IBM Accelerators
July 11, 2016
Outline

- Roadmaps of Z and Power
- Arithmetic Feature Comparison
- How to Get Performance without Frequency
z Systems - Processor Roadmap

z10
2/2008

z196
9/2010

Top Tier Single Thread Performance, System Capacity
Accelerator Integration
Out of Order Execution
Water Cooling
PCIe I/O Fabric
RAIM
Enhanced Energy Management

zEC12
8/2012

Leadership Single Thread, Enhanced Throughput
Improved out-of-order
Transactional Memory
Dynamic Optimization
2 GB page support
Step Function in System Capacity

z13
1/2015

Leadership System Capacity and Performance
Modularity & Scalability
Dynamic SMT
Supports two instruction threads
SIMD
PCIe attached accelerators (XML)
Business Analytics Optimized

Workload Consolidation and Integration Engine for CPU Intensive Workloads
Decimal FP
Infiniband
64-CP Image
Large Pages
Shared Memory
z13 Continues the CMOS Mainframe Heritage Begun in 1994

* MIPS Tables are NOT adequate for making comparisons of z Systems processors. Additional capacity planning required
** Number of PU cores for customer use
**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/queues
- Enhanced prefetching
- 64K data cache, 32K instruction cache

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

From HotChips 2013 Presentation
POWERS8 Core

Execution Improvement vs. POWER7
- SMT4 → SMT8
- 8 dispatch
- 10 issue
- 16 execution pipes:
  - 2 FXU, 2 LSU, 2 LU, 4 FPU, 2 VMX, 1 Crypto, 1 DFU, 1 CR, 1 BR
- Larger Issue queues (4 x 16-entry)
- Larger global completion, Load/Store reorder
- 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 Single Thread
- ~2x Max SMT

From HotChips 2013 Presentation
POWER7 Core Base

I$

IB

ISU

2 LS, 2 FX, 1 BR, 1 CR, 1 (FP, ALU, CX), 1 (FP, PM, DF)

LS LS

LS

32k D$

FX FX

BR CR

256k L2$

FP

ALU

CX

FP

PM

DF

32MB L3$
Enhanced POWER8 Core

- Improved Branch Prediction
- Deeper Out-of-Order Processing

1. IS$ → IB, IB
2. ISU, ISU
3. L, LS, L, LS
4. FX, FX
5. BR, CR
6. 2 LS, 2 LU, 2 FX, 2 (FP, ALU, PM, DF), 1 CR, 1 BR
7. 64k D$
8. 512k L2$
9. 96MB L3$
10. L4$ Mem Buf

- Wider dispatch & issue
- More Execution Bandwidth
- Bigger Caches
- Crypto (AES, SHA) support
z13 Instr / Execution Dataflow

additional instruction flow for higher core throughput
additional execution units for higher core throughput
new arch registers / execution units to accelerate business analytics workloads

*FXa pipes execute reg writers and support b2b execution to itself
FXb pipes execute non-reg writers and non-relative branches (needs 3w AGEN)
## Features of Recent IBM FPUs

<table>
<thead>
<tr>
<th></th>
<th>GHz/FO4</th>
<th>BFU pipe</th>
<th>Core per chip</th>
<th>DFP add</th>
<th>BFU pipes DP-SP- VDP</th>
<th>Features added</th>
<th>problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>P5</td>
<td>2.2 / 23 / 130nm</td>
<td>6</td>
<td>2</td>
<td>OOO</td>
<td>N/A</td>
<td>2 – 2 – 0</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>5.0 / 13 / 65nm</td>
<td>6</td>
<td>2</td>
<td>InO</td>
<td>2 – 4 – 2</td>
<td>13 FO4</td>
<td>inorder</td>
</tr>
<tr>
<td>P7</td>
<td>4.1 / 20 / 45nm</td>
<td>6</td>
<td>8</td>
<td>OOO</td>
<td>+10</td>
<td>2 – 4 – 4</td>
<td>OOO, more VRs</td>
</tr>
<tr>
<td>P7+</td>
<td>4.2 / 20 / 32nm</td>
<td>6</td>
<td>8</td>
<td>OOO</td>
<td>+10</td>
<td>2 – 8 – 4</td>
<td>2 SP per DP</td>
</tr>
<tr>
<td>P8</td>
<td>4.1 / 20 / 22nm</td>
<td>6</td>
<td>12</td>
<td>OOO</td>
<td>+5</td>
<td>2 – 8 – 4</td>
<td>enh DFU + CAPI + SMT8</td>
</tr>
<tr>
<td>Z9</td>
<td>1.7 / 27 / 90nm</td>
<td>5</td>
<td>1</td>
<td>InO</td>
<td>firmware</td>
<td>1 – 1 – 0</td>
<td>software DFU</td>
</tr>
<tr>
<td>Z10</td>
<td>4.4 / 15/ 65nm</td>
<td>7</td>
<td>2</td>
<td>InO</td>
<td>16b-31w</td>
<td>1 – 1 – 0</td>
<td>15 FO4</td>
</tr>
<tr>
<td>Z196</td>
<td>5.2 / 16 / 45nm</td>
<td>8</td>
<td>4</td>
<td>OOO</td>
<td>12l-7t</td>
<td>1 – 1 – 0</td>
<td>OOO screams</td>
</tr>
<tr>
<td>ZEC12</td>
<td>5.5 / 16 / 32nm</td>
<td>8</td>
<td>6</td>
<td>OOO</td>
<td>12l-7t</td>
<td>1 – 1 – 0</td>
<td>Zoned to DFU</td>
</tr>
<tr>
<td>Z13</td>
<td>5.0 / 18 / 22nm</td>
<td>8</td>
<td>8</td>
<td>OOO</td>
<td>8l – 1t</td>
<td>2 – 2 – 2</td>
<td>enh DFU, SIMD, VRs, SMT2</td>
</tr>
<tr>
<td>Model</td>
<td>Year</td>
<td>GHz/FO4</td>
<td>BFU pipe</td>
<td>Core per chip</td>
<td>SIMD</td>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>-----------</td>
<td>---------------</td>
<td>------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>2004</td>
<td>2.2 / 23 / 130nm</td>
<td>6</td>
<td>2</td>
<td>32 x 64b FPRs scalar</td>
<td>2w SMT</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>2007</td>
<td>5.0 / 13 / 65nm</td>
<td>6</td>
<td>2</td>
<td>VMX 32 x 128b VRs + FPRs</td>
<td>2w SMT</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>2010</td>
<td>4.1 / 20 / 45nm</td>
<td>6</td>
<td>8</td>
<td>VSU 64 x 128b VRs, 16bit Vec Int MPY</td>
<td>4w SMT</td>
<td></td>
</tr>
<tr>
<td>P7+</td>
<td>2012</td>
<td>4.2 / 20 / 32nm</td>
<td>6</td>
<td>8</td>
<td>VSU + RNG</td>
<td>4w SMT</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>2014</td>
<td>4.1 / 20 / 22nm</td>
<td>6</td>
<td>12</td>
<td>VSU + Crypto/AES + 32b VINT MPY</td>
<td>8w SMT</td>
<td></td>
</tr>
<tr>
<td>Z9</td>
<td>2005</td>
<td>1.7 / 27 / 90nm</td>
<td>5</td>
<td>1</td>
<td>16 x 64b FPRs scalar</td>
<td>COP- CMPR + CRYPTO</td>
<td></td>
</tr>
<tr>
<td>Z10</td>
<td>2007</td>
<td>4.4 / 15/ 65nm</td>
<td>7</td>
<td>2</td>
<td>FPRs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z196</td>
<td>2010</td>
<td>5.2 / 16 / 45nm</td>
<td>8</td>
<td>4</td>
<td>OOE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>zEC12</td>
<td>2012</td>
<td>5.5 / 16 / 32nm</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z13</td>
<td>2015</td>
<td>5.0 / 18 / 22nm</td>
<td>8</td>
<td>8</td>
<td>32 x 128b VRs +VSU (Int/FP/String)</td>
<td>COP, 2w SMT 32b V INT MPY</td>
<td></td>
</tr>
</tbody>
</table>
Decimal Floating Point Unit Evolution in HDW since 2006

- **Power6**
  - 2006
  - Inorder
  - 13 FO4

- **Power7**
  - 2010
  - Out of Order
  - 20 FO4

- **Power7+**
  - 2012

- **Power8**
  - 2014
  - SRT R16
  - QP signed BCD addition in DFU

- **Partially pipelined**
  - Z10
  - 2007
  - 15 FO4
  - Z196
  - 2010
  - OOE 2X FP perf
  - zEC12
  - 2012

- **Fully pipelined addition**
  - z990
  - GA2
  - 2006
  - Z910
  - 2007
  - 15 FO4
  - QP Binary in DFU
  - z13
  - 2015
  - SMT 2
  - SIMD
FPU Architecture Advances

- Quad Precision Hex in hardware for more years than I’ve been at IBM
- Quad Precision Binary since 1998
- Integer FMA (56 x 56 + 112) in hardware since 2003

- Z196 - 2010
  - BFP new rounding mode (FPC bit 29)
    - Truncate and OR Inexactness
    - supports SP A <= SP B + DP C with 1 rounding error
      - IEEE 754-2008 heterogeneous support
  - DFP quantum exception
    - Tells Software which is emulating a greater precision and range whether
      hardware precision (16 or 34 digits and exponent > 398 or 6176) is exceeded
    - 0.5 mask, 1.5 flag, new DXC code
    - clamped or rounded
    - replaces use of Test Data Group for every operation
  - Converts to/from Integer to BFP/DFP

- zEC12 - 2012
  - Converts to/from zoned Decimal to DFP
More Changes

- **Z13 - 2015**
  - SMT2 and double execution units
  - Non-blocking and separate divide pipeline
  - 4 X size of register file
  - SIMD integer and string
  - SIMD floating-point BFP DP
    - Move away from CC and branches
  - 128 bit integer adds and beyond (support cin and cout)
Trends

- Memory is getting relatively slower
- Frequency constant
- Power important – clock/power gating
- More parallel hardware/software
  - SMT and SIMD/Vector
  - Specialized Accelerators
  - FPGA
- Lines will blur with GPUs
- Need processing power where data stored
Future Systems

Core
Multiple
threads

GPU

FPGA

Probably
On chip

Possibly both
On chip and off chip

OpenPower

POWER8 CAPI
FPGA accelerator attachment

**CAPP: Coherently attached processor proxy**
- Provides PowerBus Surrogate
- Directory of cache lines used by accelerator
- Communicates to accelerator over industry-standard PCIe electrical.
  - Coherency with full Memory Range
  - POWER Effective Address (EA) Translation

**FPGA**
- Accelerated application
- Data
- Control

**Host Service Layer**
- Provides interface to user application layer
- Responds to snoop commands of interest from PowerBus
- Performs Address translation and table walks

**Standard System Topology Preserved**
- Accelerators do not consume a Processor Socket
- Accelerators are a System Option – Not a configuration
- Accelerators do not reduce System Memory
Performance

- Not getting any faster (GHz)
- Parallelism
- Reduce Scalar Bottlenecks
  - Scatter / Gather of memory elements
    - Need parallel cache fetch/store engines
- Mostly Parallel Execution - Predicated Execution
- Avoiding Branches
- Easy Programming model
Final Note: Reliability

- With all the parallelism something will go wrong

- Many parallel executions need to be checked \((1.00001)^{1000}\)
  - Duplication physical vs time redundancy
  - Residue and Parity checking

- Dynamic Sparing

- Or need physical replacing with common FRU