Putting Linux for System z into Production: True Stories

Erich Amrehn
Session 9249
Amrehn@de.ibm.com

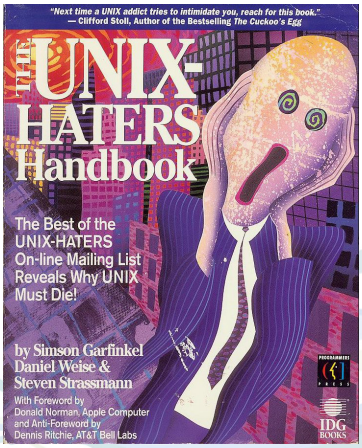
What it takes is One brave decision every day

Agenda

- HVB
- NRI
- ZIVIT
- CIO
- Japanese Story
- Swisscom
- Endress und Hausser
- Sparda
- Deutsche Bahn

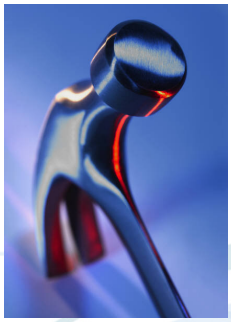
What are Performance issues ??

- The problem is not that the System z proc. is to slow !!!
  - It is the expectation and the sizing
- Avoid synthetic benchmarks
  - Best results are with real application test/benchmark
- So it Depends



Mainframe consideration System z

- One size does not fit all
- The mainframe has never been for everyone, and that has not changed !!



However, the mainframe is the best solution for a number of environments when all factors are considered

Define the race – then pick the vehicle

- Typical industry benchmarks for Linux and UNIX tend to measure the performance of a single server running a single application
- Results tend to be highly dependent upon processor speed
- Stand alone processor may run a higher speed than mainframes, hence they look better in typical industry benchmarks
- Mainframes distinguish themselves through outstanding capacity, usually not measured by typical industry benchmarks
- The work performed by multiple stand alone servers is a good candidate for consolidation when:
  - The servers are lightly to moderate loaded
  - The servers do not peak at the same time

The Race ???

- TCO or TCA what more expensive ??
  - The Learjet is a great deal less expensive !!
- Therefore all airlines should start buying Learjet's instead 747's
- Of course it is not that simple. It depends

The Race ???

- A race car goes faster than a trailer truck, BUT if the contest is which one can move 100 refrigerators across the country fastest, bet on the truck

The Economics of Workload Consolidation

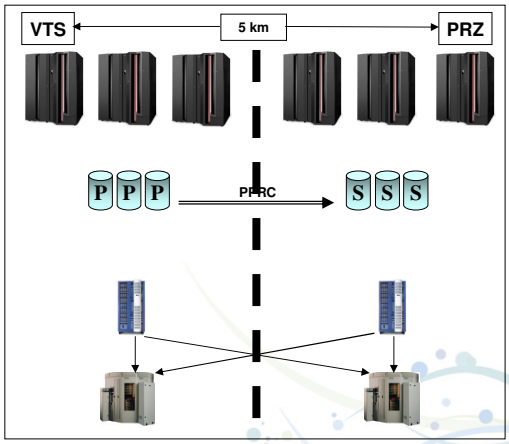
- Distributed servers often run at average utilization levels in the range of 5% to 20%[1]
  - Production servers, development servers, test servers
- Virtualization and workload management enable consolidation on the mainframe
  - Run multiple images on fewer processors
  - Achieve utilization levels of 85% or more
- Become Leaner, Greener, Cheaper, Simpler, and more through IT consolidation and simplification

1. Average Utilization includes Production, Development, and Test servers

HVBInfo

Decision for a new Tivoli-Platform

HVBInfo Mainframe Architekture „Multisite“



Vier Sysplexe:
A-Plex: Produktion
Q-Plex: Quality Assurance
E-Plex: Anwendungs-Entwicklung
T-Plex: Test für System Programmierung

DASD farm (based on IBM)
A-Plex: 72 TB (mirrored)
Q-Plex: 34 TB (single)
E-Plex: 13 TB (mirrored)
T-Plex: 3 TB (mirrored)

TAPE farm (based on STK)
A-Plex: 4 VSMs, 78 TAPE drives, 478 TB, 4 silos (mirrored)
Q-Plex: 3 VSMs, 24 TAPE drives, 8 TB, 1 silo (single)
E-Plex: 2 VSMs, 16 TAPE drives, 60 TB, 2 silos (mirrored)
T-Plex: none

PoC for the decision

Mainframe and Intel Platform in comparison PoC; Many Tests done in the BB IBM Labor with all configuration

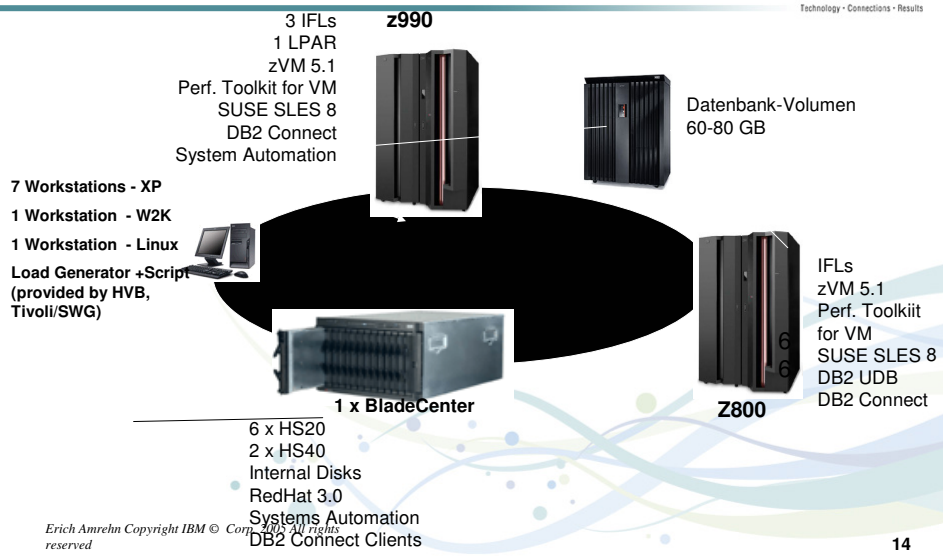
Result: both platforms could do the job

Intel-Platform with 55 Systems:
- Higher throughput
- use existing skills
- less hardware cost (TCA)

Mainframe with two Systems:
- 7x24 available
- concurrent Service
- low operating cost
- New Concept: „Server in Minutes“

Decision : and the winner is ?? .

BB Lab PoC Configuration



After the PoC comparison of the two platforms

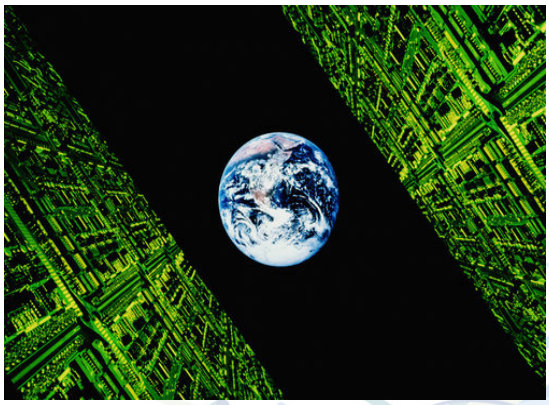
Criteria	Weight				
High availability	9,17%	2	0,18	4	0,37
Furture needs	4,59%	3	0,14	2	0,09
Smal risk for Migration	14,68%	4	0,59	2	0,29
DR	9,17%	2	0,18	4	0,37
Concurrent Service	18,35%	1	0,18	4	0,73
Administration	14,68%	2	0,29	3	0,44
Scalability of the Architekture	5,50%	3	0,17	2	0,11
Throughput “Events”	9,17%	5	0,46	1	0,09
Scalability of the licence management	5,50%	3	0,17	3	0,17
Virtualization	4,59%	1	0,05	4	0,18
Test, development and quality env.	4,59%	2	0,09	3	0,14
Result:			**2,5**		**2,98**

1=(low, minimal) to 5=(high, max)

HVBInfo

SHARE

Technology · Connections · Results



Tivoli-Monitoring goes new Ways

55 virtual LINUX-Server with unique Options

NRI Japan

LinuxWorld Tokyo key note speech by Toru Kanazawa, Managing Director, Group IT Strategy Department Nomura Holdings,inc.
http://www.computerworld.jp/topics/srv/41121.html

Nomura Research Institute adopted solutions rebuilding the backend mission critical database servers by Linux(Novell SUSE Linux), IBM mainframe (IBM System z9), and Oracle Real Application Clusters. This mission critical system has already been running, and Mr. Kanazawa says "the system shows the performance to process 1,000 transactions per second now." He also says that "the performance will reach over 2,000 transactions per second by judging from the current CPU usage."

17

ZIVIT Zentrum fuer Informationsverarbeitung und Informationstechnik

- What is ZIVIT and where ?
- ZIVIT mainframe landscape
- z/VM setup and architecture
- HA setup of Linux guest
- Tools used and written by ZIVIT

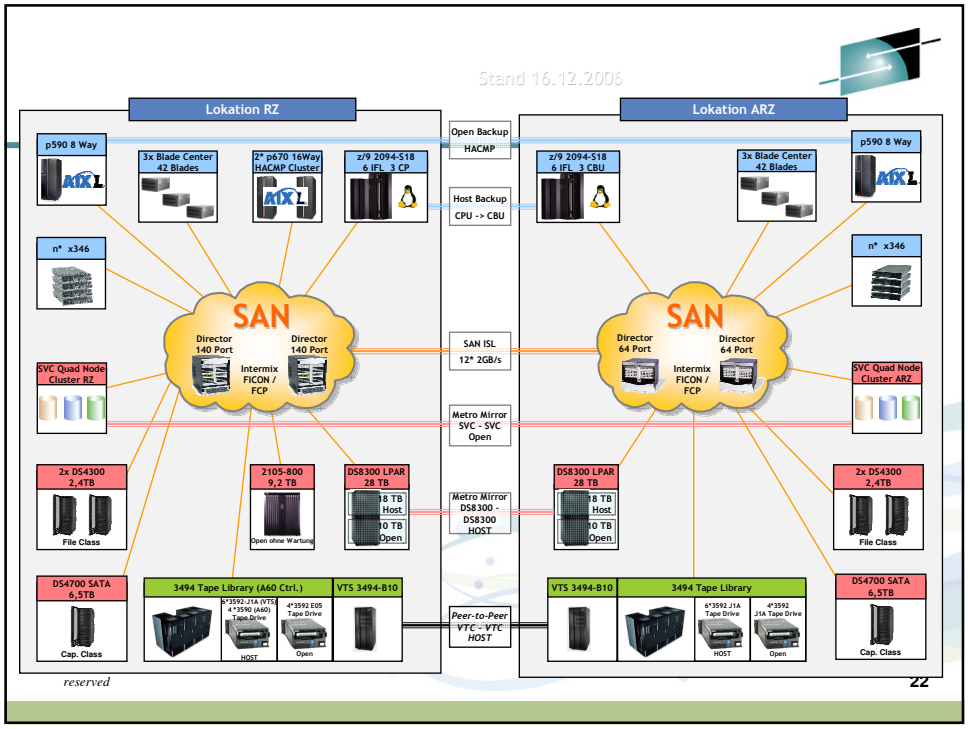
- Thanks to Armin Arbinger (ZIVIT) and Martin Grimm (Millenux)

Build Jan.2006 from IT of Bundeszollverwaltung (Central Customs Government) and Bundesamt fuer Finanzen (Central Finance Government)

The ZIVIT works for the Government as well as the citizen.
They have about 1000 employees (400 software developer, 370 IT service on 7 different location)
Applications:
•Personal payment systems for the German Government (Bundesverwaltung)

Hosting of application for 700 inland revenue office (Landesfinanzverwaltungen) about 120.000 User
• Provider of Internet- and Intranet service for Federal Fiscal (Bundesfinanzverwaltung) Information portal
www.bzst.de, www.bundesliegenschaften.de, www.finanzamt.de, www.bundesimmobilien.de, www.dz-portal.de, www.zivit.de
"E-Payment„: Online-payment

• IT-support for other government institution.
Auswärtiges Amt (Gehaltsabrechnung)
Bundesanstalt für Immobilienagelegenheiten (Liegenschafts-kataster, Holzbewirtschaftung)



System z9 config

- System 1
- 3CPs / 4 LPARs
  - Production z/OS
  - Production z/OS DB2
  - 2 test LPARs
- 6 IFLs / 2 LPARs
  - Production z/VM 5.3
  - Test z/VM 5.3
  - Plan 8IFLs - 2 prod-LPARs
- 96 GB Memory

- System 2
- 3 CBU / 4 predefined LPARs
  - Backup prod. z/OS
  - Backup prod z/OS DB2
  - 2 backup test LPARs
- 6 IFLs / 2 LPARs
  - Production z/VM 5.3
  - Test z/VM 5.3
  - Plan 8IFLs - 2 prod-LPARs
- 96GB Memory

z/OS Overview

- z/OS Version 1.7
  - Adabas and DB2
  - Cobol and Natural application
  - CICS
- Critical applications
  - HKR for Bundeshaushalt
  - KIDICAP salary for Bundesbeschaeftigte
  - Central Tax application and DBs
- About 200.000 User

z/VM and Linux for System z

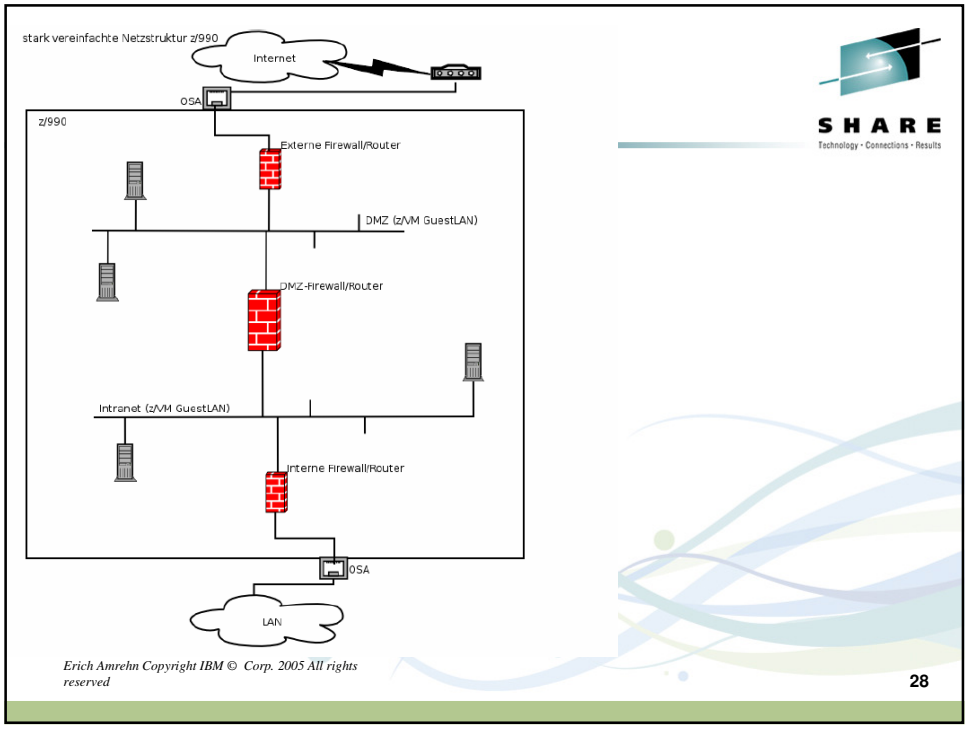
- z/VM 5.3
  - RACF (VSWITCH, DASD and VM-logon)
  - PerfKit
  - Split User Direct
    - Individual SYSTEM DIRECT for each VM instance
    - Merged LINUX DIRECT with all Linux guests
- Linux Kernel 2.6 (64 bit, z optimized)
  - 31 bit comp. Mode
- Debian GNU Linux 3.1
  - Own repository plus (IBM-Java, udev, Tomcat5, PHP5, Kernel)
- About 160 production guests

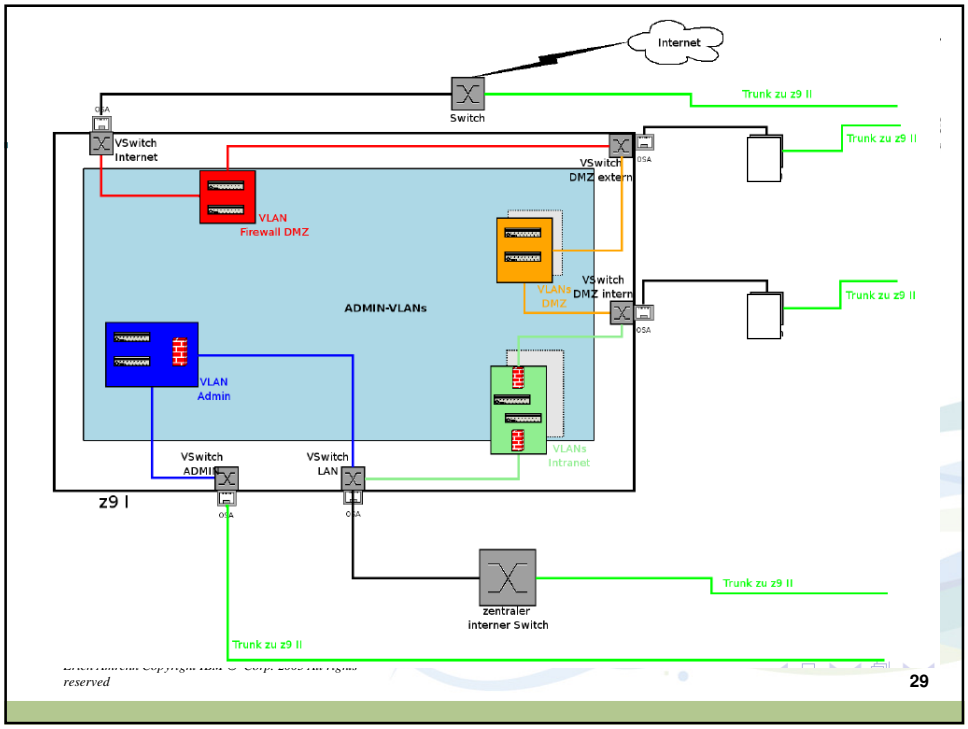
z/VM and Linux applications

- Firewalls (IPtables, HA with keepalived)
- VPN (OpenVPN)
- Web appl.Server (Apache, Tomcat, Jboss, Zope)
  - ePayment, foreign tax id-numbers
- Proxy server (squid with load balancer)
- Mail with Spam and Virus scanner (Exim, Postfix, ClamAV, AMaViS, SpamAssassin, greylistd)

z/VM and Linux applications

- MySQL DB server
- Communication/Groupware
  - Groupware server (KOLAB)
  - Instant Messaging (Jabber)
  - Mailing listen (mailman)
  - Foren Server WIKIs (phpBB, mediawiki)
  - Data Transfer (Virtuelle Poststelle VPS, MACH5)
  - Ticket system OTRS
- Name server (Bind)
- Timeserver (Open NTP)
- Software version management SVN (Subversion)


stark vereinfachte Netzstruktur z/990
Internet
OSA
z/990
Externe Firewall/Router
DMZ (z/VM GuestLAN)
DMZ-Firewall/Router
Intranet (z/VM GuestLAN)
Interne Firewall/Router
OSA
LAN
SHARE
Technology · Connections · Results
Erich Amrehn Copyright IBM © Corp. 2005 All rights reserved
28


Internet
Switch
Trunk zu z9 II
VSwitch Internet
VSwitch DMZ extern
Trunk zu z9 II
VLAN Firewall DMZ
VSwitch DMZ intern
VLANs DMZ
Trunk zu z9 II
ADMIN-VLANs
VLAN Admin
VSwitch ADMIN
VSwitch LAN
VLANs Intranet
z9 I
Trunk zu z9 II
zentraler interner Switch
Trunk zu z9 II
reserved
29

Network setup description

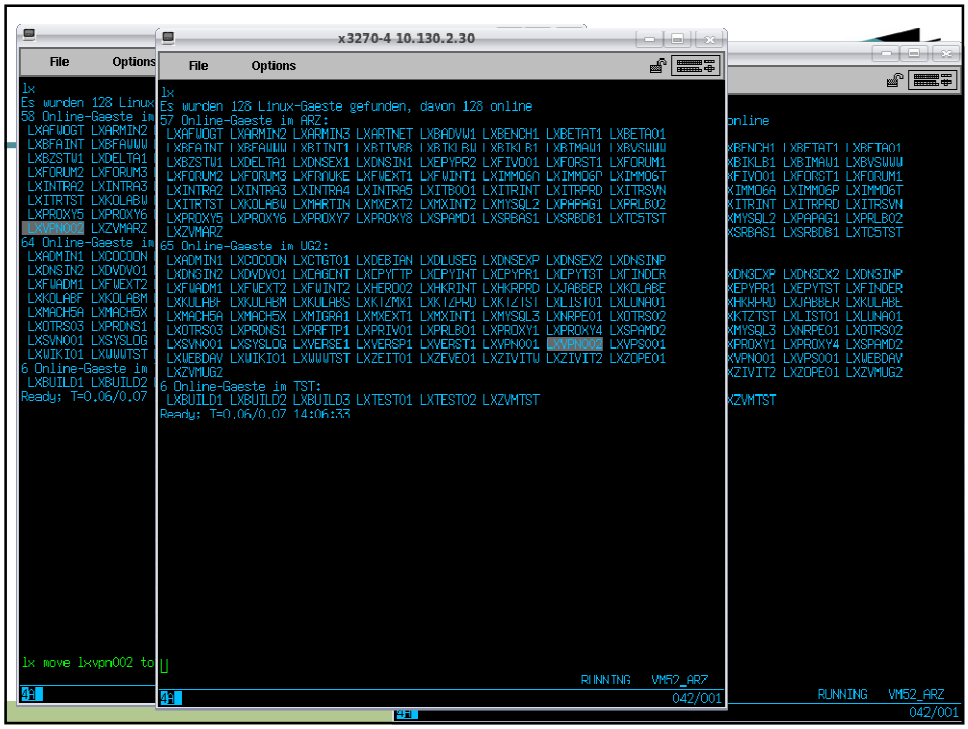
- Every VSWITCH has 2 OSAs as Trunk between the 2 z9 systems
- VLANs of the VSWITCH are available on both z9
- Each Guest/Group has there own VLAN
  - Access defined through RACF
  - Firewalls: VLAN = Interface
- Separate ADMIN VSWITCH
  - Administration of Guest (Build, installation, Update, Loghost, Monitoring)
  - All guest have there own VLAN connection with Firewall
  - SSH access only allowed through the Admin Net

Application requirements and solution's

- Different req. from the applications
  - Continuous availability, HA (production)
  - Fast recovery and reset (test and development)
  - Load balancing between the guest's and CECs
- Setup with one or multiple guest's
- Availability on protocol level
  - DNS (Master-Slave)
  - Mail (MX entries)
- Failover with multiple guest's
  - Heartbeat, keepalived
- Load balancer to distribute load between the systems
  - Apache as load balancer in front of Tomcat
  - Linux Virtual Server (LVS)
    - Load distribution and failover

z/VM - Tools

- REXX-script‘s and System management
- GSYNC
  - program/exec run‘s on all z/VM systems to control Linux guest‘s
- LX
  - REXX exec uses GSYNC to manage Linux guest‘s
- LXSETUP
  - To define new guest‘s (LINUX DIRECT, RACF etc)
- LINSTALL
  - Automated Linux installation (FAI)



Linux Tools

- Debian FAI
  - Configuration of network, dasd, software packages, user using easy configuration files
  - Automated installation and setup
- Monreader – read monitor data from z/VM using a special Linux guest
- NAGIOS to monitor all guest‘s and environment
  - system management

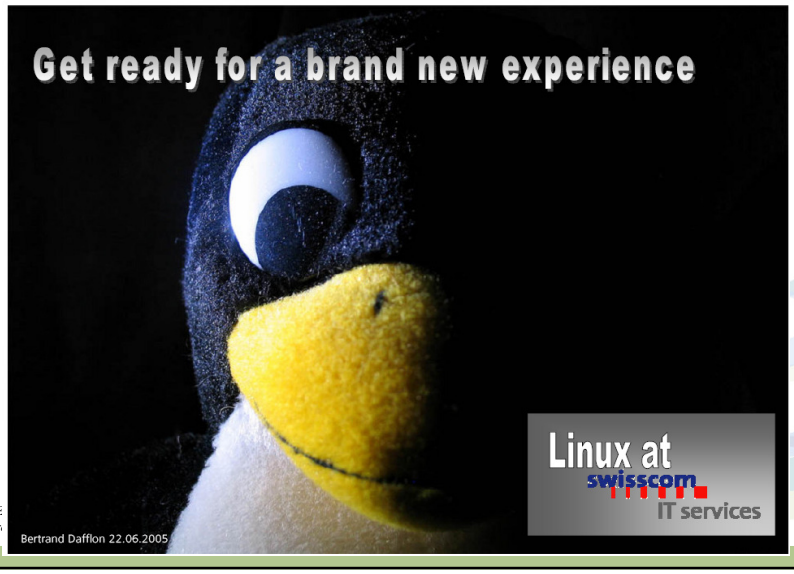
TCO- view--Hardware to expensive ??

- z/Linux uses the existing Hardware for z/OS as a base
- Pure Linux-part on both z/990-Systemen (2006)
  - Total of about ca. 1.000.000,- EURO
- Minus the saving for a additional/new Backup solution

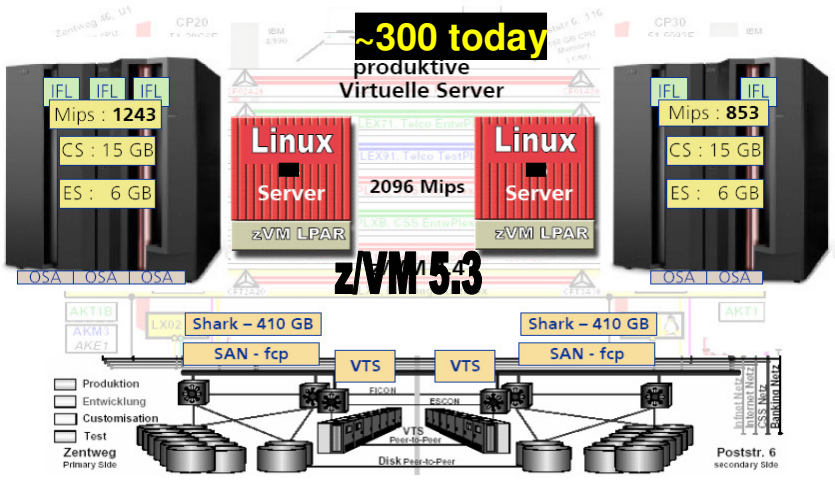
The Real cost TCO vs. TCA

- Comparison z990 with a Blade solution
- Cost for 4 Blade-Center
  - Totals to about 1.200.000,- EURO
- Plus additional cost (which normally don't show up in TCA !!)
  - Complex network infrastructure (ca. 100.000 to 150.000 EURO)
  - Complex SAN-Infrastructure (x * 100.000,- EURO to replace the directors and switches)
  - High use of Energy
    - z/Series: 10,6 kW
    - Blade-Center: 20 kW
  - Additional cost for cooling and space (??? EURO)
  - Higher amount for FTE to Administer the solution

Swisscom IT services



Linux for zSeries Architecture



Swisscom IT today

- Moved to 2 x z9 5 IFLs (3IFLs)
- Total number of virt. Linux Servers 280 -300 (~180 prod.SRV)
- z/VM 5.2 & SUSE SLES, SLES9, SLES10 as well as RedHat (test)
- Order Management System (EJBs) (2003)
  - 1500 User 600.000 (450.000) TX / Day (~450 MIPs)
- Swisscom billing system (Orbix) (2002)
  - 4000 User 450.000 (350.000) TX / Day (~80MIPs / 12%-58% cost reduction 1.Year)
- Swisscom billing Mediator (C, pearl, FCP-SAN) (2005)
  - (~180 MIPs / 45% cost reduction in less than 1.Year)
- **Migrated over night 300+ Linux Server from z990 to z9**

Swisscom IT today

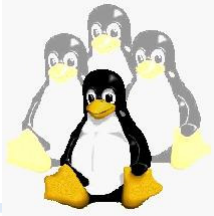
- MAPLIN Internet online system Fixnet (J2EE) (2004), 3TB Oracle DB completed
  - 17 new Linux for zSeries Server running Oracle 9.2 (DB 40-80 GB each)
  - Replaced 21 HP Server (4100 True64)
  - (~90MIPs / 26 % - 70 % cost reduction 1.Year)
- Web application in Virtual Secure Zone for Enterprise Customers to administrate all
  - Swisscom Group Bills
- Server Consolidation Project -Replacement of proprietary Swisscom Applications ongoing
  - Deployment depends on the SLA with the customer (Blade or Linux on System z)
- Re-Hosting of Swiss.com Swiss-Airline website within the swisscom infrastructure
  - Most servers where possible to re-locate to Linux on system z

Challenges Today / Future

- Charge back Method across Swisscom
  - First customer charged
- Group scheduling
  - Multiple WAS (4-6) make up one application (Total 500MIPs)
- FCP DR design
  - First discussion and test implementation still in PoC mode
- Standalone Dump to DASD
  - No tapes allocated to the z/VM LPARS
  - Open FITS req.
- Swiss need CICS TX-series for Linux on System z

50 years Endress+Hauser

50 Years of KnowHow
1953-2003



LINUX as a strategic Plattform for Future IT at E&H

E & H Migration / growth path

08/2002	12/2002	06/2003	11/2003	2/2004	03/2006

2 x z900

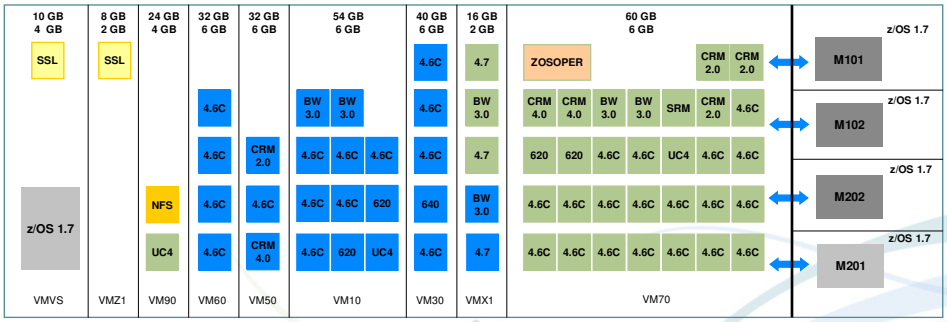
z9-109 S54

37 IFLs ~ 13.500 MIPs

512 GB Mem.

IBM System z9-109 S54 SAP System Landscape in Detail

Z9 – 109 Model S54 16 CP's, 37 IFL's, 1 CF



- SSL Server VM
- SAP production system
- SAP T, D- and Q-systems
- IBM/DB2 V7.1 for production systems
- IBM/DB2 V7.1 for T, D- Q-systems
- Hypersocket

Operating Systems: z/VM 5.3

SUSE SLES9 SP3

testing on SUSE SLES10 SP3

SAP Server consolidation

Why ?

High-Availability, better Scalability of the total System / solution

Higher Performance on Batch Processing

Disaster Recovery optimized

Faster and highly secure access to DB2 using Hipersockets

Build fast and effective new SAP Systems (provisioning / cloning)

Cost savings in operations and Systems-management (TCO)

Implementation effort ?

2 FTE z/VM and LINUX Skill available at the customer

IBM support (2 MM)

Challenges Today

- Standalone Dump support for DASD
- Linux for System z SAP DB
- Want to move all z/OS workload to Linux on system z

4 years Linux on z in production

Sparda-Datenverarbeitung eG

offers it‘s service to

28 Bank-Companies with

7900 Employees and a total

Balance of 66,2 Bn. €

23,600,000 Accounts

5,000,000 Customers

oliver.roethinger@spb.de

Sparda-Datenverarbeitung eG

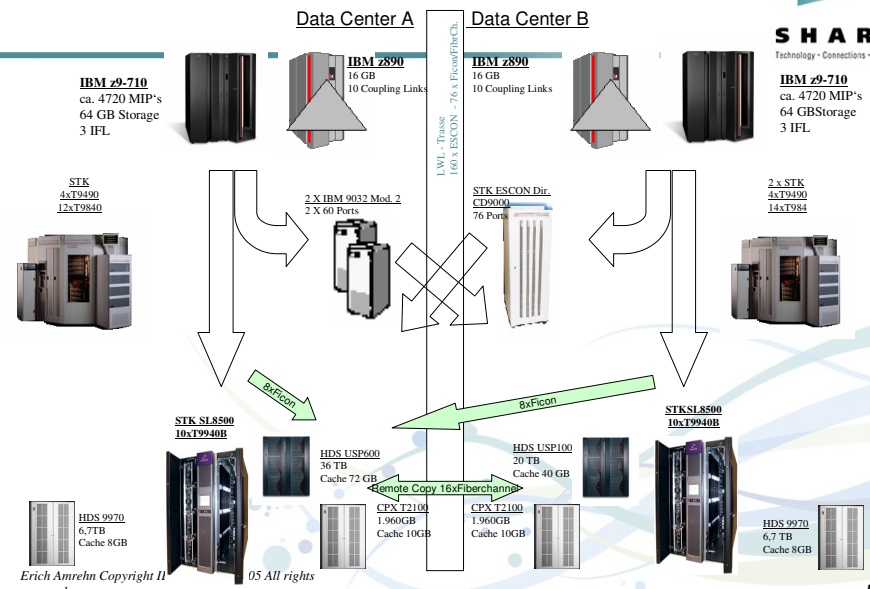
supports with 360 employees

about 10.000 PC‘s
650 servers
2 mainframes

394 locations allover Germany

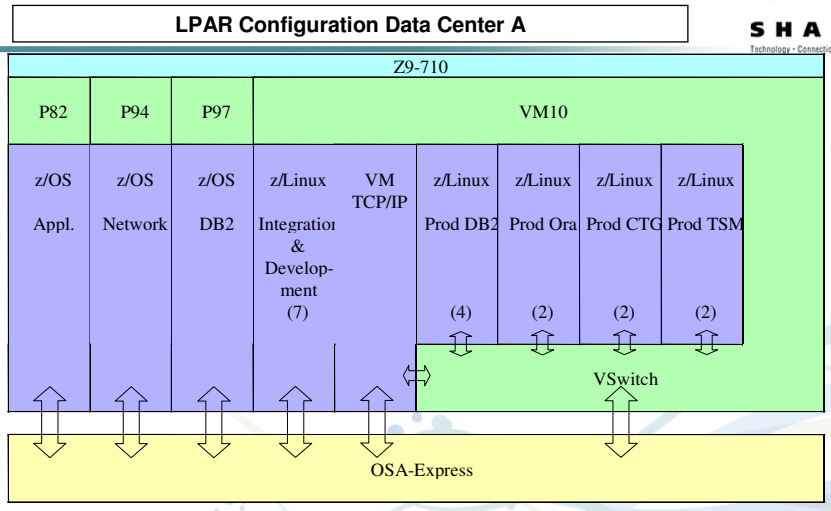
959 cash machines
1.200 account statement printers
1.700 machines for different tasks

Host Infrastructure



Current Infrastructure (Mainframe)

- 2x z9 (IBM 2094-710) - each 64 GB; 40GB for each z/VM
- 2x z890 (IBM 2086 / 2 Engines CF / each 16 GB)
- Coupling Links (XCF): ISC-D (Card) and ISC-3 (Port)
- 2x HDS USP (USP600 & USP100) together 56 TB
- 2x HDS 9970 together 13 TB
- DASD: Direct attached Ficon
- Tape: STK SL8500/9940B, STK 9310/9840 both at each DataCenter
- OS: zOS 1.8, z/VM 5.2 and SLES9



The world before Linux on z

- Big Intel based Server Farms with Red Hat.

- Expensive Infrastructure for cooling, power, physical space and administration.

- Delayed projects, because linux administrators were very busy.

- Some Servers had very high idle times. CPU could not be used for other servers.

- TSM was running in z/OS and needed 1 CPU. This is really expensive, because
  all Software costs will grow the same way as TSM CPU consumption grows.

- Ordering hardware for stand-alone servers takes several weeks.

Why we decided to use linux on z

- Most applications at a banking company have an availability of
  24 * 7 * 365. IBM mainframes are known as high reliable.

- Nearly all applications must be ready for disaster recovery.
  With the Virtualization of z/VM we expect to reduce infrastructure complexity.

- Many applications need a high I/O rate. This is one of the main advantages of a mainframe.

- The expensive z/OS CPU´s should be used for core applications like CICS and Adabas.

- Building a new Linux image can be done very fast:

  - Installing a new image takes about 3 hours.
  - With our cloning concept we need about 30 minutes for a new image.

The beginning

- 1st week in May 2004 workshop in Montpellier.

- 3rd week in May 2004 test installation z/VM.

- 4th week in May 2004 the first Linux system with DB2 database.

- June 2004 several tests were done on the zLinux system by the development.
The application programmers were really satisfied but the first they said:

"We need more Linux systems and when will we get them?"

Our answer was:

"You can have more systems and you will get them the next day".

- This was the birth of Linux on z at the Sparda Datenverarbeitung e.G.

Project TSM under Linux on zSeries

- This was the first "big deal".

- Costs!!!
TSM was using 10% of our CPU. This amount could be saved from the software pricing.

- New technologies were not available for z/OS TSM Version.

- The lack of storage in the library forced a fast decision.

Project TSM under linux on zSeries Challenges

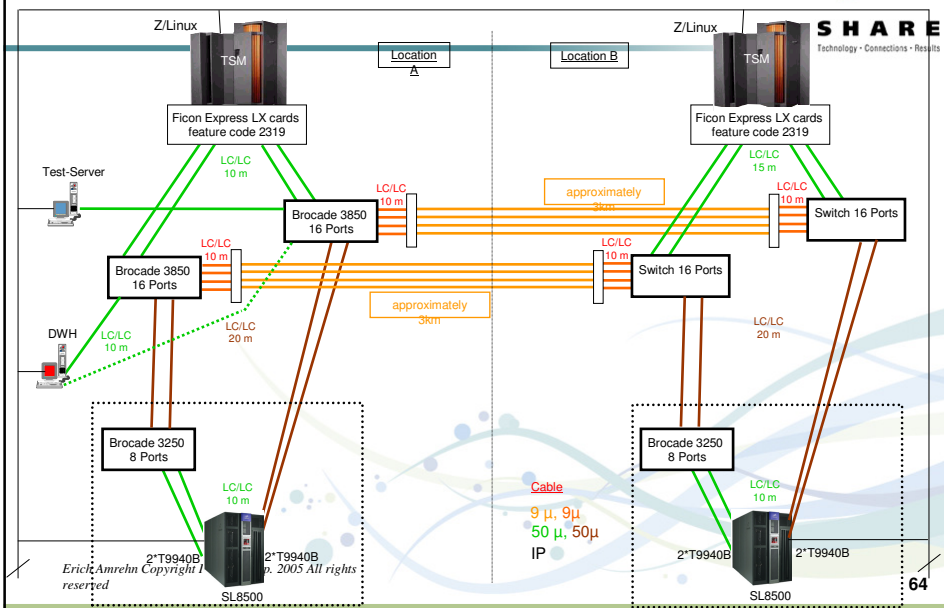
Which backups do we need?

- about 250 servers (Windows, Linux and Unix) with a capacity of 1 TB.
The backup must be finished at 05.00am every day.

- Backup of database logs are causing high availability of the TSM server.

- LAN Free Backup Data-Warehouse.

Infrastructure TSM with Linux



The next challenges

- The linux images are growing rapidly.
  We need a tool for software distribution (SUN UCE is in discussion).

- Performance monitoring is needed.
  We use the z/VM performance toolkit.

- Maintenance concept is necesary.
  This could be also done with SUN UCE.

- Cloning concept must be developed.
  No tool is needed, we use z/VM DDR and do the changes manuelly.

- The new platform should save money not causing costs.
  We have to use existing software.

- Organization barriers and turf wars.

Usage of existing z/OS components

- Backup of z/VM system disks with DFDSS.

- Backup of z/VM minidisks in z/OS; sheduled with OPC (now TWS).

- z/Linux performance data is available in z/OS TDS (Tivoli Decision Support;)
  but we prefer perfomance toolkit data.

- Archiving system logs with BETA Systems Software.

- Backup of z/Linux system disks with DFDSS.

Our applications on Linux on z

The most systems we are running are database servers.
We are not running CPU intensive applications.

- As mentioned earlier TSM. It's an I/O intensive application.

- Our Internet Home Banking. This application has the highest availability.
  We are running systems with DB2 and for connection to CICS we are using CICS Transaction Gateway.

- Brokerage between Bank-Companies. This application is running with critical data.

Challenges

- Device Driver depends on special kernel versions.
  With the latest kernel versions this problem is less important.

- Knowledge of z/VM and z/Linux in the same department is useful.

- As we installed the system z9 (November 2005) we had 8 outages of the VSWITCH. The issue was a faulty microcode on the OSA.

- Sometimes it takes some time till the latest software versions are available on z.

- Some tools need a special version on system z (for instance tcpdump).

Summary

- We save money with Linux on z.

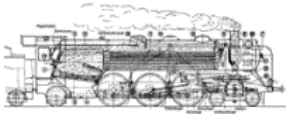
- We have reliable hardware.

- We have reliable software.

- We save time.

- We have flexible solutions for disaster recovery and maintenance.

Die Bahn DB



Use case at DB Systems

- ***Business need***
  The customer wanted to consolidated its distributed scattered backup infrastructure.

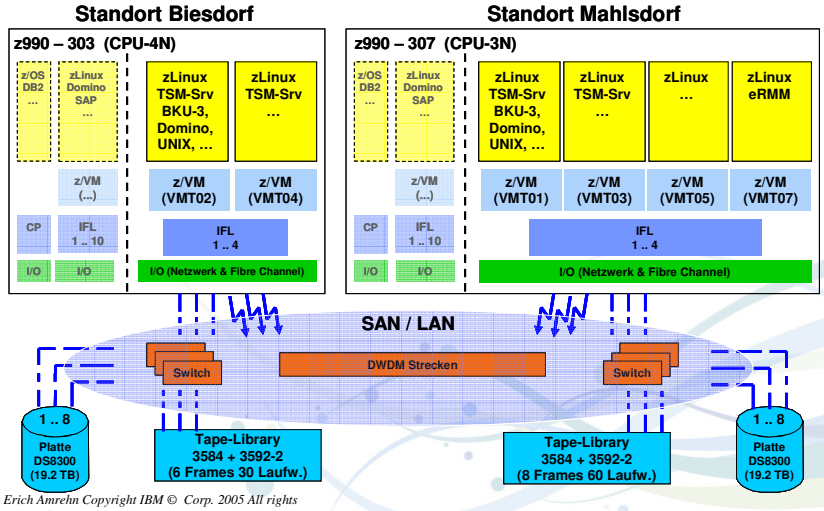
- ***Proposed solution by eRMM (now* IBM Integrated Removable Media Manager** (IRMM)

  IBM offered a farm of TSM servers running on zLinux where IRMM manages the tape resources for the TSM server farm.

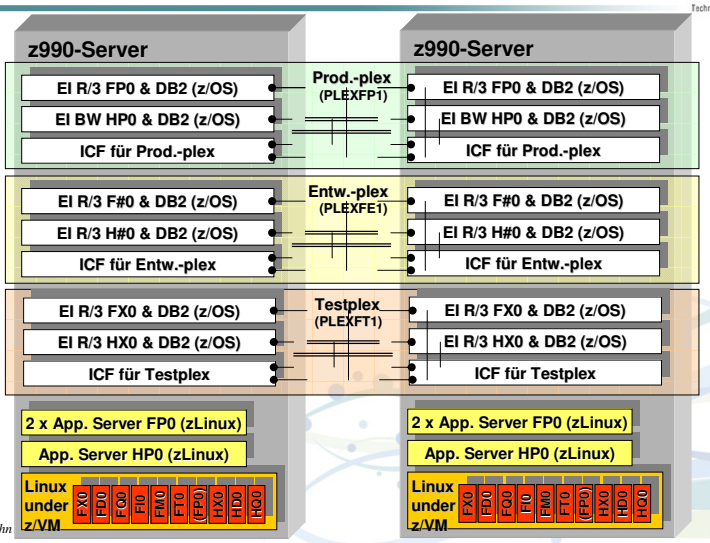
- ***Business value***
  The customer builds a backup infrastructure which is very flexible and scalable. The advanced management functions of zLinux, z/VM, and IRMM allow to add new TSM Servers and new TSM Storage Agents as needed.

IRMM's advanced tape management capabilities help to master the complexity of this frequently changing setup.

Customer setup at DB Systems



SAP R/2 to R/3 and build of M11 LPAR

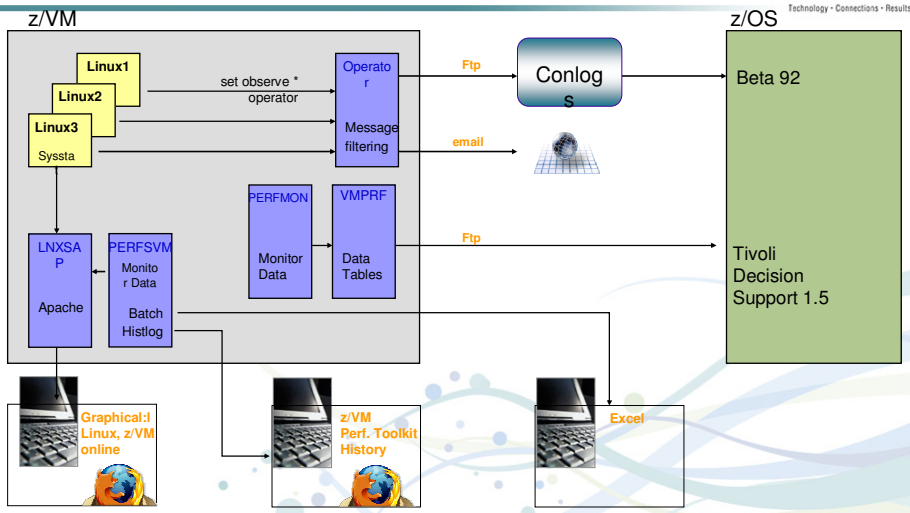


Monitoring and Accounting
Which Data to collect?

•z/VM	•LINUX on zSeries

• Performance / Capacity
  • CPU utilization in percent (%CP)
  • User CPU utilization in percent (%US)
  • Total CPU load, in percent (%CPU)
  • SSCH and RSCH rate per second (I/O/s)
  • Average users active in monitor interval (ACT)
  • Total page rate, per second (PG/s)
  • Percent of expanded storage in use (%XS)
  • Page migrate rate's, expanded storage to dasd (MIGR/s)
• Availabilty
  • Logon / Logout user
  • Active time/user
• Accounting
  • Total CPU / guest
  • Total WSS / guest
	• Performance / Capacity
	  • Load average (ldavg-1,ldavg-5, ldavg-15)
	  • SWAP (kbswpused, pswpin/s, pswpout/s)
	  • CPU utilization (%user, %system, idle)
	  • Storage / Application
	  • Memory (kbmemused)
	  • Buffers (kbbuffers)
	  • Cache (kbcached)
	• Availabilty
	  • Uptime Linux Nativ (Uptime command)
	  • Uptime Linux Guest (Uptime command)
	• Consol Logs
	 

Monitoring and Accounting
Consol Logs, Performance Monitoring Realtime and Histlog



Monitoring and Accounting
Sysstat, Accounting Data, Erep/Ontop

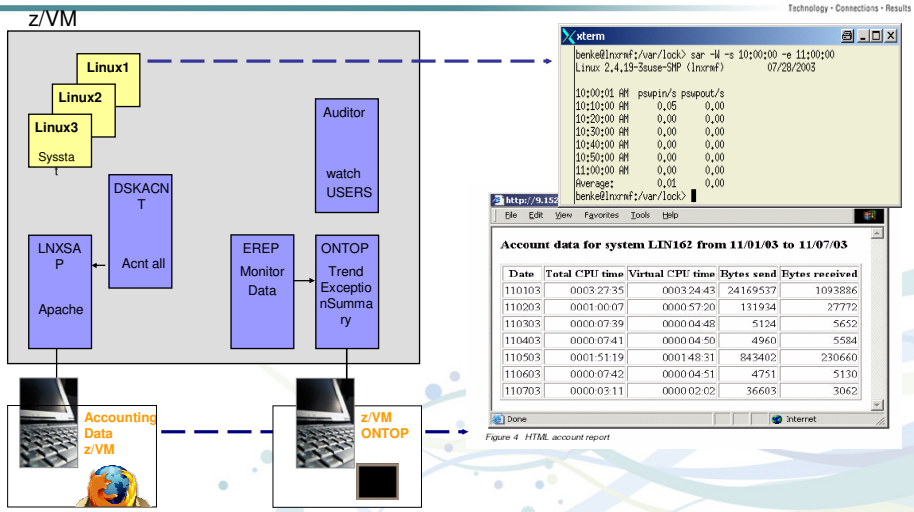


Figure 4 HTML account report

Challenges Today

- How to manage 19 z/VM Systems
  - Shared Directory and Service Concept
- Linux for System z Lifecycle Management
  - Patch-level management
- Charge Back for used capacity
  - CPU, Mem, I/O, Network

Question's?

End

ENABLING BUSINESS. A THROUGH Z.

Erich Amrehn
Executive IT Specialist
IBM Germany, Boeblingen
AMREHN@DE.IBM.COM