04 Gbps	1.5Gbps** (per SPE)	11x
	TDES	0.12 Gbps	0.16 Gbps** (per SPE)	10x
	DES	0.43 Gbps	0.49 Gbps** (per SPE)	9x
	SHA-1	0.85 Gbps	1.98 Gbps** (per SPE)	18x
video processing	mpeg2 decoder (CIF)	----	1267 fps* (per SPE)	--
	mpeg2 decoder (SDTV)	354 fps (IA32)	365 fps** (per SPE)	8x
	mpeg2 decoder (HDTV)	----	73 fps* (per SPE)	--

Notes: \* Hardware measurement

\*\* Simulation results

# Cell BE Blade Overview

## Blade

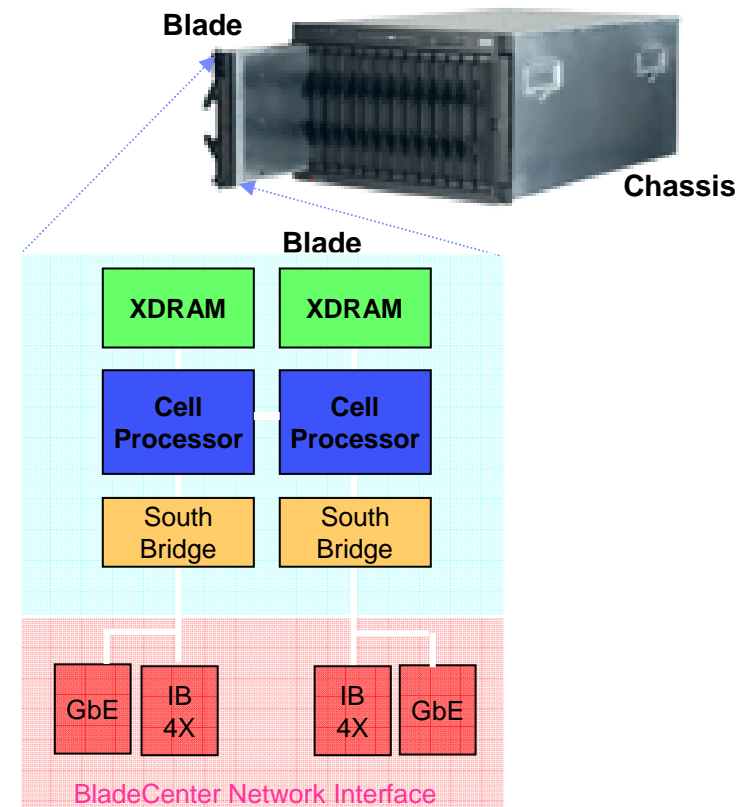
- Two Cell Processors (SMP) and Support Logic
- 1GB XDRAM
- BladeCenter Interface ( Based on IBM JS20)
- Infiniband 4x (10Gbps) interconnect

## Chassis

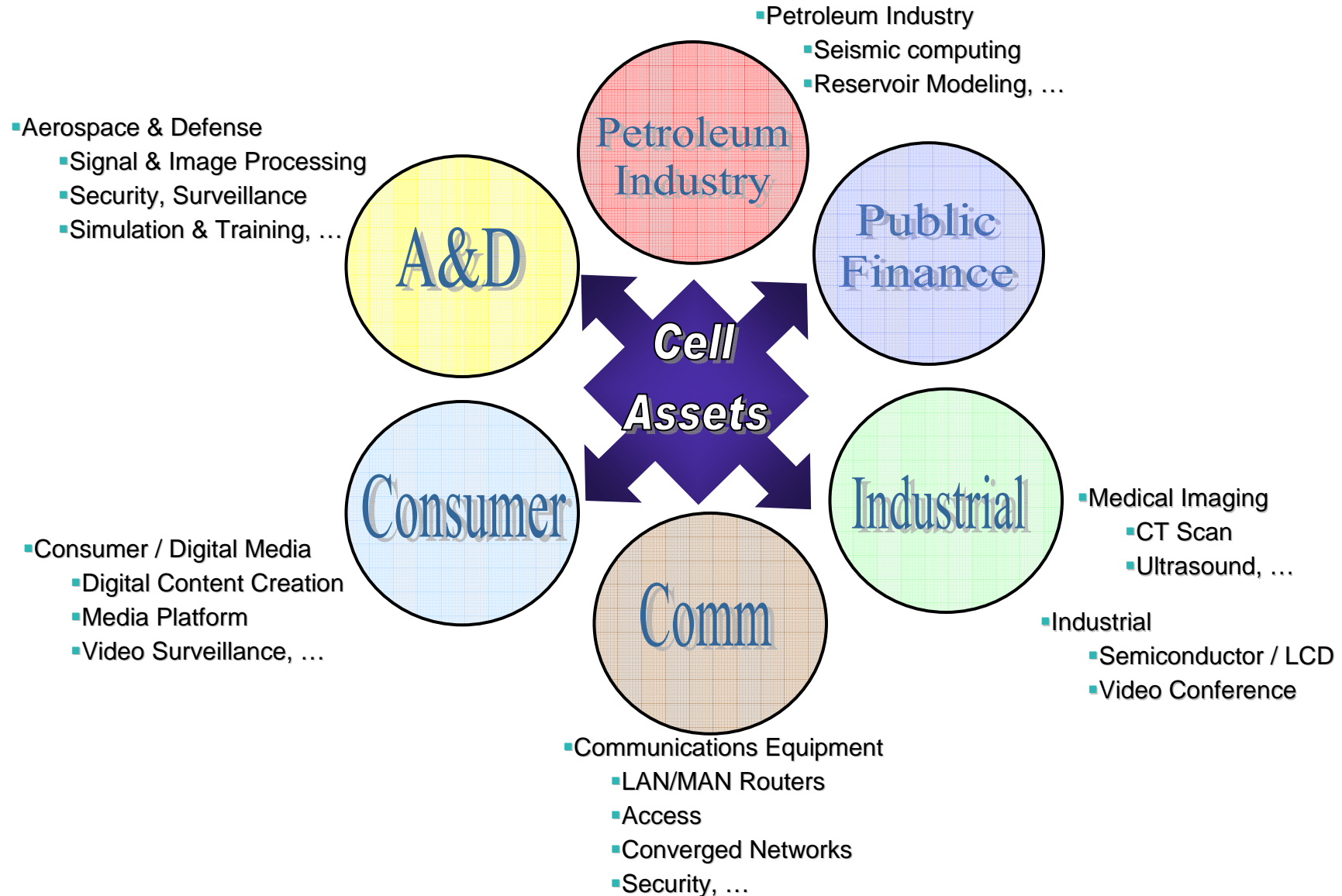
- Standard IBM BladeCenter form factor with:
  - 7 Blades (for 2 slots each) with full performance
  - 2 switches (1Gb Ethernet) with 4 external ports each
- Updated Management Module Firmware.
- External Infiniband Switches with optional FC ports.

## Typical Configuration (available today from E&TS)

- eServer 25U Rack
- 7U Chassis with Cell BE Blades, OpenPower 710
- Nortel GbE switch
- GCC C/C++ (Barcelona) or XLC Compiler for Cell (alphaworks)
- SDK Kit on <http://www-128.ibm.com/developerworks/power/cell/>



# Opportunities for Cell BE Blade



# Target applications leverage Cell's disruptive capabilities

## Cell Disruptors

- Non-homogeneous coherent multi-Processor
  - Dual-threaded control-plane processor
  - 8 independent data-plane processors
  - Thread-level parallelism
- SIMD processing architecture
  - 128-entry, 128-bit register files
  - Pipelined execution units
  - Branch hint
  - Data-level parallelism
- Rich integer instruction set
  - Word, halfword, byte, bit
  - Boolean
  - Shuffle
  - Rotate, shift, mask
- Single-precision floating point
- Double-precision floating point
- 256KB SPU local stores
  - Asynchronous DMA/main memory interface
  - Channel interface
  - Single-cycle load/store to/from registers
- High-bandwidth internal bus
  - 96 bytes transferred per clock
  - 100+ outstanding transfers supported
- Coherent bus interface
  - Up to 30GB/s out, 25 GB/s in
  - Direct attach of another Cell
  - Can be configured as non-coherent
- Non-coherent bus interface
  - Up to 10GB/s out, 10 GB/s in
- 25+ GB/s XDR memory interface

## Accelerated Functions

- Signal processing
- Image processing
- Audio resampling
- Noise generation
- Sound oscillation
- Digital filtering
- Curve and surface evaluation
- FFT
- Matrix mathematics
- Vector mathematics
- Game Physics / Physics simulation
- Video compression / decompression
- Surface subdivision
- Transform-light
- Graphics content creation
- Security encryption / decryption
- Pattern matching
- Language parsing
- TCP/IP offload
- Encoding / decoding
- Parallel processing
- Real time processing
- ...

## Target Applications

- Medical imaging / visualization
- Drug discovery
- Petroleum reservoir modeling
- Seismic analysis
- Avionics
- Air traffic control systems
- Radar systems
- Sonar systems
- Training simulation
- Targeting
- Defense and security IT
- Surveillance
- Secure communications
- LAN/MAN Routers
- Network processing
- XML and SSL acceleration
- Voice and pattern recognition
- Video conferencing
- Computational chemistry
- Climate modeling
- Data mining and analysis
- Media server
- Digital content creation
- Digital content distribution
- ....



# Summary

- **Cell is current example of next generation systems thinking**
  - High level of integration
  - Massively parallel processing
- **Flexible/Scaleable architecture required to meet needs of tomorrow**
  - Flexibility - application based customization/optimization
  - Cost/power considerations
  - Processing power to support next generation UI
  - Scalability to support massive parallelization and connectivity
- **Greater understanding of gaming style interaction and user base will have an impact on mainstream computing in the future**
  - New market spaces, applications and uses
- **Cell Blades: Product IBM Blades Q20**

# Publicly Available Information

- Introduction to the Cell Broadband Engine White Paper
  - Cell Broadband Engine Public Registers Guide (subset of CDA version)
  - Cell Broadband Engine Linux Reference Implementation Application Binary Interface Specification
  - SPU C/C++ Language Extensions Software Reference Manual
  - SPU Application Binary Interface Specifications
  - SPU Assembly Language Specifications
  - Broadband Engine Linux Application Binary Interface Specification
  - Cell Broadband Engine SDK Libraries, Overview and User's Guide
  - Cell Broadband Engine Architecture
  - Cell Broadband Engine Datasheet
  - SPU Instruction Set Architecture Specifications
  - Cell Broadband Engine Processor Full System Simulator
  - XLC Alpha Edition for Cell Broadband Engine
  - IBM Cell Broadband Engine Software Sample and Library Source Code
  - GCC Toolchain for Cell Broadband Engine
  - Cell Broadband Engine SPE Management Library
  - Linux Kernel patch for Cell Broadband Engine
  - SDK Installation script
- 
- Introduction to the Cell Microprocessor, Article
  - A 4.8GHz Fully Pipelined Embedded SRAM in the Streaming Processor of a Cell Processor, Article
  - A Double-Precision Multiplier with Fine-Grained Clock-Gating Support for a First-Generation Cell Processor, Article
  - A Streaming Processing Unit for a Cell Processor, Article
  - The Design and Implementation of a First-Generation Cell Processor, Article
  - Microprocessor Report - Cell Moves into the Limelight, Analyst Report
  - Microprocessor Reports - 2004 Technology Awards, Analyst Report

(c) Copyright International Business Machines Corporation 2005.  
All Rights Reserved. Printed in the United States September 2005.

The following are trademarks of International Business Machines Corporation in the United States, or other countries, or both.

IBM

IBM Logo

Power Architecture

Other company, product and service names may be trademarks or service marks of others.

All information contained in this document is subject to change without notice. The products described in this document are NOT intended for use in applications such as implantation, life support, or other hazardous uses where malfunction could result in death, bodily injury, or catastrophic property damage. The information contained in this document does not affect or change IBM product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of IBM or third parties. All information contained in this document was obtained in specific environments, and is presented as an illustration. The results obtained in other operating environments may vary.

While the information contained herein is believed to be accurate, such information is preliminary, and should not be relied upon for accuracy or completeness, and no representations or warranties of accuracy or completeness are made.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED ON AN "AS IS" BASIS. In no event will IBM be liable for damages arising directly or indirectly from any use of the information contained in this document.

IBM Microelectronics Division  
1580 Route 52, Bldg. 504  
Hopewell Junction, NY 12533-6351

The IBM home page is <http://www.ibm.com>  
The IBM Microelectronics Division home page is  
<http://www.chips.ibm.com>