IBM zEnterprise

zEnterprise.
A New Dimension in Computing

IBM zEnterprise 196

Dr. Stephan Kammerer
IBM Deutschland Research & Development
September 2010
The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

<table>
<thead>
<tr>
<th>AIX*</th>
<th>IMS</th>
<th>S/390*</th>
<th>z10 EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPN*</td>
<td>InfiniBand*</td>
<td>Sysplex Timer*</td>
<td>z/Architecture*</td>
</tr>
<tr>
<td>BladeCenter*</td>
<td>Multiprise*</td>
<td>System/390*</td>
<td>z/OS*</td>
</tr>
<tr>
<td>CICS*</td>
<td>OS/2*</td>
<td>System p*</td>
<td>z/VM*</td>
</tr>
<tr>
<td>DB2*</td>
<td>Parallel Sysplex*</td>
<td>System x*</td>
<td>z/VSE</td>
</tr>
<tr>
<td>DataPower*</td>
<td>Power*</td>
<td>System z*</td>
<td>zEnterprise</td>
</tr>
<tr>
<td>DS8000*</td>
<td>POWER*</td>
<td>System z9*</td>
<td>zSeries*</td>
</tr>
<tr>
<td>e business (logo)*</td>
<td>POWER7</td>
<td>System z10*</td>
<td></td>
</tr>
<tr>
<td>ESCON*</td>
<td>Power Architecture*</td>
<td>VSE/ESA</td>
<td></td>
</tr>
<tr>
<td>eServer</td>
<td>PowerVM</td>
<td>WebSphere*</td>
<td></td>
</tr>
<tr>
<td>FICON*</td>
<td>PR/SM</td>
<td>X-Architecture*</td>
<td></td>
</tr>
<tr>
<td>GDPS*</td>
<td>Resource Link*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HiperSockets</td>
<td>Redbooks*</td>
<td>z9*</td>
<td></td>
</tr>
<tr>
<td>IBM*</td>
<td>REXX</td>
<td>z10 Business Class</td>
<td></td>
</tr>
<tr>
<td>IBM (logo)*</td>
<td>RMF</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries. Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

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.

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

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

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:
Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

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

IBM does not develop the software products used in this presentation. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

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

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.
Von 'System/360' (S/360) zu ESA/390 und z/-Architektur

- 1964 S/360
  - CISC, 24bit Adressierung, 'Real Storage', Uniprozessoren
  - Amdahl, G.M., Blaauw, G.A., and Brooks, F.P.: *Architecture of the IBM System/360*

- 1971 S/370
  - 'Virtual Storage', Multiprozessor-Unterstützung, ...

- 1981 S/370 XA (Extended Architecture)
  - 31bit Adressierung (2GB), 'Expanded Storage' (>2GB), 'Channel Subsystem'
  - 'Interpretive Execution': Basis für Logische Partitionierung ('LPAR')

- 1988 ESA/370
  - ESA = Enterprise Systems Architecture, Logische Partitionierung
  - Ausbau der Speicher-Zugriffsmethoden: Mehr als ein 'address space'

- 1990 ESA/390
  - 'ESCON' (Enterprise Systems Connection Architecture) Glasfasertechnologie ...
  - DatenkompRESSION, KryptoGRAPHie, LPAR Erweiterungen

- 1994 Parallel Sysplex, Übergang von Bipolar zu CMOS Technologie
  - 'Coupling Facility', Cluster von bis zu 32 x 16-way MultiProzessoren
  - 'FICON' (Fiber Channel Connectivity), Ausbau der Glasfasertechnologie

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
Do GHz matter?

- GHz does matter
  - It is the "rising tide that lifts all boats"
  - It is especially important for CPU-intensive applications

- GHz is not the only dimension that matters
  - System z focus is on balanced system design across many factors
    - Frequency, pipeline efficiency, energy efficiency, cache / memory design, I/O design

- System performance is not linear with frequency
  - Need to use LSPR+ System z capacity planning tools for real client / workload sizing

- System z has been on consistent path while others have oscillated between extremes
  - Growing frequency steadily, with occasional jumps/step functions (G4 in 1997, z10 in 2008)
Industry’s Approach to Integrated Systems Performance

- Microprocessor frequency will no longer be the dominant driver of system level performance
- Scale-out and small SMPs will continue to outpace scale-up growth
- Systems will increasingly rely on modular components for continued performance leadership
- Systems will be designed with the ability to dynamically manage and optimize power
- Integration over the entire stack, from semiconductor technology to end-user applications, will replace scaling as the major driver of increased system performance
IBM z196 Instruction Set Architecture

- Continues line of upward-compatible mainframe processors
  - Application compatibility since 1964
  - Supports all z/Architecture-compliant OSes
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 performance and scalability

Think Inside the box!  Think System z Qualities of Service!

- zEnterprise ensembles – Multiple nodes
  - Node – z196 with or without zBX
- zEnterprise Unified Resource Manager
- Multiple architectures
  - z/Architecture®
  - Power Architecture® - POWER7
  - X-Architecture®

- z196
  - Largest z196 model
    1.6x compared to z10 EC E64
  - Equivalent n-ways
    1.4x compared to z10 EC
  - With compiler optimization additional 30% additional for some CPU intensive work

Federated capacity from multiple architectures
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
256 GB
64 GB
32-way
64-way
54-way
80-way
PCI for 1-way
1202

Processors

96 GB/sec
24 GB/sec
920
3 TB**
1202
172.8 GB/sec*
288 GB/Sec*

* Servers exploit a subset of its designed I/O capability
** Up to 1 TB per LPAR
PCI - Processor Capacity Index

z196
z10 EC
z9 EC
zSeries 990
zSeries 900
Scalability: System-Structures optimized for data

The key problem of current microprocessor-systems: Memory access does not scale with CPU-cycletime!
z196 Overview

Machine Type
- 2817

5 Models
- M15, M32, M49, M66 and M80

Processor Units (PUs)
- 20 (24 for M80) PU cores per book
- Up to 14 SAPs per system, standard
- 2 spares designated per system
- Dependant on the H/W model - up to 15,32,49,66 or 80 PU cores available for characterization
  - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z Application Assist Processors (zAAPs), System z Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
- Sub-capacity available for up to 15 CPs
  - 3 sub-capacity points

Memory
- System Minimum of 32 GB
- Up to 768 GB per book
- Up to 3 TB for System and up to 1 TB per LPAR
  - Fixed HSA, standard
  - 32/64/96/128/256 GB increments

I/O
- Up to 48 I/O Interconnects per System @ 6 GBps each
- Up to 4 Logical Channel Subsystems (LCSSs)

STP - optional (No ETR)
z196 – Under the covers (Model M66 or M80)

- Internal Batteries (optional)
- Power Supplies
- 2 x Support Elements
- I/O cage
- I/O drawers
- Fiber Quick Connect (FQC) Feature (optional)
- Processor Books, Memory, MBA and HCA cards
- Ethernet cables for internal System LAN connecting Flexible Service Processor (FSP) cage controller cards
- InfiniBand I/O Interconnects
- FICON & ESCON FQC
- 2 x Cooling Units
- Flexible Service Processor (FSP) cage controller cards
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
z196 Multi-Chip Module (MCM) Packaging

- **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
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 1st level cache for instructions, 128KB 1st level cache of data
  - 1.5MB L2 cache containing both instructions and data
## z10 EC MCM vs. z196 MCM Comparison

<table>
<thead>
<tr>
<th>Z10 EC MCM</th>
<th>Z196 MCM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MCM</strong></td>
<td><strong>MCM</strong></td>
</tr>
<tr>
<td>– 96mm x 96mm in size</td>
<td>– 96mm x 96mm in size</td>
</tr>
<tr>
<td>– 5 PU chips per MCM</td>
<td>– 6 PU chips per MCM</td>
</tr>
<tr>
<td>• Quad core chips with 3 or 4 active cores</td>
<td>• Quad core chips with 3 or 4 active cores</td>
</tr>
<tr>
<td>• PU Chip size 21.97 mm x 21.17 mm</td>
<td>• PU Chip size 23.5 mm x 21.8 mm</td>
</tr>
<tr>
<td>• 4.4 GHz</td>
<td>• 5.2 GHz</td>
</tr>
<tr>
<td>• Superscalar, In order execution</td>
<td>• Superscalar, OOO execution</td>
</tr>
<tr>
<td>• L1: 64K I / 128K D private/core</td>
<td>• L1: 64K I / 128K D private/core</td>
</tr>
<tr>
<td>• L1.5: 3M I+D private/core</td>
<td>• L2: 1.5M I+D private/core</td>
</tr>
<tr>
<td>• L2: 2 x 24 M = 48 M L2 per book</td>
<td>• L3: 24MB/chip - shared</td>
</tr>
<tr>
<td>• SC Chip size 21.11 mm x 21.71 mm</td>
<td>• SC Chip size 24.4 mm x 219.6 mm</td>
</tr>
<tr>
<td>– Power 1800 Watts</td>
<td>– Power 1800 Watts</td>
</tr>
</tbody>
</table>
Evolution of System z Specialty Engines

Building on a strong track record of technology innovation with specialty engines – DB Compression, SORT, Encryption, Vector Facility

Eligible for zIIP:
- DB2 remote access and BI/DW
- ISVs
- New! IPSec encryption
- z/OS XML
- z/OS Global Mirror*

 Eligible for zAAP:
- Java execution environment
- z/OS XML

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

*SOD: IBM plans to enhance z/VM in a future release to support the new System z10 EC capability to allow any combination of CP, zIIP, zAAP, IFL, and ICF processor-types to reside in the same z/VM LPAR.
Processing Units

- Modern processor offer many specialist processor types. Most of the additional types are not counted as full processors for software charging
  - Central Processor (CP) - A full z/OS processor
  - System Assistance Processor (SAP) - Used for the I/O subsystem – each machine has at least one
  - Integrated Facility for Linux (IFL) - Special processor for Linux - optional
  - zAAP – Used for JAVA code and XML – optional
  - zIIP – Used for DB2 processing, XML, IPSec & Global Mirror – optional
  - Integrated Coupling Facility (ICF) - For coupling facilities
  - Spares

- SAPs, IFLs, ICFs, zIIPs, and zAAPs are not counted in the model number and or against software costs
### z196 Compression and Cryptography Accelerator

- **Data compression engine**
  - Static dictionary compression and expansion
  - Dictionary size up to 64KB (8K entries)
    - Local 16KB cache per core for dictionary data

- **CP Assist for Cryptographic Function (CPACF)**
  - Enhancements for new NIST standard
  - Complemented prior ECB and CBC symmetric cipher modes with XTS, OFB, CTR, CFB, CMAC and CCM
  - New primitives (128b Galois Field multiply) for GCM

- **Accelerator unit shared by 2 cores**
  - Independent compression engines
  - Shared cryptography engines
z196 SC Chip Detail

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

- **Chip Area**
  - 478.8mm^2
  - 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**
System z Cache Topology – z10 EC vs. z196 Comparison

**z10 EC**

- **L1**: 64K + 128K
  - 8w Set Associative DL1
  - 4w Set Associative IL1
  - 256B cache line size

- **L1.5**: 3MB Inclusive of L1.5
  - 12w Set Associative
  - 256B cache line size

- **L2**: 48MB Excl Inclusive + XI Dir
  - 24w Set Associative
  - 256B cache line size

**z196**

- **L1**: 64K + 128K
  - 8w DL1, 4w IL1
  - 256B cache line size

- **L2**: Private 1.5MB Inclusive of L1s
  - 12w Set Associative
  - 256B cache line size

- **L3**: Shared 24MB Inclusive of L2s
  - 12w Set Associative
  - 256B cache line size

- **L4**: 192MB Inclusive
  - 24w Set Associative
  - 256B cache line size

**Cache Topology**

- **z10 EC** has 4 L2 Caches:
  - 48MB Shared L2

- **z196** has 4 L4 Caches:
  - 192MB Shared eDRAM L4
z196 PU chip, SC chip and MCM

**z196 Quad Core PU CHIP**

**96 MB SC CHIP**

**MCM**

Front View

Front View Fanouts
z196 Book Layout

- Backup Air Plenum
- 16X DIMMs 100mm High
- MCM @ 1800W Refrigeration Cooled or Water Cooled
- 8 I/O FAN OUT (HCA2-C, HCA-O)
- 2 FSP
- 3x DCA
- 11 VTM Card Assemblies
  - 8 Vertical
  - 3 Horizontal
- 14X DIMMs 100mm High
- DCA Power Supplies
- Memory
- Memory
- Cooling from/to MRU
- Fanout Cards

Rear
Front
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**
• **Out of order yields significant performance benefit through**
  
  – Re-ordering instruction execution
    
    • Instructions stall in a pipeline because they are waiting for results from a previous instruction or the execution resource they require is busy
    
    • In an in-order core, this stalled instruction stalls all later instructions in the code stream
    
    • In an out-of-order core, later instructions are allowed to execute ahead of the stalled instruction
  
  – Re-ordering storage accesses
    
    • Instructions which access storage can stall because they are waiting on results needed to compute storage address
    
    • In an in-order core, later instructions are stalled
    
    • In an out-of-order core, later storage-accessing instructions which can compute their storage address are allowed to execute
  
  – Hiding storage access latency
    
    • Many instructions access data from storage
    
    • Storage accesses can miss the L1 and require 10 to 500 additional cycles to retrieve the storage data
    
    • In an in-order core, later instructions in the code stream are stalled
    
    • In an out-of-order core, later instructions which are not dependent on this storage data are allowed to execute
z196 New Instruction Set Architecture

- Re-compiled code/apps get further performance gains through 100+ new instructions
- High-Word Facility (30 new instructions)
  - Independent addressing to high word of 64-bit GPRs
  - Effectively provides compiler/software with 16 additional 32-bit registers
- Interlocked-Access Facility (12 new instructions)
  - Interlocked (atomic) load, value update and store operation in a single instruction
  - Immediate exploitation by Java
- Load/Store-on-Condition Facility (6 new instructions)
  - Load or store conditionally executed based on condition code
  - Dramatic improvement in certain codes with highly unpredictable branches
- Distinct-Operands Facility (22 new instructions)
  - Independent specification of result register (different than either source register)
  - Reduces register value copying
- Population-Count Facility (1 new instruction)
  - Hardware implementation of bit counting ~5x faster than prior software implementations
- Integer to/from Floating point converts (21 new instructions)
z196 New CP Architecture

- **New trunc and OR inexactness Binary Floating Point rounding mode**

- **New Decimal Floating Point quantum exception**
  - Eliminates need for test data group for every operation

- **Virtual Architecture Level**
  - Allows the z/VM Live Guest Relocation Facility to make a z196 behave architecturally like a z10 system
  - Facilitates moving work transparently between z196 and z10 systems for backup and capacity reasons

- **On Non-quiescing SSKE:**
  - Significant performance improvement for systems with large number CPUs (typically 30+)
  - Improves SMP scaling for OS images
  - Up to 10% performance increase when exploited by the operating system
  - Exploited by all z/OS 1.10 and above (with PTF for 1.10 and 1.11)

- **Other minor architecture**
  - RRBM, Fast-BCR-Serialization Facility, Fetch-Store-Access Exception Indicator, CMPSC Enhancement Facility
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
Performance Drivers with z196

State benefits if you re-compile
(Run-time modules or to make use of out-of-order execution)

- **Hardware**
  - Memory subsystem
    - focus on keeping data “closer” to the processor unit
    - new chip-level shared cache
    - much larger book-level shared cache
  - Processor
    - Out-of-Order execution
    - new instructions to allow for
      - reduced processor quiesce effects
      - reduced cache misses
      - reduced pipeline disruption
    - up to 80 configurable processor units
    - Total of 4 different uni speeds, 3 sub-capacity

- **HiperDispatch**
  - exploits new cache topology
  - reduced cross-book “help”
  - better locality for multi-task address spaces
System z overall RAS Strategy
.....Continuing our RAS focus helps avoid outages

Sources of Outages
Pre z9
-Hrs/Year/Syst-

<table>
<thead>
<tr>
<th>Impact of Outage</th>
<th>Prior Servers</th>
<th>z9 EC</th>
<th>Z10 EC</th>
<th>z196</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unscheduled Outages</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 requirements</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Power &amp; Thermal Management</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Temperature = Silicon Reliability Worst Enemy
Wearout = Mechanical Components Reliability Worst Enemy.
z196 Redundant Array of Independent Memory (RAIM)

- **System z10 EC memory design:**
  - Four Memory Controllers (MCUs) organized in two pairs, each MCU with four channels
  - DIMM technology is Nova x4, 16 to 48 DIMMs per book, plugged in groups of 8
  - 8 DIMMs (4 or 8 GB) per feature – 32 or 64 GB physical memory per feature
    Equals 32 or 64 GB for HSA and customer purchase per feature
  - 64 to 384 GB physical memory per book = 64 to 384 GB for use (HSA and customer)

- **z196 memory design:**
  - Three MCUs, each with five channels. The fifth channel in each z196 MCU is required to implement memory as a Redundant Array of Independent Memory (RAIM). This technology adds significant error detection and correction capabilities. Bit, lane, DRAM, DIMM, socket, and complete memory channel failures can be detected and corrected, including many types of multiple failures.
  - DIMM technology is SuperNova x81, 10 to 30 DIMMs per book, plugged in groups of 5
    5 DIMMs (4, 16 or 32 GB) per feature – 20, 80 or 160 GB physical RAIM per feature
    Equals 16, 64 or 128 GB for use per feature. **RAIM takes 20%. (There is no non-RAIM option.)**
  - 40 to 960 GB RAIM memory per book = 32 to 768 GB of memory for use
    (Minimum RAIM for the M15 is 60 GB = 48 GB = 16 GB HSA plus 32 GB customer memory)

- **For both z196 and z10**
  - The Hardware System Area (HSA) is 16 GB fixed, outside customer memory
  - In some cases, offering granularity can prevent purchase of all available memory in a book
z196 – Memory Overview

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

- **Key Cache**
- **MCU 0**
- **MCU 1**
- **MCU 2**

**Level 4 Cache**

- **DATA**
- **CHECK ECC RAIM Parity**

*Extra column provides RAIM function*
Subcapacity CPs may be ordered on ANY z196 model with 1 to 15 CPs. If 16 or more CPs are ordered all must be full 7xx capacity.

- All CPs on a z196 CPC must be the same capacity.
- All specialty engines run at full capacity. The one for one entitlement to purchase one zAAP and one zIIP for each CP purchased is the same for CPs of any capacity.
- Only 15 CPs can have granular capacity but other PU cores may be characterized as full capacity specialty engines.
- The z196 is capable of over 2 million 4k byte read I/O operations per second. This measurement was done using a z196 4 book 14 SAP configuration with 104 FICON Express8 channels connected to 11 DS8000 *Storage systems using zHPF protocols.
- Processor Unit Value (PUV) for z196 is 120.

**CP MSU Capacity Relative to Full Capacity**

<table>
<thead>
<tr>
<th>xx</th>
<th>Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>7xx</td>
<td>100%</td>
</tr>
<tr>
<td>6xx</td>
<td>~64%</td>
</tr>
<tr>
<td>5xx</td>
<td>~49%</td>
</tr>
<tr>
<td>4xx</td>
<td>~20%</td>
</tr>
</tbody>
</table>

**MSU Sub Capacity**
z196 System Upgrades

- **z196 to higher hardware z196 model**
  - Upgrade of z196 Models M15, M32, M49 and M66 to M80 is disruptive
  - When upgrading to z196 all the Books are replaced
  - Upgrade from Air to Water cooled not available

- **Any z9 EC to any z196**
- **Any z10 EC to any z196**
The z196 Books are fully interconnected in a point to point topology as shown in the diagram.

Data transfers are direct between Books via the Level 4 Cache chip in each MCM.

Level 4 Cache is shared by all PU chips on the MCM.
Connected to your world

- Improved performance and flexibility for connectivity
- Broad set of options to meet your needs

**Within the server**
- HiperSockets™
  - Multi Write Facility
  - Layer2 support
- Integrated console controller
- Integrated communications controller support

**To the Network**
- OSA-Express3
  - 10 Gigabit Ethernet
- OSA-Express2
  - 1000BASE-T Ethernet
  - Gigabit Ethernet LX and SX
  - 10 Gigabit Ethernet LR

**For Clustering**
- InfiniBand Coupling Links
- ISC-3
- Support for n-2 and above servers
Connectivity for Coupling and I/O

- **Up to 8 fanout cards per book**
  - 2 ports per fanout
  - Up to 16 ports per book
    - 48 Port System Maximum

- **Fanout cards – Two-port InfiniBand host channel adapters dedicated to function**
  - HCA2-C fanout – I/O Interconnect
    - Supports FICON, ESCON, OSA, ISC-3 and Crypto Express3 cards in I/O drawer and I/O cage domains. Always plugged in pairs.
  - HCA2-O fanout – 12x InfiniBand coupling links
    - CHPID type – CIB for Coupling
      - Fiber optic external coupling link – 150 m
  - HCA2-O LR fanout – 1x InfiniBand coupling links – Long Reach
    - CHPID type – CIB for Coupling
    - Fiber optic external coupling link – 10 km (Unrepeated), 100 km repeated

Up to 16 CHPIDs – across 2 ports
z196 I/O Cages and Drawers

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

<table>
<thead>
<tr>
<th>Current I/O cage</th>
<th>Current I/O drawer</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 I/O Slots – Supports all current I/O</td>
<td>8 I/O Slots – Supports all current I/O</td>
</tr>
</tbody>
</table>

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**
- **FRAME 1-32 slots**
  - Z: IBF, IBF
  - A: IBF, Air Backup
  - BPA, CPC
  - I/O Drawer
    - Slots 1-8
    - Slots 9-16
    - Slots 17-24
    - Slots 25-32

**ONE I/O Cage**
- **FRAME 33-44 slots**
  - Z: IBF, IBF
  - A: IBF, Air Backup
  - BPA, CPC
  - I/O Drawer
    - Slots 1-8
    - Slots 9-16
    - Slots 17-24
    - Slots 25-32
    - Slots 29-36
    - Slots 37-44

**TWO I/O Cages**
- **FRAME 45-72+ required slots**
  - Z: IBF, IBF
  - A: IBF, Air Backup
  - BPA, CPC
  - I/O Drawer
    - Slots 1-8
    - Slots 9-16
    - Slots 17-24
    - Slots 25-32
    - Slots 29-36
    - Slots 37-44
    - Slots 57-64
    - Slots 65-72

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

**Supported Channel Types**

- **I/O Channels**
  - FICON Express8
  - FICON Express4 (Carry forward on upgrade)
  - ESCON – Migrate Away (240 or fewer)

- **Networking**
  - OSA-Express3
    - 10 Gigabit Ethernet LR and SR
    - Gigabit Ethernet LX and SX
    - 1000BASE-T Ethernet
  - OSA-Express2 (Carry forward on upgrade)
    - 1000BASE-T Ethernet
    - Gigabit Ethernet LX and SX
    - HiperSockets (Define only)

- **Coupling Links**
  - InfiniBand Coupling Links
    - 12x InfiniBand
    - 1x InfiniBand
  - ISC-3 – Migrate Away (Peer mode only)
  - IC (Define only)

- **Crypto**
  - Crypto Express3
    - Configurable Coprocessor / Accelerator

**Non-Supported Channel Types**

- **I/O Channels**
  - FICON (before FICON Express4)
  - FCV – ESCD Model 5 Bridge Card

- **Networking**
  - OSA-Express2 10 GbE LR
  - OSA-Express (pre OSA-Express2)

- **Coupling Links**
  - ICB-4 and earlier ICB

- **Crypto**
  - Crypto Express2 and earlier

- **ETR**
  - Sysplex Timer® (ETR) Attachment
More than 240 ESCON channels, more than 72 I/O cards
  – Migrate and consolidate to FICON Express8, consider OEM FICON conversion technology if required
  – Eliminate parallel control units or consider OEM FICON conversion technology
FICON Express2 – Migrate and consolidate to FICON Express8
FICON Express – Ends support for FICON Conversion with ESCON Director Model 5 Bridge Card
  – Migrate and consolidate to FICON Express8, consider OEM FICON conversion technology if required
OSA-Express2 10 GbE – Migrate to OSA-Express3 10 GbE
Crypto Express2 – Migrate to Crypto Express3
Parallel Sysplex discontinuations
  – Sysplex with z990, z890 – Limit z196 sysplex coexistence to z9 and z10
  – Sysplex Timer 9037 (ETR) – Migrate to STP and ensure only z9, z10, and z196 in the CTN
  – Dynamic ICF Expansion – Change production CF to dedicated ICFs (shared ICFs test only)
  – ICB-4 – Migrate to 12x InfiniBand coupling links
  – Note: ISC-3 is supported but consider consolidation/migration to 12x InfiniBand for up to 150m or 1x InfiniBand >150m and up to 10Km unrepeated
ESCON Future Support

- **ESCON channels to be phased out (HW Announcement, April 28, 2009)**
  - System z10 EC and System z10 BC will be the last server to support greater than 240 ESCON channels
  - Currently, 1024 channels are supported on System z10 EC and 480 channels on the System z10 BC

- **FICON bridge has been phased out and 9032-005 directors are approaching End Of Service**

- **Customer Options**
  - Run unsupported
  - Replace control units with FICON interfaces and FICON channels
  - Utilize Optica PRIZM solution, FICON to ESCON converter

- **Similar approach was utilized to phase out of Parallel Channels (i.e. Bus and Tag)**
  - Optica was IBM's strategic partner utilizing many ESCON to parallel converters (which still exist in the field).

*All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.*
z196 – New OSA-Express3 CHPID types OSX and OSM

- **Two new OSA-Express3 CHPID types to support new types of z196 networks**
- **A z196 System can have up to 6 types of OSA-Express3 CHPID’s**
  - External (customer managed) networks
    - Defined as OSC, OSD, OSE, & OSN
    - Existing customer provided and managed OSA ports used for access to the current customer external networks - no changes
  - Intranode management network (INMN)
    - Defined as CHPID type OSM, OSA-Express for Unified Resource Manager
      - When the PCIe adaptor on 1000BASE-T is defined as CHPID type OSM, the second port cannot be used for anything else
    - OSA-Express3 1000BASE-T configured as CHPID type OSM for connectivity to INMN from z196 to Unified Resource Manager functions
    - OSA connection via the Bulk Power Hub (BPH) on the z196 to the Top of the Rack (TORs) switches on zBX
  - Intraensemble data networks (IEDN)
    - Defined as CHPID OSX, OSA-Express for zBX
    - OSA-Express3 10 GbE configured as CHPID type OSX for connectivity and access control to IEDN from z196 to zBX

- **Functions Supported:**
  - Dynamic I/O support
  - HCD
  - CP Query capabilities
  - Ensemble Management for these new channel paths and their related subchannels.

- **z/VM 5.4 – CHPID types OSX and OSM cannot be varied online**
Three subchannel sets per LCSS for z196

- Addition of a subchannel set of 64K devices to the existing two subchannel sets

- Value
  - Extends the amount of addressable storage for the largest System z customers
  - Provides a means to provide consistent device address definitions to help simplify addressing scheme for congruous devices
    - Can utilize same device number in different subchannel sets

- The first subchannel set (SS 0) allow definitions of any type of device as is allowed today, (i.e. bases, aliases, secondaries, and those other than disk that do not implement the concept of associated aliases or secondaries)

- Second and third subchannel sets (SS1 and SS2) now be designated for use for disk alias devices (of both primary and secondary devices) and/or Metro Mirror secondary devices only

- CHPID types: ESCON - CNC, FICON - FC (both native FICON and zHPF paths)

- Features: ESCON, FICON Express8, FICON Express4

- z/OS V1.12

- z/OS V1.10 and V1.11 with PTFs
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
InfiniBand coupling links on z196

<table>
<thead>
<tr>
<th>Type</th>
<th>Speed</th>
<th>Distance</th>
<th>Fanout</th>
<th>Cabling</th>
</tr>
</thead>
<tbody>
<tr>
<td>12x InfiniBand</td>
<td>6 or 3 GBps</td>
<td>150 meters</td>
<td>HCA2-O</td>
<td>50µ MM (OM3) fiber</td>
</tr>
<tr>
<td>1x InfiniBand</td>
<td>5 or 2.5 GBps</td>
<td>10 km</td>
<td>HCA2-O LR</td>
<td>9µ SM fiber</td>
</tr>
</tbody>
</table>

Up to 16 CHPIDs – across 2 ports

- Ports exit from the front of a book
  Does not use I/O card slots

- 12x InfiniBand – z196, z10, z9
  - DDR at 6 GBps
    - z196 and z10
  - SDR at 3 GBps
    - z196 & z10 to z9
    - First addition to z9 is disruptive
    - z9 to z9 connection not supported

- 1x InfiniBand – z196 and z10 (not z9)
  - DDR at 5 Gbps
  - SDR at 2.5 Gbps (if DWDM requires)

DDR = double data rate, SDR = single data rate
z196 Coupling Links

- Fanout, not I/O slot, used for InfiniBand
- ICB-4 – No longer supported
- ETR – No longer supported
- All coupling links support STP
- Sysplex Coexistence – z10 EC and BC
z196 Parallel Sysplex coexistence of Servers/CFs and coupling connectivity

InfiniBand

z9 to z9 NOT supported

z9 EC and z9 BC S07

InfiniBand, ISC-3

z800, z900, z890 and z990

Not supported!

z196

12x InfiniBand-SDR Up to 150 meters

ISC-3 Up to 100 KM

1x InfiniBand DDR 10/100 KM

z10 EC and z10 BC

InfiniBand, ISC-3,

12x InfiniBand DDR 150 meters

ISC-3 Up to 100 KM

Note: ICB-4s and ETR NOT supported on z196
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
The right level of business continuity protection for your business …

GDPS family of offerings

- **Continuous Availability of Data within a Data Center**
  - Single Data Center
  - Applications remain active
  - Near-continuous availability to data

- **Continuous Availability / Disaster Recovery Metropolitan Region**
  - Two Data Centers
  - Systems remain active
  - Automated D/R across site or storage failure
  - No data loss

- **Disaster Recovery at Extended Distance**
  - Two Data Centers
  - Automated Disaster Recovery
  - “seconds” of Data Loss

- **Continuous Availability Regionally and Disaster Recovery Extended Distance**
  - Three Data Centers
  - Data availability
  - No data loss
  - Extended distances

GDPS®/PPRC
HyperSwap Manager

GDPS/PPRC
HyperSwap Manager
GDPS/PPRC

GDPS/GM
GDPS/XRC

GDPS/MGM
GDPS/MzGM
z196 – Basics of CoD

Capacity on Demand

- Permanent Upgrade
- Temporary Upgrade
- Replacement Capacity
- Billable Capacity (On/Off CoD)
- Pre-paid
- Post-paid

Using pre-paid unassigned capacity up to the limit of the HWM
No expiration
Capacity
- MSU %
- # Engines

On/Off CoD with tokens
No expiration
Capacity
- MSU %
- # Engines
Tokens
- MSU days
- Engine days

On/Off CoD
180 days expiration
Capacity
- MSU %
- # Engines
Tokens
- MSU days
- Engine days

On/Off CoD with tokens
180 days expiration
Capacity
- MSU %
- # Engines
Tokens
- MSU days
- Engine days
CoD Provisioning Architecture

- Orders downloaded from System Support electronically or by IBM Service

- Enforce terms and conditions
- Enforce physical model limitations

- Up to 8 temporary capacity offerings
- Each record represents an individual offering
- Customer assigns in any combination

- Base model
- Change permanent capacity via MES order

Customer defined policy or manual operations

Capacity Provisioning Manager

HMC

API
query, activate, deactivate

Authorization Layer

R1 R2 R... R8

Dormant Capacity
CIU - CBU – On/Off CoD – CPE

Permanent Capacity
Energieaufteilung im Rechenzentrum

Steigerung der Effizienz

- Effizientere Kühlung und Energieversorgung
- Optimierte Server Hardware und Design, Energie Management
- Verbesserter Prozessor Design, Energie Management
- Verringerung Ungenutzter Kapazitäten, Virtualisierung

Quelle: Technologietrends und Smarter Planet, IBM 2010
zEnterprise – Three fundamentals of energy management

Measure/Trend Power Consumption
- Determine the power being consumed now
- Trending energy and thermals over extended periods of time

Allocate Power Correctly
- Power consumed is a function of the HW configuration, environment, application mix and system utilization.
- Allocate power based on past history using power measurements
- Rightsizing of power and cooling allocations
- Enables deployment of more servers within the physical limits of a data center

Reduce power consumed
- Reduce power in periods of low utilization to limit energy cost
- Allows reduction of power budget to either
  - Reduce energy footprint of data center
  - Dynamically increase power budget other system(s)
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 operates 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)
z196 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

Air cooling back-up mode for maximum robustness if complete loss of chilled water occurs!
z196 Water cooling infrastructure

2 x Exhaust Air Heat Exchangers

2 x Water Cooling Units

Rear View – A Frame

Note: This is not the Rear Door Heat Exchanger design!
Water Conditioning Unit (WCU) detail

- Internal Water Reservoir
- Pump
- Water to Water Heat Exchanger
- Temperature Control Valve (controls flow rate)
- Customer water
z196 Water Cooling Units

Connections to the Books
(Internal system water closed-loop)
System z – The Ultimate Virtualization Resource

- Massive, robust consolidation platform; virtualization is built in, not added on
- Up to 60 logical partitions on PR/SM; 100’s to 1000’s of virtual servers on z/VM
- Virtual networking for memory-speed communication, as well as virtual layer 2 and layer 3 networks supported by z/VM
- Most sophisticated and complete hypervisor function available
- Intelligent and autonomic management of diverse workloads and system resources based on business policies and workload performance objectives
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*

**Evolving & Emerging Workload Components**

- Java
- XML
- Analytics
- SOA
- Sensors
- Encryption
- Networking
- Digital Media
- Events
- Search
- Data Protection

**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

Integration will be critical

Both General and Special Purpose capabilities are needed
Today’s Environment

Today… Many System z shops run their core, data-sensitive business processes on System z but also maintain distributed servers to accommodate a variety of processing that interacts with mainframe resources.

- **The mainframe** has hosted most of the world’s business data and executed a majority of core business transactions.
- **The mainframe is best suited for applications that require the highest levels of:**
  - Reliability
  - Security
  - Performance
  - Availability
  - Service
  - Scalability
- **However, the mainframe is not optimized for all workloads**

- There are particular industry application where **distributed topology is the standard.**
- **Distributed environments are suited for applications where:**
  - High levels resource sharing is not required
  - The workload is
    - CPU intensive
    - Does not require high levels of data sharing
    - Highest levels of availability are not required
- **Other Distributed System Characteristics**
  - Multiple management interfaces
  - Redundancy required for availability
  - On-site assembly
  - End-user service for problem resolution
IBM zEnterprise System
The integration of Superior technologies

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

**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
## Management stack

**Building an architectural construct of hardware, software, services**

<table>
<thead>
<tr>
<th>Service Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility, Control and Automation for Applications, Transactions, Databases and Data Center Resources</td>
</tr>
<tr>
<td>End-to-End Workload Management and Service Level Objectives that Align IT Management with Business Goals</td>
</tr>
<tr>
<td>Common Usage and Accounting for business accounting</td>
</tr>
<tr>
<td>Dynamic/Centralized Management of Application Workloads based on Policies</td>
</tr>
<tr>
<td>Business Resilience for multi-site recovery</td>
</tr>
<tr>
<td>End-to-end Enterprise Security</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Platform Management</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Hardware Management</th>
</tr>
</thead>
</table>

--- **Extending with Unified Resource Manager** ---

- Hypervisor management and creation of virtual networks
- Operational controls, service and support for hardware / firmware
- Network management of private and secure data and support networks
- Energy monitoring and management
- Workload awareness and platform performance management
- Virtualization management – single view of virtualization across the platform
Focused, collaborative innovation

A “complete systems” approach
Moving to Integrated System with Accelerators and Co-Processors

OLTP
Distributed
Integrated

Values
- Reduced latency
- Simple operations
- Reduced points of failure
- Simpler and faster problem determination
- Reduced communication path lengths and costs
- Coordinated management and resource sharing
- Supports both componentisation and simplification
System z “Specialty Engine” Evolution to the zEnterprise Ensemble

IBM Smart Analytics Optimizer
WebSphere DataPower Appliance*

Optimizers

IBM Blades
POWER7
System x*

IFL
Partitioned z Engines for Linux

zAAP

zIIP

CP

z/OS LPAR

z196

zBX

Application Inventory

Better

Price-Performance

Worse

Less

More

*Statement of Direction, 1H 2011
IBM zEnterprise System

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

IBM zEnterprise Unified Resource Manager (zManager)
z196 and zBX Hardware Components

- **New System z Machine Types:**
  - CPC = 2817
  - zBX = 2458 Model 002
  - Customer supplied IBM System x* and POWER7 Blades

- **Key points:**
  1. CPC Availability and Integrity
  2. Independent Lifecycles
     - Asynchronously Upgradeable
     - Enhancements not tied to CPC HW
  3. Data availability not tied to single CPC
     - Shareable between CPCs
  4. Growth
  5. System z service

*All statements regarding IBM's plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.*
z196 zBX Hardware Components

**zBX Infrastructure**

- Rack
- Top-of-Rack Switch
- Blade Center Chassis
- Ethernet & FC Cables
- BC Switches (ESM, FC)
- Power Dist. Units
- Opt: Heat Exchanger, Power cord types

**Blades**

- IBM Smart Analytics Optimizer
- POWER7 Blades
- System x Blades*

WebSphere DataPower Appliance*

*All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.*
z196 and zBX – Unified Resource Management

System z196

- Change Management
- Configuration Management
- Operations Management
- Performance Management

Serviceability Management
- Problem Management
- Service Reporting

zBX Infrastructure

- Rack
- Top-of-Rack Switch
- Blade Center Chassis
- Ethernet & FC Cables
- BC Switches (ESM, FC)
- Power Dist. Units
- Opt: Heat Exchanger, Power cord types
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
- z10 can access the Optimizers, but can’t be part of the managed ensemble.
zEnterprise System – 196 + zBX + Unified Resource Manager

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

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

Intraensemble data network

*All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.
Unified Resource Manager and IBM Director Positioning

- **Unified Resource Manager** is a unique implementation of IBM Director, VMControl and native z management capabilities for the hardware and platform management of integrated zEnterprise resources.

- **Unified Resource Manager and System Director** are complementary with some small areas of overlap.

- **Unified Resource Manager** is like System Director and VMControl in that it provides:
  - Configuration management
  - Operational controls
  - Service and support
  - Lifecycle management
  … but for zEnterprise hardware components firmware, and virtual resources

  And with zEnterprise platform-unique functionality and differentiation (workload centric, goal oriented, single point of control via zHMC)

- **System Director/VMControl** provides:
  - External management APIs for all IBM platforms
  - Common cross-platform consistency and federation
  - Virtual server image management and deployment
  - Network and storage management

- **Use Unified Resource Manager** to manage zEnterprise ensembles

- **Use IBM Director/VMControl** to manage everything else (non-ensemble z196, other distributed IBM systems)

- **Only IBM** has optimized approaches for workload management across heterogeneous cross platform resources
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
Extending System z Qualities of Service to Make Clouds Ready for Business

**Security**
industry leading security at the core of an integrated infrastructure
*Identifies potential fraud in Real Time*

**Virtualization**
Centralize Management of virtual servers across a heterogeneous pool
*Enable thousands of virtual servers within a single integrated system*

**Availability**
Resiliency management and fewer points of failure
*Centralized workload management aligned to business priorities*

**Efficiency**
Economies of scale for Labor, software and environmental costs
*Reduce labor, energy, and development costs*

**Scalability**
Ability to meet massive demands from users and data
*Unmatched scalability with the highest transaction processing capacity*
IBM zEnterprise System: A revolutionary change has come to IT

- Redefining IT frameworks to bring change to operational silos and extend System z governance to POWER7 and System x* blades
- Driving business decisions based on insight rather than hindsight
- Improving agility to compete with consolidation and simplification
- Delivering consistent business controls across applications and platforms
- Focused on integration and collaboration to fuel business growth

*All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.
Service levels to match your business needs
Increased flexibility for your multi-architecture strategy when data is on z/OS

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

**TCA Focus**
- Distributed Systems

**SCALABILITY, SECURITY, DYNAMIC WORKLOAD MANAGEMENT**
- Lower
- Higher

**zEnterprise System**
- 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

**Select IBM Blades in zBX**
- Expanded ISV support for enterprise applications
- Targeted for applications that interact with mainframe data and transactions
- Provisioned and managed by System z
- 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®**

© 2010 IBM Corporation
### We know, you hear it all the time: Mainframe Misperceptions

<table>
<thead>
<tr>
<th>MYTH</th>
<th>REALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Antiquated technology&quot; Only for legacy systems, not modern workloads</td>
<td>• Can provision new virtual servers (images) in minutes rather than days/weeks, add capacity on the fly</td>
</tr>
<tr>
<td></td>
<td>• Runs modern technologies such as Java®, C/C++, XML, Linux, Web 2.0, SOA</td>
</tr>
<tr>
<td></td>
<td>• Unrivaled availability for mission critical workloads</td>
</tr>
<tr>
<td></td>
<td>• And now zEnterprise.. So who is the dinosaur?</td>
</tr>
<tr>
<td>&quot;Too Expensive&quot;</td>
<td>• System z Solution Editions package HW, SW maintenance services at very attractive prices, close or equivalent to distributed, including for SAP</td>
</tr>
<tr>
<td></td>
<td>• The most efficient server for large scale consolidation. System z can reduce per-core licensing costs up to 28 to 1 compared to x86 environments</td>
</tr>
<tr>
<td></td>
<td>• With unrivaled virtualization and scale, System z can be significantly less cost than distributed systems for running multiple, diverse workloads</td>
</tr>
<tr>
<td></td>
<td>• Low cost server specialty engines (zAAP, zIIP, IFL) enable many applications (e.g. Java®, XML, Linux) and large scale consolidation on the platform</td>
</tr>
<tr>
<td>&quot;No Apps Support&quot;</td>
<td>• Over 6,500+ applications available for System z</td>
</tr>
<tr>
<td></td>
<td>• Over 3,000+ Linux applications are supported on System z; 18% growth in 2008</td>
</tr>
<tr>
<td></td>
<td>• Over 1,700+ ISVs building applications for System z</td>
</tr>
<tr>
<td>&quot;Skills Shortage&quot;</td>
<td>• Over 60,000 students educated worldwide, adding 5,000 more in China by 2010</td>
</tr>
<tr>
<td></td>
<td>• Over 600 schools enrolled globally</td>
</tr>
<tr>
<td></td>
<td>• 50,000 Students completed courses to date</td>
</tr>
</tbody>
</table>
Thank You