Performance Characteristics of the POWER8™ Processor

Alex Mericas
Systems Performance
IBM Systems & Technology Group Development
Designed for Big Data - optimized for analytics performance

- **Processors**: flexible, fast execution of analytics algorithms
- **Memory**: large, fast workspace to maximize business insight
- **Data Bandwidth**: bring massive amounts of information to compute resources in real-time

Optimized for a broad range of data and analytics:

- IBM Predictive Customer Intelligence
- Industry Solutions
  - Retail
  - Government
  - Healthcare
  - Telecom
  - Banking
- 5X Faster
POWER8 Processor

Technology
• 22nm SOI, eDRAM, 15 ML 650mm2

Caches
• 512 KB SRAM L2 / core
• 96 MB eDRAM shared L3
• Up to 128 MB eDRAM L4 (off-chip)

Memory
• Up to 230 GB/s sustained bandwidth

Bus Interfaces
• Durable open memory attach interface
• Integrated PCIe Gen3
• SMP Interconnect
• CAPI (Coherent Accelerator Processor Interface)

Energy Management
• On-chip Power Management Micro-controller
• Integrated Per-core VRM
• Critical Path Monitors

Cores
• 12 cores (SMT8)
• 8 dispatch, 10 issue, 16 exec pipe
• 2X internal data flows/queue
• Enhanced prefetching
• 64K data cache, 32K instruction cache

Accelerators
• Crypto & memory expansion
• Transactional Memory
• VMM assist
• Data Move / VM Mobility

Shown at Hot Chips 25
**POWER8 Core**

### Execution Improvement vs. POWER7
- SMT4 → SMT8
- 8 instruction dispatch
- 10 instruction issue
- 16 execution pipes:
  - 2 Fixed Point, 2 Ld/Store, 2 Ld
  - 4 Floating Point, 2 Vector
  - 1 Crypto, 1 Decimal Floating Point
  - 1 Conditional, 1 Branch
- Larger Issue queues (4 x 16-entry)
- Larger completion table (28 groups)
- Larger Ld/Store reorder (128 / thrd)
- Improved branch prediction
- Improved unaligned storage access

### Larger Caching Structures vs. POWER7
- 2x L1 data cache (64 KB)
- 2x outstanding data cache misses
- 4x translation Cache

### Wider Load/Store
- 32B → 64B L2 to L1 data bus
- 2x data cache to execution dataflow

### Enhanced Prefetch
- Instruction speculation awareness
- Data prefetch depth awareness
- Adaptive bandwidth awareness
- Topology awareness

### Core Performance vs. POWER7
- ~1.6x Thread
- ~2x Max SMT

Shown at Hot Chips 25
POWER8 Memory Organization

- Up to 8 high speed channels, each running up to 9.6 Gb/s for up to 230 GB/s sustained
- Up to 32 total DDR ports yielding 410 GB/s peak at the DRAM
- Up to 1 TB memory capacity per fully configured processor socket (at initial launch)
POWER8 On Chip Caches

- L2: 512 KB 8 way per core
- L3: 96 MB (12 x 8 MB 8 way Bank)
- “NUCA” Cache policy (Non-Uniform Cache Architecture)
  - Scalable bandwidth and latency
  - Migrate “hot” lines to local L2, then local L3 (replicate L2 contained footprint)
- Chip Interconnect: 150 GB/sec x 16 segment per direction per segment
Cache Bandwidths

- GB/sec shown assuming 4 GHz
  - Product frequency will vary based on model type

- Across 12 core chip
  - 4 TB/sec L2 BW
  - 3 TB/sec L3 BW
Scale-Out Processor Version (Announced April 2014)

- Scale-UP Processor (Shown at Hot Chips 25)
  - Optimized for Large SMP
  - 22nm SOI, eDRAM, 15 ML 650mm2
  - 12 Core Chip
  - 32x PCIe Gen3 (16x CAPI)
  - Large memory capacity and bandwidth

- Scale-Out Processor (1 module per socket)
  - Optimized for Scale-OUT systems
  - 2 x 6-Core Chip (362mm2 each)
  - 48x PCIe Gen3 (32x CAPI)
  - Same core, L2, L3, etc
New Power Scale-out systems built with open innovation to put data to work

- **Power S822L**
  - 1 or 2 sockets
  - 10 or 12 cores/socket

- **Power S824 or Power S814**
  - 1 or 2 sockets
  - 6, 8, 10 or 12 cores/socket

- **Power S822**
  - 1 or 2 sockets
  - 6, 8, 10 or 12 cores/socket

**Designed for Big Data**

**Superior Cloud Economics**

**Open Innovation Platform**

OpenPOWERTM
New Power Scale-out systems detailed features

- **2 Sockets** (1 socket upgradeable)
- Up to **24 cores** (192 threads)
- Up to **1 TB memory** capacity
- Hot Plug PCIe gen 3 Slots
- SR-IOV support (statement of direction)
- Ethernet: Quad 1 Gbt / (x8 slot)
- Native I/O
  - USB (3), Serial (2), HMC (2)
- Internal Storage
  - Up to 18 SFF Bays
  - Up to 8 1.8” SSD Bays (Easy Tier)
  - DVD
- Power Supplies: (200-240 AVC)
POWER8 Performance Characteristics
### POWER8 CPI Stack

- Introduced with PowerPC970, the CPI stack uniquely identifies components of CPI (Cycles Per Instruction).
- Enhanced every generation to add detail and eliminate “other” category.
- POWER8 splits dependency chains within a group to separate cause and effect (e.g. long latency load feeding 1 cycle add).
- Items in blue are new with POWER8.

#### Stalled Cycles

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stall due to BR or CR</td>
<td>Stall due to Branch</td>
</tr>
<tr>
<td>Stall due to CR</td>
<td>Stall due to CR</td>
</tr>
<tr>
<td>Stall due to Fixed Point</td>
<td>Stall due to Fixed-Point long</td>
</tr>
<tr>
<td>Stall due to Fixed Point (other)</td>
<td>Stall due to Fixed-Point (other)</td>
</tr>
<tr>
<td>Stall due to Vector</td>
<td>Stall due to Vector long</td>
</tr>
<tr>
<td>Stall due to (other)</td>
<td>Stall due to Vector (other)</td>
</tr>
<tr>
<td>Stall due to Scalar</td>
<td>Stall due to Scalar long</td>
</tr>
<tr>
<td>Stall due to Scalar (other)</td>
<td>Stall due to Scalar (other)</td>
</tr>
<tr>
<td>Stall due to Load/Store</td>
<td>Stall due to Load/Store</td>
</tr>
<tr>
<td>Stall due to Load/Store (other)</td>
<td>Stall due to Load/Store (other)</td>
</tr>
<tr>
<td>Stall due to LSU Reject</td>
<td>Stall due to LSU Reject</td>
</tr>
<tr>
<td>Stall due to LSQ Reject</td>
<td>Stall due to LSQ Reject</td>
</tr>
<tr>
<td>Stall due to Store Finish</td>
<td>Stall due to Load Finish</td>
</tr>
<tr>
<td>Stall due to Store Forward</td>
<td>Stall due to Store Forward</td>
</tr>
<tr>
<td>Stall due to Load/Store (other)</td>
<td>Stall due to Load/Store (other)</td>
</tr>
<tr>
<td>Stall due to Next-To-Complete Flush</td>
<td>Stall due to Next-To-Complete Flush</td>
</tr>
</tbody>
</table>

#### Cycles

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waiting to Complete</td>
<td>Blocked due to LWSYNC</td>
</tr>
<tr>
<td>Thread Blocked</td>
<td>Blocked due to HWSYNC</td>
</tr>
<tr>
<td>Thread Blocked</td>
<td>Blocked due to ECC Delay</td>
</tr>
<tr>
<td>Thread Blocked</td>
<td>Blocked due to Flush</td>
</tr>
<tr>
<td>Thread Blocked</td>
<td>Blocked due to COQ Full</td>
</tr>
<tr>
<td>Thread Blocked (Other)</td>
<td></td>
</tr>
<tr>
<td>Completion Table Empty</td>
<td>Completion Table Empty due to Icache L3 Miss</td>
</tr>
<tr>
<td>Completion Table Empty</td>
<td>Completion Table Empty due to Icache L3 Miss (Other)</td>
</tr>
<tr>
<td>Completion Table Full</td>
<td>Dispatch Held due to Mapper</td>
</tr>
<tr>
<td>Dispatch Held due to Store Queue</td>
<td>Dispatch Held due to Store Queue</td>
</tr>
<tr>
<td>Dispatch Held due to Issue Queue</td>
<td>Dispatch Held due to Issue Queue</td>
</tr>
<tr>
<td>Dispatch Held (Other)</td>
<td></td>
</tr>
</tbody>
</table>

#### Completion Cycles

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion Table Empty</td>
<td>Dispatch Held due to Mapper</td>
</tr>
<tr>
<td>Completion Table Empty</td>
<td>Dispatch Held due to Store Queue</td>
</tr>
<tr>
<td>Completion Table Empty</td>
<td>Dispatch Held due to Issue Queue</td>
</tr>
<tr>
<td>Completion Table Empty</td>
<td>Dispatch Held (Other)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>
Sampled Instruction Event Register (SIER)

- Augments sampling-based performance analysis and profiling
- Detailed information is collected for sampled instruction
  - Instruction type
  - CPI Stack
  - Branch prediction
  - Cache access
  - Translation
Additional Performance Monitor Enhancements

• Sample Filtering
  – “Needle in haystack” problem
  – Reduces number of samples presented to software by filtering out un-interesting ones

• Hotness table
  – Hardware keeps track of recently sampled addresses and generates an interrupt if the address is “hot”

• Branch History Rolling Buffer
  – Rolling list of recent branches
  – Can be used to detect branch prediction problems
  – Can be used as a call trace leading up to Performance Monitor interrupt

• Event-Based Branches (User Mode Interrupts)
  – Allows user-mode programs to catch Performance Monitor alerts
  – Reduces overhead for user-mode programs to monitor themselves
POWER7 SMT Design

- Divided into two thread sets
  - Static mapping between thread number and thread set
  - Moving to lower SMT level requires
    Move execution to appropriate thread(s)
    Nap remaining thread(s)
    Request SMT level change
  - OS tries to keep threads balanced between thread sets by moving execution to appropriate thread
POWER8 SMT Design

- Divided into two thread sets
  - Dynamic mapping between thread number and thread set
  - Moving to lower SMT level requires
    - Nap the idle thread
    - Hardware will shift to the appropriate SMT level
  - Hardware monitors active threads and balances threads between the thread sets

POWER8 automatically tunes itself
# POWER8 Vector/Scalar Unit (VSU)

<table>
<thead>
<tr>
<th></th>
<th>POWER7</th>
<th>POWER8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base SIMD</td>
<td>1X Simple</td>
<td>2X Simple (FX and Logical)</td>
</tr>
<tr>
<td></td>
<td>1X Permute</td>
<td>2X Permute (byte shuffling manipulation)</td>
</tr>
<tr>
<td></td>
<td>1X Complex</td>
<td>2X Complex (integer multiplication)</td>
</tr>
<tr>
<td></td>
<td>W/DW aligned support</td>
<td>Byte aligned support</td>
</tr>
<tr>
<td>Integer SIMD</td>
<td>32 bit integer</td>
<td>64 bit integer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>128 bit integer extension/bit permute</td>
</tr>
<tr>
<td>Compression /Unstructured data/Parallel Bit Stream Processing</td>
<td>-</td>
<td>On-Chip Accelerator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vector CLZ, Vector Gather bits</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GR-VR Direct Move</td>
</tr>
<tr>
<td>Crypto</td>
<td>-</td>
<td>On-Chip Accelerator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AES/SHA User level instructions</td>
</tr>
<tr>
<td>RAID CRC/syndrome (Check sum calculation)</td>
<td>-</td>
<td>Vector Polynomial Multiply</td>
</tr>
<tr>
<td>Binary Floating Point</td>
<td>8 DP Flops/cyc</td>
<td>8 DP Flops/cyc</td>
</tr>
<tr>
<td></td>
<td>8 SP Flops/cyc</td>
<td>16 SP Flops/cyc</td>
</tr>
<tr>
<td>Decimal</td>
<td>Non-Pipelined</td>
<td>Pipeline</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hardware Encryption

- On-Chip Hardware Accelerators introduced with POWER7+
  - POWER8 has same accelerators
  - Offload encryption for OS-based large messages (encrypted file systems, etc)

- POWER8 includes user-mode instructions to accelerate common algorithms

---

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>POWER7+ On-Chip</th>
<th>POWER8 On-Chip</th>
<th>POWER8 In-Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES-GCM</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AES-CTR</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AES-CBC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AES-ECB</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SHA-256</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SHA-512</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RNG</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CRC</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>POWER7+[+] (SW)</th>
<th>POWER8 (HW)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single Thread</td>
<td>Multi Thread</td>
</tr>
<tr>
<td>SHA512</td>
<td>35</td>
<td>10.7</td>
</tr>
<tr>
<td>AES-128-ENC</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>AES-256-ENC</td>
<td>21</td>
<td>5.5</td>
</tr>
</tbody>
</table>
POWER8 Batch Performance
POWER8 Reduces Batch Window Requirements

- **56% lower response time and 2.3x more throughput** with POWER8 (Single Thread mode) than POWER7+ (Single Thread Mode)
- **82% lower response time and 1.4x more throughput** with POWER8 (Single Thread mode) than POWER7+ (SMT4)
- **31% lower response time and 2.9x more throughput** with POWER8 (SMT8) than POWER7+ (SMT4)

POWER8 vs. POWER7+ processor performance on an IBM internal workload that emulates batch tasks performing compression where response time is important.

POWER7+ 740 - 16C
POWER8 S824 - 16C
POWER8 Socket Performance

- POWER7+ 740 - 16C
- POWER8 S824 - 24C

- HotChips 2013 Scale Up Estimate
- S824 Scale Out Measured
Up to 2.7x performance across key workloads vs. other 24-core Scale-Out Systems

1) Results are based on best published results on Xeon E5-2697 v2 from the top 5 Intel system vendors.
2) SPECjbb2013 results are valid as of 7/7/2014. For more information go to http://www.specbench.org/jbb2013/results
3) SPECcpu2006 results are submitted as of 4/22/2014. For more information go to http://www.specbench.org/cpu2006/results/
POWER8

Designed for Big Data - optimized for analytics performance

Processors
flexible, fast execution of analytics algorithms

Memory
large, fast workspace to maximize business insight

Data Bandwidth
bring massive amounts of information to compute resources in real-time

Optimized for a broad range of data and analytics:

IBM Predictive Customer Intelligence

Industry Solutions
Retail  Government  Healthcare  Telecom  Banking

5X Faster

© 2014 International Business Machines Corporation
Thank You!
Definitions

• eDRAM = embedded DRAM
• SMP = Simultaneous Multi-Processing
• SMT = Simultaneous Multi-Threading
• SR-IOV = Single Root I/O Virtualization
• HMC = Hardware Management Console
• SFF = Small Form Factor
Special notices

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquiries, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

IBM hardware products are manufactured from new parts or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.
Special notices (cont.)

IBM, the IBM logo, ibm.com AIX, AIX (logo), AIX 5L, AIX 6 (logo), AS/400, BladeCenter, Blue Gene, ClusterProven, DB2, ESCON, i5/OS, i5/OS (logo), IBM Business Partner (logo), IntelliStation, LoadLeveler, Lotus, Lotus Notes, Notes, Operating System/400, OS/400, PartnerLink, PartnerWorld, PowerPC, pSeries, Rational, RISC System/6000, RS/6000, THINK, Tivoli, Tivoli (logo), Tivoli Management Environment, WebSphere, xSeries, z/OS, zSeries, Active Memory, Balanced Warehouse, CacheFlow, Cool Blue, IBM Watson, IBM Systems Director VMControl, pureScale, TurboCore, Chipopper, Cloudscape, DB2 Universal Database, DS4000, DS6000, DS8000, EnergyScale, Enterprise Workload Manager, General Parallel File System, GPFS, HACMP, HACMP/6000, HASM, IBM Systems Director Active Energy Manager, iSeries, Micro-Partitioning, POWER, PowerLinux, PowerExecutive, PowerVM, PowerVM (logo), PowerHA, Power Architecture, Power Everywhere, Power Family, POWER Hypervisor, Power Systems, Power Systems (logo), Power Systems Software, Power Systems Software (logo), POWER2, POWER3, POWER4, POWER4+, POWER5, POWER5+, POWER6, POWER6+, POWER7, POWER7+, POWER8, POWER7 Systems, System i, System p, System p5, System Storage, System z, TME 10, Workload Partitions Manager and X-Architecture are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries.

A full list of U.S. trademarks owned by IBM may be found at: http://www.ibm.com/legal/copytrade.shtml.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linux is a registered trademark of Linus Torvalds in the United States, other countries or both.

PowerLinux™ uses the registered trademark Linux® pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the Linux® mark on a worldwide basis.

Microsoft, Windows and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries or both.

SPECint, SPECfp, SPECjbb, SPECweb, SPECjAppServer, SPEC OMP, SPECviewperf, SPECapc, SPECjvm, SPECmail, SPECimap and SPECsfs are trademarks of the Standard Performance Evaluation Corp (SPEC).

The Power Architecture and Power.org wordmarks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org.

UNIX is a registered trademark of The Open Group in the United States, other countries or both.

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