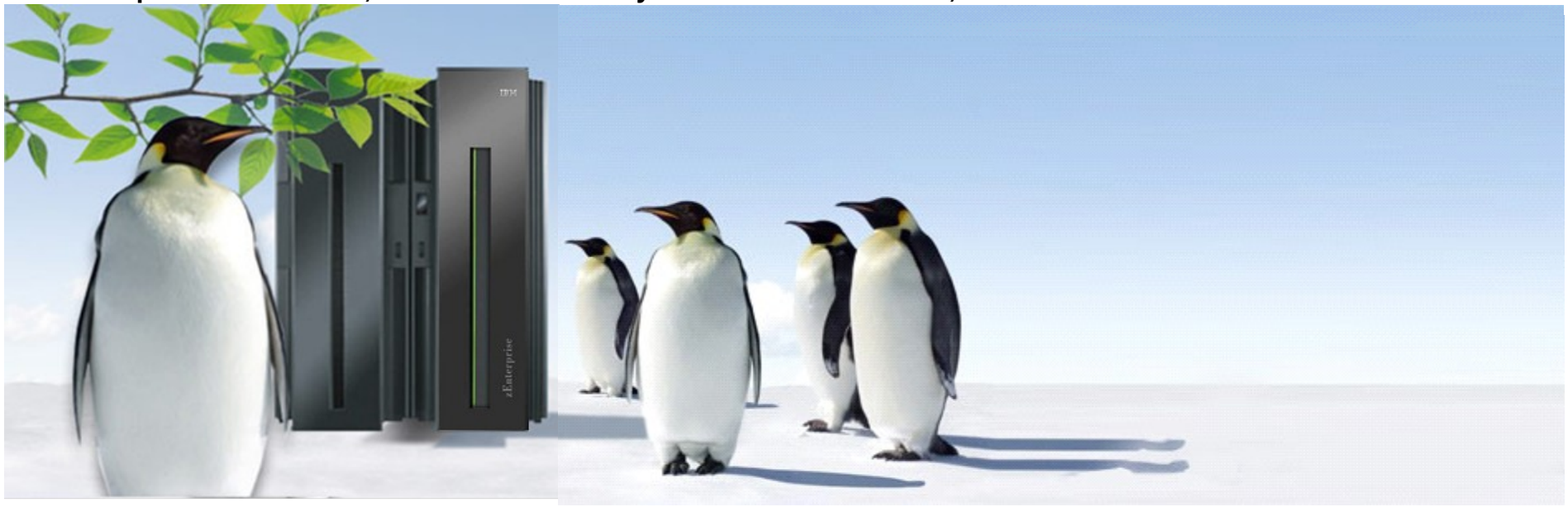
Stefan Haberland, Linux on System z Development <stefan.haberland@de.ibm.com>

Linux on System z Update:
Current & Future Linux on System z Technology

Session: VM01 / IS04

5th European TU for z/VSE, z/VM and Linux on System z -24 -26 Oct. 2011, Berlin



How Linux on System z is developed

Let's see what IBM is doing in the Labs to develop Linux

IBM collaborates with the Linux community

- ...has been an active participant since 1999
- ...is one of the leading commercial contributors to Linux
- ...has over 600 full-time developers working with Linux and open source

Linux Kernel & Subsystem Development

Kernel Base

Security

Systems Mgmt

Virtualization

Filesystems,

and more...

Expanding the Open Source Ecosystem

Apache

Eclipse

Mozilla Firefox

OpenOffice.org,

and more...

Promoting Open Standards & Community Collaboration

The Linux Foundation

Linux Standards Base

Common Criteria certification,

and more...

Foster and Protect the Ecosystem

Software Freedom Law Center

Free Software Foundation (FSF),

and more...

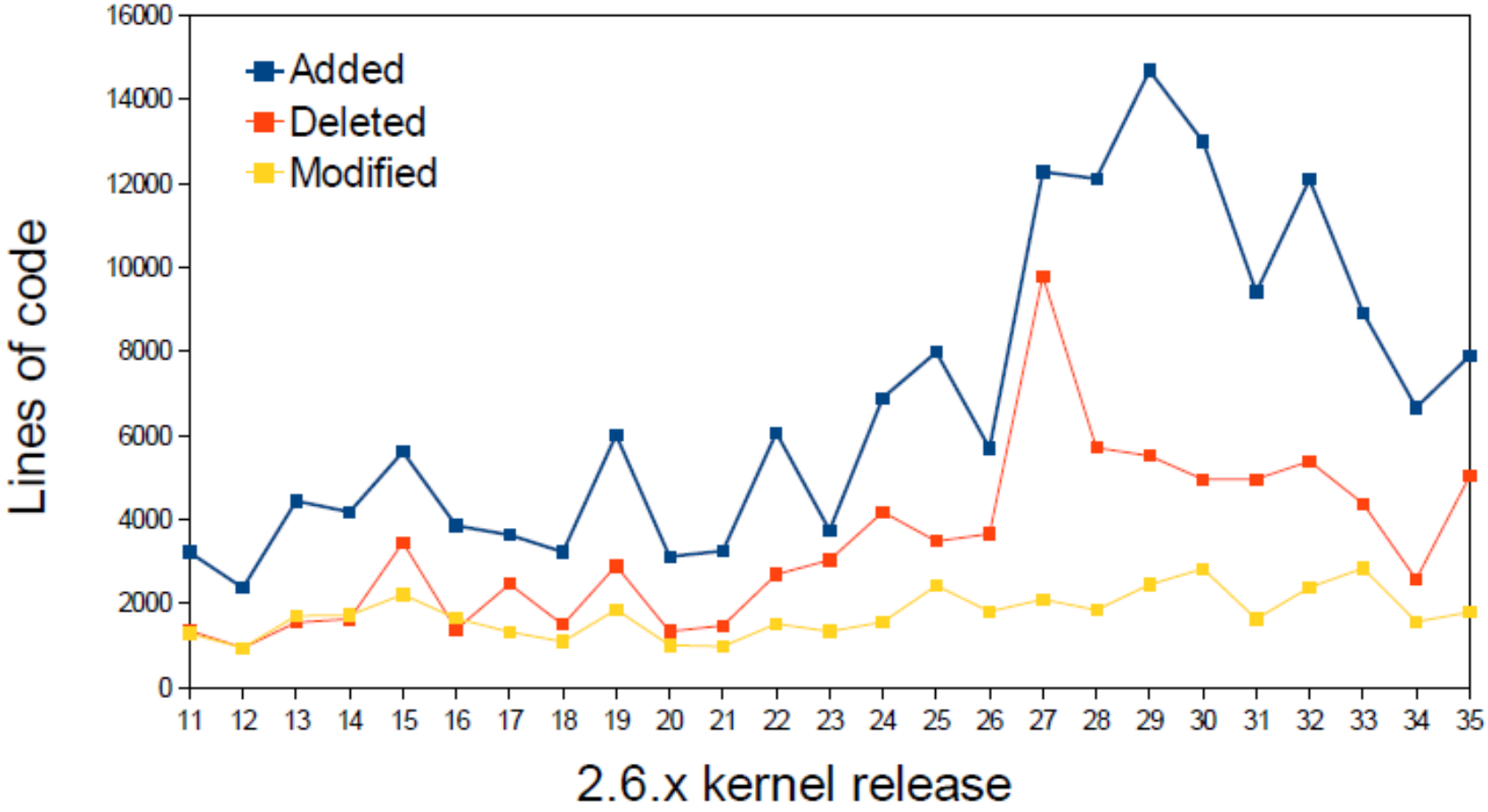
Facts on Linux

- Last year, **75%** of the Linux code was developed by **programmers working for corporations**.
- **$7.37 billion:** projected cost to produce the 283 million lines of code which are contained in Linux Distribution **in a commercial environment.**
- IDC forecasts show that **Linux server revenue will grow by 85.5%** between 2008 and 2012 **in the non-x86 server space** equalling a four year compound annual growth rate of 16.7%.
- **Linux is Linux**, but …features, properties and quality differ dependent on your platform

Source: Intelligence Slideshow: 40 Fast Facts on Linux http://www.baselinemag.com/c/a/Intelligence/40-Fast-Facts-on-Linux-727574/
http://www.internetnews.com/dev-news/article.php/3659961
http://public.dhe.ibm.com/software/au/downloads/IBM_zLinux_DAG_FINAL.pdf

Linux kernel development: Rate of Change

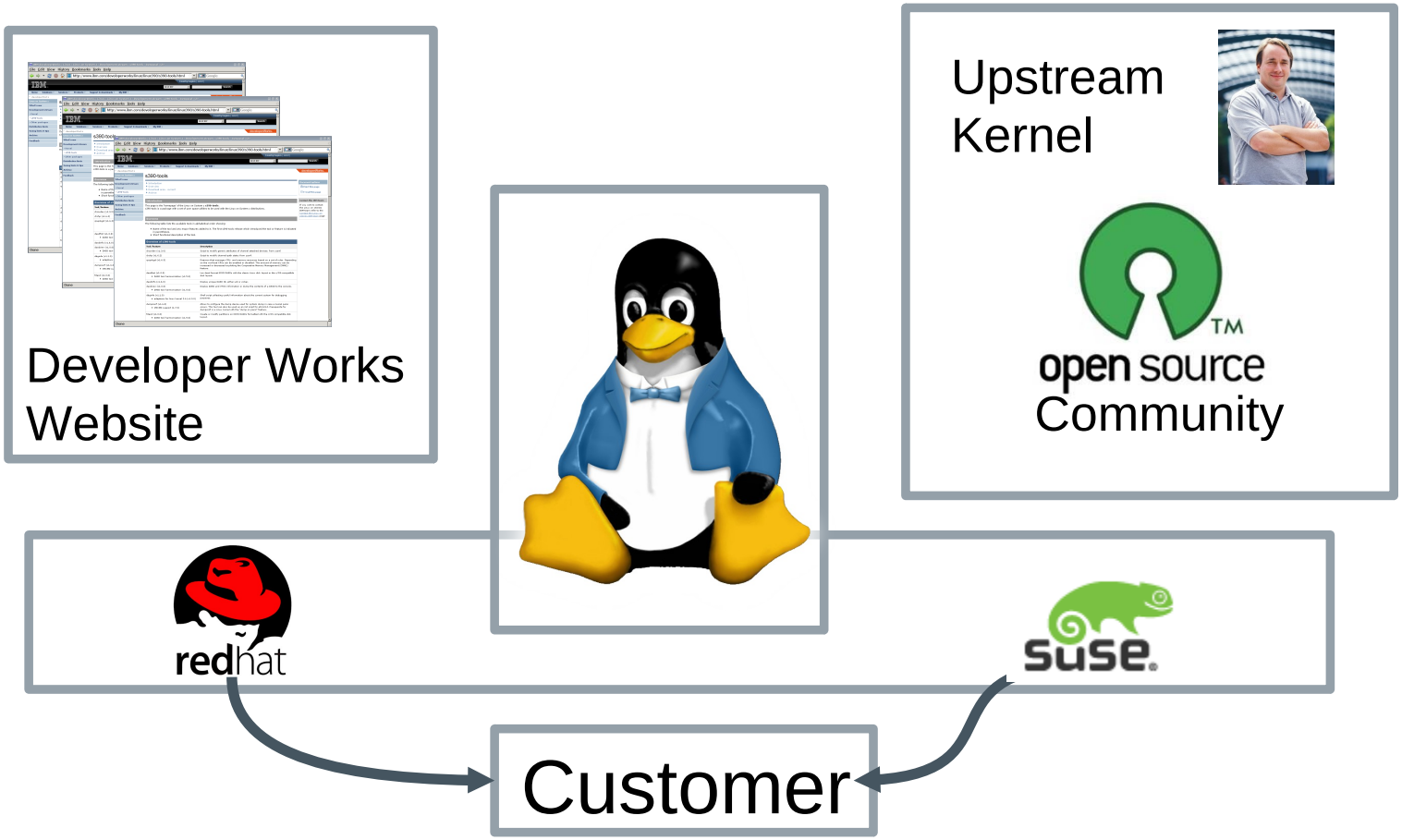
Average: 6683 lines added, 3774 lines removed, 1797 lines changed every day for the last 5 1/2 years.



Source: http://www.linuxfoundation.org/docs/lf_linux_kernel_development_2010.pdf

IBM Linux on System z Development

IBM Linux on System z Development contributes in the following areas: Kernel, s390-tools, Open Source Tools (e.g. eclipse, ooprofile), GCC, GLIBC, Binutils



....the code you use is the result of the efforts of an anonymous army of blue penguins involved in developing, testing, documenting, ...

Linux on System z distributions (Kernel 2.6 based)

- **SUSE Linux Enterprise Server 9 (GA 08/2004)**
  - Kernel 2.6.5, GCC 3.3.3, Service Pack 4 (GA 12/2007)
- **SUSE Linux Enterprise Server 10 (GA 07/2006)**
  - Kernel 2.6.16, GCC 4.1.0, Service Pack 4 (GA 05/2011)
- **SUSE Linux Enterprise Server 11 (GA 03/2009)**
  - Kernel 2.6.27, GCC 4.3.3, Service Pack 1 (GA 06/2010), Kernel 2.6.32
- **Red Hat Enterprise Linux AS 4 (GA 02/2005)**
  - Kernel 2.6.9, GCC 3.4.3, Update 9 (GA 02/2011)
- **Red Hat Enterprise Linux AS 5 (GA 03/2007)**
  - Kernel 2.6.18, GCC 4.1.0, Update 6 (GA 05/2011)
- **Red Hat Enterprise Linux AS 6 (GA 11/2010)**
  - Kernel 2.6.32, GCC 4.4.0 Update 1 (GA 05/2011)
- **Others**
  - Debian, Slackware,
  - Support may be available by some third party

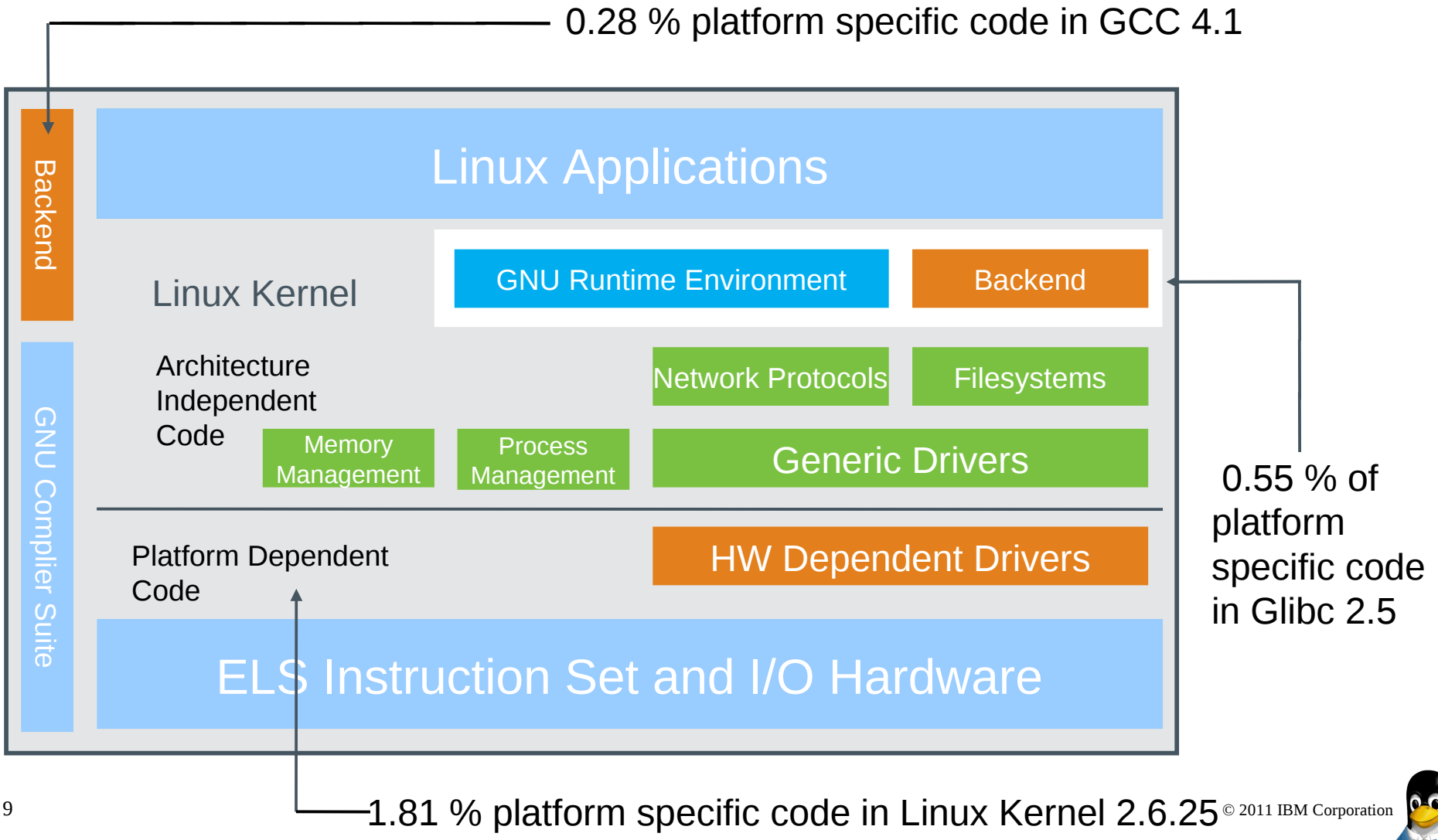
IBM Supported Linux Distributions for System z

	z196	z10	z9	zSeries
RHEL 6	✓	✓	✓	✗
RHEL 5	✓	✓	✓	✓
SLES 10	✓	✓	✓	✓
SLES 11	✓	✓	✓	✗

http://www-03.ibm.com/systems/z/os/linux/resources/testedplatforms.html

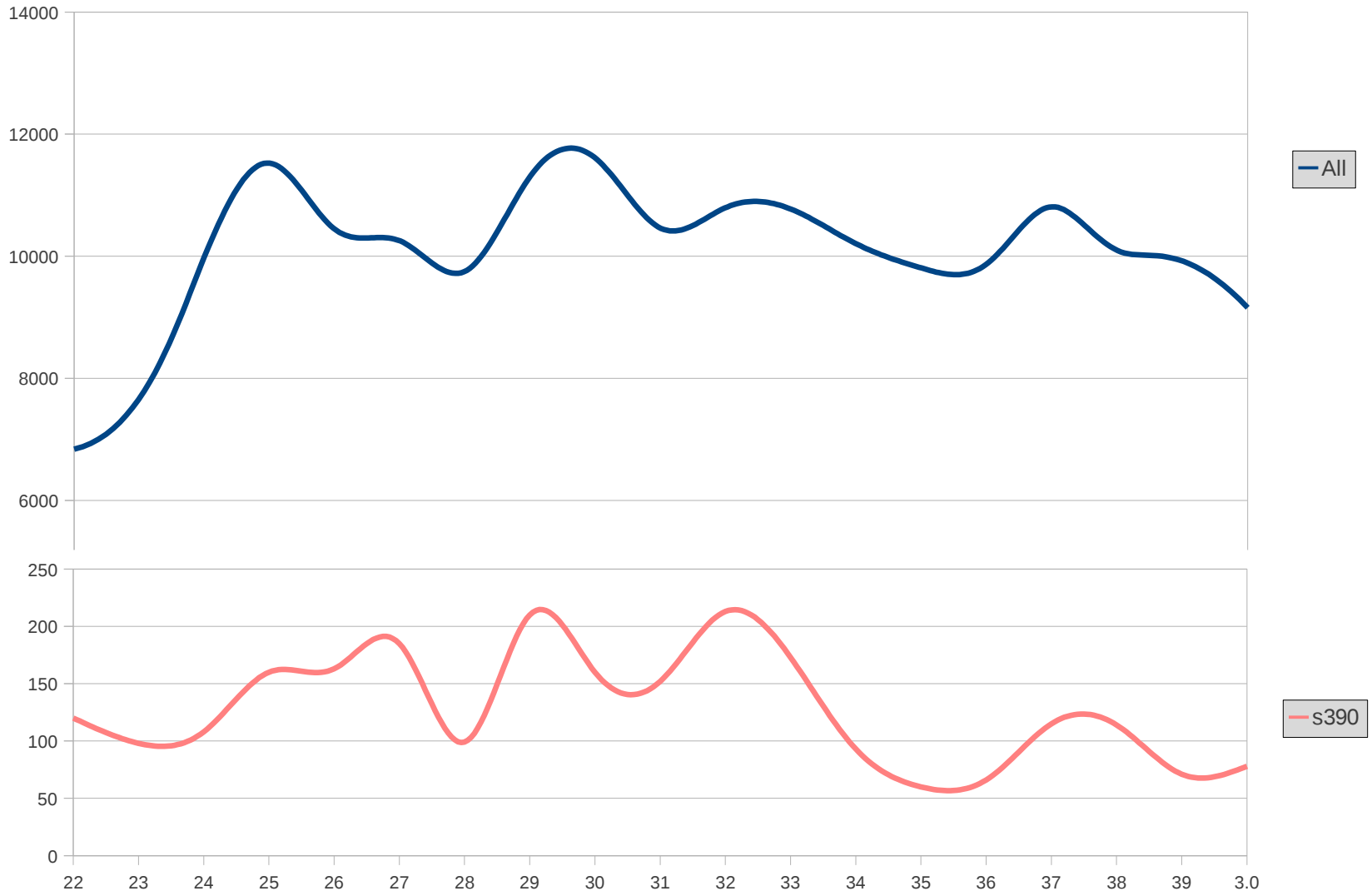
Structure of Linux on System z

Many Linux software packages did not require any code change to run on Linux on System z

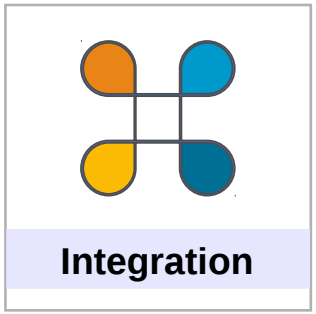


Linux kernel development: System z contributions

Changesets per 2.6.x/3.x kernel release



Linux on System z Development Focus



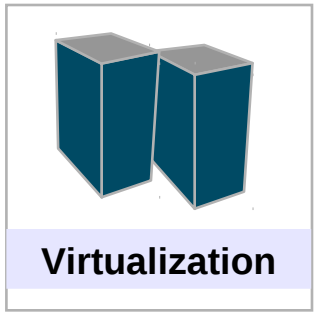
Integration

Application Serving

- z/OS & z/VSE integration

Data Hub

- Database Consolidation



Virtualization

Virtualization &

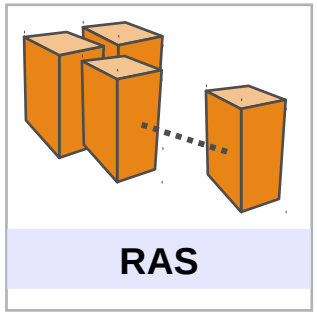
Virtualization Management

- Ease of Use
- Serviceability
- Hosting capacity

Security

Security

- Certifications
- Data security & privacy

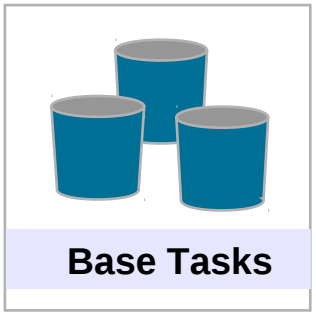


RAS

Continuous Availability

& Data Replication

- RAS
- Differentiation for mission critical workloads



Base Tasks

Customer Requirements

- Address customer observed deficiencies

Competitiveness

- Close competitive gaps
- Differentiation / innovation that matters

Hardware Support

- Exploitation of new System z HW
- Storage exploitation

Linux

- Maintainership & code currency

Your Linux on System z Requirements?

Are you missing a certain feature, functionality or tool? **We'd love to hear from you!**

We will evaluate each request and (hopefully) develop the additional functionality you need.

Send your input to stefan.haberland@de.ibm.com

Current Linux on System z Technology

Features & Functionality contained in the Novell & Red Hat Distributions

Kernel Core

- **Breaking event address for user space programs (kernel 2.6.35)**
  - Remember the last break in the sequential flow of instructions
  - Valuable aid in the analysis of wild branches

- **z196 enhanced node affinity support (kernel 2.6.37)**
  - Allows the Linux scheduler to optimize its decisions based on the z196 topology

- **Performance indicator bytes (kernel 2.6.37)**
  - Display capacity adjustment indicator introduced with z196 via /proc/sysinfo.

How can you read files on a CMS disk with Linux?

About the CMS user space file system (fuse) support


11.1


6.1

- Allows to mount a z/VM minidisk to a Linux mount point
- z/VM minidisk needs to be in the enhanced disk format (EDF)
- The cmsfs fuse file system transparently integrates the files on the minidisk into the Linux VFS, no special command required

```
root@larsson:~> cmsfs-fuse /dev/dasde /mnt/cms
root@larsson:~> ls -la /mnt/cms/PROFILE.EXEC
-r--r----- 1 root root 3360 Jun 26 2009
/mnt/fuse/PROFILE.EXEC
```

- By default no conversion is performed
  - Mount with '-t' to get automatic EBCDIC to ASCII conversion

```
root@larsson:~> cmsfs-fuse -t /dev/dasde /mnt/cms
```

- Use fusermount to unmount the file system again

```
root@larsson:~> fusermount -u /mnt/cms
```


How can you read & write files on a CMS disk with Linux?

About the CMS user space file system (fuse) support

- Write support is also available in RHEL 6.1
  - use “vi” to edit PROFILE.EXEC anyone ?

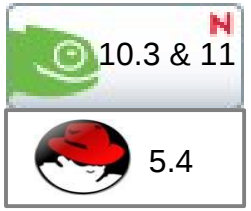
Virtualization

- **TTY terminal server over IUCV**

  Provide central access to the Linux console for the different guests of a z/VM.

  Fullscreen applications like *vi* are usable on the console.

  Access Linux instances with no external network because IUCV is independent from TCP/IP

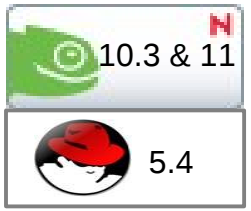


- **Dynamic memory attach/detach**

  Allows to attach/detach memory for Linux as a guest without needing to reipl.

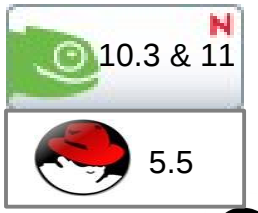
- **Extra kernel parameter via VMPARM**

  Allows to use z/VM VMPARM variable to add or substitute the kernel command line.



- **Provide CMS script for initial IPL**

  Avoids having to create an script to start a new installation under z/VM.

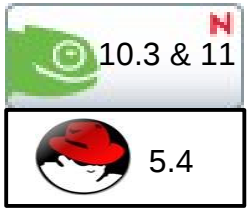


Note: this list comprises selected items only

Virtualization (cont.)

- **Exploitation of DCSSs above 2G**
  Solves restriction to use DCSS above or greater than 2GB.

- **Provide service levels of HW & Hypervisor in Linux**
  Improves serviceability by providing uCode and z/VM levels via /proc interface




```
root@larsson:~> cat /proc/service_levels
VM: z/VM Version 5 Release 2.0
service level 0801(64-bit)
qeth: 0.0.f5f0 firmware level 087d
```

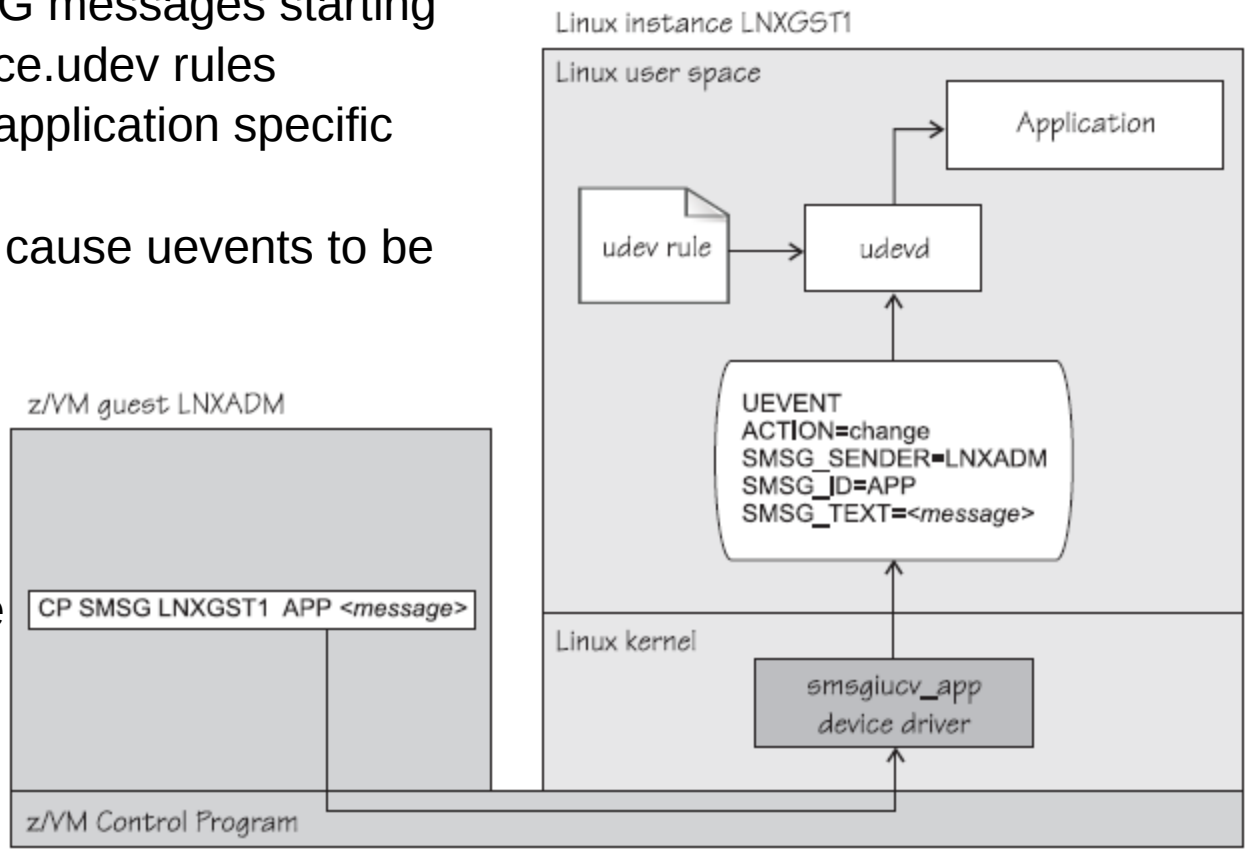

Note: this list comprises selected items only

Deliver z/VM CP special messages as uevent

Allows to forward SMSG messages starting with “APP” to user space.udev rules can be used to trigger application specific actions

The special messages cause uevents to be generated.

See “Writing udev rules for handling CP special messages” on page 229 in the Device Driver Book for information about handling the uevents.



hyptop: display hypervisor utilization data

- The hyptop command is a top-like tool that displays a dynamic real-time view of the hypervisor environment
  - It works with both the z/VM and LPAR hypervisor
  - Depending on the available data it can display information about CPU and memory
  - running LPARs or z/VM guest operating systems
- **The following is required to run hyptop**
  - The debugfs file system must be mounted
  - The hyptop user must have read permission for the required debugfs files:
    - z/VM: <debugfs mount point>/s390_hypfs/diag_2fc
    - LPAR: <debugfs mount point>/s390_hypfs/diag_204
  - To monitor all LPARs or z/VM guests your instance requires additional privileges
    - For z/VM: The user ID requires privilege class B
    - For LPAR: The global performance data control box in the LPAR activation profile needs to be selected

hyptop: Display hypervisor utilization data

Example of z/VM utilization data

```
10:11:56 CPU-T: UN(16)                                                ?=help
system    #cpu     cpu     Cpu+    online memuse memmax wcur
(str)      (#)     (%)     (hm)     (dhm)  (GiB)  (GiB)  (#)
T6360003     6  506.92 3404:17 44:20:53   7.99   8.00  100
T6360017     2  199.58    8:37 29:23:50   0.75   0.75  100
T6360004     6   99.84  989:37 62:00:00   1.33   2.00  100
T6360005     2    0.77    0:16  5:23:06   0.55   2.00  100
T6360015     4    0.15    9:42 18:23:04   0.34   0.75  100
T6360035     2    0.11    0:26  7:18:15   0.77   1.00  100
T6360027     2    0.07    2:53 62:21:46   0.75   0.75  100
T6360049     2    0.06    1:27 61:17:35   0.65   1.00  100
T6360010     6    0.06    5:55 61:20:56   0.83   1.00  100
T6360021     2    0.06    1:04 48:19:08   0.34   4.00  100
T6360048     2    0.04    0:27 49:00:51   0.29   1.00  100
T6360016     2    0.04    6:09 34:19:37   0.30   0.75  100
T6360008     2    0.04    3:49 47:23:10   0.35   0.75  100
T6360006     2    0.03    0:57 25:20:37   0.54   1.00  100
NSLCF1       1    0.01    0:02 62:21:46   0.03   0.25  100
VTAM         1    0.00    0:01 62:21:46   0.01   0.03  100
T6360023     2    0.00    0:04  6:21:20   0.46   0.75  100
PERFSVM      1    0.00    2:12  7:18:04   0.05   0.06    0
AUTOVM       1    0.00    0:03 62:21:46   0.00   0.03  100
FTPSERVE     1    0.00    0:00 62:21:47   0.01   0.03  100
TCPIP        1    0.00    0:01 62:21:47   0.01   0.12 3000
DATAMOVE     1    0.00    0:06 62:21:47   0.00   0.03  100
VMSERVU      1    0.00    0:00 62:21:47   0.00   0.03 1500
OPERSVMP     1    0.00    0:00 62:21:47   0.00   0.03  100
```

hyptop: Display hypervisor utilization data

Example of single LPAR utilization data

```
10:16:59 H05LP30 CPU-T: IFL(18) CP(3) UN(2)                          ?=help
cpuid  type    cpu  mgm visual
(#)   (str)    (%)  (%) (vis)
0       IFL  29.34 0.72 |#############                                |
1       IFL  28.17 0.70 |#############                                |
2       IFL  32.86 0.74 |################                               |
3       IFL  31.29 0.75 |##############                               |
4       IFL  32.86 0.72 |################                               |
5       IFL  30.94 0.68 |##############                               |
6       IFL   0.00 0.00 |                                             |
7       IFL   0.00 0.00 |                                             |
8       IFL   0.00 0.00 |                                             |
9       IFL   0.00 0.00 |                                             |
=:V:N       185.46 4.30
```

RAS

- **Suspend / resume support (kernel 2.6.31)**

  Add the ability to stop a running Linux system and resume operations later on. The image is stored on the swap device and does not use any system resource while suspended.

  Only suspend to disk is implemented, suspend to RAM is not supported.

- **Add Call Home data on halt and panic if running in LPAR (kernel 2.6.32)**

  Report system failures (kernel panic) via the service element to the IBM service organization. Improves service for customers with a corresponding service contract. (by default this features is deactivated)

- **Large image dump on DASD**

  Solves restriction to dump only 48GB of memory to DASD. Now up to 32 ECKD DASDs can be used in a multiple volume configuration

Note: this list comprises selected items only

Suspend / resume support

- Ability to stop a running Linux on System z instance and later continue operations
- Memory image is stored on the swap device specified with a kernel parameter: **`resume=/dev/dasd<x>`**
- Lower the swap device priority for the resume partition

```
root@larsson:~> grep swap /etc/fstab
/dev/dasdb1 swap swap pri=-1 0 0
/dev/dasdc1 swap swap pri=-2 0 0
```

- Suspend operation is started with a simple echo:

```
root@larsson:~> echo disk > /sys/power/state
```

- Resume is done automatically on next IPL
- Use signal quiesce to automatically suspend a guest

```
ca::ctrlaltdel:/bin/sh -c "/bin/echo disk > \
/sys/power/state || /sbin/shutdown -t3 -h now"
```


System z kernel features – Storage FICON

- **Unit check handling (kernel 2.6.35)**
  - Improve handling of unit checks for internal I/O started by the common-I/O layer
  - After a unit check certain setup steps need to be repeated, e.g. for PAV

- **Dynamic PAV toleration (kernel 2.6.35)**
  - Tolerate dynamic Parallel Access Volume changes for base PAV
  - System management tools can reassign PAV alias device to different base devices.

- **Tunable default grace period for missing interrupts in DASD (kernel 2.6.36)**
  - Provide a user interface to specify the timeout for missings interrupts for standard I/O operations on DASD

System z kernel features – Storage FCP

- **Store I/O status and initiate logging (SIOSL) (kernel 2.6.36)**
  - Enhance debug capability for FCP attached devices
  - Enables operating system to detect unusual conditions on an FCP channel

- **SAN utilities (trace, ping, ..) (kernel 2.6.36, lib-zfcp-hbaapi 2.1)**
  - Two new utilities have been added: zfcp_ping and zfcp_show
  - They are useful to discover a storage area network

SAN Utilities: zfcp_show

- **Query Fibre Channel nameserver about ports available for my system:**

```
root@larsson:~> zfcp_show -n
Local Port List:
       0x500507630313c562 / 0x656000 [N_Port] proto = SCSI-FCP  FICON
       0x500507640122 41e4 / 0x656100 [N_Port] proto = SCSI-FCP
       0x5005076401221b97 / 0x656400 [N_Port] proto = SCSI-FCP
```

- **Query SAN topology, requires FC management server access:**

```
 root@larsson:~> zfcp_show
Interconnect Element Name       0x100000051e4f7c00
Interconnect Element Domain ID  005
Interconnect Element Type       Switch
Interconnect Element Ports      256
      ICE Port 000  Online
            Attached Port [WWPN/ID] 0x50050763030b0562 / 0x650000 [N_Port]
      ICE Port 001  Online
            Attached Port [WWPN/ID] 0x500507640122 41e5 / 0x650100 [N_Port]
      ICE Port 002  Online
            Attached Port [WWPN/ID] 0x5005076303008562 / 0x650200 [N_Port]
      ICE Port 003  Offline
     ...
```

SAN Utilities: zfcp_show

- **Query Fibre Channel nameserver about ports available for my system:**

```
root@larsson:~> zfcp_ping 0x5005076303104562
Sending PNG from BUS_ID=0.0.3c00 speed=8 GBit/s
       echo received from WWPN (0x5005076303104562) tok=0 time=1.905 ms
       echo received from WWPN (0x5005076303104562) tok=1 time=2.447 ms
       echo received from WWPN (0x5005076303104562) tok=2 time=2.394 ms

---------- ping statistics -----------
min/avg/max = 1.905/2.249/2.447 ms
--------------------------------------
```

- **zfcp_show and zfcp_ping are part of the zfcp-hbaapi 2.1 package:**

http://www.ibm.com/developerworks/linux/linux390/zfcp-hbaapi-2.1.html

System z kernel features – Networking

- **Offload outbound checksumming (kernel 2.6.35)**
  - Move calculation of checksum for non-TSO packets from the driver to the OSA network card

- **OSX (OSM) CHPIDs for hybrid data network (kernel 2.6.35)**
  - The OSA cards for the zBX Blade Center Extension will have a new CHPID type
  - Allows communication between zBX and Linux on System z

- **NAPI support for QDIO and QETH (kernel 2.6.36)**
  - Convert QETH to the NAPI interface, the “new” Linux networking API
  - NAPI allows for transparent GRO (generic receive offload)

System z kernel features – Networking

- **Support for assisted VLAN null tagging (kernel 2.6.37)**
  - Close a gap between OSA and Linux to process null tagged frames correctly
  - z/OS may sent null-tagged frames to Linux

- **Configuration tool for System z network devices (s390-tools 1.8.4)**
  - Provide a shell script to ease configuration of System z network devices

znetconf network device configuration tool

- Allows to list, add, remove & configure System z network devices
- For example: list all potential network devices:

```
root@larsson:~> znetconf -u
Device Ids                       Type Card Type CHPID Drv.
------------------------------------------------------------
0.0.f500,0.0.f501,0.0.f502 1731/01 OSA (QDIO) 00 qeth
0.0.f503,0.0.f504,0.0.f505 1731/01 OSA (QDIO) 01 qeth
```

- Configure device 0.0.f503

```
root@larsson:~> znetconf -a 0.0.f503
```

- Configure device 0.0.f503 in layer2 mode and portname “myport”

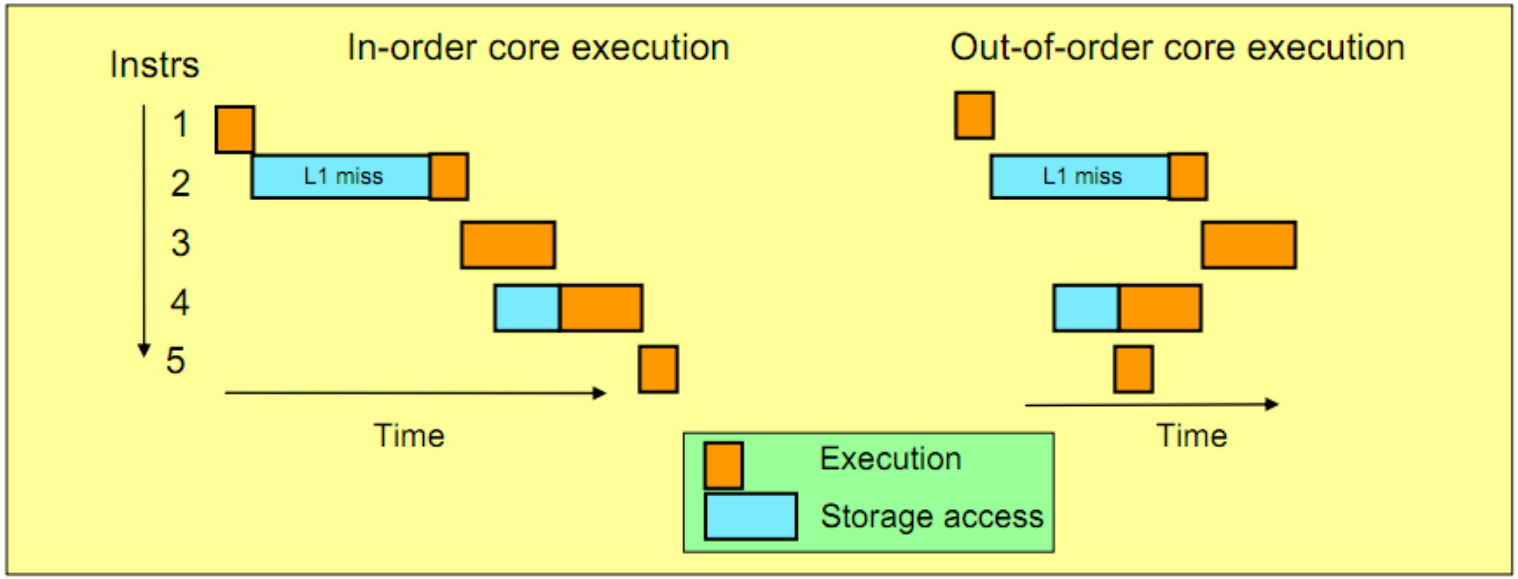
```
root@larsson:~> znetconf -a 0.0.f503 -o layer2=1 -o
portname=myport
```

- Remove network device 0.0.f503

```
root@larsson:~> znetconf -r 0.0.f503
```

System z toolchain

- **zEnterprise 196 exploitation (gcc 4.6)**
  - Use option -march=z196 to utilize the new instructions added with z196
  - Use -mtune=z196 to schedule the instruction appropriate for the new out-of-order pipeline of z196
  - Re-compiled code/apps get further performance gains through 110+ new instructions



Future Linux on System z Technology

*Software which has already been developed and integrated into the Linux Kernel – but is **not** yet available in any Enterprise Linux Distribution*

Kernel news – Common code

Linux version 2.6.35 (2010-08-01)

- Filesystems: btrfs improvements, XFS delayed logging
- Support for multiple multicast route tables
- Support for Layer 2 Tunneling Protocol L2TP Version 3
- Memory compaction

Linux version 2.6.36 (2010-10-20)

- Tilera architecture support
- Concurrency-managed workqueues
- Improve VM-related desktop responsiveness
- Integration of AppArmor
- New out-of-memory killer (OOM)

Linux version 2.6.37 (2011-01-04)

- Filesystems: better SMP scalability for ext4, XFS scalability improvements
- Removal of the BKL: Big Kernel Lock
- I/O throttling support for process groups
- Jump labels: performance optimization for disabled tracepoints

Linux version 2.6.38 (2011-03-14)

- Automatic process grouping (SCHED_AUTOGROUP)
- RCU-based path name lookup (dcache scalability)
- Transparent huge pages
- Transmit packet steering (XPS) for multiqueue devices

Kernel news – Common code

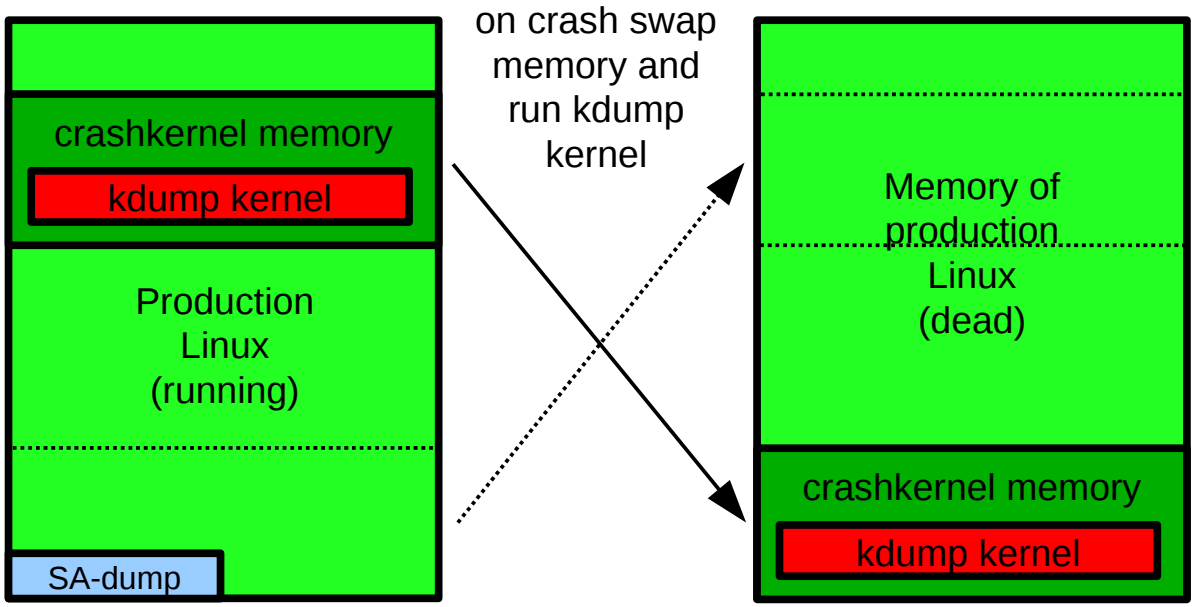
- **Linux version 2.6.39 (2011-05-18)**
  - Ext4 SMP scalability
  - IPset network resource groups
  - Transcendent memory
  - Unicore32 architecture
- **Linux version 3.0 (2011-07-21)**
  - New kernel version numbering scheme
  - Cleancache (was transcendent memory) support for ext4, btrfs and XFS
  - Preemptible mmu_gather for reduced latency
  - Enhancements for the memory cgroup controller

System z Kernel Features - Core

- **QDIO outbound scan algorithm (kernel 2.6.38)**
  - Improved scheduling of QDIO tasklets, OSA / HiperSockets / zfcp need different thresholds.
- **Two stage dumper / kdump support (> kernel 3.0)**
  - Enhanced dump support that is able to reduced dump size, share disk space, dump to network, etc.
  - Integrated into the System z stand-alone dump tools and shutdown actions framework

System z kernel features – two stage dumper / kdump support

- Use a preloaded crashkernel to run in case of a system failure
- Can be triggered either as panic action or by the stand-alone dumper
- Use the makedumpfile tool to filter the memory of the crashed system



Networking

- **Toleration of optimized latency mode (kernel 2.6.35)**
  - OSA devices in optimized latency mode can only serve a small number of stacks / users. Print a helpful error message if the user limit is reached.
  - Linux does not exploit the optimized latency mode
- **New default qeth configuration values (kernel 2.6.39)**
  - Receive checksum offload, generic receive offload & number of inbound buffers
- **QETH debugging per single card (kernel 2.6.36)**
  - Split some of the global QETH debug areas into separate per-device areas
  - Simplifies debugging for complex multi-homed configurations
- **IPv6 support for the qetharp tool (kernel 2.6.38)**
  - Extend the qetharp tool to provide IPv6 information in case of a layer 3 setup.
  - This is required for comminucation with z/OS via HiperSockets using IPv6.

Note: this list comprises selected items only

New Linux on System z Storage Features (FICON)

- **Query DASD reservation status (kernel 2.6.37)**
  - New DASD ioctl to read the 'Sense Path Group ID' data
  - Allows to determine the reservation status of a DASD in relation to the current Linux
- **Multi-track extension for HPF (kernel 2.6.38)**
  - Allows to read from and write to multiple tracks with a single CCW
- **Improve handling of stolen DASD reservation (kernel 2.6.38)**
  - Provide alternatives to handle unit checks that indicate stolen reservations
    - Fail any request to a device until it is set offline
    - Queue I/O until reservation is released again
- **Access to raw ECKD data from Linux (kernel 2.6.38)**
  - This item allows to access ECKD disks in raw mode
  - Use the 'dd' command to copy the disk level content of an ECKD disk
  - Storage array needs to support the read-track and write-full-track commands

Note: this list comprises selected items only

System z kernel features – Storage FICON

- **Automatic menu support in zipl (s390-tools 1.11.0)**
  - Zipl option that will create a boot menu for all eligible non-menu sections in the zipl configuration file
- **reIPL from device-mapper devices (s390-tools 1.12.0)**
  - The automatic re-IPL function only works with a physical device
  - Enhance the zipl support for device-mapper devices to provide the name of the physical device if the zipl target is located on a logical device

System z kernel features – Usability / RAS

- **CHPID reconfiguration handling (kernel 2.6.37)**
  - Update data structures after channel-path related information change
  - Inform device drivers about relevant changes
- **4096 bit RSA fast path (kernel 2.6.38)**
  - Make use of 4096 bit RSA acceleration available with Crypto Express 3 GA2 cards
- **Address space randomization (kernel 2.6.38)**
  - Enable flexible mmap layout for 64 bit
  - Randomize start address for the runtime stack and the mmap area
- **New libica APIs for supported crypto modes**
  - Provide a programmatic way to query for supported crypto ciphers, modes and key sizes.
  - Deliver information wheter the cryptographic features are implemented in hardware or in software

System z kernel features – Usability / RAS

- **Get CPC name (kernel 2.6.39)**
  - Useful to identify a particular hardware system in a cluster
  - The CPC name and HMC network name are provided
- **CP ACF exploitation of System z196 (kernel 2.6.39)**
  - Add support for new HW crypto modes:
    * cipher feedback mode (CFB), output feedback mode (OFB),
    * counter mode (CTR), Galois counter mode (GCM),
    * XEX based Tweaked Code Book with Cipher Text Stealing (XTS),
    * cipher based message authentication mode (CMAC), and counter with cipher block chaining message authentication (CCM)
- **Removal of data execution protection**
  - “no execute” support relies on the secondary space mode for data separation
  - With System z10 the new instructions LRL, LGRL and LGFRL for pc-relative data access have been added
  - These new instructions access the memory operand in the same address space from where the instructions has been fetched.

System z kernel features – Usability / RAS

- **Dump on panic – prevent reipl loop (s390-tools 1.8.4)**
  - Delay arming of automatic reipl after dump. Avoids dump loops where the restarted system crashes immediately.
- **Add support for makedumpfile tool (kernel 2.6.34, s390-tools 1.9.0)**
  - Convert Linux dumps to the ELF file format, use the makedumpfile tool to remove user data from the dump. Multi-volume tape dump will be removed.
- **Breaking event address for user space (kernel 2.6.35)**
  - Store the breaking-event-address for user space programs, it is a valuable aid in the analysis of wild branches,
- **Precise process accounting (> kernel 2.6.36)**
  - Extend the taskstats interface to provide better process accounting values. Quality goal is a resolution of 10ths of microseconds.

System z toolchain

- **64 bit register in 31 bit compat mode**
  Make use of 64 bit registers in 31 bit application running in z/Architecture mode.
  Allows to use instruction operating on 64 bits, e.g. 64 bit multiplication
  Needs kernel support for asynchronous signals

- **Oprofile hardware customer mode sampling**
  Provide CPU measurement data to applications for performance tuning
  Based on hardware counters and samples built into the CPU
  Use oprofile to communicate the information to user space programs

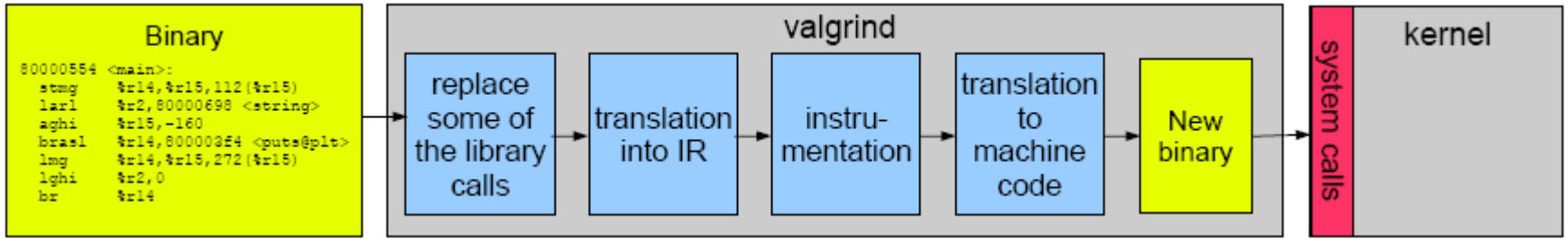
- **Valgrind System z support**
  Valgrind is a generic framework for creating dynamic analysis tools
  Valgrind is in essence a virtual machine using just-in-time (JIT) compilation techniques
  Valgrind can be used for memory debugging, memory leak detection, and profiling (e.g. cachegrind)

Valgrind System z support

- **valgrind –tool=memcheck [--leak-check=full] [--track-origins] <program>**
  Detects if your program accesses memory it shouldn't
  Detects dangerous uses of uninitialized values on a per-bit basis
  Detects leaked memory, double frees and mismatched frees

- **valgrind –tool=cachegrind**
  Profile cache usage, simulates instruction and data cache of the cpu
  Identifies the number of cache misses

- **valgrind –tool=massif**
  Profile heap usage, takes regular snapshots of program's heap
  Produces a graph showing heap usage over time



Where to Find More Information

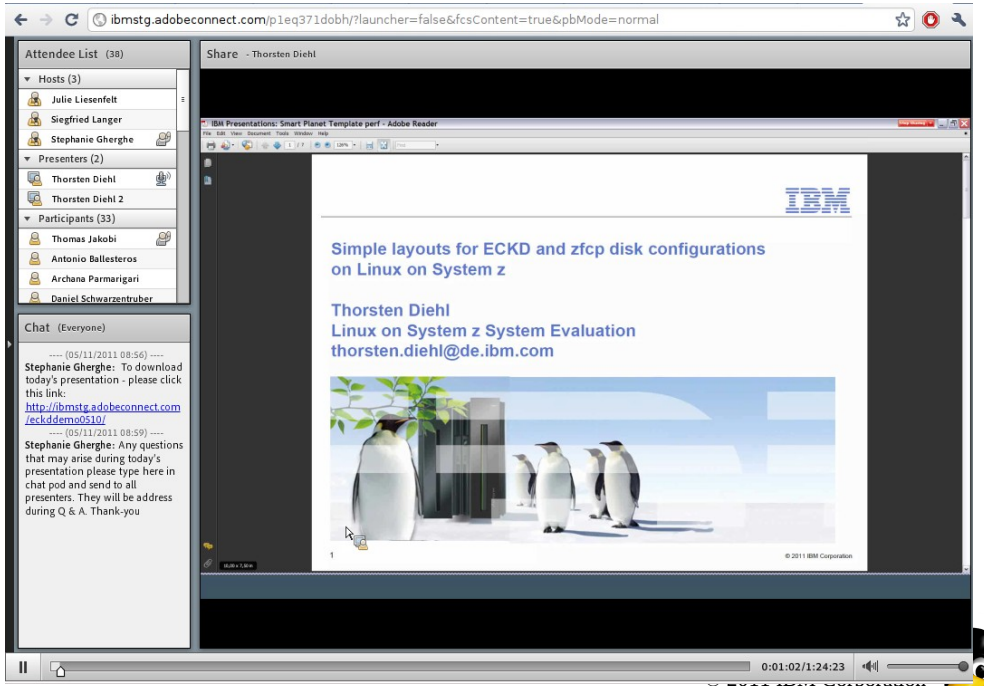
Live Virtual Classes for z/VM and Linux

http://www.vm.ibm.com/education/lvc/

IBM offers education on a variety of z/VM, Linux on System z and z/VSE topics in the form of 'Live Virtual Classes' (LVC) available on the Internet for Customers, Business Partners and IBMers

The day of the LVC broadcast, you can see the charts and listen to the speaker 'live'. In addition, you are able (and are encouraged) to ask questions of the speaker during a Q&A session following the prepared presentation.

* The day following each LVC, we post the the charts in PDF format.
* Shortly thereafter we provide a replay where you can read the charts, hear the recording and the Q's and A's in MP3 Format
*. You are welcome to read the charts or listen to the replay without registration when you can't participate 'live' or even if you wish to hear it all again.

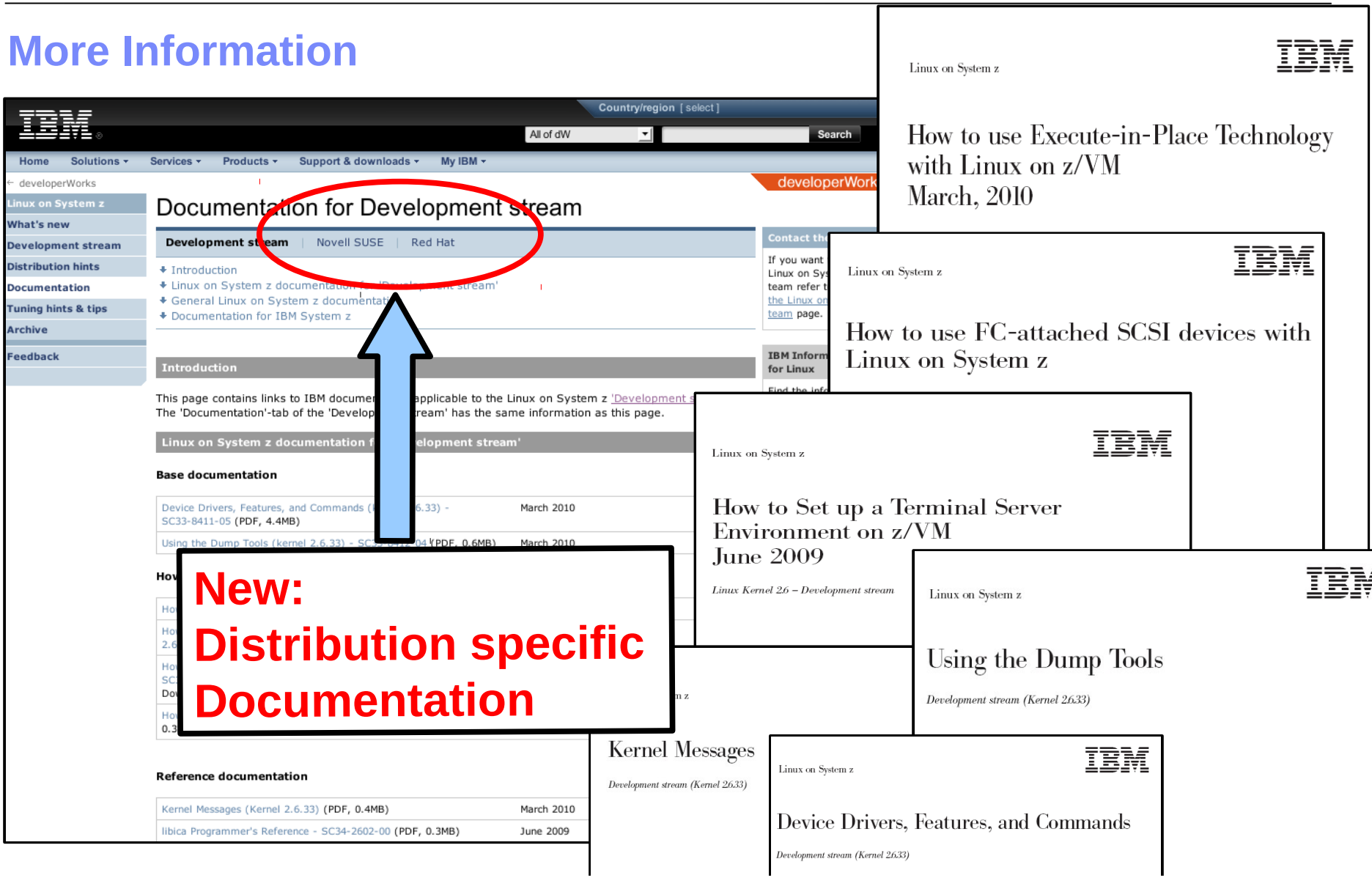


Live Virtual Classes for z/VM and Linux

- July 13, 2011
  - Backing Up and Restoring z/VM and Linux with IBM Solutions
- June 22, 2011
  - Automating Operations on z/VM and Linux with IBM Solutions
- May 10/11, 2011
  - Live Demo: Setup of simple and multipathed disk I/O configurations of ECKD and zfcp Volumes on Linux on System z
- April 6/7 2011
  - Problem Reporting and Analysis Linux on System z - How to survive a Linux critical situation !
- March 16/17, 2011
  - Red Hat Enterprise Server Performance Report for Linux on System z
- February 16/17, 2011
  - Lessons learned from putting Linux on System z in Production
- January 26, 2011
  - Best Practices for WebSphere Application Server on System z Linux
- December 15, 2010
  - What's new in RHEL 6 for Linux on System z
- November 17, 2010
  - Introduction to the new Linux on System z Terminal Server Using IUCV

[….]

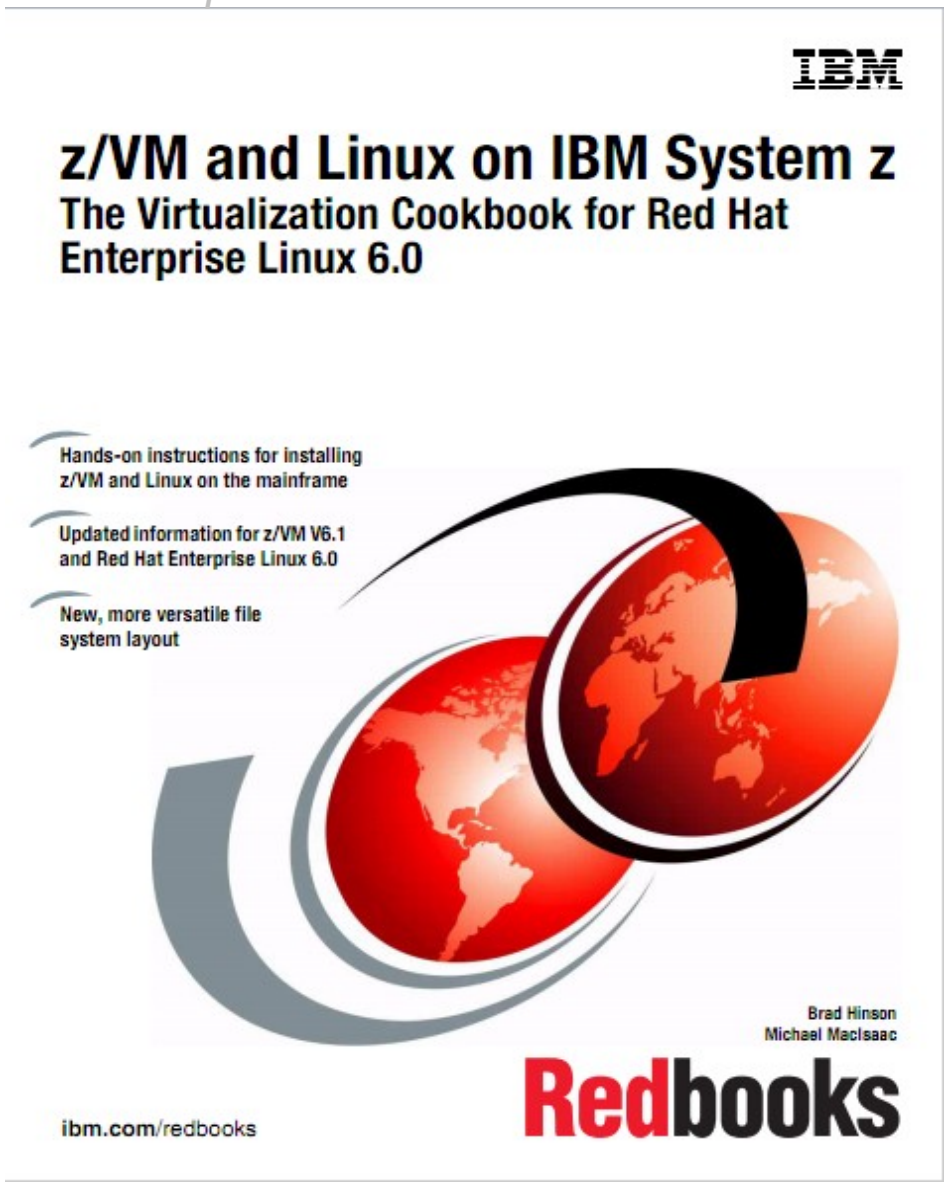
More Information

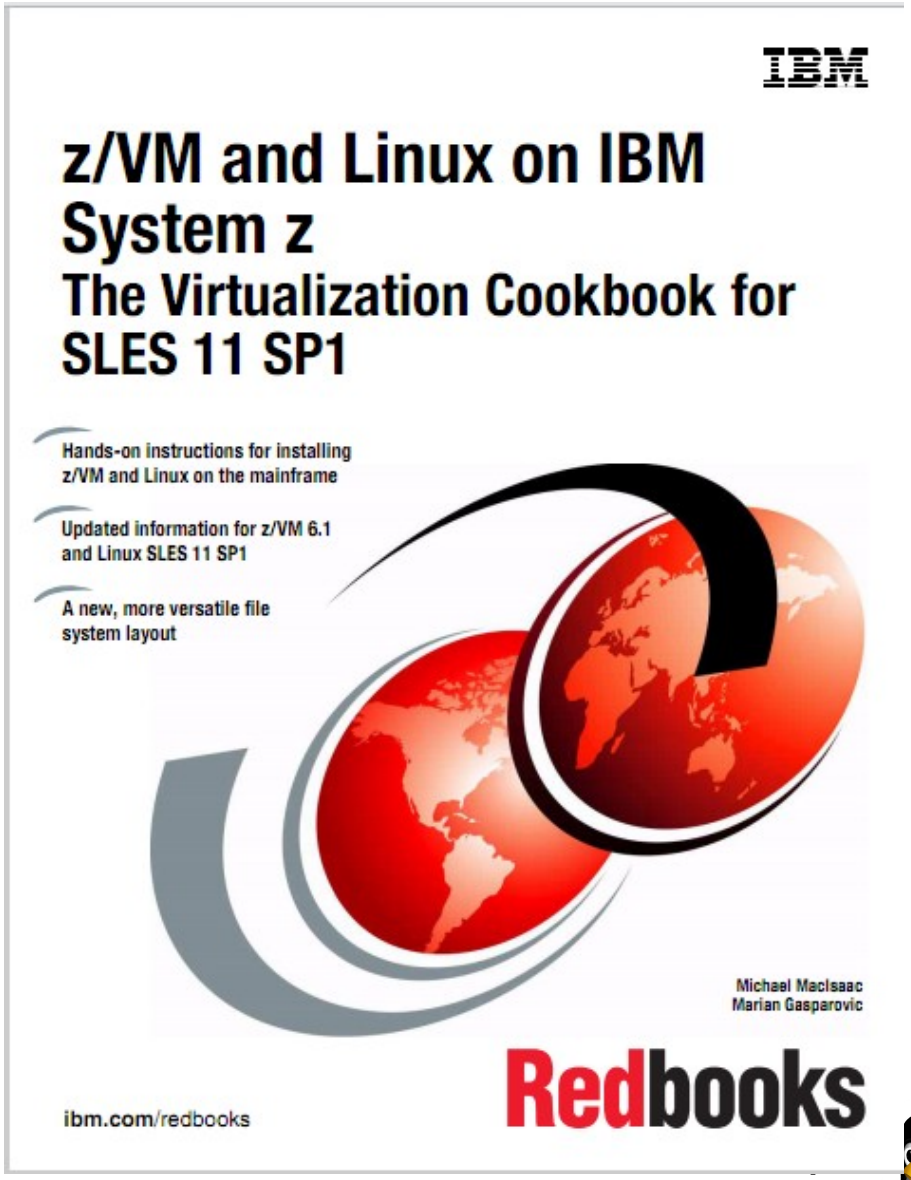


http://www-03.ibm.com/systems/z/os/linux/resources/index.html

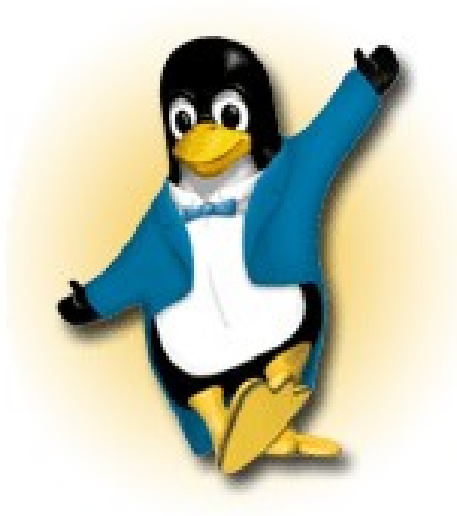
2 new Redbooks have just been released!

Visit http://www.redbooks.ibm.com





Questions?



Stefan Haberland

Linux on System z Development

IBM Deutschland Research & Development GmbH
Schönaicher Strasse 220
71032 Böblingen, Germany

Phone +49 (0)7031 - 161760
stefan.haberland@de.ibm.com

Trademarks & Disclaimer

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DB2, e-business logo, ESCON, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, System Storage, System z9, VM/ESA, VSE/ESA, WebSphere, xSeries, z/OS, zSeries, z/VM.

The following are trademarks or registered trademarks of other companies

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries. LINUX is a registered trademark of Linux Torvalds in the United States and other countries. UNIX is a registered trademark of The Open Group in the United States and other countries. Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation. SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC. Intel is a registered trademark of Intel Corporation. * 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. References in this document to IBM products or services do not imply that IBM intends to make them available in every country. Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use. The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice. Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.