POWER7+™
Outline

- POWER Processor History
- Design Overview
- Performance Benchmarks
- Key Features
  - Scale-up / Scale-out
  - The new accelerators
  - Advanced energy management
- Summary

* Statements regarding Power7+™ features do not imply that IBM will introduce a system with this capability
20+ Years of POWER Processors

- **POWER1** - AMERICA's
  - 1990
- **RS64I**
  - 1990
  - 0.5um
  - Cobra A10 - 64 bit
  - RS64I Pulsar
  - RS64I North Star
  - RS64I Apache BiCMOS
- **Power2™ P2SC**
  - 0.35um
  - 1995
  - Muskie A35
  - 0.72um
- **Power3™-630**
  - 0.25um
  - 2000
  - 0.35um
- **Power4™-SMT**
  - 0.22um
  - 2001
  - 0.25um
  - 0.5um
- **Power5™-Ultra High Frequency**
  - 180nm
  - 2006
  - 0.25um
  - 0.35um
- **Power6™-Multi-core-EDRAM**
  - 130nm
  - 2010
  - 0.25um
  - 0.35um
- **Power7+™-accelerators**
  - 45nm
  - 2012
  - 0.22um
  - 0.6um
- **Power7+™-Multi-core-EDRAM**
  - 32nm
  - 2012
  - 0.25um
  - 0.35um

**Major POWER® Innovation**
- 2001 Dual Core Processors
- 2001 Large System Scaling
- 2001 Shared Caches
- 2003 On Chip Memory Control
- 2003 SMT
- 2006 Ultra High Frequency
- 2006 Dual Scope Coherence Mgmt
- 2006 Decimal Float/VMX
- 2006 Processor Recovery/Sparing
- 2010 Balanced Multi-core Processor
- 2010 On Chip EDRAM
- 2012 On chip Accelerators
- 2012 Massive L3 cache
- 2012 Power Gating

* Dates represent approximate processor power-on dates, not system availability
POWER7+ Processor Chip

- Area: 567mm2
- Eight processor cores
  - 12 execution units per core
  - 4 Way SMT per core
  - 32 Threads per chip
  - 256KB L2 per core
- Scalability up to 32 Sockets
  - 360GB/s SMP bandwidth/chip
  - 20,000 coherent operations in flight
- Technology: 32nm lithography, Cu, SOI, eDRAM, 13 metal levels
- 2.1B transistors
  - Equivalent function of 5.4B
- 80MB on chip eDRAM shared L3
- Accelerators
- Enhanced Power management
- Binary Compatibility with POWER6/7
An Improved Core

- Up to 25% frequency gain due to mapping into 32nm technology and power management improvements.
- Increase of L3 memory capacity by 2.5x
- Doubled single precision floating-point performance
- Added Power Gating regions for Core/L2 & L3 regions

![Diagram showing Core/L2 Power-Gating and L3 Power-Gating regions](image)
Optimized in Two Dimensions

Increased Clock Speed → Thread Strength
2.5x L3 cache

Single Chip Module

POWER7+

Thread Strength
SMP Scaling

Scale UP

2x Cores / socket → Throughput
5x L3 cache → Bandwidth Amplifier

5x L3 cache

Single Chip Module

POWER7

Dual Chip Module

Scale OUT
Performance View

Normalized POWER7 vs POWER7+ Comparison

<table>
<thead>
<tr>
<th>Workload</th>
<th>Power7 SCM</th>
<th>Power7+ SCM</th>
<th>Power7+ DCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERP</td>
<td>1.2</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Integer</td>
<td>1.4</td>
<td>1.8</td>
<td>2.2</td>
</tr>
<tr>
<td>OLTP</td>
<td>1.6</td>
<td>2.4</td>
<td>2.8</td>
</tr>
<tr>
<td>Java</td>
<td>1.3</td>
<td>1.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>
POWER7+ Accelerators

- Provide CPU off-load and workload speedup for SSL, encrypted file system, and active memory expansion (AME).
  - Asymmetric Math Functions (AMF)
    - RSA cryptography
    - ECC (elliptic curve cryptography)
  - Advanced Encryption Standard (AES)/Secure Hash Algorithm (SHA)
    - Symmetric-key cryptography with combinational modes
  - Random Number Generator (RNG) – True hardware entropy generator
    - Cannot be algorithmically reverse engineered
  - 842 proprietary compression algorithm
    - High bandwidth, area efficient

- Integrated across silicon, ISA, hypervisor, and OS
Accelerators Details

- Advanced Encryption Standard engine
  - Modes: ECB, CBC, CTR, CCM, CCA, GCM, GCA, GMAC, CM, F8, XBC-MAC-96
  - Key lengths: 128b, 192b, 256b
  - Three engines

- Secure Hash Algorithm engine
  - Modes: SHA-1, SHA-256, SHA-512, MD5
  - HMAC supported for SHA
  - Three engines

- Asymmetric Math Functions
  - Modular math functions for RSA (Rivest, Shamir, Adleman) and ECC (elliptic curve cryptography): mod add, mod subtract, mod inverse, mod reduction, mod multiplication, mod exponentiation, mod exponentiation CRT (integer only)
  - Point functions for ECC GF(p) and GF(2m): point add, point double, point multiply
  - RSA lengths: 512b, 1024b, 2048b, 4096b
  - ECC GF(p) lengths: 192b, 224b, 256b, 384b, 521b (SuiteB)
  - ECC GF(2m) lengths: 163b, 233b, 283b, 409b, 571b (SuiteB)

- Random Number Generator
  - All digital design which produces 64b random numbers accessible by MMIO load instructions
  - Correctness verified against the NIST Random Number Generator Test Suite

- Active Memory Expansion
  - IBM-proprietary algorithm with 8B-, 4B-, and 2B-phrase parsings
  - Throughput: Up to 8 bytes of compression or 8 bytes of decompression per bus cycle.

- MCD
  - Hardware to predict whether memory access is on-node or off-node.
POWER7+ Sleep & Winkle Overview

**Nap (per core)**
- Stop clocks to only processor core execution engines.
- Leave all caches running.
- Saves ~ 10% power with ~5us Latency

**Sleep (per core)**
- Power OFF the core plus private L2 cache.
- Requires restore/re-init to wakeup.
- Leave shared L3 cache running.
- Saves ~ 80% power with ~3ms Latency

**Winkle (per chiplet)**
- Power OFF the entire chiplet.
- Requires restore/re-init to wakeup.
- Takes offline 1/8 of the shared L3 cache.
- Saves > 95% power with < 6ms Latency
Save Energy When Idle

Three idle states were implemented to optimize power vs. latency

- **Nap** (Continued POWER7 support)
  - Optimized for wake-up time
  - Turn off clocks to execution units only
  - Caches remain coherent

- **Sleep** (Improved from POWER7)
  - More savings at increased latency
  - Purge and power off core plus L2 caches
  - Leave shared L3 cache running

- **Winkle** (New for POWER7+)
  - Maximum savings at higher latency
  - Purge and power off entire chiplet
  - Takes eighth of chip L3 cache offline

![Diagram showing processor idle states and energy reduction]

- Processor Energy Reduction (compared to Idle Loop)
- Typical Wake-Up Latency (ms)
Real Time Chip Guardband

- **Conventional guardband**
  - Static, conservative voltage margins for potential worst-case conditions
  - Causes unnecessary loss of energy efficiency during typical server usage

- **Critical Path Monitor (CPM)**
  - Real Time detection of available circuit timing margin

**Components of Chip Guardband**

- Static Guardband
  - Test inaccuracy
  - Uncertainty (margin)
  - Reliability (wearout)
  - Voltage variation
  - Thermal variation
Real Time Guardband – DPLL/CPM feedback loop

Reference frequency ➔ DPLL ➔ Clock Buffering ➔ Sensor Aggregator ➔ CPM

CPM data encoding:
“11111” = large margin
“11110” = some margin
“11100” = ideal margin
“11000” = margin too small
“10000” = not enough margin

Per Core - Clock Distribution

CPM Fmax Delay = f(Tj, V)
POWER7+ Core CPM Infrastructure

CPMs are strategically placed in known hot spots typically near micro-architecture critical paths.

The real time feedback from CPMs can reduce how much margin is needed for various guardband components.

Real-time guardbanding will allow for greater energy efficiency.

Components of Chip Guardband

- **No CPM**
  - Test inaccuracy
  - Uncertainty (margin)
  - Reliability (wearout)
  - Voltage variation
  - Thermal variation

- **With CPM**
  - Remove Reclaimed Margin
  - Reduce
  - No Change

- = CPM (Critical Path Monitor)
- = AND Buffer
Advanced 32nm Technology

- 32nm High-K Metal Gate (HKMG) SOI based logic technology
  - 3 logic transistor threshold voltages ($V_t$) optimizes power/performance
  - 13-layer BEOL metal stack minimizes cross die latency
    - 1x, 2x, 4x, 8x, & ultra-thick metal layers
  - eDRAM provides 3-4x density advantage over SRAM
- Advanced features make the node effectively perform as one with sub-32nm features
POWER7+: The next major step in IBM’s roadmap

- Brings significant improvement to both scale up & scale out systems.
- The new accelerators optimize specific functions while offloading CPU.
- Advanced energy management greatly improves data center efficiency.
Acknowledgements:

Many Thanks to the entire POWER7+ design, manufacturing and product teams.

Questions: Send email to: sctaylor@us.ibm.com