IBM System z10

Technologie

Michael Stoerchle, System z IT Specialist
Stoerchle@de.ibm.com
The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

<table>
<thead>
<tr>
<th>Trademark</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CICS*</td>
<td>HiperSockets</td>
</tr>
<tr>
<td>DB2*</td>
<td>HyperSwap</td>
</tr>
<tr>
<td>DFSMS</td>
<td>IBM*</td>
</tr>
<tr>
<td>DFSORT*</td>
<td>IBM Certified Used Equipment</td>
</tr>
<tr>
<td>DRDA*</td>
<td>IBM eServer</td>
</tr>
<tr>
<td>DS6000</td>
<td>IBM logo*</td>
</tr>
<tr>
<td>DS8000</td>
<td>NetView*</td>
</tr>
<tr>
<td>ESCON*</td>
<td>Parallel Sysplex*</td>
</tr>
<tr>
<td>FICON*</td>
<td>PR/SM</td>
</tr>
<tr>
<td>FlashCopy*</td>
<td>pureXML</td>
</tr>
<tr>
<td>GDPS*</td>
<td>Resource Link</td>
</tr>
<tr>
<td>RETAIN*</td>
<td>Scalable Architecture for Financial Reporting</td>
</tr>
<tr>
<td>System Storage</td>
<td>z9*</td>
</tr>
<tr>
<td>System z</td>
<td>z10 BC</td>
</tr>
<tr>
<td>System z9*</td>
<td>z10 EC</td>
</tr>
<tr>
<td>System z10</td>
<td>z/Architecture*</td>
</tr>
<tr>
<td>Systems Director Active Energy Manager</td>
<td>z/VM*</td>
</tr>
<tr>
<td>Tivoli*</td>
<td>z/VSE</td>
</tr>
<tr>
<td>TotalStorage*</td>
<td>zSeries*</td>
</tr>
<tr>
<td>VSE/ESA</td>
<td></td>
</tr>
<tr>
<td>WebSphere*</td>
<td></td>
</tr>
</tbody>
</table>

* Registered trademarks of IBM Corporation

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.

InfiniBand is a trademark and service mark of the InfiniBand Trade Association.

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.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. 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.
Agenda

- System z Architektur
- Virtualisierung mit IBM System z
- Was ist mit Unix auf IBM System z?
Agenda

- System z Architektur
- Virtualisierung mit IBM System z
- Was ist mit Unix auf IBM System z?
System z10 EC

- **Machine Type**
  - 2097
- **5 Models**
  - E12, E26, E40, E56 and E64
- **Processor Units (PUs)**
  - 17 (17 and 20 for Model E64) PU cores per book
  - Up to 11 SAPs per system, standard
  - 2 spares designated per system
  - Dependant on the H/W model - up to 12, 26, 40, 56 or 64 PU cores available for characterization
    - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)
- **Memory**
  - System Minimum of 16 GB
  - Up to 384 GB per book
  - Up to 1.5 TB GB for System and up to 1 TB per LPAR
    - Fixed HSA, standard
    - 16/32/48/64 GB increments
- **I/O**
  - Up to 48 I/O Interconnects per System @ 6 GBps each
  - Up to 4 Logical Channel Subsystems (LCSSs)
- **ETR Feature, standard**
z10 EC – Under the covers (Model E56 or E64)

- Internal Batteries (optional)
- Power Supplies
- 2 x Support Elements
- 3x I/O cages
- HMC
- 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
- 2 x Cooling Units
- IBM Systems
z10 EC Book Layout
z10 EC Multi-Chip Module (MCM)

- **96mm x 96mm MCM**
  - 103 Glass Ceramic layers
  - 17 and 20 way MCMs

- **CMOS Technology**
  - 65 nm
  - 5 PU chips/MCM – Each up to 4 cores
    - 21.97 mm x 21.17 mm
    - 994 million transistors/PU chip
    - L1 cache/PU core
      - 64 KB I-cache
      - 128 KB D-cache
    - L1.5 cache/PU core
      - 3 MB
      - 4.4 GHz
  - 2 Storage Control (SC) chip
    - 21.11 mm x 21.71 mm
    - 1.6 billion transistors/chip
    - L2 Cache 24 MB per SC chip (48 MB/Book)
Siblings, not identical twins

Share lots of DNA
- IBM 65nm Silicon-On-Insulator (SOI) technology
- Design building blocks:
  - Latches, SRAMs, regfiles, dataflow elements
- Large portions of Fixed Point Unit (FXU), Binary Floating-point Unit (BFU), Hardware Decimal Floating-point Unit (HDFU), Memory Controller (MC), I/O Bus Controller (GX)
- Core pipeline design style
  - High-frequency, low-latency, mostly-in-order
- Many System z and System p® designers and engineers working together

Different personalities
- Very different Instruction Set Architectures (ISAs)
  - very different cores
- Cache hierarchy and coherency model
- SMP topology and protocol
- Chip organization
- IBM z10 EC Chip optimized for Enterprise Data Serving Hub
System z10 BC

- **Machine Type**
  - 2098

- **Single Model – E10**
  - Single frame, air cooled
  - Non-raised floor option available

- **Processor Units (PUs)**
  - 12 PU cores per System
  - 2 SAPs, standard
  - Zero spares when all PUs characterized
  - Up to 10 PUs available for characterization
    - Central Processors (CPs), Integrated Facility for Linux (IFLs), Internal Coupling Facility (ICFs), System z10 Application Assist Processors (zAAPs), System z10 Integrated Information Processor (zIIP), optional - additional System Assist Processors (SAPs)

- **Memory**
  - System Minimum of 4 GB
  - Up to 128 GB for System, including HAS
  - 8 GB Fixed HSA, standard, up to 120 GB for customer use
    - 4, 8 and 32 GB increments

- **I/O**
  - Up to 12 I/O Interconnects per System @ 6 GBps each
  - 2 Logical Channel Subsystems (LCSSs)
  - Fiber Quick Connect for ESCON and FICON LX
  - New OSA-Express3 Features
  - ETR feature, standard
z10 BC – Under the covers Front View

Power Supplies

Internal Battery (optional)

CPC (SCMs, Memory, MBA, HCA and FSP) Drawer

2 x Support Elements

I/O Drawer #3

I/O Drawer #2

I/O Drawer #1

I/O Drawer #4

HMC
System z10 BC CPC Drawer
**z10 BC CPC Drawer Components**

- **Up to 32 DIMMS**

- **CPC Drawer**
  - **DCA Power**
  - **I/O Hub for fanout slots**
  - **2 x OSC/ETR Cards**

- **PU chip, SC chips, Land Grid Array (LGA) socket, Indium foil**

- **4 x PU and 2 x SC pluggable SCMs**
z10 BC – Enterprise Quad Core Processor Chip on a Single Chip Module (SCM)
z10 BC SCM Vs z10 EC MCM Comparison

**z10 BC SCMs**
- **PU SCM**
  - 50mm x 50mm in size – fully assembled
  - Quad core chip with 3 active cores at 3,5 GHz
  - 4 PU SCMs per System with total of 12 cores
  - PU Chip size 21.97 mm x 21.17 mm
- **SC SCM**
  - 61mm x 61mm in size – fully assembled
  - 2 SC SCMs per System
  - 24 MB L2 cache per chip
  - SC Chip size 21.11 mm x 21.71 mm

**z10 EC MCM**
- **MCM**
  - 96mm x 96mm in size
  - 5 PU chips per MCM
    - Quad core chips with 3 or 4 active cores at 4 GHz
    - PU Chip size 21.97 mm x 21.17 mm
  - 2 SC chips per MCM
    - 24 MB L2 cache per chip
    - SC Chip size 21.11 mm x 21.71 mm
  - Up to 4 MCMs for System
Evolution der System z Specialty Engines

“Kaufmännische Lösung” für New Workload Wachstum auf System z

- Internal Coupling Facility (ICF) 1997
- Integrated Facility for Linux (IFL) 2000
- System z Application Assist Processor (zAAP) 2004
- System z Integrated Information Processor (zIIP) 2006

Anwendungsbeispiele für zIIP:
- DB2 remote access und BI/DW
- IPSec encryption
- z/OS XML System Services
- z/OS Global Mirror

Anwendungsbeispiele für zAAP:
- Java execution environment
- z/OS XML System Services
**z/-Architektur-Elemente**

**Hauptspeicher**
- Byte-weise adressierbar
- 64bit Adressierbarkeit
- 'shared' von allen CPU'S

**I/O Anschlüsse**
- 'alt': parallel (Kupferkabel), 4.5MB/sec
- 1990': seriell (Glasfaser), 17MB/sec
- 1999: FiCON (Glasfaser), 100+MB/sec

**ESCON & FICON 'Directoren'**
- 'Schalter', 'switches'

**Control Units (CU)**
- Steuereinheiten

**Endgeräte, 'Devices'**
- Festplatten(HD), Bandeinheiten (Tapes), Drucker (Printer), ...

**Netzwerk (GbE, ...)**
Von 'System/360' (S/360) zu ESA/390 und z/-Architektur

- 1964 S/360
  - CISC, 24bit Adressierung, 'Real Storage', Uniprozessoren
  - Amdahl, G.M., Blauw, G.A., and Brooks, F.P.: *Architecture of the IBM System/360*
  - IBM Journal of Research and Development 8,2 (April 1964)

- 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

- 2000 z/-Architektur (64-bit)
  - Hardware 2000-2006: zSeries z900/z800 & z990/z890, z9, z10
Vom Uniprozessor zum Parallel Sysplex

- Einheitlicher Design von IBM Hardware und Software
- Basis für zukünftige Erweiterungen

Parallel Sysplex
- Continuous Availability
- Flexible Growth
  - Granularity
- Scalable
  - Nearly Unlimited
- Data Sharing
- Dynamic Workload Balancing
- Single System Image
Agenda

- System z Architektur
- Virtualisierung mit IBM System z
- Was ist mit Unix auf IBM System z?
IBM Virtualization Evolution
40 Jahre kontinuierliche Innovation

Der IBM Mainframe ist Pionier und Perfektionist auf dem Gebiet der Virtualisierung
System i & p haben fortschrittliche Hypervisor mit PR/SM-ähnlicher Funktionalität
Andere Server hinken mit ihren Virtualisierungsmechanismen nach
System z Virtualization Technology

**Start Interpretive Execution**
- Establish architecture for guest systems
- Maintain status
- Invoke SIE assists

**LPAR Zoning**: each partition has a zero-origin address space, allowing I/O access to memory without hypervisor intervention

**PR/SM** – SIE – EAL 5

**LPAR** – Up to 60 Logical Partitions

**z/VM** – SIE – EAL 3+ – 100s of Virtual Machines – Shared Memory

**Most sophisticated and functionally complete hypervisors**
Able to host z/OS, Linux, z/VSE, z/TPF, and z/VM-on-z/VM
Shared everything architecture
Highly granular resource sharing (less than 1% utilization)
Any virtual CPU can access any virtual I/O path within the attached logical channel subsystem
z/VM can simulate devices not physically present
Application integration with HiperSockets and VLANs
Intelligent and autonomic workload management

**Shared resources per mainframe footprint**
Up to 64 OS-configurable CPUs
Up to 10 SAP processors
Up to 1.5 TB of memory
Up to 1024 channel paths
Up to 16 internal HiperSockets networks

**HW (LPAR) and SW (z/VM) hypervisors**
Hardware support, SIE, microcode assist
Virtualization is transparent for Op Sys execution
Hardware-enforced isolation

The potential performance impact of the Linux server farm is isolated from the other LPARs
IBM System z Virtualization Architecture

- **Multi-dimensional virtualization technology**
  - System z provides logical (LPAR) and software (z/VM) partitioning
  - PR/SM enables highly scalable virtual server hosting for LPAR and z/VM virtual machine environments
  - IRD coordinates allocation of CPU and I/O resources among z/OS and non-z/OS LPARs*

* Excluding non-shared resources like Integrated Facility for Linux processors
Basic Direct HW virtualization

System z with SIE  
(Start Interpretive Instruction Execution)

* System z runs ALWAYS in PR/SM-LPAR mode under SIE
* LPAR is the “only” game in town, meaning performance items and other functionalities is developed accordingly

zVM invokes SIE to run VM’s (SIE under SIE)

* Efficient for performance - and new version of OS and Hypervisor

Positioning of System z & Intel/AMD

- **System z:**
  the requirement for the underlying HW support is NOT an issue for System z, - since the basic System z Architecture & HW design has implemented this for “decades”.

- **Intel and AMD**
  is developing some virtualization HW support based on a similar structure like the SIE architecture as it was externally documented 25 years ago.

  The amount of HW and SIE assist functionalities are seemingly rather limited at this point.

  No “SIE” under “SIE”
Funktionalität von z/VM

- **Guest Support**
  - Applications for Linux, z/VSE™, z/VM, z/OS®, and System z Transaction Processing (z/TPF)
  - Parallel Sysplex® Support
  - FICON®, FICON Express and Virtual FICON CTCA
  - Enhanced memory utilization using VMRM between z/VM and Linux guests
  - Enhanced memory management for Linux guests (CMMA)
  - Guest LAN sniffer
  - IPL from SCSI disks
  - Dynamic memory upgrade support
  - Enhanced specialty engine support

- **Application Development**
  - XL C/C++ for z/VM Compiler
  - IBM Debug Tool for z/VM
  - REXX™
  - CMS Pipelines
  - POSIX
  - REXX Sockets
  - Reusable Server Kernel
  - Binder/Loader
  - Converged C Sockets Libraries

- **Security**
  - LDAP server and client services
  - RACF® data and file protection
    - Longer passwords
    - Change logging
  - QDIO data connection isolation

- **Systems Management**
  - Automated Operations
  - OSA/SF
  - Systems management APIs
  - Systems management using the HMC
  - Coordination of DirMaint™ and RACF changes
  - Enhanced directory management
  - **Performance Toolkit for VM™**
  - VM Resource Manager (VMRM)

- **Communications**
  - TCP/IP
  - RSCS FL540
  - OSA-Express/Express2/Express3
  - OSA-Express2/3 OSA for NCP (OSN)
  - **VSWITCH** support for Link Aggregation
  - HiperSockets™/TELNET IPv6
  - z/VM guest LAN

- **Data**
  - Distributed Relational Databases
  - Shared File System
  - Storage Management Subsystem
  - Byte File System
  - Network File System

---

V5.4

z/VM

IBM Systems
Agenda

- System z Architektur
- Virtualisierung mit System z
- Was ist mit Unix auf IBM System z?
Was ist Linux?

- Ein “UNIX-ähnliches” Betriebssystem
  - Source-Code ist offengelegt
  - Wird von einer Community entwickelt
    - ‘Master Repository’ gepflegt von Linus Torvalds
    - ‘Experimental Repository’ gepflegt von Andrew Morton
    - ‘System z Subsystem Repository’ gepflegt von Martin Schwidetsky, IBM Lab BB
    - ‘Steering Committee’ betreut die Projekte

- Erhältlich für viele Architekturen
  - x86, POWER, System z…
  - IBM Chipshopper

- In der Regel wird eine Linux Distribution auf Basis einer “Support Subscription Fee” von Linux Distribution Partners (LDP) gekauft
  - Novell und Red Hat dominant
Was ist Linux on System z?

*Linux ist Linux ...*

- Kein spezielles Linux
  - Keine Änderung am Look&Feel
  - Etwa 1% Source-Code sind customized
- Reine ASCII umgebung
  - Keine EBCDIC Codepage
- Linux ist Linux ist Linux...
  - ...aber Leistung, Eigenschaften und Qualitäten hängen von der Hardware ab
- Linux-Only Mainframe ist möglich
- Unterstützt die speziellen Palttform-Features des Mainframe, inklusive FCP!
- Keine Ablösung für bestehende Betriebs-systeme auf System z

... *und Linux on System z bietet einzigartigen Mehrwert!*
Linux on System z - System Struktur

- GNU C Compiler
- GNU Binutils
- Linux Applications
  - GNU Runtime Environment
  - Network Protocols
  - File systems
  - Generic Drivers
  - Memory Management
  - Process Management
  - arch
  - System z Instructionset and I/O Hardware
  - System z dependent Code
  - Backend

Backend

System z dependent Code
z/OS USS – Unix System Services

- Eigenständige Komponente unter z/OS
- EBCDIC-Codepage
- POSIX/XPG konform
- Sehr enge Integration in und mit anderen z/OS Komponenten (RACF, WLM...)
- Kann andere z/OS Funktionen und Module aufrufen sowie von diesen auch aufgerufen werden
z/OS – USS Komponenten

Shell and Utilities
- Applications
- C / C++ for z/OS
  - Compiler
- DBX
  - Debugger

DFSMS
- HFS
- NFS

Language Environment
for z/OS

z/OS UNIX Kernel
- Process Management
- File Systems
- Communications
- Daemons

z/OS Base

Assembler and
REXX Interface
Thank You