SCSI over FCP for Linux on System z Roundup

Dr. Holger Smolinski
IBM Germany Research & Development GmbH

2010-08-03
9222

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market. Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml

The following are trademarks or registered trademarks of other companies.

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.

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

Notes:
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.

Agenda

- Introduction to FCP on System z
- FCP with Linux on System z
- IPL over FCP
- SCSI dump
- Multipathing
  - Multipathing for root file system
- NPIV

FCP in a Nutshell

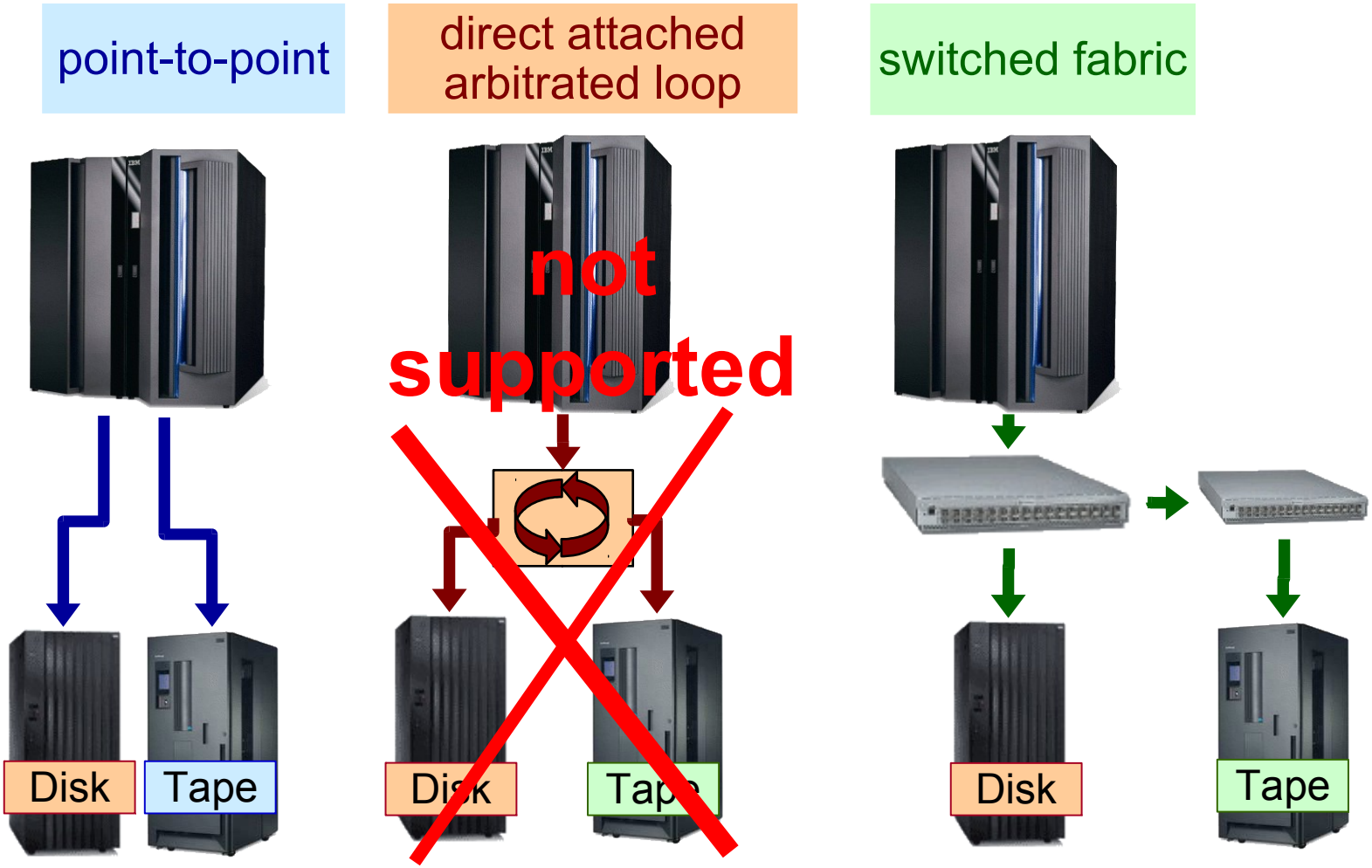
- Storage Area Networks (SANs) are specialized networks dedicated to the transport of mass storage data
- Today the most common SAN technology used is Fibre Channel Protocol (FCP)
- With this technology the SCSI protocol is used to address and transfer raw data between the servers and the storage device
- Each server is equipped with a least one adapter which provides the physical connection to the SAN
- For System z any supported FCP adapter, such as FICON Express or FICON Express2 can be used for this purpose.
- The Fibre Channel (FC) standard was developed by the National Committee of Information Technology Standards (NCITS)

Why FCP?

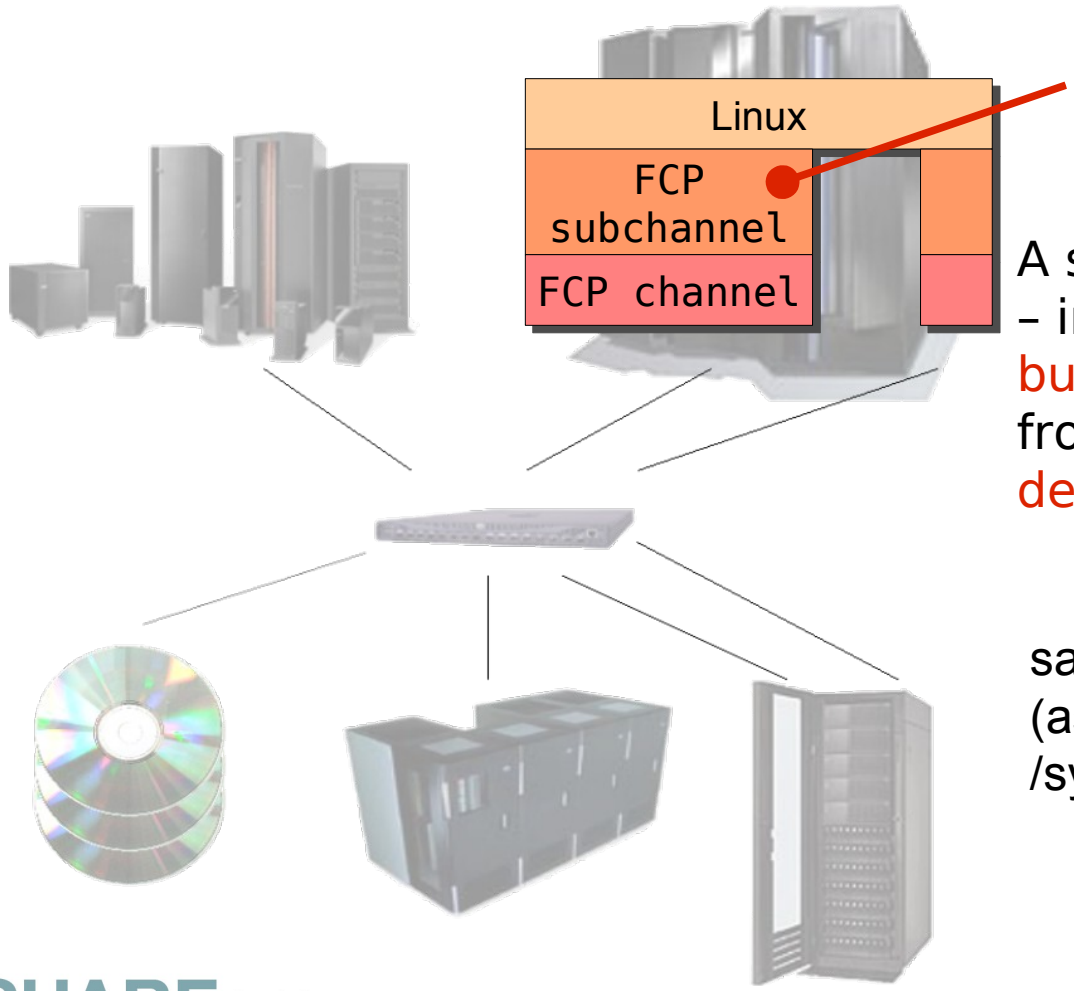
- Performance advantages
  - concurrent I/O to same device
  - no ECKD emulation/ no FICON protocol
- No disk size restrictions
- Up to 15 partitions (16 minor numbers per device)
- SCSI disks do not waste disk space
  (no low-level formatting)

- System z integration in existing FC SANs
- Use of existing FICON infrastructure
  - FICON Express adapter cards
  - FC switches / Cabling
  - Storage subsystems
- Dynamic configuration
  - Adding of new storage subsystems possible without IOCDS change
- Does NOT require more CPU than FICON

SAN topologies and System z



FCP channel and subchannel

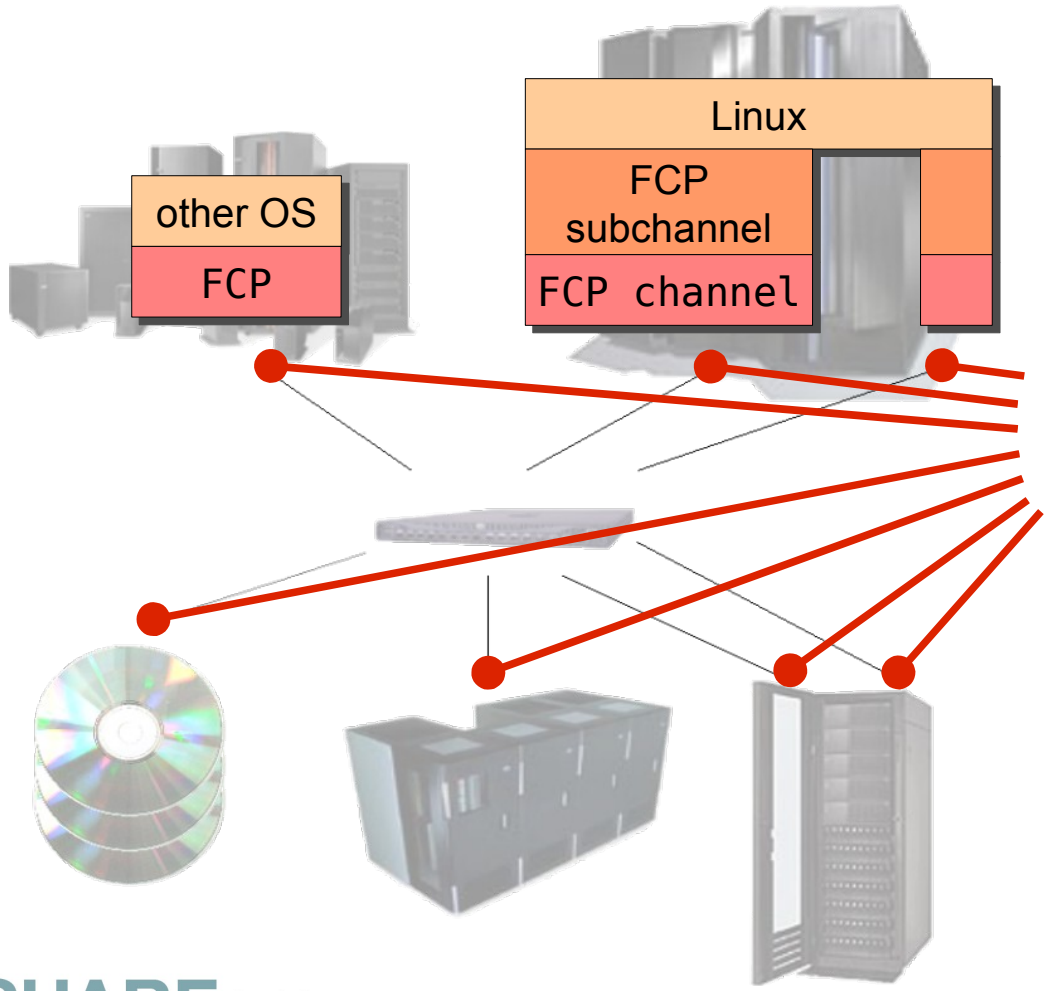


Linux connects through FCP subchannels to FCP attached storage.

A subchannel is identified – in Linux - by its bus identifier which is derived from the subchannel's device number.

sample FCP subchannel (as seen in Linux):
/sys/bus/ccw/drivers/zfcp/0.0.1900

World Wide Port Names (WWPNs)



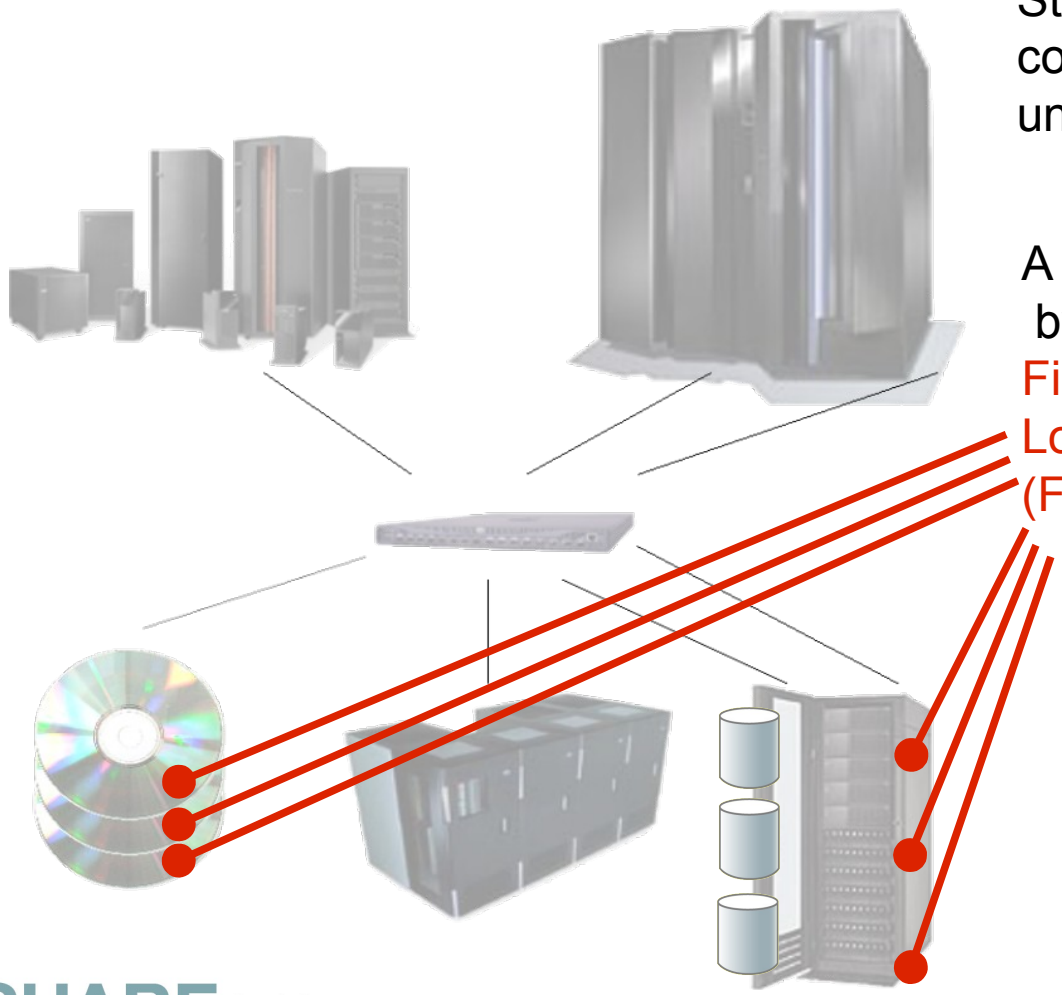
Storage devices and servers attach through Fibre Channel ports (called N_Ports).

An N_Port is identified by its World-Wide Port Name (WWPN).

For redundancy, servers or storage may attach through several N_Ports.

sample WWPN:
0x500507630300104

Logical Unit Numbers (LUNs)

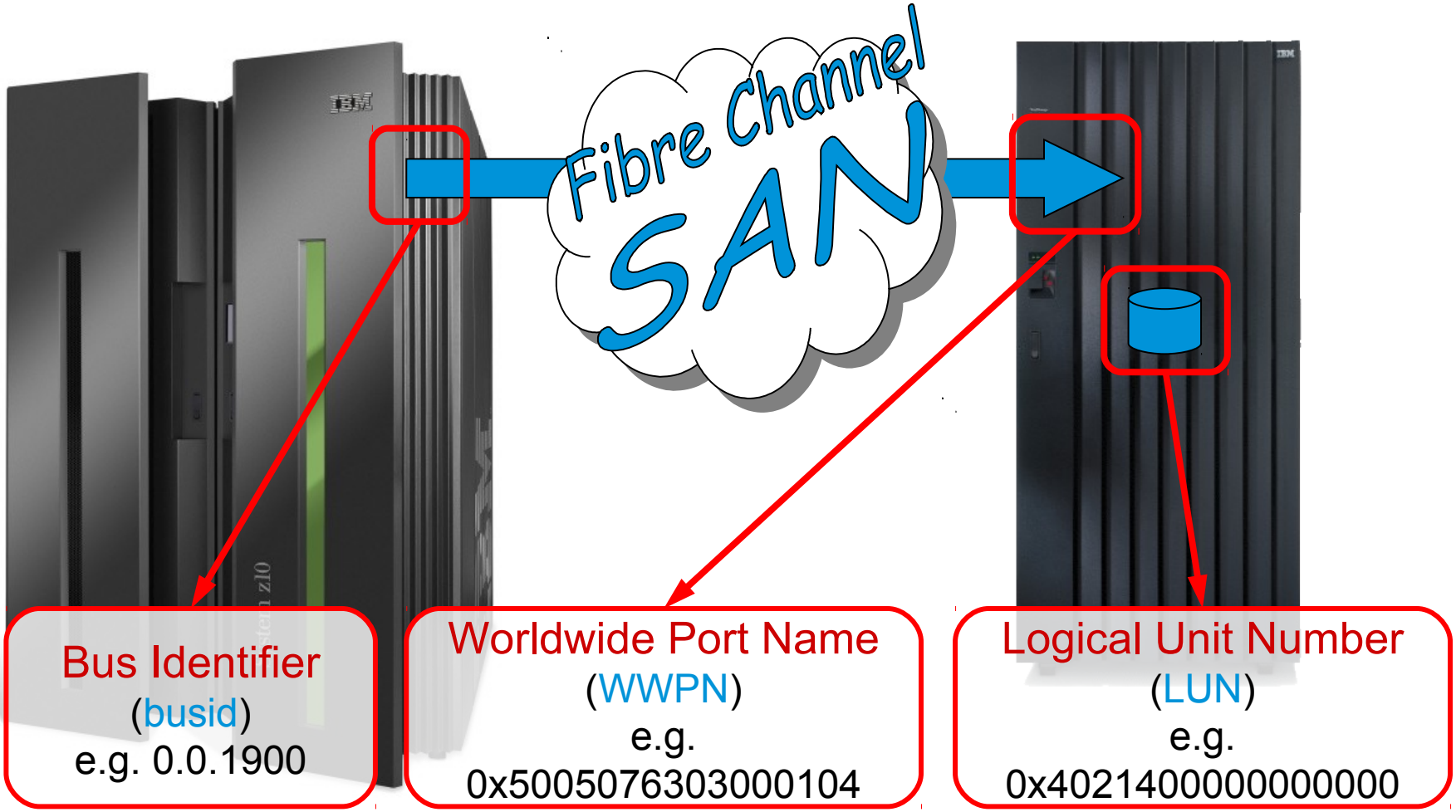


Storage devices usually comprise many logical units (volumes, tape drives, ...).

A logical unit is identified by its Fibre Channel Protocol Logical Unit Number (FCP LUN).

sample FCP LUN:
0x**40**2**1**4**00**00000000
Beware of LUN translation!

Navigating in a SAN



SCSI compared to Channel I/O

- SCSI / FCP
  - adapter defined in System z I/O configuration
  - Ports and LUNs attachment handled in Operating Systems
  - Multipathing handled in Operating System
  - No disk size restrictions for SCSI disks
  - Additional configuration outside System z necessary
    - Zoning in the SAN fabric
    - LUN masking on the storage server
- Channel I/O
  - device defined in System z I/O configuration
  - Ports attachment handled in System z I/O config
  - Multipathing handled in System z firmware
  - Disk size restrictions to Mod 54 / Mod 224
  - Switch configuration via System z I/O config

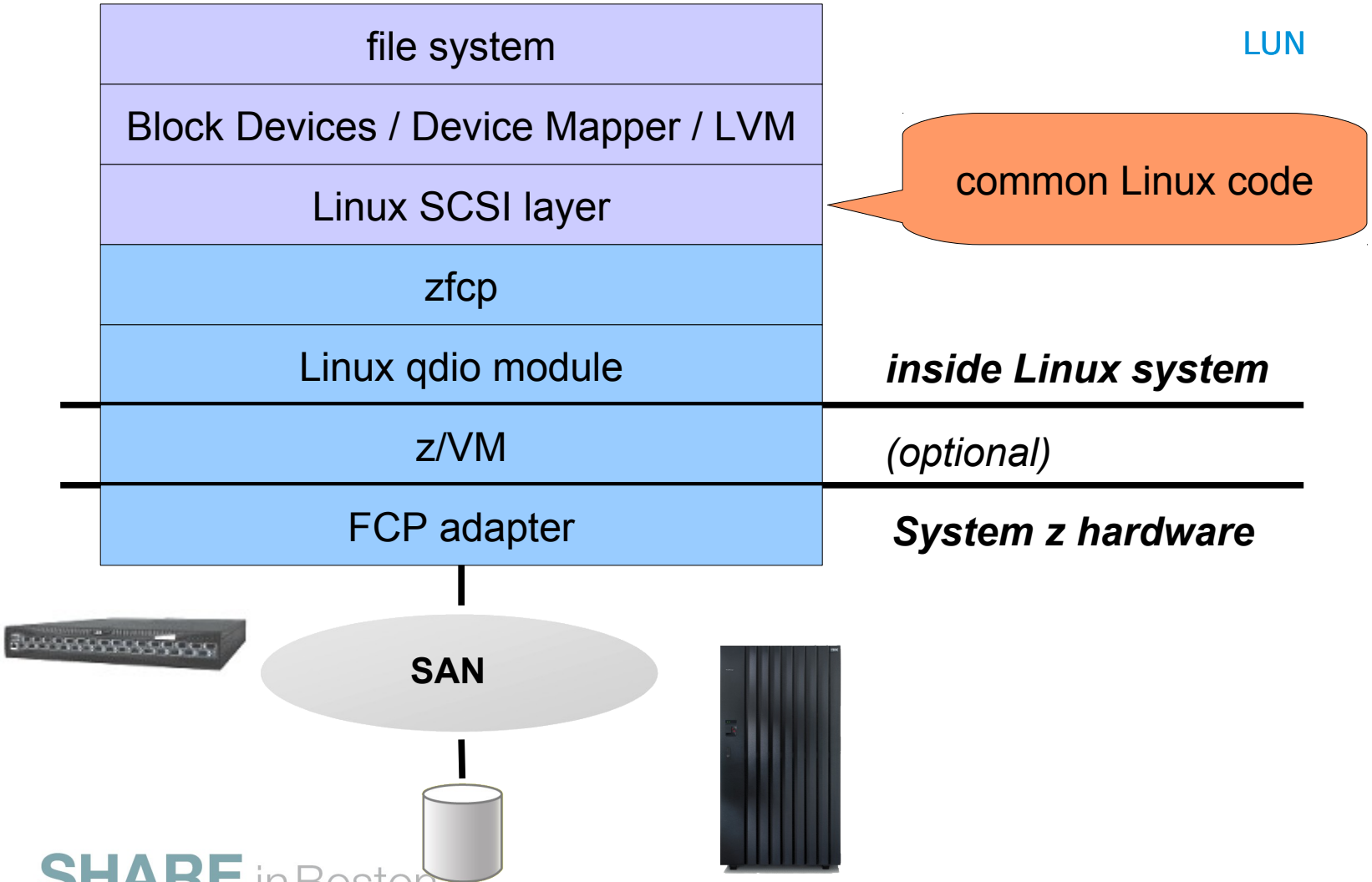
zfcp: Getting started

- Configure a Fibre Channel host adapter within the mainframe (I/O Definition File).
- Configure zoning for the Fibre Channel host adapter to gain access to desired target ports within a SAN.
  - Segmentation of a switched fabric is achieved though zoning. It can be used to partition off certain portions of the switched fabric, allowing only the members of a zone to communicate with that zone.
- Configure LUN masking for the Fibre Channel host adapter at the target device to gain access to desired LUNs.
  - A LUN represents a portion of a controller, such as a disk device. With the use of LUNs, a controller can be logically divided into independent partitions. Access to this LUNs can be restricted to distinctive WWPNs as part of the controller configuration
- In Linux, configure target ports and LUNs of the SCSI device at the target port for use of zfcp.
- Note: If the Fibre Channel host adapter is directly attached to a target device (point-to-point connection), step 2 is not needed.

Hardware: Define FCP adapter in IOCDS

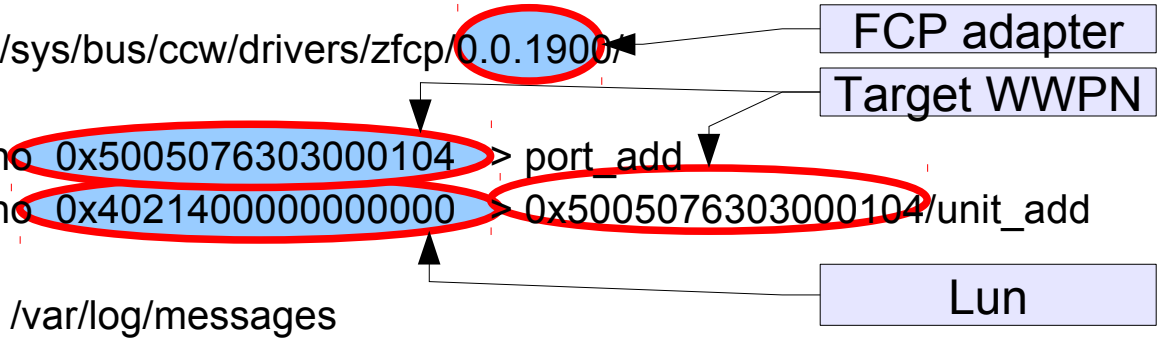
```
CHPID PATH=(CSS(0,1,2,3),51),SHARED,                                  *
      NOTPART=((CSS(1),(TRX1),(=)),(CSS(3),(TRX2,T29CFA),(=)))*
      ,PCHID=1C3,TYPE=FCP
CNTLUNIT CUNUMBR=3D00,                                                *
      PATH=((CSS(0),51),(CSS(1),51),(CSS(2),51),(CSS(3),51)), *
      UNIT=FCP
IODEVICE ADDRESS=(3D00,001),CUNUMBR=(3D00),UNIT=FCP
IODEVICE ADDRESS=(3D01,007),CUNUMBR=(3D00),                           *
      PARTITION=((CSS(0),T29LP11,T29LP12,T29LP13,T29LP14,T29LP*
      15),(CSS(1),T29LP26,T29LP27,T29LP29,T29LP30),(CSS(2),T29*
      LP41,T29LP42,T29LP43,T29LP44,T29LP45),(CSS(3),T29LP56,T2*
      9LP57,T29LP58,T29LP59,T29LP60)),UNIT=FCP
IODEVICE ADDRESS=(3D08,056),CUNUMBR=(3D00),                           *
      PARTITION=((CSS(0),T29LP15),(CSS(1),T29LP30),(CSS(2),T29*
      LP45),(CSS(3),T29LP60)),UNIT=FCP
```

I/O stack for SCSI and Linux



zfcp: Configuration

```
# chccwdev -e 0.0.1900

# cat /var/log/messages
zfcp: The adapter 0.0.1900 reported the following characteristics:
WWNN 0x5005076400c3c03f, WWPN 0x5005076401a28753, S_ID 0x00687700,
adapter version 0x4, LIC version 0xb02, FC link speed 4 Gb/s
zfcp: Switched fabric fibrechannel network detected at adapter 0.0.1900.

# cd /sys/bus/ccw/drivers/zfcp/0.0.1900/

# echo 0x5005076303000104 > port_add
# echo 0x4021400000000000 > 0x5005076303000104/unit_add

# cat /var/log/messages
zfcp: Switched fabric fibrechannel network detected at adapter 0.0.1900.
  Vendor: IBM       Model: 2107900          Rev: 1.50
  Type:   Direct-Access                     ANSI SCSI revision: 05
scsi 0:0:0:1: Attached scsi generic sg0 type 0
SCSI device sda: 10485760 512-byte hdwr sectors (5369 MB) .....
```



Not required when using SLES11 due to auto port scanning

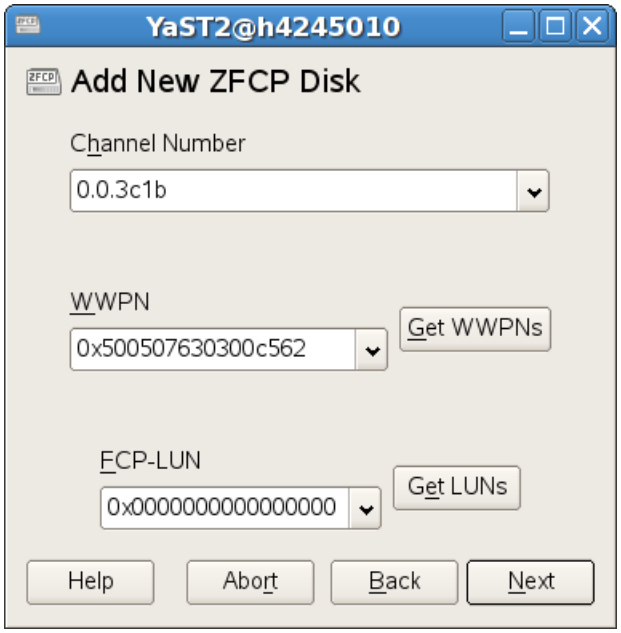
zfcp: Configuration (cont'd)

```
# lszfcp -D
0.0.1900/0x5005076303000104/0x4021400000000000 0:0:0:1
# lsscsi
[0:0:0:1]    disk    IBM      2107900          1.50  /dev/sda
```

Manually disabling a scsi device from current configuration

```
# echo 1 > /sys/bus/scsi/devices/0:0:0:1/delete
# echo 0x4021400000000000
  > /sys/bus/ccw/drivers/zfcp/0.0.1900/0x5005076303000104/unit_remove
# echo 0x5005076303000104
  > /sys/bus/ccw/drivers/zfcp/0.0.1900/port_remove
# chccwdev -d 1900
```

SLES: GUI-Setup



- zfcp dialog in YaST simplifies setup of SAN attached devices
- Auto detects available FCP subchannels, WWPNs, and LUNs
- *copy&paste* WWPNs and FCP_LUNs from configuration file obtained from SAN management tools or administrator

- alternatively on command line
  - SLES 10: /etc/sysconfig/hardware/hwcfg-zfcp-bus-ccw-0.0.*
  - SLES 11: zfcp_{host|disk}_configure → /etc/udev/rules.d/51-zfcp-0.0.*.rules

RHEL: GUI-Setup

- Ignore subsequent complaints in case of DASD-less system.
- GUI only available during installation. Later define FCP devices in /etc/zfcp.conf for permanent addition.


```
# cat /etc/zfcp.conf
0.0.170e 0x5005076300c18154 0x4010402000000000
# cat /etc/modprobe.conf
[...]
alias scsi_hostadapter zfcp
# /sbin/zfcpconf.sh
```

zfcp: toolchain

- lsscsi
  - Uses information in sysfs to list scsi devices (or hosts) currently attached to the system

```
[0:0:0:0]disk    IBM      2107900         1.50  /dev/sda
```

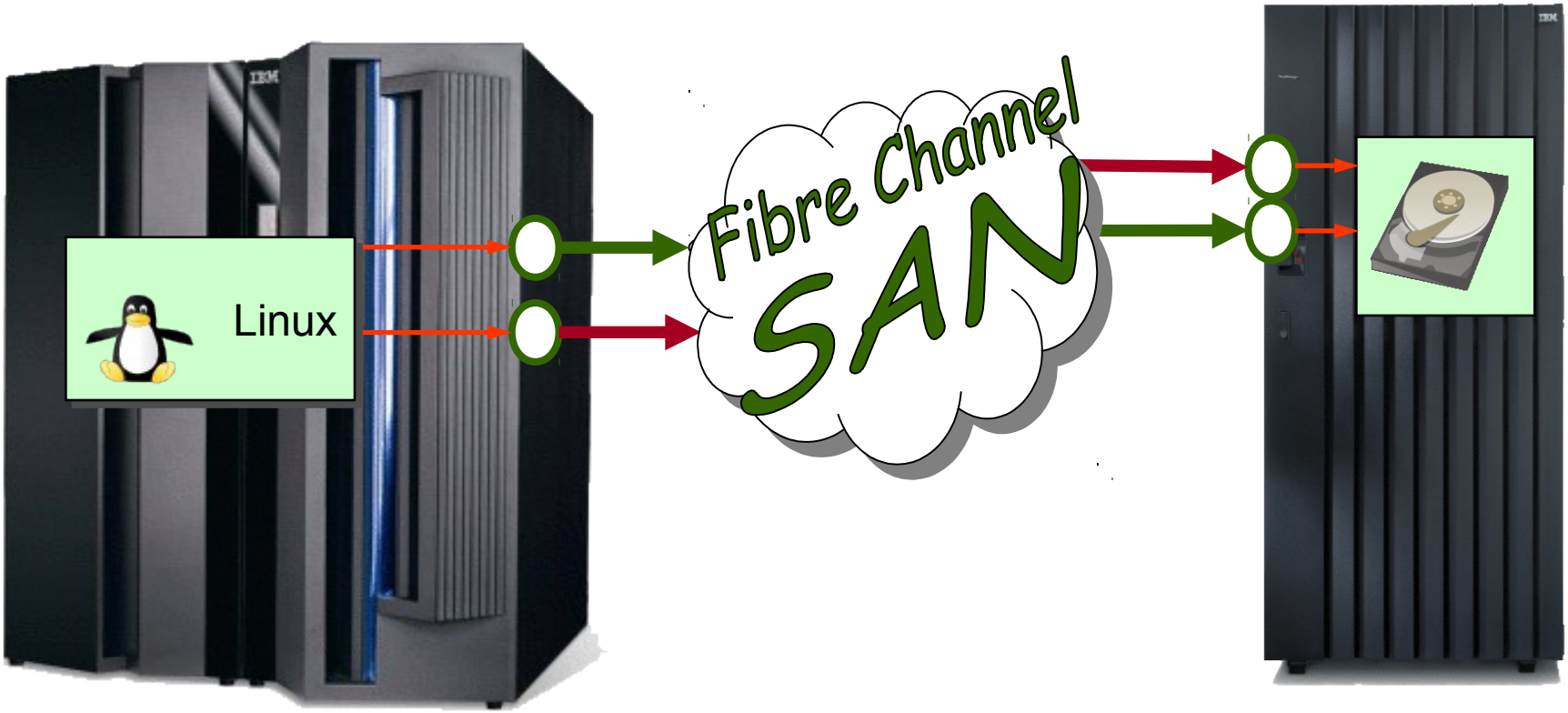
- lszfcp
  - lszfcp provides information contained in sysfs about zfcp adapters, ports and units and its associated scsi_hosts, fc_hosts, fc_remote_ports and scsi_devices.
  - The default is to list busids of all zfcp adapters and their corresponding SCSI host name

```
# lszfcp -H                                          shows information about hosts
0.0.170e host0
# lszfcp -P                                          shows information about ports
0.0.170e/0x500507630300c562 rport-0:0-0
# lszfcp -D                                          shows information about SCSI devices
0.0.170e/0x500507630300c562/0x4010402000000000 0:0:0:0
```

zfcp: SCSI Disk Usage

```
# fdisk /dev/sda

Command (m for help): p

Disk /dev/sda: 5368 MB, 5368709120 bytes
166 heads, 62 sectors/track, 1018 cylinders
Units = cylinders of 10292 * 512 = 5269504 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1               1        1018     5238597   83  Linux
# mke2fs -j /dev/sda1
```

FCP Multipathing



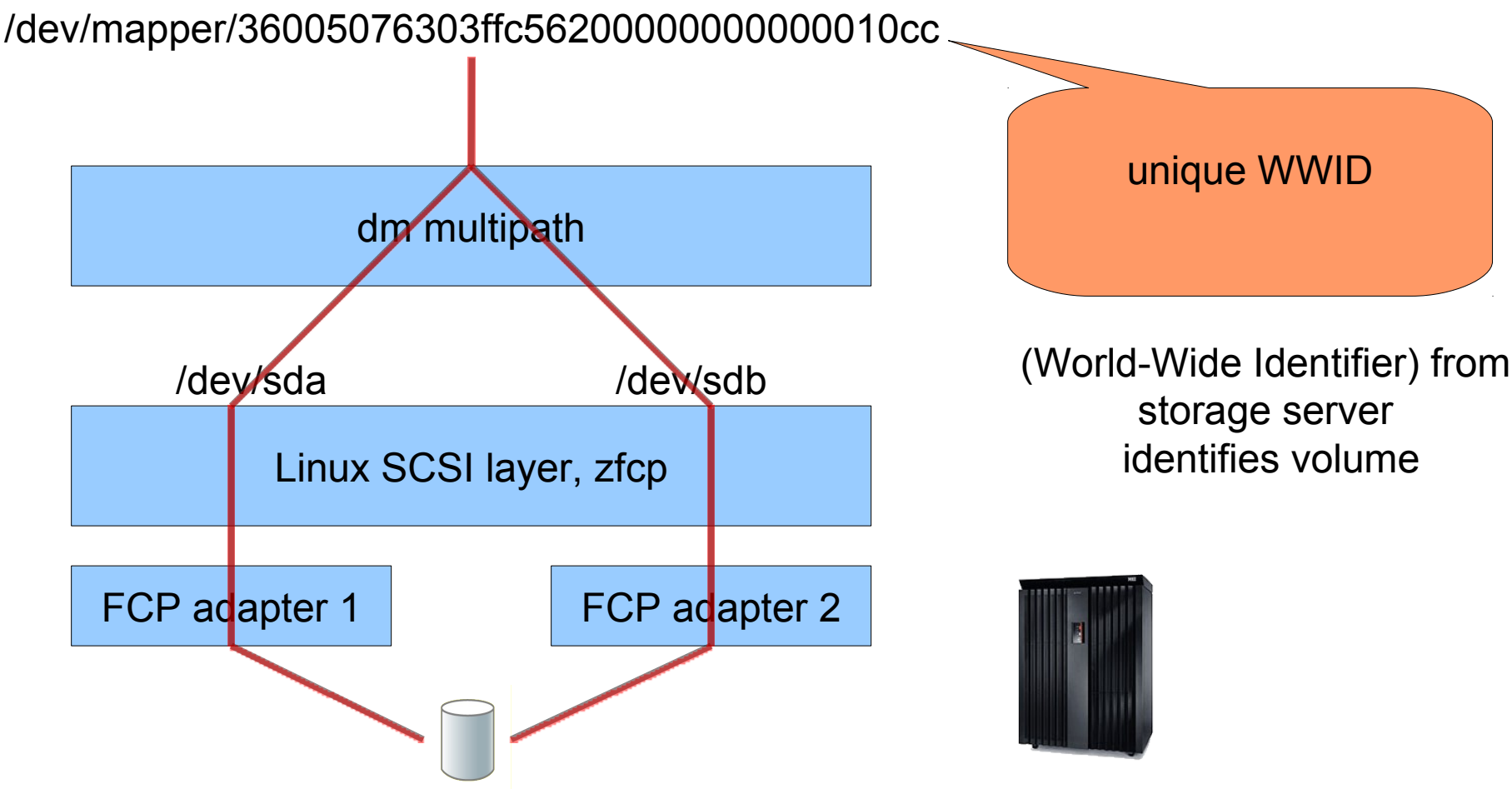
2 paths to disk through independent FCP adapters and independent controllers.

Multipathing for disks

- Multipathing is mandatory for FCP-attached SCSI disks
- In general there are two reasons for establishing multiple paths to a device
  - failover and failback capabilities for high availability
    - each controller or node might be unavailable
      - *hardware maintenance*
      - *microcode updates*
      - *internal resets*
  - load balancing for high performance (throughput)
    - spread I/O load across available paths
- device-mapper (kernel) multipathing
  - Included with standard distributions (SLES and RHEL)
  - supports more than 2 paths
- multipathd daemon
  - reads configuration and establishes setup
    - identifies and groups available paths automatically
  - reestablishes paths (failback)
    - checks paths periodically
- multipath tool that allows the user to configure and manage multipathed devices.
- kpartx for partitions on multipath devices

Multipathing for disks - Linux device mapper

The device mapper creates one block device for the LUN /dev/mapper/xxx



zfcp setup for multipathing

- Have multiple paths to one disk
- Avoid shared components in different paths

```
# cd /sys/bus/ccw/drivers/zfcp/
# echo 1 > 0.0.3c00/online
# echo 1 > 0.0.3d00/online
# echo 0x500507630313c562 > 0.0.3c00/port_add
# echo 0x500507630303c562 > 0.0.3d00/port_add
# echo 0x401040cc00000000 > 0.0.3c00/0x500507630313c562/unit_add
# echo 0x401040cc00000000 > 0.0.3d00/0x500507630303c562/unit_add
```

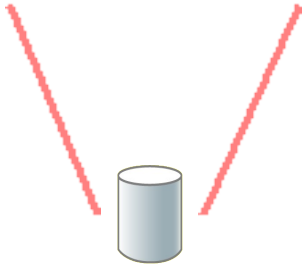
usually same FCP LUN (check on storage server)

different adapters and different ports to avoid single points of failures

zfcp setup for multipathing (cont'd)

- zfcp and SCSI report each path as device
- multipathing happens on higher layer

```
# lsscsi
[0:0:0:0]    disk    IBM      2107900          2.27  /dev/sda
[1:0:1:0]    disk    IBM      2107900          2.27  /dev/sdb
# lszfcp -D
0.0.3c00/0x500507630313c562/0x401040cc00000000 0:0:0:0
0.0.3d00/0x500507630303c562/0x401040cc00000000 1:0:1:0
```



Multipathing for disks - SLES 10 and SLES 11

- add all paths to system
  - YaST or edit /etc/sysconfig/hardware/hwcfg-zfcp-* (SLES 10)
  - hwup zfcp-bus-ccw-0.0.3c00
  - zfcp_{host|disk}_configure (SLES 11)
- cp /usr/share/doc/packages/multipath-tools/multipath.conf.synthetic /etc/multipath.conf
  - Make sure there is an appropriate device entry for your SAN
- enable device scanning and multipathd
  - chkconfig multipathd on
  - chkconfig boot.multipath on
- reboot or manually start multipath scripts
  - /etc/init.d/boot.multipath start
  - /etc/init.d/multipath start

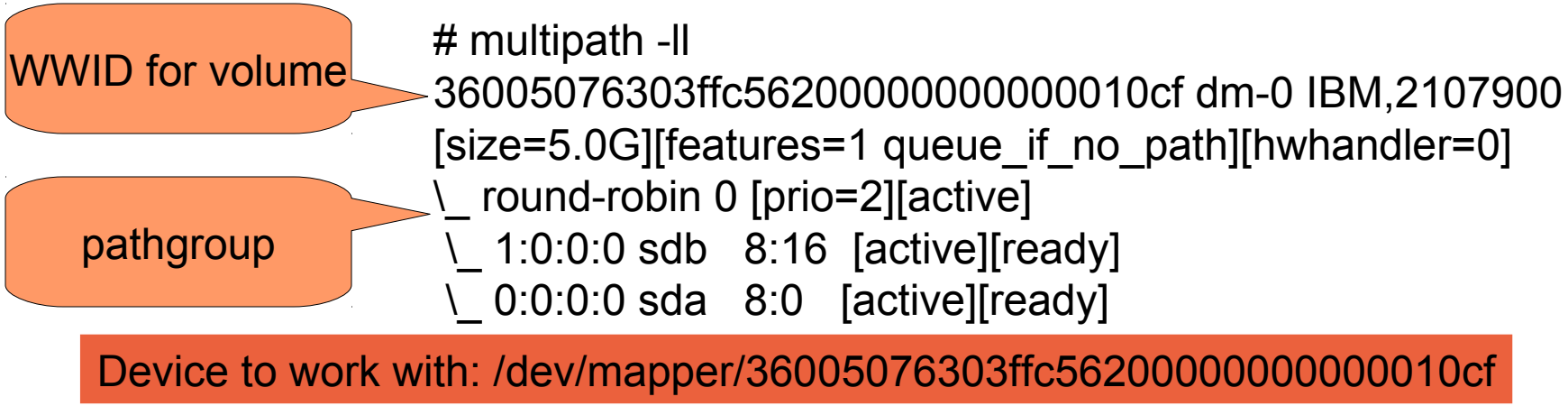
Multipathing for disks - RHEL5

- attach all paths to system
  - /etc/zfcp.conf
  - /sbin/zfcpconf.sh or reboot
- Adjust /etc/multipath.conf
- chkconfig --add multipathd
- /etc/init.d/multipathd start

- user_friendly_names and aliases
  - /dev/mapper/mpath0 instead of /dev/mapper/36005076303ffc56200000000000010ce
- But: WWID is unique, alias maybe not
  - mapping depends on config file
- Recommendation: Use WWIDs

```
# cat /etc/multipath.conf
…
blacklist {
        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
        devnode "^hd[a-z][[0-9]*]"
        devnode "^cciss!c[0-9]d[0-9]*[p[0-9]*]"
        devnode "^dasd[a-z]+[0-9]*"
}
…
```

DM multipathing status



- No config file necessary to get started
- Defaults are good for availability
  - Storage Controller specific settings used as defaults
  - can be overwritten e.g. for load balancing

Multipathing - policies

- failover
  - First path is used as long as it is available – no failback
  - Recommended for DS8000
  - consider load balancing during configuration
- multibus / round robin
  - All paths are used alternatively at same priority
  - Round robin parameter adjustable
  - May imply congestion on selected paths.
- group_by_prio
  - A priority_callout is used to determine priority of each path
  - Default for DS6000 preferred pathing (via ALUA callout)
  - Can be (ab)used for load distribution

Root Filesystem on Multipath

Required for root filesystem on SCSI disk

Multipath setup has to be available early from initrd

Starting with RHEL 5.2 and SLES 11
installers support install on multipath device

Partly requires special boot flags in parm file on IPL of installer
→ please see distro documentation on installation and multipath storage

Issues:

For older distros, where installers don't support install on multipath device:

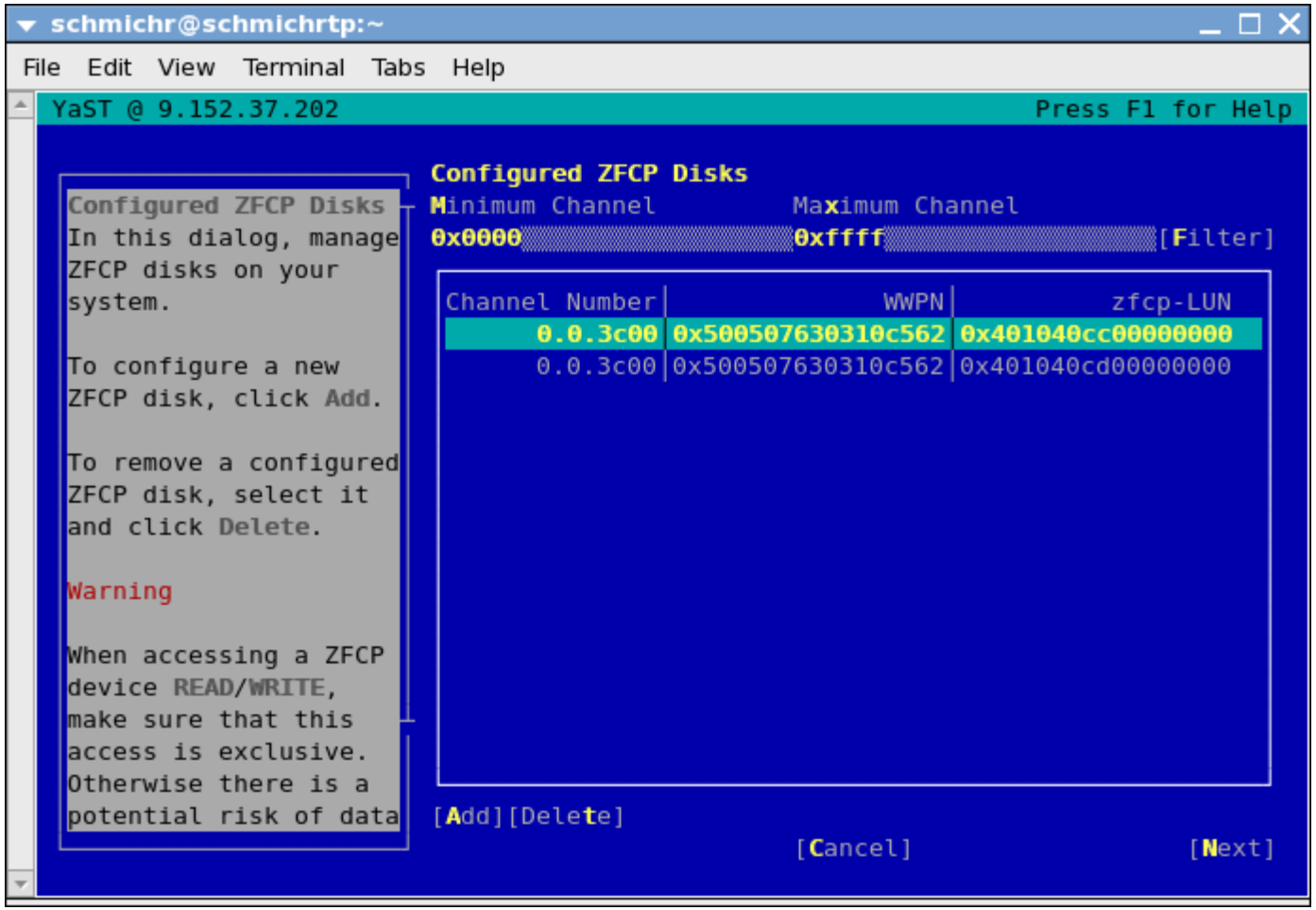
install on single path and change setup later

zipl does not work on multipath device

Use additional single-path device for /boot
(SCSI or DASD)

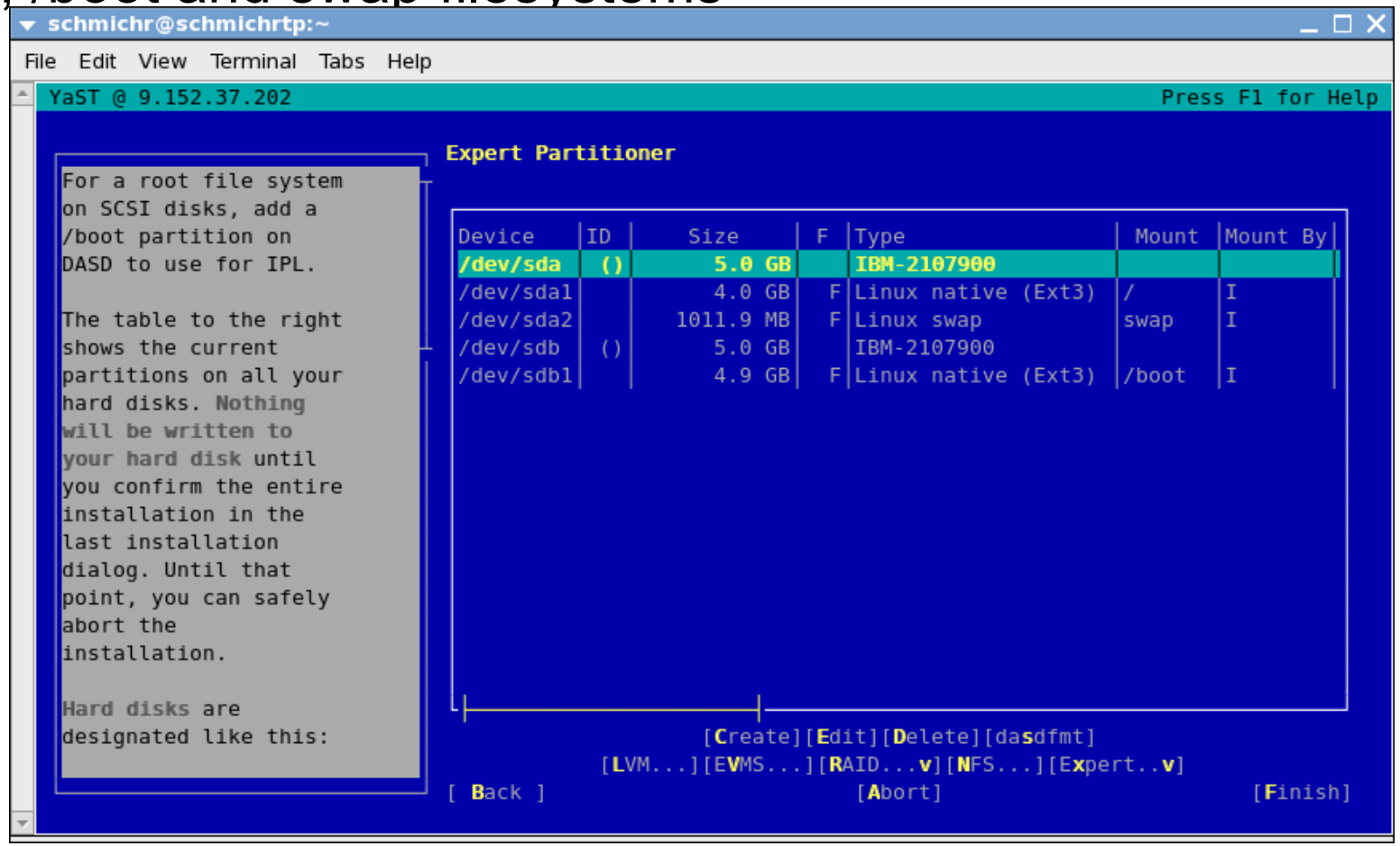
Example: Install SLES10 on multipath root

2 devices for / and /boot



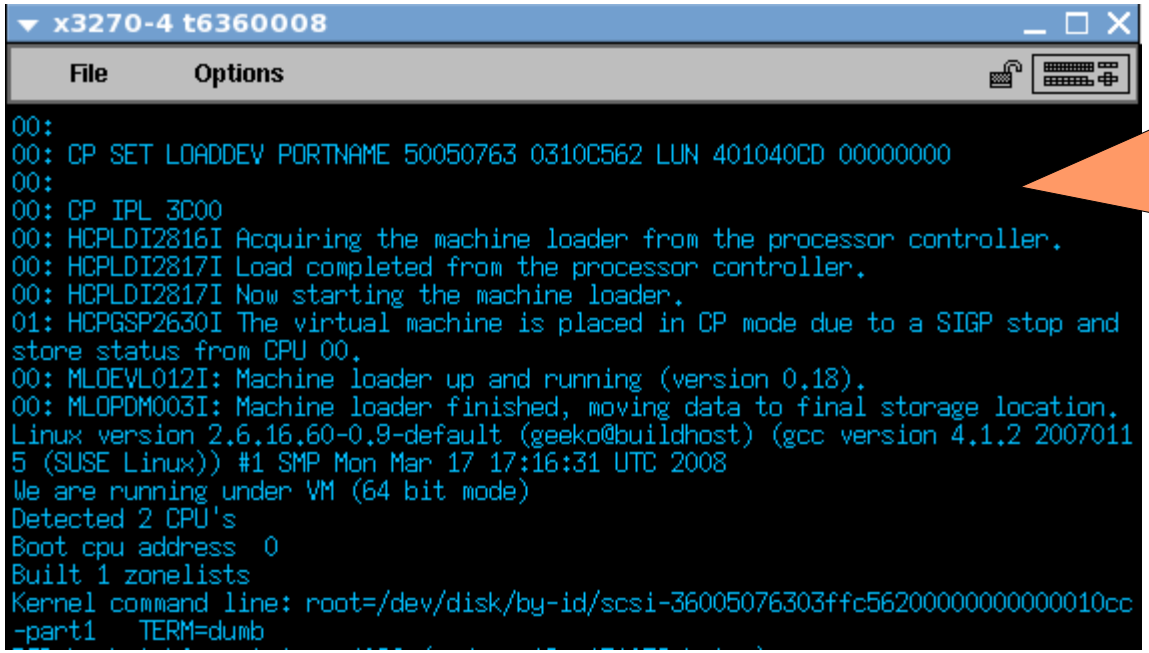
Example: Install SLES10 on multipath root

/, /boot and swap filesystems



Example: Install SLES10 on multipath root

initial boot via disk for /boot



set loaddev for port and lun,

ipl from FCP adapter

Example: Install SLES10 on multipath root

system with single path setup after installation

dedicated disk for /boot

```
# mount
/dev/sda1 on / type ext3 (rw,acl,user_xattr)
/dev/sdb1 on /boot type ext3 (rw,acl,user_xattr)
[...]

# lsscsi
[0:0:0:1087127568]disk    IBM      2107900          2.27  /dev/sda
[0:0:0:1087193104]disk    IBM      2107900          2.27  /dev/sdb

# lszfcp -D
0.0.3c00/0x500507630310c562/0x401040cc00000000 0:0:0:1087127568
0.0.3c00/0x500507630310c562/0x401040cd00000000 0:0:0:1087193104
```

Example: Install SLES10 on multipath root

add second path for root filesystem

create /etc/sysconfig/hardware/hwcfg-zfcp-bus-ccw-0.0.3d00

[...]
ZFCP_LUNS="
0x500507630310c562:0x401040cc00000000"

attach second path (trigger hwup scripts or reboot)

```
# chccwdev -d 3d00
Setting device 0.0.3d00 offline
Done
# modprobe vmcp
# vmcp det 3d00
FCP 3D00 DETACHED
# vmcp att 3d00 \*
FCP 3D00 ATTACHED TO T6360008 3D00
```

Example: Install SLES10 on multipath root

enable multipath services for next reboot

```
# chkconfig --add boot.multipath
boot.multipath          0:off  1:off  2:off  3:off  4:off  5:off  6:off  B:on
# chkconfig --add multipathd
multipathd              0:off  1:off  2:off  3:on   4:off  5:on   6:off
```

make system use new multipath device

adjust root and swap in /etc/fstab, but don't touch /boot

```
/dev/mapper/36005076303ffc56200000000000010cc-part1 /                [...]
/dev/mapper/36005076303ffc56200000000000010cc-part2 swap             [...]
/dev/disk/by-id/scsi-36005076303ffc56200000000000010cd-part1 /boot   [...]
```

change kernel parameters line in /etc/zipl.conf

```
    parameters = "root=/dev/mapper/36005076303ffc56200000000000010cc-part1 TERM=dumb"
[...]
    parameters = "root=/dev/mapper/36005076303ffc56200000000000010cc-part1 TERM=dumb 3"
```

Example: Install SLES10 on multipath root

switch boot process to use multipath device for root

create new initrd with multipath tools
mkinitrd -f mpath

don't forget to run zipl
zipl

reboot

multipath -ll
36005076303ffc56200000000000010cc dm-0 IBM,2107900
[size=5.0G][features=1 queue_if_no_path][hwhandler=0]
_ round-robin 0 [prio=2][active]
 _ 1:0:0:1087127568 sdc 8:32 [active][ready]
 _ 0:0:0:1087127568 sda 8:0   [active][ready]

t6360008:~ # mount
/dev/mapper/36005076303ffc56200000000000010cc-part1 on / type ext3
(rw,acl,user_xattr)
[...]

Example: Install SLES10 on multipath root

```
cp q loaddev
PORTNAME 50050763 0310C562    LUN  401040CD 00000000    BOOTPROG 0
BR_LBA   00000000 00000000

cp ipl 3c00
00: HCPLDI2816I Acquiring the machine loader from the processor controller.
00: HCPLDI2817I Load completed from the processor controller.
00: HCPLDI2817I Now starting the machine loader.
01: HCPGSP2630I The virtual machine is placed in CP mode due to a SIGP stop and
store status from CPU 00.
00: MLOEVL012I: Machine loader up and running (version 0.18).
00: MLOPDM003I: Machine loader finished, moving data to final storage location.
Linux version 2.6.16.60-0.9-default (geeko@buildhost) (gcc version 4.1.2 2007011
5 (SUSE Linux)) #1 SMP Mon Mar 17 17:16:31 UTC 2008
We are running under VM (64 bit mode)
Detected 2 CPU's
Boot cpu address  0
Built 1 zonelists
Kernel command line: root=/dev/mapper/36005076303ffc5620000000000010cc-part1
TERM=dumb
```

disk for /boot, used for zipl

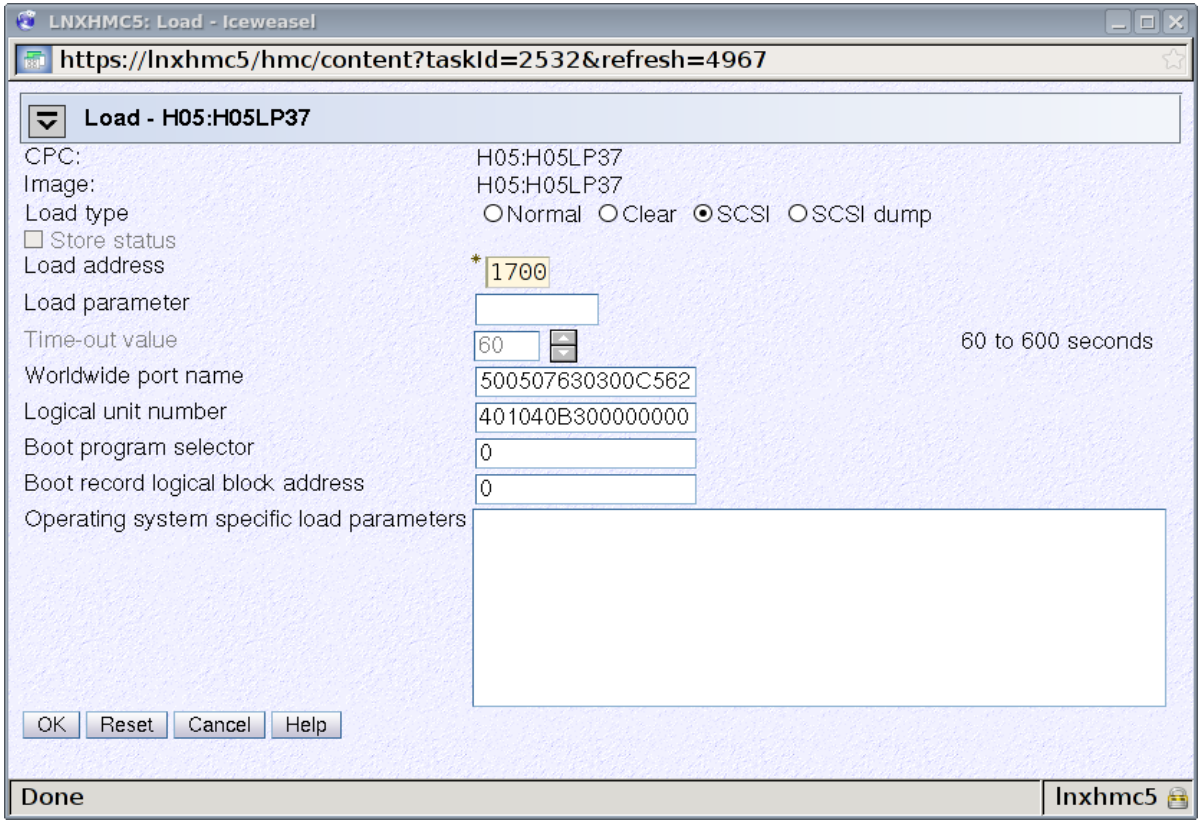
multipath device for /

```
Setup multipath devices: ok.
Waiting for device /dev/mapper/36005076303ffc5620000000000010cc-part1 to appear
:  ok
rootfs: major=253 minor=1 devn=64769
fsck 1.38 (30-Jun-2005)
[/bin/fsck.ext3 (1) -- /] fsck.ext3 -a /dev/mapper/36005076303ffc562000000000000
10cc-part1
/dev/mapper/36005076303ffc5620000000000010cc-part1: clean, 91021/525888 files,
550917/1050241 blocks
fsck succeeded. Mounting root device read-write.
Mounting root /dev/mapper/36005076303ffc5620000000000010cc-part1
```

SCSI IPL

- The traditional initial program load (IPL) process relies on accessing a device using System z channel attachment
- For IPL from a FCP-attached device, this is not possible
- SCSI IPL expands the set of IPL'able devices
  - SCSI disks as Linux boot file system possible
- New set of IPL parameters
- Requires to address the SCSI disk
  - FCP adapter id
  - Remote port
  - LUN
- LPAR and z/VM guests supported
- SCSI (IPL) with z/VM
    - *z/VM Version 4.4 (PTF UM30989) or newer*
    - *z/VM Version 5.3 (current version)*

SCSI-IPL example LPAR



SCSI IPL: z/VM

Note the hexadecimal format with a blank separating the first 8 from the final 8 digits

```
set loaddev port 50050763 0300C562 lun 40104020 00000000
Ready; T=0.01/0.01 22:11:01
```

WWPN

LUN

```
query loaddev
PORTNAME 50050763 0300C562    LUN  40104020 00000000    BOOTPROG 0
BR_LBA   00000000 00000000
Ready; T=0.01/0.01 22:11:06
```

5021 is the device number of the FCP subchannel that provides access to the SCSI boot disk.

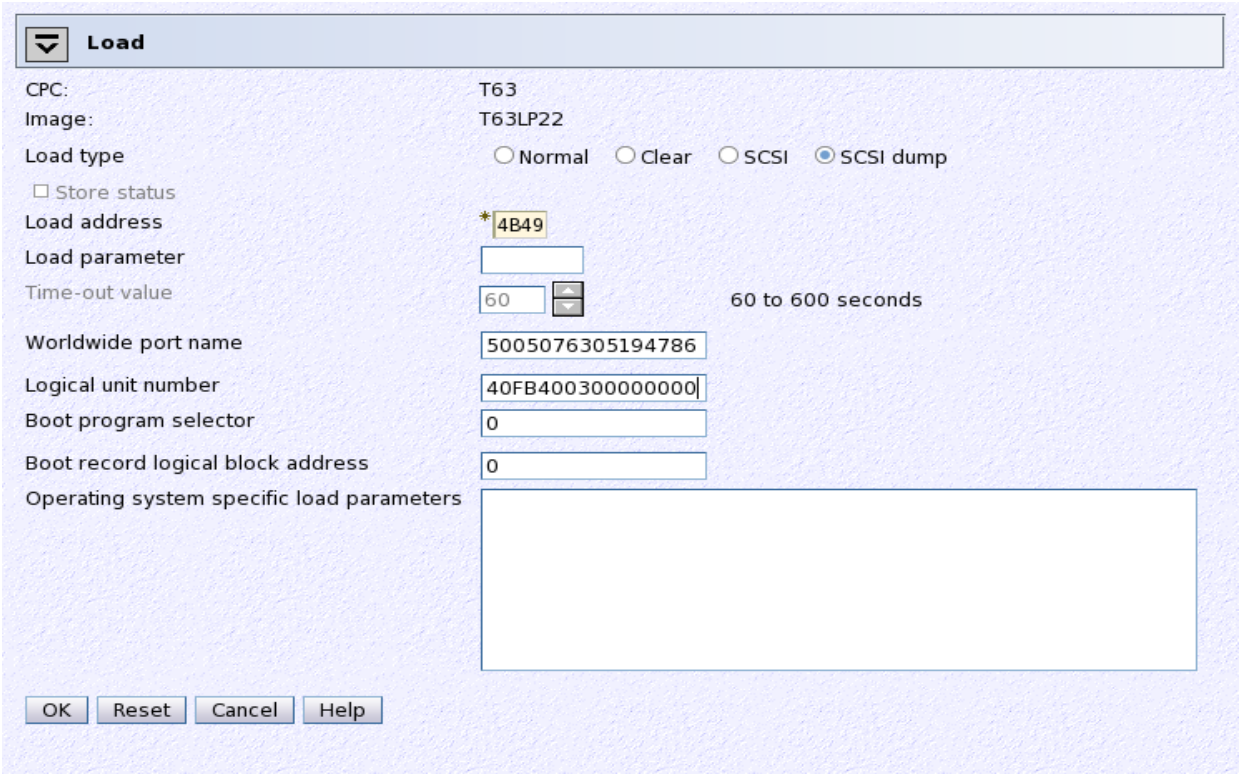
```
i 5021
00: HCPLDI2816I Acquiring the machine loader from the processor controller.
00: HCPLDI2817I Load completed from the processor controller.
00: HCPLDI2817I Now starting the machine loader.
00: MLOEVL012I: Machine loader up and running (version 0.18).
00: MLOPDM003I: Machine loader finished, moving data to final storage location.
Linux version 2.6.16-18.x.20060403-s390xdefault (wirbser@t2944002) (gcc version
4.1.0) #1 SMP PREEMPT Mon Apr 3 09:56:54 CEST 2006
We are running under VM (64 bit mode)
Detected 4 CPU's
Boot cpu address  0
Built 1 zonelists
Kernel command line: dasd=e960-e962 root=/dev/sda1 ro noinitrd zfcp.device=0.0.3d21,
0x500507630300c562,0x401040ee00000000
```

SCSI dump

- Dump memory of one LPAR to disk for problem analysis
- Similar to VMDUMP and dump to DASD
- SCSI dump supported for LPARs and as of z/VM 5.4
- Preparation summary:
  - large SCSI disk (at least system memory + 11 MB)
  - fdisk /dev/sda
  - mke2fs /dev/sda1
  - mount /dev/sda1 /mnt
  - zipl -D /dev/sda1 -t /mnt
  - umount /mnt

SCSI dump from HMC

- Select CPC image for LPAR to dump
- Goto Load panel
- Issue SCSI dump
  - FCP device ID
  - WWPN
  - LUN



SCSI dump under z/VM

- SCSI dump from z/VM is supported as of z/VM 5.4
- Issue SCSI dump

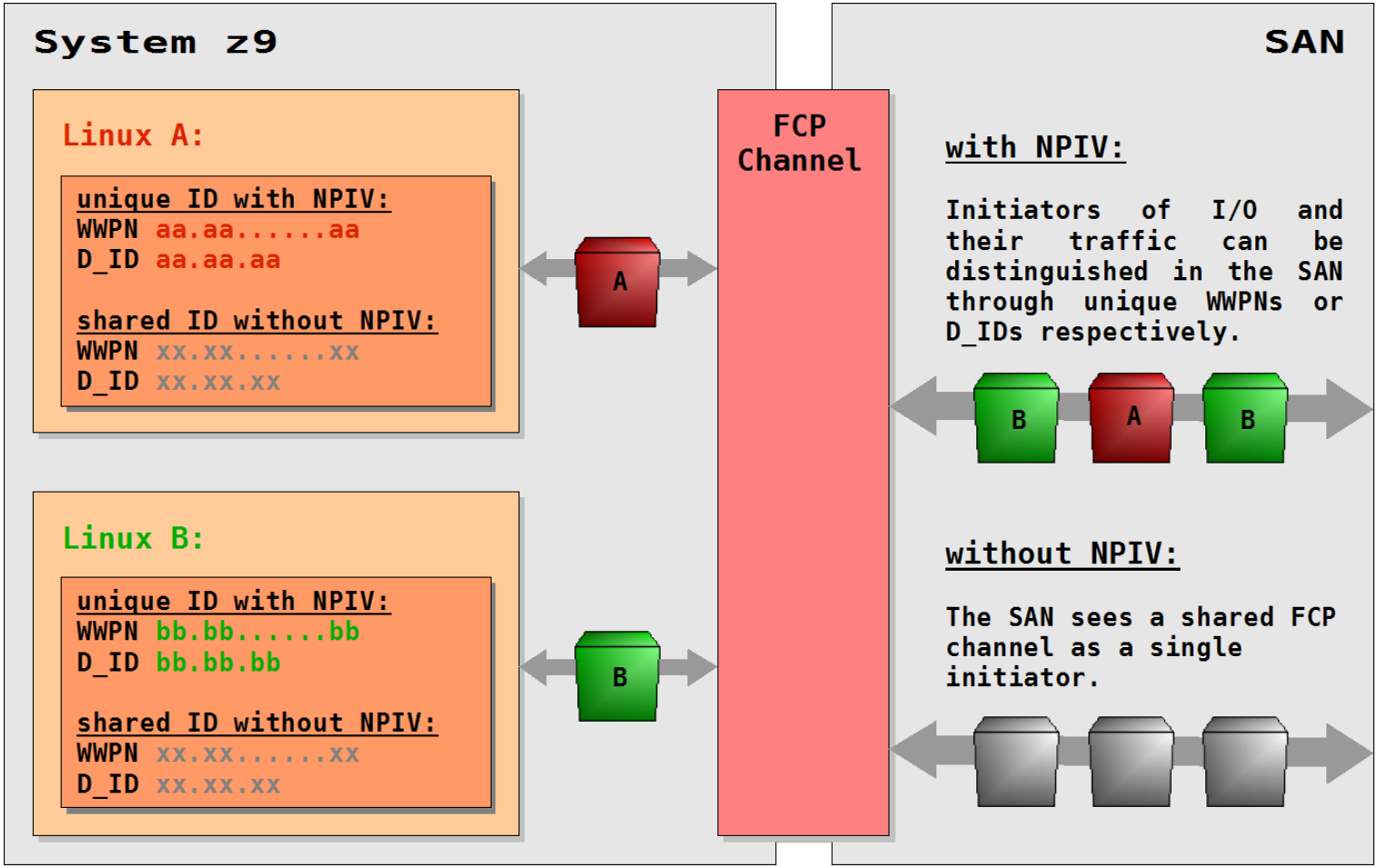
```
#cp cpu all stop

#cp cpu 0 store status

#cp set dumpdev portname 47120763 00ce93a7 lun 40104020
  00000000 bootprog 0

#cp ipl 4b49 dump
```

- To access the dump, mount the dump partition

NPIV

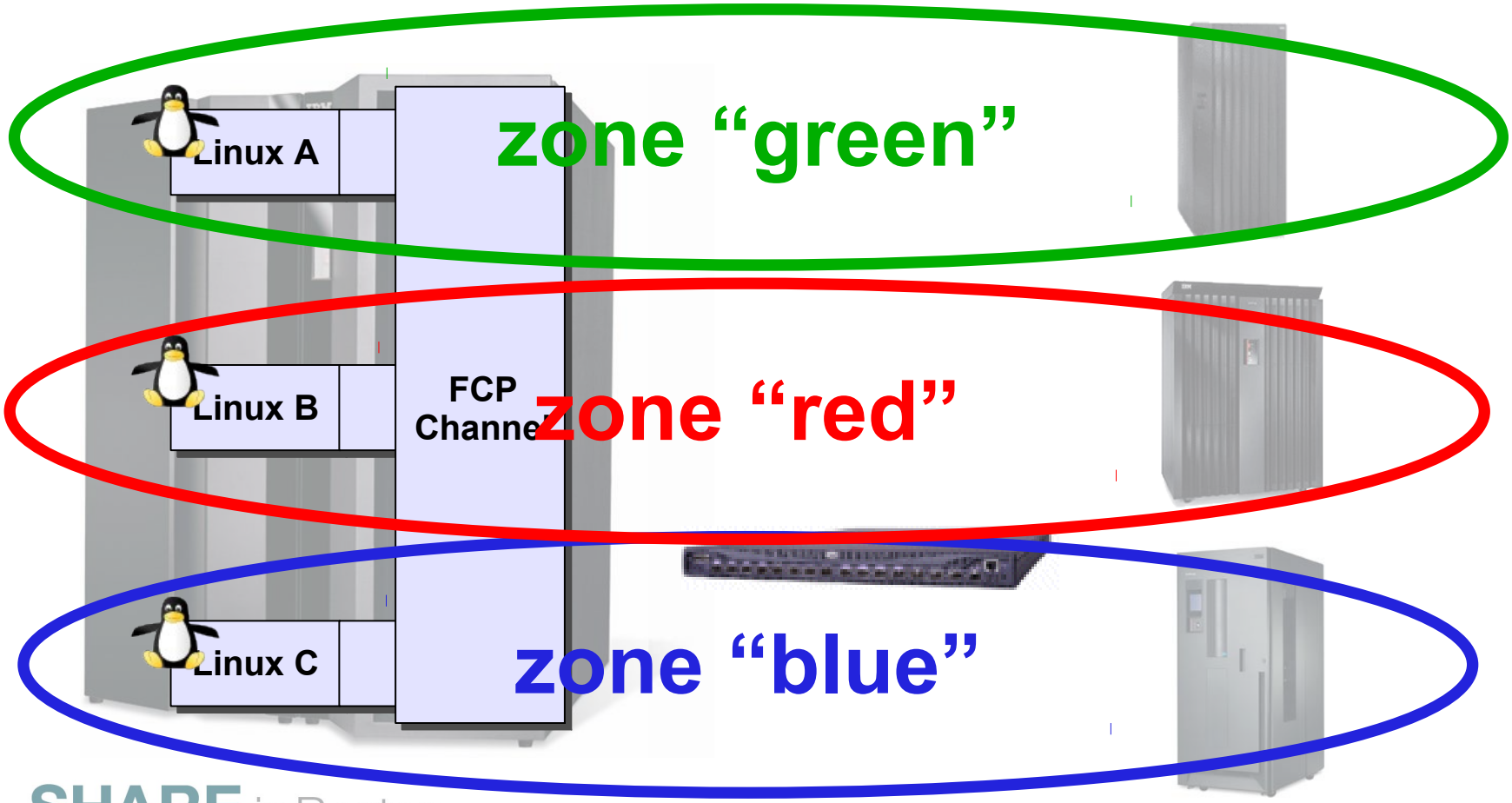
- N_Port Identifier Virtualization (NPIV) is a Fibre Channel facility allowing multiple WWPNs to share a single physical WWPN.
  - without NPIV: one WWPN for FCP channel
  - with NPIV: unique WWPN for each FCP subchannel
- enables
  - proper zoning in SAN fabrics
  - proper LUN masking in storage devices
- security
- access control

NPIV – Unique SAN Identities!



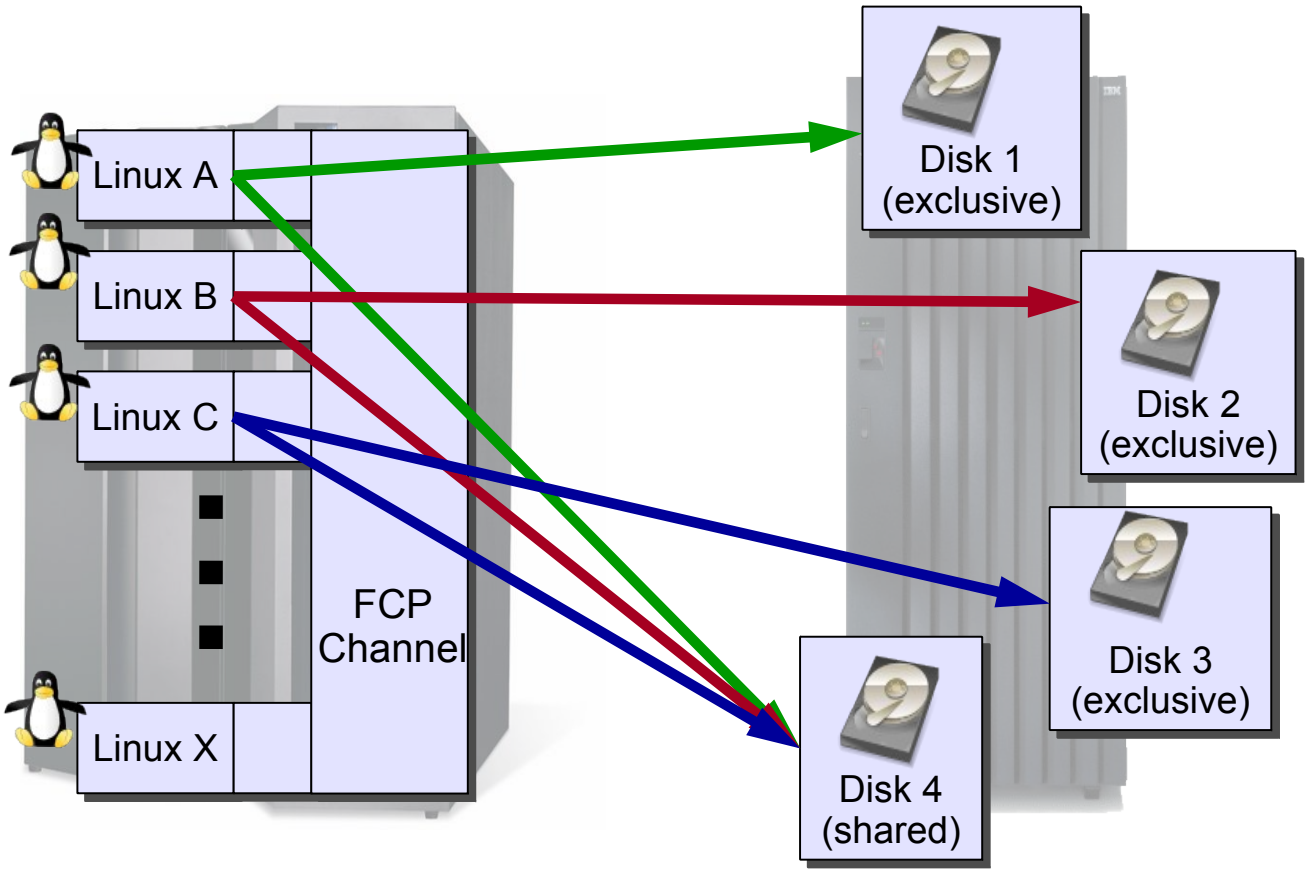
SAN zoning with NPIV

Different Linux guests in different zones



LUN masking with NPIV

Storage server can identify Linux guests via WWPNs



NPIV requirements


- NPIV is available on System z servers.
  - FICON Express 2 adapter running with MCL003 on EC J99658
- z/VM
  - z/VM 5.2 or 5.3
  - z/VM 5.1 with the PTF for APAR VM63744
- Linux Distribution
  - Currently SLES9 SP3/4, SLES10, RHEL5, SLES 11
- NPIV-Capable Switch
  - only required for switch adjacent to System z

NPIV: Do's and don'ts

- Do not use more than 32 FCP devices per physical channel in NPIV mode.
- Zone each NPIV WWPN individually. This can reduce fabric traffic.
- Multipathing remains mandatory (performance and availability).
- Enable NPIV on the SAN switch before enabling it on the System z9 server.
- Be aware that each login from a NPIV-mode FCP device into a storage subsystem counts as a separate host login. There are limits at storage side.
- Switches typically limit the number of supported N_Port IDs.
- Some switches limit the number of N_Port IDs that can be assigned to a physical port.
- FCP microcode MCL003 on EC J99658 requires a special activation procedure. All FCP PCHIDs should be configured off before activating the MCL.

Device Support

IBM I/O connectivity website

http://www-03.ibm.com/systems/z/connectivity/products/fc.html

http://www-03.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss

Switches	Disks	Tape
IBM	IBM DS8000	IBM 3590 drive
Brocade	IBM DS6000	IBM 3592 drive
Cisco	IBM XIV	IBM 3494 libr.
CNT	IBM SVC	IBM 3584 libr.
McData		IBM TS 7510
	Vendor Disks*	Vendor Devices & libraries *

* if Vendor & Software support the attachment

Summary of FCP

- available for IBM zSeries and System z
- based on existing Fibre Channel infrastructure
- runs on all available z/VM and RHEL/SLES versions
- integrates System z into standard SANs
- connects to switched fabric or point-to-point
- multipathing for SCSI disks is mandatory
- SCSI tape is the only tape attachment supported by Backup/Archive middleware such as TSM
- gives you new storage device choices
- usually performs better than FICON
- buys you flexibility at the cost of complexity
- tooling available, receiving better integration

ECKD and SCSI Comparison

	ECKD DASD	SCSI Disk
Configuration	IOCDS/VM (operator)	IOCDS/VM & SAN & Linux (operator & SAN admin & Linux admin)
Access Method	SSCH/CCW	QDIO
Block Size (Byte)	512, 1K, 2K, 4K	512
Disk Size	3390 Model 3/9/27/54	any
Formatting (low level)	dasdfmt	not necessary
Partitioning	fdasd	fdisk
File System	mke2fs (or others)	
Access	mount	

More Information

I/O Connectivity on IBM zSeries mainframe servers:

www.ibm.com/systems/z/connectivity/

Supported Attachments of IBM Storage to IBM Servers

www-03.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss

Linux on zSeries: Fibre Channel Protocol Implementation Guide

www.redbooks.ibm.com/redpapers/pdfs/redp0205.pdf

How to use FC-attached SCSI devices with Linux on System z

download.boulder.ibm.com/ibmdl/pub/software/dw/linux390/docu/l26cts00.pdf

Linux for IBM System z

www.ibm.com/developerworks/linux/linux390/

Linux for IBM System z Device Drivers Book and other documentation

www.ibm.com/developerworks/linux/linux390/october2005_documentation.html

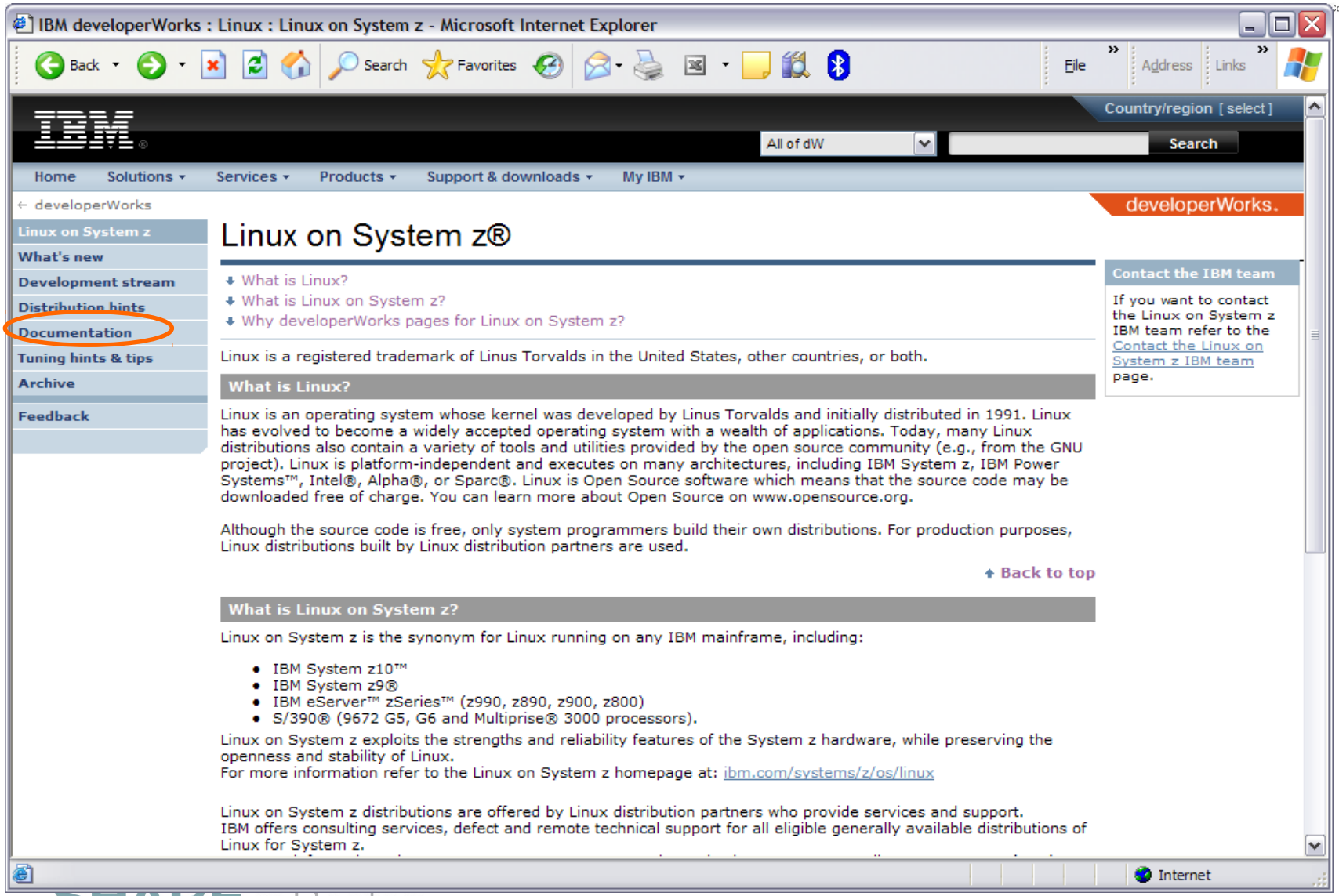
Questions?

SCSI over FCP for Linux on System z Roundup

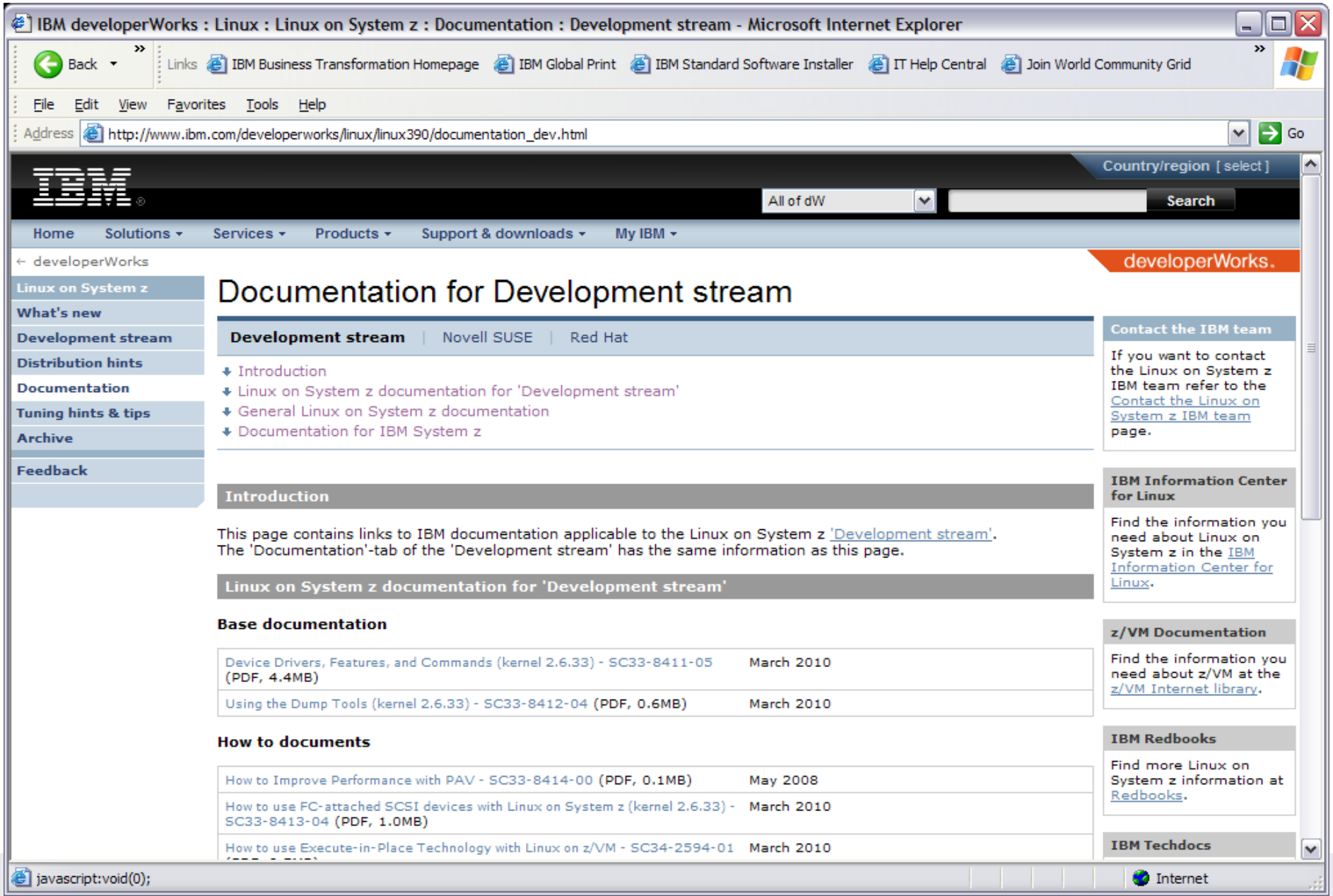
Dr. Holger Smolinski
IBM Germany Research & Development GmbH

2010-08-03
9222

developerWorks – entry page for documentation



Development stream – Novell SUSE – Red Hat documentation



More information

ibm.com/systems/z/linux

www.vm.ibm.com

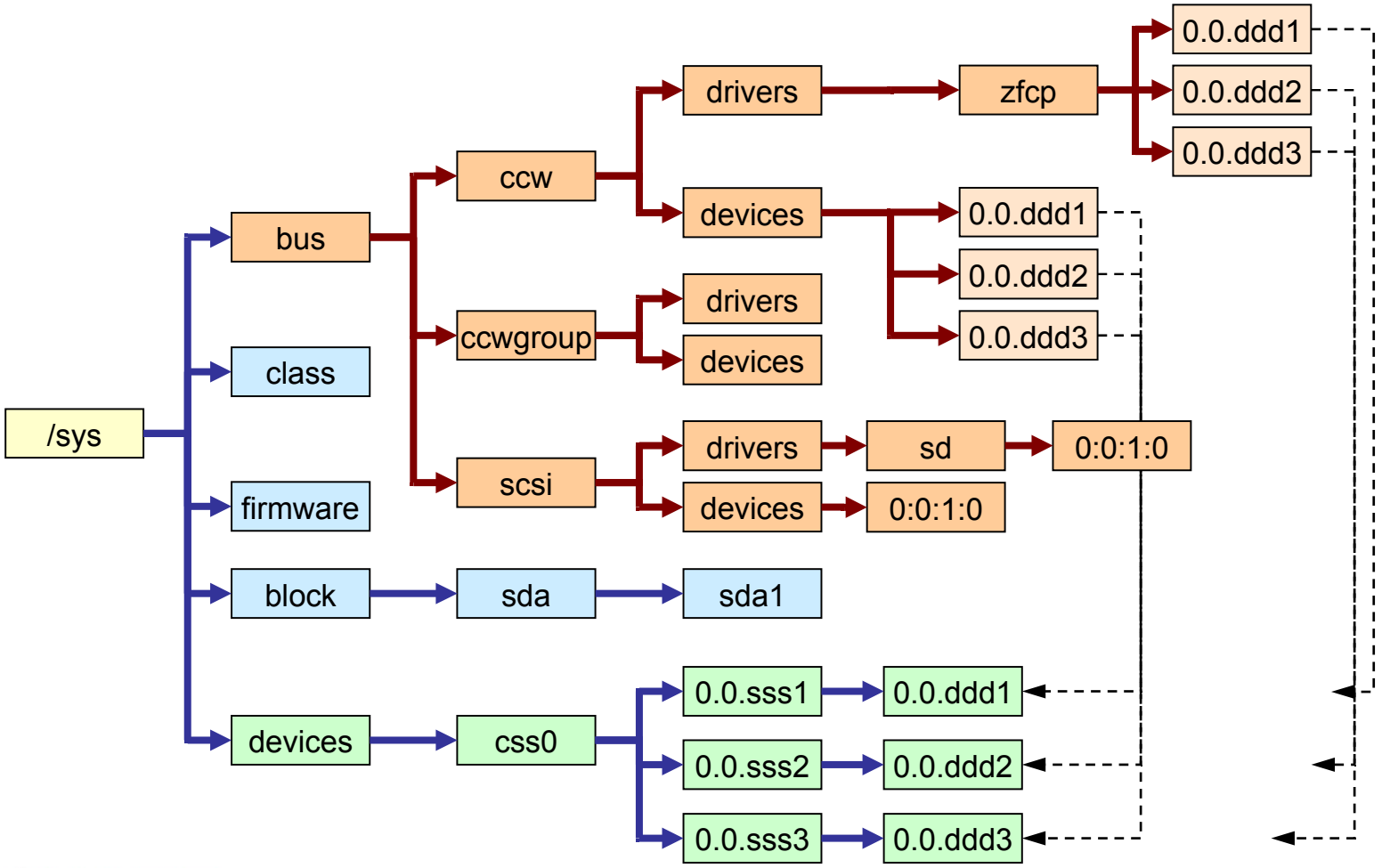
Appendix

Where to find information

The Linux on System z documentation can be found at these key locations:

IBM developerWorks	ibm.com/developerworks/linux/linux390/documentation_dev.html ibm.com/developerworks/linux/linux390/perf/index.html
IBM Redbooks	http://www.redbooks.ibm.com
IBM Techdocs	http://www.ibm.com/support/techdocs/atsmastr.nsf/Web/Techdocs
z/VM Internet Library	http://www.vm.ibm.com/library/
IBM Information Center for Linux	http://publib.boulder.ibm.com/infocenter/lnxinfo/v3r0m0/index.jsp

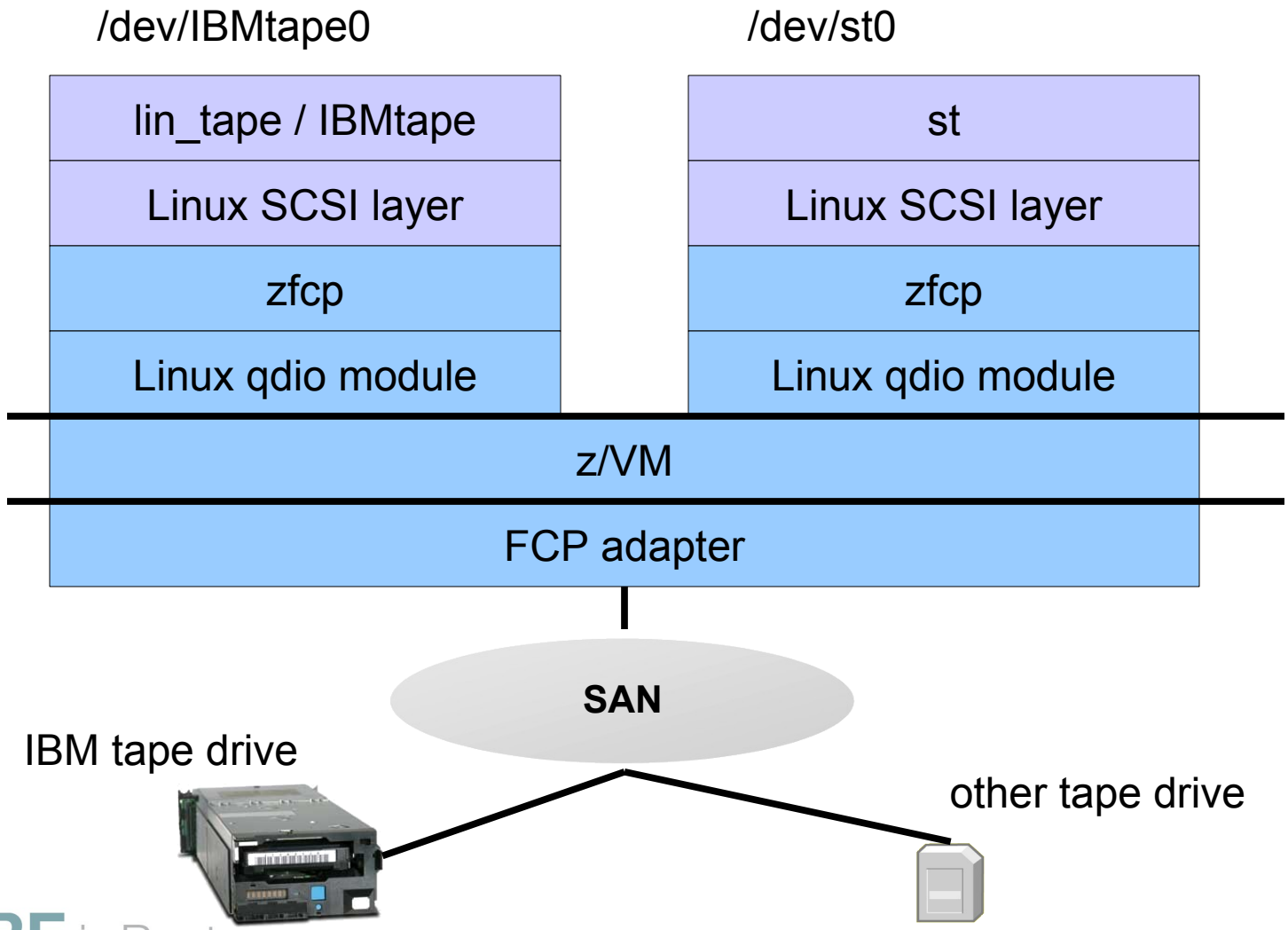
Sysfs


/sys
bus
class
firmware
block
devices
ccw
ccwgroup
scsi
drivers
devices
drivers
devices
drivers
devices
zfcp
sd
0:0:1:0
sda
sda1
css0
0.0.sss1
0.0.sss2
0.0.sss3
0.0.ddd1
0.0.ddd2
0.0.ddd3
0.0.ddd1
0.0.ddd2
0.0.ddd3
0.0.ddd1
0.0.ddd2
0.0.ddd3
0:0:1:0

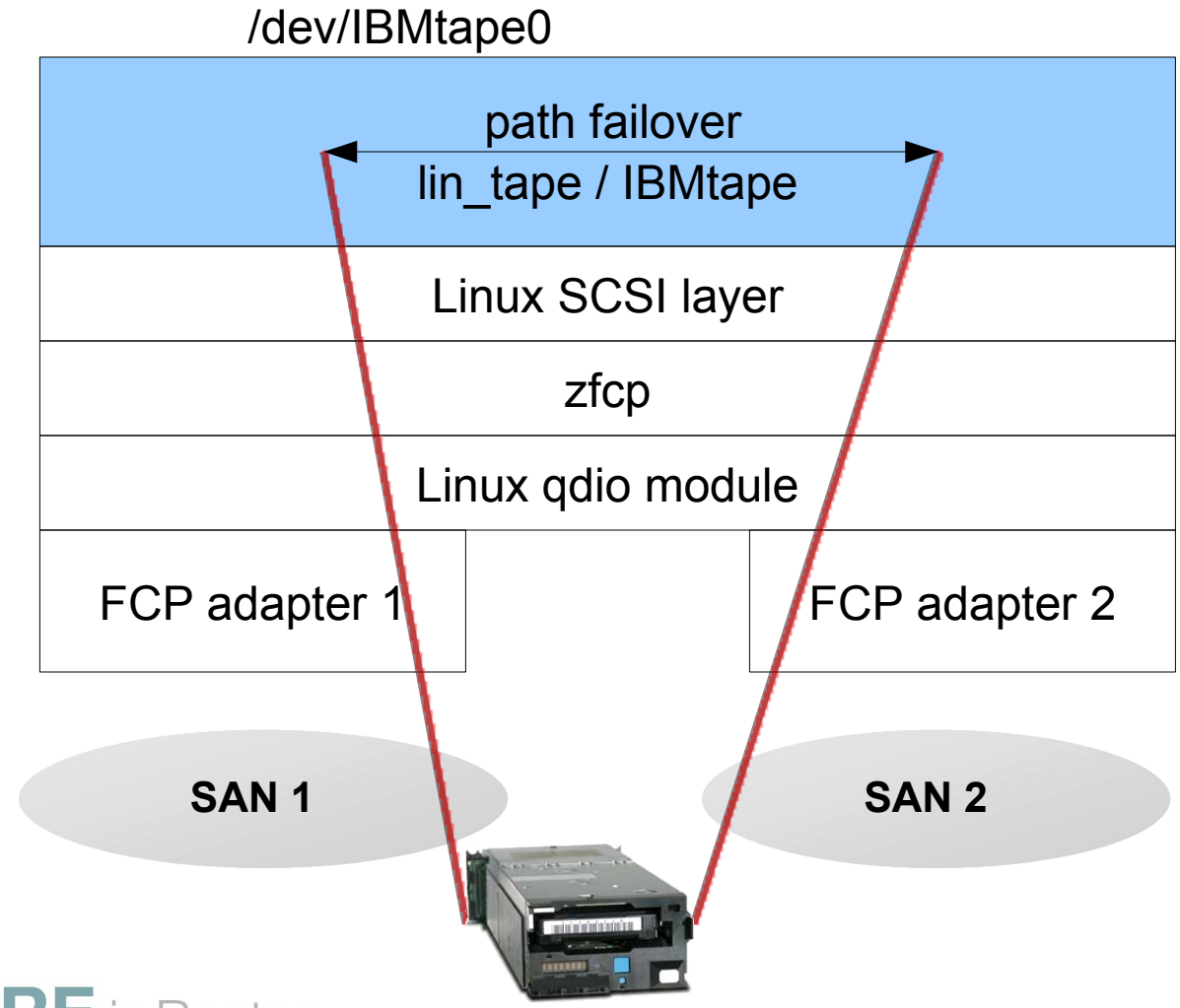
Backing up data using TSM?

* “stand-alone” Linux backup solution, no assistance from z/OS required
* TSM supports many SCSI tape devices, including OEM devices (System z only supports SCSI tape devices from IBM so far)
* both TSM client and TSM server are available for Linux on System z

Multipathing for IBM tapes (1)



Multipathing for IBM tapes (2)



Multipathing for IBM tapes (3)

- Multipathing provided by IBM tape device driver
  - lin_tape (formerly IBMtape)
- Supported together with tape drive
- Capable of failover and failback, no load balancing
- Does not cover data mirroring
  - responsibility of backup and media management applications

Multipathing for IBM tapes (4)

Setup:

- enable via module parameter in /etc/modprobe.conf.local
  - options lin_tape alternate_pathing=1
- attach all paths to tape drive