IBM zEnterprise

Lahnstein,
September 30, 2010

Dr. Jürgen Probst
IBM Labor Böblingen
jprobst@de.ibm.com
Penetrating the Technology Wall

Intelligence Beats Speed
CMOS Power Issue: Active vs. Passive Power

Power components:
- Active power
- Passive power
- Gate leakage
- Source – Drain sub-Vt leakage
Technology Discontinuity: Bipolar Power Crisis

Start of Water Cooling

Module Heat Flux (watts/cm²)

Year of Announcement

Bipolar

CMOS

Maintenance: up to 65% reduced
Energy: up to 97% reduced
Area: up to 91% reduced
Frequency Scaling to 11 nm

Optimizing for maximum performance for each core

Constant performance improvement, 20% per gen.

Constant power density 25 W/cm²

D. Frank, C. Tyberg
Workload Optimized Systems

A Workload Optimized System:

- provides value for particular workload or set of workloads important to the client
- provides unique functionality or differentiated performance
- reduces cost of deployment and operation
- is accomplished through co-design of HW, SW and services
System z, a Workload Optimized System
since 45 years
IBM zEnterprise System
The integration of Superior technologies

zEnterprise Unified Resource Manager
- Unifies resources, extending System z qualities of service across the infrastructure
- Install, Monitor, Manage, Optimize, Diagnose & Service

zEnterprise 196
- The industry’s fastest and most scalable enterprise server
- Ideally suited for large scale data and transaction serving and mission critical enterprise applications

zEnterprise BladeCenter Extension

IBM Blades
- Runs app unchanged and supports what you know. Logical device integration between System z and distributed resources

Optimizers
- Workload specific accelerators to deliver a lower cost per transaction, appliance for example IBM Smart Analytics Optimizer
IBM System z: System Design Comparison

System I/O Bandwidth

288 GB/Sec*

Balanced System
CPU, nWay, Memory,
I/O Bandwidth*

Memory
3 TB**

1.5 TB**

512 GB

172.8 GB/sec*

16-way

32-way

54-way

64-way

80-way
Processors

PCI for
1-way

1202

PCI - Processor Capacity Index

* Servers exploit a subset of its designed I/O capability
** Up to 1 TB per LPAR

z196
z10 EC
z9 EC
zSeries 990
zSeries 900

© 2010 IBM Corporation
z196 Continues the CMOS Mainframe Heritage

- **G4** – 1st full-custom CMOS S/390®
- **G5** – IEEE-standard BFP; branch target prediction
- **G6** – Copper Technology (Cu BEOL)
- **z900** – Full 64-bit z/Architecture
- **z990** – Superscalar CISC pipeline
- **z9 EC** – System level scaling
- **z10 EC** – Architectural extensions
- **z196** – Additional Architectural extensions and new cache structure
z196 Architecture

- **Continues line of upward-compatible mainframe processors**

- **Rich CISC Instruction Set Architecture (ISA)**
  - 984 instructions (762 implemented entirely in hardware)
  - 24, 31, and 64-bit addressing modes
  - Multiple address spaces robust inter-process security
  - Multiple arithmetic formats
  - Industry-leading virtualization support
    - High-performance logical partitioning via PR/SM
    - Fine-grained virtualization via z/VM scales to 1000’s of images
  - Precise, model-independent definition of hardware/software interface

- **Architectural extensions for IBM z196**
  - 110+ new instructions added to improve compiled code efficiency
  - Decimal floating point quantum exceptions
  - New crypto functions and modes
  - Virtual architectural level
  - Non-quiescing SSKE
z196 Water cooled – Under the covers (Model M66 or M80) front view

- Internal Batteries (optional)
- Power Supplies
- Support Elements
- I/O cage
- I/O drawers
- Ethernet cables for internal System LAN connecting Flexible Service Processor (FSP) cage controller cards
- Processor Books, Memory, MBA and HCA cards
- InfiniBand I/O Interconnects
- 2 x Water Cooling Units
96mm x 96mm MCM
- 103 Glass Ceramic layers
- 8 chip sites
- 7356 LGA connections
- 20 and 24 way MCMs
- Maximum power used by MCM is 1800W

CMOS 12s chip Technology
- PU, SC, S chips, 45 nm
- 6 PU chips/MCM – Each up to 4 cores
  - One memory control (MC) per PU chip
  - 23.498 mm x 21.797 mm
  - 1.4 billion transistors/PU chip
  - L1 cache/PU core
    - 64 KB I-cache
    - 128 KB D-cache
  - L2 cache/PU core
    - 1.5 MB
  - L3 cache shared by 4 PUs per chip
    - 24 MB
  - 5.2 GHz
- 2 Storage Control (SC) chip
  - 24.427 mm x 19.604 mm
  - 1.5 billion transistors/SC chip
  - L4 Cache 96 MB per SC chip (192 MB/Book)
  - L4 access to/from other MCMs
- 4 SEEPROM (S) chips
  - 2 x active and 2 x redundant
  - Product data for MCM, chips and other engineering information
- Clock Functions – distributed across PU and SC chips
  - Master Time-of-Day (TOD) function is on the SC
z196 Quad Core PU Chip Detail

- **Up to Four active cores per chip**
  - 5.2 GHz
  - L1 cache/ core
    - 64 KB I-cache
    - 128 KB D-cache
  - 1.5 MB private L2 cache/ core
- **Two Co-processors (COP)**
  - Crypto & compression accelerators
    - Includes 16KB cache
    - Shared by two cores
- **24MB eDRAM L3 Cache**
  - Shared by all four cores
- **Interface to SC chip / L4 cache**
  - 41.6 GB/sec to each of 2 SCs
- **I/O Bus Controller (GX)**
  - Interface to Host Channel Adapter (HCA)
- **Memory Controller (MC)**
  - Interface to controller on memory DIMMs
  - Supports RAIM design

- **12S0 45nm SOI Technology**
  - 13 layers of metal
  - 3.5 km wire
- **1.4 Billion Transistors**
- **Chip Area** – 512.3mm²
  - 23.5mm x 21.8mm
  - 8093 Power C4’s
  - 1134 signal C4’s
- **I/O Bus Controller (GX)**
  - Interface to Host Channel Adapter (HCA)
- **Memory Controller (MC)**
  - Interface to controller on memory DIMMs
  - Supports RAIM design
z196 SC Chip Detail

- **12S0 45nm SOI Technology**
  - 13 layers of metal

- **Chip Area**
  - 478.8mm²
  - 24.4mm x 19.6mm
  - 7100 Power C4’s
  - 1819 signal C4’s

- **1.5 Billion Transistors**
  - 1 Billion cells for eDRAM

- **eDRAM Shared L4 Cache**
  - 96 MB per SC chip
  - 192 MB per Book

- **6 CP chip interfaces**
- **3 Fabric interfaces**
- **2 clock domains**
- **5 unique chip voltage supplies**
z196 Book Layout

- Backup Air Plenum
- 16X DIMMs 100mm High
- MCM @ 1800W Refrigeration Cooled or Water Cooled
- Refrigeration Cooled or Water Cooled
- 8 I/O FAN OUT 2 FSP
- 2 FSP
- 8 I/O FAN OUT
- MCM
- 3x DCA
- 11 VTM Card Assemblies 8 Vertical 3 Horizontal
- 14X DIMMs 100mm High
- Cooling from/to MRU
- Memory
- DCA Power Supplies
- Fanout Cards
z196 PU core

- Each core is a superscalar, out of order processor with these characteristics:
  - Six execution units
    - 2 fixed point (integer), 2 load/store, 1 binary floating point, 1 decimal floating point
  - Up to three instructions decoded per cycle (vs. 2 in z10)
  - 211 complex instructions cracked into multiple internal operations
    - 246 of the most complex z/Architecture instructions are implemented via millicode
  - Up to five instructions/operations executed per cycle (vs. 2 in z10)
  - Execution can occur out of (program) order
    - Memory address generation and memory accesses can occur out of (program) order
    - Special circuitry to make execution and memory accesses appear in order to software
  - Each core has 3 private caches
    - 64KB 1<sup>st</sup> level cache for instructions, 128KB 1<sup>st</sup> level cache of data
    - 1.5MB L2 cache containing both instructions and data
z196 Out of Order (OOO) Value

- OOO yields significant performance benefit for compute intensive apps through
  - Re-ordering instruction execution
    - Later (younger) instructions can execute ahead of an older stalled instruction
  - Re-ordering storage accesses and parallel storage accesses
- OOO maintains good performance growth for traditional apps
z196 vs. z10 hardware comparison

- **z10 EC**
  - CPU
    - 4.4 Ghz
  - Caches
    - L1 private 64k instr, 128k data
    - L1.5 private 3 MBs
    - L2 shared 48 MBs / book
    - book interconnect: star

- **z196**
  - CPU
    - 5.2 Ghz
    - Out-of-Order execution
  - Caches
    - L1 private 64k instr, 128k data
    - L2 private 1.5 MBs
    - L3 shared 24 MBs / chip
    - L4 shared 192 MBs / book
    - book interconnect: star
System z overall RAS Strategy
.....Continuing our RAS focus helps avoid outages

<table>
<thead>
<tr>
<th>Unscheduled Outages</th>
<th>Prior Servers</th>
<th>z9 EC</th>
<th>Z10 EC</th>
<th>z196</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Scheduled Outages</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Planned Outages</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Preplanning</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Temperature = Silicon Reliability Worst Enemy
Wearout = Mechanical Components Reliability Worst Enemy.
Layers of Memory Recovery

**ECC**
- Powerful 90B / 64B Reed Solomon code

**DRAM Failure**
- Marking technology; no half sparing needed
- 2 DRAM can be marked
- Call for replacement on third DRAM

**Lane Failure**
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

**DIMM Failure (discrete components, VTT Reg.)**
- CRC with Retry
- Data – lane sparing
- CLK – RAIM with lane sparing

**DIMM Controller ASIC Failure**
- RAIM Recovery

**Channel Failure**
- RAIM Recovery
z196 Redundant Array of Independent Memory (RAIM) Structure

Extra column provides RAIM function
z196 I/O Infrastructure

- **Book 0**
  - Memory
  - HCA
  - IFB-MP
  - FBC/L4 Cache
  - RII
  - 2 GBps mSTI
  - Channels
  - FICON Express8
  - 2/4/8 Gbps

- **Book 1**
  - Memory
  - HCA
  - IFB-MP
  - FBC/L4 Cache
  - RII
  - 500 MBps mSTI
  - ISC-3

- **Book 2**
  - Memory
  - HCA
  - IFB-MP
  - FBC/L4 Cache
  - RII
  - 333 MBps mSTI
  - ESCON

- **Book 3**
  - Memory
  - HCA
  - IFB-MP
  - FBC/L4 Cache
  - RII
  - 2 GBps mSTI
  - OSA-Express3
  - 10 GbE

**Connections**
- HCA2-C fanout
- 1st level Copper Cables
  - 6 GBps
- 2nd level Embedded
**z196 I/O Cages and Drawers**

*z196 I/O infrastructure will support I/O cages (z10 EC) and I/O drawers (z10 BC)*

**I/O cage and Drawer considerations**

- Can’t order cages or drawers, have to order I/O and/or Crypto features, eConfig will deliver the correct mix of I/O drawers and I/O cages
- There is NO Plan Ahead option available due to the concurrent nature of the I/O drawers
- The I/O drawer can be concurrently added or removed (non-disruptively)
- I/O cage additions and removals are disruptive
- Air cooled models will have a max of 2 x PSC, Water cooled models will have a max of 1 x PSC

**Consolidation and Pre-Planning**

- For customers with large I/O requirements, focus on consolidating to fewer then 72 slots, reducing to 64 slots or less would be ideal as it would leave room for future I/O expansion.
z196 New Build and Box MES I/O offerings

NO I/O Cage

<table>
<thead>
<tr>
<th>FRAME</th>
<th>Z</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>IBF</td>
<td>IBF</td>
</tr>
<tr>
<td>41</td>
<td>IBF</td>
<td>Air Backup</td>
</tr>
<tr>
<td>40</td>
<td>BPA</td>
<td>CPC</td>
</tr>
<tr>
<td>39</td>
<td>I/O Drawer</td>
<td>MRU</td>
</tr>
<tr>
<td>38</td>
<td>Slots 1-8</td>
<td>Slots 17-24</td>
</tr>
<tr>
<td>37</td>
<td>Slots 9-16</td>
<td>Slots 25-32</td>
</tr>
<tr>
<td>36</td>
<td>Slots 1-32</td>
<td></td>
</tr>
</tbody>
</table>

ONE I/O Cage

<table>
<thead>
<tr>
<th>FRAME</th>
<th>Z</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>IBF</td>
<td>IBF</td>
</tr>
<tr>
<td>41</td>
<td>IBF</td>
<td>Air Backup</td>
</tr>
<tr>
<td>40</td>
<td>BPA</td>
<td>CPC</td>
</tr>
<tr>
<td>39</td>
<td>I/O Drawer</td>
<td>MRU</td>
</tr>
<tr>
<td>38</td>
<td>Slots 17-24</td>
<td>Slots 37-44</td>
</tr>
<tr>
<td>37</td>
<td>Slots 25-32</td>
<td>Slots 57-64</td>
</tr>
<tr>
<td>36</td>
<td>Slots 1-32</td>
<td>Slots 29-56</td>
</tr>
<tr>
<td>35</td>
<td>Slots 9-16</td>
<td>Slots 65-72</td>
</tr>
<tr>
<td>34</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Slots 9-16</td>
<td></td>
</tr>
</tbody>
</table>

33-44 slots

Note: Drawings Not to Scale

TWO I/O Cages

<table>
<thead>
<tr>
<th>FRAME</th>
<th>Z</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>IBF</td>
<td>IBF</td>
</tr>
<tr>
<td>41</td>
<td>IBF</td>
<td>Air Backup</td>
</tr>
<tr>
<td>40</td>
<td>BPA</td>
<td>CPC</td>
</tr>
<tr>
<td>39</td>
<td>I/O Drawer</td>
<td>MRU</td>
</tr>
<tr>
<td>38</td>
<td>Slots 1-32</td>
<td>Slots 29-56</td>
</tr>
<tr>
<td>37</td>
<td>Slots 9-16</td>
<td>Slots 65-72</td>
</tr>
<tr>
<td>36</td>
<td>Slots 17-24</td>
<td>Slots 29-56</td>
</tr>
<tr>
<td>35</td>
<td>Slots 25-32</td>
<td>Slots 1-28</td>
</tr>
<tr>
<td>34</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Slots 9-16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Slots 17-24</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Slots 25-32</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Slots 1-32</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Slots 9-16</td>
<td></td>
</tr>
</tbody>
</table>

45-72* required slots

64 or fewer slots IDEAL

Note: Drawings Not to Scale
### z196 Channel Type and Crypto Overview

#### Supported Channel Types

<table>
<thead>
<tr>
<th><strong>I/O Channels</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>FICON Express8</td>
</tr>
<tr>
<td>FICON Express4 (Carry forward on upgrade)</td>
</tr>
<tr>
<td>ESCON – Migrate Away (240 or fewer)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Networking</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>OSA-Express3</td>
</tr>
<tr>
<td>• 10 Gigabit Ethernet LR and SR</td>
</tr>
<tr>
<td>• Gigabit Ethernet LX and SX</td>
</tr>
<tr>
<td>• 1000BASE-T Ethernet</td>
</tr>
<tr>
<td>OSA-Express2 (Carry forward on upgrade)</td>
</tr>
<tr>
<td>• 1000BASE-T Ethernet</td>
</tr>
<tr>
<td>• Gigabit Ethernet LX and SX</td>
</tr>
<tr>
<td>HiperSockets (Define only)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Coupling Links</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>InfiniBand Coupling Links</td>
</tr>
<tr>
<td>• 12x InfiniBand</td>
</tr>
<tr>
<td>• 1x InfiniBand</td>
</tr>
<tr>
<td>ISC-3 – Migrate Away (Peer mode only)</td>
</tr>
<tr>
<td>IC (Define only)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Crypto</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crypto Express3</td>
</tr>
<tr>
<td>• Configurable Coprocessor / Accelerator</td>
</tr>
</tbody>
</table>

#### Non-Supported Channel Types

<table>
<thead>
<tr>
<th><strong>I/O Channels</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>FICON (before FICON Express4)</td>
</tr>
<tr>
<td>FCV – ESCD Model 5 Bridge Card</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Networking</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>OSA-Express2 10 GbE LR</td>
</tr>
<tr>
<td>OSA-Express (pre OSA-Express2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Coupling Links</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ICB-4 and earlier ICB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Crypto</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crypto Express2 and earlier</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>ETR</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sysplex Timer® (ETR) Attachment</td>
</tr>
</tbody>
</table>
z196 – Parallel Sysplex (No ICB-4 Link Support)

The "intermediate" CFs can provide a 'bridge' to connect to z196
- No IBC-4 Links to z196
- Can intermix existing ICB-4 and InfiniBand link technology if using z9 or z10 Coupling Facilities
z196 HiperSockets – doubled the number

- High-speed “intraserver” network
- Independent, integrated, virtual LANs
- Communication path – system memory
- Communication across LPARs
  - Single LPAR - connect up to 32 HiperSockets
- Support for multiple LCSS's & spanned channels
- Virtual LAN (IEEE 802.1q) support
- HiperSockets Network Concentrator
- Broadcast support for IPv4 packets
- IPv6
- HiperSockets Network Traffic Analyzer (HS NTA)
- No physical cabling or external connections required
z196 Capacity per Watt improvements

<table>
<thead>
<tr>
<th></th>
<th>15 years of CMOS: G2 to z196</th>
<th>Net Effect: G2 to z196</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power Increase:</td>
<td>17% per year</td>
<td>Performance increased by:</td>
</tr>
<tr>
<td>Performance increase:</td>
<td>46% per year</td>
<td>Performance / kWatt increased by:</td>
</tr>
<tr>
<td>Power density increase:</td>
<td>13% per year</td>
<td>Performance / sq ft increased by:</td>
</tr>
</tbody>
</table>

Note: z196 data is best available 8/27/2010
z196 Water cooling infrastructure

Rear View – A Frame

2 x Water Cooling Units

2 x Exhaust Air Heat Exchangers

Rear View – A and Z Frame
Water Conditioning Unit (WCU) detail

- Internal Water Reservoir
- Pump
- Water to Water Heat Exchanger
- Temperature Control Valve (controls flow rate)
- Customer water
Water cooling option

- Water cooled cold plate on processor MCM in each processor book
- N+1 Water Conditioning Unit (WCU) with independent chilled water connections
- One WCU can support system
- Heat Exchanger (HX) removes heat from exhaust air at back of both frames
- Target to remove 60-65% of air heat load from the System z
- Results in ~10kW system air heat load max (5kW per frame)
- Input energy savings of ~2-3kW/system for 3 and 4 book system.
- Additional power savings in data center for reduced air cooling heat load
- Air cooling back-up mode for maximum robustness if lose chilled water to system

Internal, closed, conditioned water loop

Chilled water flow is function of heat load on WCU & chilled water temp.

Not to scale, not physically representative

Demo purposes only, XAHX mounted in rear not on top of frames

Air cooling back-up mode for maximum robustness if complete loss of chilled water occurs!
zEnterprise Static Power Save Mode

- **Main use cases**
  - Periods of low utilization
  - CBU Systems: Systems used for disaster recovery

- **Base mechanism**
  - Build upon existing RAS functions (frequency/voltage variation) implemented originally for MRU failures (since z900)
  - Use frequency and voltage reduction to reduce energy consumption of system
  - System continues to operate with MCM refrigeration (most power efficient)
  - Only explicitly triggered by customer. No autonomic changes done “under the cover”

- **Customer Controls**
  - Controls implemented in HMC, SE and Active Energy Manager
  - Granularity of saving steps: one power saving mode

- **Power Save Mode expectations**
  - Frequency reduction: ~ 17%
  - Processor voltage reduction: ~ 9% voltage reduction
  - Expected system power savings: ~ 10%-20% power savings (configuration dependent)
zEnterprise: A natural evolution, and a Virtualization Revolution

- **Application specialty engines**
  - Dedicated processors for key environments (e.g. Linux, Java™)
  - Improved price / performance for new workloads
  - Very low cost of large scale consolidation

- **Expanding the specialty engine concept to enable more applications**
  - Integrated / networked attached resources to optimization for cost, performance and quality of service
  - Take advantage of innovative new technologies

- **Next Generation: Integrated Virtual Server Management**
  - Integrated Platform Management across diverse platforms from a single control point to lower cost and improve service
  - Workload management of enterprise applications across virtual servers to improve quality of service
Emerging Applications with Special-Purpose Capabilities

*Future objectives include extended application integration and optimization*

**General Purpose Enterprise Systems**
- Optimized for a broad set of applications or components

**Special Purpose Systems and Optimizers**
- Optimized for a specific set of applications or components

**Evolving & Emerging Workload Components**
- Traditional Workload Components
  - Java
  - XML
  - Analytics
  - SOA
  - Sensors
  - Encryption
  - Networking
  - Digital Media
  - Search
  - Events
  - Data Protection

Integration will be critical

Both General and Special Purpose capabilities are needed
## IBM zEnterprise System

### The integration of Superior technologies

<table>
<thead>
<tr>
<th><strong>zEnterprise</strong></th>
<th><strong>Unified Resource Manager</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unifies resources, extending System z qualities of service across the infrastructure</td>
</tr>
<tr>
<td></td>
<td>Install, Monitor, Manage, Optimize, Diagnose &amp; Service</td>
</tr>
</tbody>
</table>

### zEnterprise 196

- The industry's fastest and most scalable enterprise server
- Ideally suited for large scale data and transaction serving and mission critical enterprise applications

### zEnterprise BladeCenter Extension

### IBM Blades

- Runs app unchanged and supports what you know. Logical device integration between System z and distributed resources

### Optimizers

- Workload specific accelerators to deliver a lower cost per transaction, appliance for example IBM Smart Analytics Optimizer

---

*Integration and centralized management by System z*
IBM zEnterprise System

IBM zEnterprise 196 (z196)  IBM zEnterprise BladeCenter Extension (zBX)

IBM zEnterprise Unified Resource Manager (zManager)
zEnterprise System – 196 + zBX + Unified Resource Manager

IBM Blades
- Linux on System x Blades*
- AIX® on POWER7 Blades

Optimizers
- IBM Smart Analytics Optimizer
- Future Offering
- Future Offering
- WebSphere DataPower Appliance*

BLADE HW RESOURCES
- Blade Virtualization
- Blade Virtualization

zBX

System z PR/SM™
- Support Element

z196
- z/OS
- z/TPF
- z/VSE
- Linux on System z
- Linux on System z
- z/VM

Customer Network
- Ensemble Management Firmware
- Intraensemble data network
- Intranode management network
- Customer Network

*All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.
Smart Analytics Optimizer - New DB2 Resource Manager

Applications

DBA Tools, z/OS Console, ...

Application Interfaces (standard SQL dialects)

Operation Interfaces (e.g. DB2 Commands)

Data Manager
Buffer Manager
IRLM
Log Manager

z/OS

Smart Analytics Optimizer

DB2
What is a zEnterprise Ensemble?

- A zEnterprise ensemble is a collection of 1 to 8 z196 CPCs with/without zBX managed collectively by the Unified Resource Manager as a single logical virtualized system using the HMC
- A zEnterprise node is a z196 CPC with 0 to 4 racks up to 2 BladeCenters per rack
  - zEnterprise nodes are deployed within a single site
  - A zEnterprise node can be a member of at most one ensemble
- z196 CPCs are deployed within a single site
- Blade based fit-for-purpose Solutions
- Integrated Advanced Virtualization Management
- Implements well-defined external interface to Data Center Service Management functions
- Virtual Resource Management and Automation
IBM Multi-Architecture Virtualization – “Fit for Purpose”
System z Multi-System, Federated Hypervisor Configuration

- **System z futures**: hosting a federation of platform management functions, including:
  - Resource monitoring
  - Workload management
  - Availability management
  - Image management
  - Energy management

- **Integrates with hardware management and virtualization functions**
- **Controls hypervisors and management agents on blades**
- **Open integration to enterprise-level management software**
Service levels to match your business needs
Increased flexibility for your multi-architecture strategy when data is on z/OS

zEnterprise System

TCO Focus

- Silo managed islands of computing
- Less dynamic than z virtualization
- Minimal resource sharing with z resources

TCA Focus

- Expanded ISV support for enterprise applications
- Targeted for applications that interact with mainframe data and transactions
- Provisioned and managed by System z

Select IBM Blades in zBX

- Extreme consolidation of servers and networking
- Superior levels of virtual server provisioning, monitoring and workload management
- Industry-best virtual I/O bandwidth and reliability
- Fewer components and reduced complexity
- System z qualities of dynamic resource management and capacity-on-demand
- Seamless integration with z/OS backup and disaster recovery solutions

Linux on z/VM®

- Extreme scalability and performance for transaction processing and data serving
- High availability and cross-system scalability with Parallel Sysplex® and GDPS®
- Leading policy-based capacity provisioning and workload management
- Pervasive, high-performance security support

z/OS

LOWER

SCALABILITY, SECURITY,
DYNAMIC WORKLOAD MANAGEMENT

HIGHER

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