POWER7 Technology for Smarter Systems
POWER7 Technology for Smarter Systems

- Smarter Innovation

- Smarter Technology → Value
  - High speed embedded DRAM → Energy Efficiency
  - Throughput: Massive, Balanced → Virtualization and Performance
  - Robust RAS capabilities → Reliability and Flexibility

- Smarter Continuity
POWER7 Technology for Smarter Systems

- **Smarter Innovation**

- Smarter Technology ➔ Value
  
  - High speed embedded DRAM ➔ Energy Efficiency
  - Throughput: Massive, Balanced ➔ Virtualization and Performance
  - Robust RAS capabilities ➔ Reliability and Flexibility

- **Smarter Continuity**
Smarter Innovation

- IBM is uniquely positioned within the industry
- Multiple businesses under one corporation
- Drives cross-disciplinary synergy and innovation across hardware and software stacks, yielding value for our customers
- No other company has this breadth
Smarter Innovation

- I am proud to work with a world-class microprocessor design team
- Most of us have been working together on Power server chips for more than 10 years
- We plan innovation across many product generations
Smarter Innovation

- Of course, world-class microprocessors require leading-edge semiconductor technology.
- Many of our competitors are becoming irrelevant because they lack this technology.

**Synergy:** Development of embedded DRAM and POWER7 cache architecture.
Smarter Innovation

- Owning both the chip and system business enables IBM to optimize across chip designs and system configurations

- Unlike some competitors, we do not simply build chips and sell them to system vendors

Synergy: Multi-chip module packaging technology

Synergy: System interconnect topology optimized for massive SMP scaling
Smarter Innovation

- Our chip and system designs are optimized to support a hypervisor that provides world class virtualization technology.

- Robust virtualized system scaling, RAS, energy mgmt, and security are enabled by this synergy.

Synergy: PowerVM has negligible performance overhead, unlike other hypervisors.
Smarter Innovation

- IBM is an operating system company, supporting
  - AIX
  - i
  - Linux

- Our hardware stack and PowerVM are optimized to support several features provided by our operating systems

**Synergy:** Active Memory Expansion
Smarter Innovation

- Finally, our middleware layers and packaged solutions also drive optimizations into the operating systems and PowerVM, as well as the chipset itself.
- We will have more and more of these cross-disciplinary innovations as the stack becomes more tightly integrated, and drives more efficiency.

IBM’s synergy uniquely enables Smarter Innovation for a Smarter Planet.
Smarter Innovation: 17 years of U.S. Patent Leadership

2009 U.S. Patent Leaders

<table>
<thead>
<tr>
<th>Company</th>
<th>Patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>4,914</td>
</tr>
<tr>
<td>Samsung</td>
<td>3,611</td>
</tr>
<tr>
<td>Microsoft</td>
<td>2,906</td>
</tr>
<tr>
<td>Canon</td>
<td>2,206</td>
</tr>
<tr>
<td>Matsushita</td>
<td>1,829</td>
</tr>
<tr>
<td>Toshiba</td>
<td>1,696</td>
</tr>
<tr>
<td>Sony</td>
<td>1,680</td>
</tr>
<tr>
<td>Intel</td>
<td>1,537</td>
</tr>
<tr>
<td>Seiko Epson</td>
<td>1,330</td>
</tr>
<tr>
<td>HP</td>
<td>1,273</td>
</tr>
<tr>
<td>SUN</td>
<td>562</td>
</tr>
<tr>
<td>Apple</td>
<td>289</td>
</tr>
<tr>
<td>EMC</td>
<td>250</td>
</tr>
<tr>
<td>Oracle</td>
<td>208</td>
</tr>
</tbody>
</table>

Source: IFI Patent Intelligence
Smarter Innovation: 20 Years of POWER Processors

- POWER1 - AMERICA's
  - 601
  - 603
  - 604e

- POWER2™ P2SC
  - 6um
  - .35um
  - .25um

- POWER3™
  - 630
  - 35um
  - 5um

- RS64I Apache BiCMOS
  - .5um

- RS64II Pulsar
  - 18um

- RS64II North Star
  - 25um

- RS64IV Sstar
  - 180nm

- POWER4™
  - 630
  - .35um

- POWER5™
  - SMT

- POWER6™
  - Multi-core
  - Ultra High Frequency

- POWER7™
  - Multi-core
  - EDRAM

- Next Generation

Major POWER® Innovation
- 1990 RISC Architecture
- 1994 Multi-processor
- 1995 Out of Order Execution
- 1996 64 Bit Enterprise Architecture
- 1997 Hardware Multi-Threadining
- 2001 Dual Core Processors
- 2001 Large System Scaling
- 2001 Shared Caches
- 2003 On Chip Memory Controller
- 2003 Simultaneous Multi-Threadining
- 2006 Ultra High Frequency
- 2006 Dual Scope Coherence Mgmt
- 2006 Decimal Float/VSX
- 2006 Processor Recovery/Sparing
- 2009 Balanced Multi-core Processor
- 2009 On Chip eDRAM

* Dates represent approximate processor power-on dates, not system availability
POWER7 Technology for Smarter Systems

- Smarter Innovation

- Smarter Technology → Value
  - High speed embedded DRAM → Energy Efficiency
  - Throughput: Massive, Balanced → Virtualization and Performance
  - Robust RAS capabilities → Reliability and Flexibility

- Smarter Continuity
Smarter Technology: The POWER7 Chip

567 mm² Chip
- 45nm lithography
- Silicon-on-insulator
- Embedded DRAM
- 11 Cu metal layers
- 1.2 billion transistors
Smarter Technology: The POWER7 Chip

567 mm² Chip
- 45nm lithography
- Silicon-on-insulator
- Embedded DRAM
- 11 Cu metal layers
- 1.2 billion transistors

8 processor cores
- 12 execution units
- 4 threads
- 32K i-cache
- 32K d-cache
- 256K L2 cache
Smarter Technology: The POWER7 Chip

567 mm² Chip
- 45nm lithography
- Silicon-on-insulator
- Embedded DRAM
- 11 Cu metal layers
- 1.2 billion transistors

8 processor cores
- 12 execution units
- 4 threads
- 32K i-cache
- 32K d-cache
- 256K L2 cache

32M shared L3 cache
- Embedded DRAM
- 8 fast local regions
Smarter Technology: High-speed SOI embedded DRAM (eDRAM)

Large on-processor-chip caching structures are typically constructed using SRAM

- SRAM uses 6 transistors to store 1 bit of information
- DRAM uses 1 transistor to store 1 bit of information

Until now, DRAM fabrication processes incompatible with high performance logic processes...

Source: ISSCC 2010
Smarter Technology: High-speed SOI embedded DRAM (eDRAM)

eDRAM enables 32M on-chip L3 cache for POWER7

**Compared to dense SRAM**
- Density: 3X less area
- Energy: 5X less standby power

**Compared to POWER6 external L3 cache chip**
- Latency: 6X lower latency to local L3 region
- Bandwidth: 2X more bandwidth per core, despite 4X as many cores
  - = 8X boost per chip
- Packaging: Eliminates interfaces to external cache chip; signal I/O re-allocated to increase memory bandwidth

Source: ISSCC 2010
**Smarter Technology: High-speed SOI embedded DRAM (eDRAM)**

**Packaging Density and Energy Efficiency Improvement:**

- 4 POWER6 chips + 4 external cache chips
Smarter Technology: High-speed SOI embedded DRAM (eDRAM)

Packaging Density and Energy Efficiency Improvement:
- 4 POWER6 chips + 4 external cache chips are replaced by
- 1 POWER7 chip
Value: Compute Density

POWER6
- Power 550 (max 8-cores)
- Power 570 (max 16-cores)

POWER7
- Power 750 (max 32-cores)
- Power 755 (max 32-cores)
- Power 770 (max 64-cores)
- Power 780 (max 64-cores)
Value: Most Energy Efficient 4-socket system on the planet

*The first Energy Star certified in server category*

**Power 750**

Most energy efficient systems

<table>
<thead>
<tr>
<th></th>
<th>Performance Per Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itanium HP rx6600</td>
<td></td>
</tr>
<tr>
<td>SPARC Sun T5440</td>
<td></td>
</tr>
<tr>
<td>x86 HP DL585</td>
<td></td>
</tr>
<tr>
<td>POWER7 Power 750 with PowerVM</td>
<td>600</td>
</tr>
</tbody>
</table>

© 2010 IBM Corporation
POWER7 Technology for Smarter Systems

- Smarter Innovation
- **Smarter Technology** → Value
  - High speed embedded DRAM → Energy Efficiency
  - Throughput: Massive, Balanced → Virtualization and Performance
  - Robust RAS capabilities → Reliability and Flexibility

- Smarter Continuity
Smarter Technology: Massive Core Throughput

- **64-bit Power ISA Architecture v2.06**

- **Execution Units**
  - 2 Fixed Point Units
  - 2 Load Store Units also do Simple FX ops
  - 4 Double Precision Floating Point Units
  - 1 Branch, 1 Condition Register, 1 Vector Unit
  - 1 Decimal Floating Point Unit
  - 6 Wide Dispatch (2 branches per group)
  - 8 Wide issue
  - Units include distributed Recovery Function

- **Out of Order Execution and 4 Thread SMT**

- **Cache Design:**
  - L1 32KB 4-way set associative I-Cache
  - L1 32KB 8-way set associative D-Cache
    - L1 cache latency reduced from 4 to 2 cycles
  - L2 256KB 8-way associative
    - Tightly coupled to core. 8 cycles away

- **POWER7 continues to support VMX / Extends SIMD support with VSX**
  - 2 VSX units that can each handle 2 Double-Precision FP instructions
  - 8 FLOPS per cycles
  - VSX units can also handle 4 Single Precision instructions per cycle
  - VSX instruction set support for vector and scalar instructions
Smarter Technology: Massive Core Throughput

Single thread out-of-order flow

4 Thread SMT out-of-order flow

Cycles

Thread 0 executing
Thread 1 executing
Thread 2 executing
Thread 3 executing
POWER7 has lower core frequency than POWER6, but POWER7 has higher core performance than POWER6.
Smarter Technology: Massive Chip Throughput

**Throughput / Chip**
- 8 cores
- 32 threads
- 96 execution units
- 48 dispatch / cycle
- 64 issue / cycle

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
POWER7 Technology for Smarter Systems

Smarter Technology: Massive Chip Balance

Throughput / Chip
- 8 cores
- 32 threads
- 96 execution units
- 48 dispatch / cycle
- 64 issue / cycle

Balance / Chip
- 32M shared L3
- 500+ GB/s chip interconnect bandwidth

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive Chip Balance

Throughput / Chip
- 8 cores
- 32 threads
- 96 execution units
- 48 dispatch / cycle
- 64 issue / cycle

Balance / Chip
- 32M shared L3
- 500+ GB/s chip interconnect bandwidth
- 100+ GB/s sustained memory bandwidth

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive Chip Balance

**Throughput / Chip**
- 8 cores
- 32 threads
- 96 execution units
- 48 dispatch / cycle
- 64 issue / cycle

**Balance / Chip**
- 32M shared L3
- 500+ GB/s chip interconnect bandwidth
- 100+ GB/s sustained memory bandwidth
- 360 GB/s raw SMP link bandwidth
- 50 GB/s raw I/O link bandwidth

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Throughput

Max 32-socket system

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Throughput

Max 32-socket system

- Local SMP links interconnect up to 4 POWER7 chips

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Throughput

Max 32-socket system

- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Throughput

Max 32-socket system
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Throughput / System
- 256 cores
- 1024 threads

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Balance

**Max 32-socket system**
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

**Throughput / System**
- 256 cores
- 1024 threads

**Balance / System**
- 20,000 coherence operations in flight
  - Relaxed updates
  - System Topology

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Balance

Max 32-socket system
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Throughput / System
- 256 cores
- 1024 threads

Balance / System
- 20,000 coherence operations in flight
  - Relaxed updates
  - System Topology
- Dual Scope Coherence Protocol
  - Boosts peak Coherence bandwidth from 480 GB/s to up to 15,000 GB/s

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Balance

Max 32-socket system
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Throughput / System
- 256 cores
- 1024 threads

Balance / System
- 20,000 coherence operations in flight
  - Relaxed updates
  - System Topology
- Dual Scope Coherence Protocol
  - Boosts peak Coherence bandwidth from 480 GB/s to up to 15,000 GB/s

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Balance

Max 32-socket system
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Throughput / System
- 256 cores
- 1024 threads

Balance / System
- 20,000 coherence operations in flight
  - Relaxed updates
  - System Topology
- Dual Scope Coherence Protocol
  - Boosts peak Coherence bandwidth from 480 GB/s to up to 15,000 GB/s

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Smarter Technology: Massive SMP System Balance

Max 32-socket system
- Local SMP links interconnect up to 4 POWER7 chips
- Remote SMP links interconnect up to 8 local groups

Throughput / System
- 256 cores
- 1024 threads

Balance / System
- 20,000 coherence operations in flight
  - Relaxed updates
  - System Topology
- Dual Scope Coherence Protocol
  - Boosts peak Coherence bandwidth from 480 GB/s to up to 15,000 GB/s

Speculative Chip Coherence Broadcast

Note: Chip/System throughput/balance metrics assume a system that fully exploits chip capabilities
Value: Virtualization

- High System Throughput
- Robust, Large Caches
- High Memory Bandwidth
- Robust SMP Scaling
- Dynamic optimization of SMT usage and core frequency
- Synergy between PowerVM, chipset team, and system team
Value: Highest Performing 4-socket system on the planet

Power 750

SPECint_rate

Itanium HP rx6600  SPARC Sun T5440  x86 HP DL585  POWER7 Power 750 with PowerVM
More SAP performance than any system in the industry
20% more performance ... one-fourth the number of cores vs. Sun M9000

37,000 SAP users on SAP SD 2 Tier

Power 780 with DB2®

Systems are listed with processor chips/core/threads under system name; IBM Power System 780, 8p / 64-c / 256-t, POWER7, 3.8 GHz, 1024 GB memory, 37,000 SD users, dialog resp.: 0.98s, line items/hour: 4,043,670, Dialog steps/hour: 12,131,000, SAPS: 202,180, DB time (dialog/update): 0.013s / 0.031s, CPU utilization: 99%; OS: AIX 6.1, DB2 9.7, cert# 2010013; SUN M9000, 64p / 256-c / 512-t, 1156 GB memory, 32,000 SD users, SPARC64 VIII, 2.88 GHz, Solaris 10, Oracle 10g, cert# 2009046; All results are 2-tier, SAP EHP 4 for SAP ERP 6.0 (Unicode) and valid as of 4/1/2010; Source: http://www.sap.com/solutions/benchmark/2tier.epx - See Power 780 benchmark details for more information
More TPC-C performance per core than any system in the industry

4.6 to 7.5 times more performance per core than HP Itanium and Sun Enterprise T5440 cluster respectively

Best results listed for IBM POWER, HP, and Sun/Oracle systems over 1M tpmC. Source: http://www.tpc.org as of 4/1/08. See Power 780 benchmark details for specific results.
POWER7 Technology for Smarter Systems

- Smarter Innovation

- **Smarter Technology** ➔ **Value**
  - High speed embedded DRAM ➔ Energy Efficiency
  - Throughput: Massive, Balanced ➔ Virtualization and Performance
  - Robust RAS capabilities ➔ Reliability and Flexibility

- Smarter Continuity
Smarter Technology: Robust RAS Capabilities and Value

**Dynamic Oscillator Failover**
- OSC0
- OSC1

**SMP Interface**
- ECC protected
- Node hot add /repair

**Processor Core**
- Checkpoint and Recovery
- Stacked latches to improve SER
- Partition isolation on checkpoints

**L3 eDRAM**
- ECC protected
- Auto-purge, Auto-delete
- Spare rows and columns

**I/O Bus**
- ECC protected
- Hot add

**InfiniBand® Interface**
- Redundant paths

**64 Byte ECC on Memory**
- Corrects full chip kill on X8 dimms

**Spare DRAM devices implemented**
- Dual memory chip failures do not cause outage

**Selective memory mirroring capability**
- Recover partition from memory failures

**HW assisted scrubbing**
- Dynamic sparing on channel interface
- PowerVM Hypervisor protected from full dimm failures

*Statements regarding SMP servers do not imply that IBM will introduce a system with this capability.*
POWER7 Technology for Smarter Systems

- Smarter Innovation
- Smarter Technology → Value
  - High speed embedded DRAM → Energy Efficiency
  - Throughput: Massive, Balanced → Virtualization and Performance
  - Robust RAS capabilities → Reliability and Flexibility
- Smarter Continuity
Smarter Continuity

- IBM Power has consistent roadmap
- Competitors have failed to deliver
Thank you!

I have enjoyed this opportunity to speak with you.

William J. Starke

POWER7 Chief Storage Hierarchy and SMP Architect