SCSI IPL for IBM zSeries Server

Volker Sameske (sameske@de.ibm.com)
Linux on zSeries Development
IBM Lab Boeblingen, Germany

Share Anaheim, California
February 27 - March 4, 2005
Session 9259
Agenda

- Requirements
- New IPL type for IBM zSeries
- New IPL I/O devices
- SCSI IPL of an LPAR
- SCSI IPL of a z/VM guest
- SCSI IPL parameters
- SCSI disk installation and preparation
- SCSI dump
Hardware Requirements

- IBM zSeries Server
  - 800, 890, 900 or 990
- Separately orderable feature
- Requires enablement by FC9904
- Requires FCP channels
  (FICON or FICON Express adapter card)
- Requires FC attached SCSI disks
- z800, z900 require IML
Software Requirements

- SCSI IPL under z/VM requires z/VM version 4.4 (PTF UM30989 installed) or higher

- SUSE Linux Enterprise Server 8 (SLES8)
  - Service Pack 3
  - Submarine update

- SUSE Linux Enterprise Server 9 (SLES9)

- Red Hat Enterprise Linux 3 (RHEL3)
  - Update 3
Challenge

- Without SCSI IPL:
  - IPL from CCW-based devices (Tape, ECKD, FBA)
  - OS installation on DASD
  - SCSI devices used as data devices

- With SCSI IPL:
  - SCSI devices usable as IPL devices
  - Linux root file system on a SCSI disk
  - SCSI-only system
New IPL Type for IBM zSeries

zSeries Server

ESCON / FICON

IPL device

FCP (QDIO)

SAN (Fibre Channel)
SCSI IPL versus CCW IPL

- Traditional CCW type IPL
  - I/O controlled by channel programs
  - Devices configured within the IOCDS (I/O configuration data set)
  - CCW = Channel command word
    - Contains a command to perform a read, write or control operation
    - Channel program is a chain of CCWs
    - Executed in a channel by channel engines
    - Running independently of the CPUs.
SCSI IPL versus CCW IPL

- Traditional CCW type IPL
  - IPL only for CCW based I/O devices supported
  - I/O devices are identified by a two-byte device number.
  - 24 bytes IPL
    - One PSW and two CCWs read from disk
    - First CCW copies more boot loader code from disk
    - Second PSW executes the restart PSW
    - PSW executes the copied boot loader code
SCSI IPL versus CCW IPL

- **SCSI IPL**
  - Completely new IPL method
  - Expands the set of IPL I/O devices → SCSI disk
  - Impractical to extend CCW type IPL
  - SCSI IPL has to
    - Login to a FC fabric
    - Send SCSI commands and associated data
    - Maintain a connection through the SAN
  - Enhanced set of parameters
  - Configuration not only within the IOCDS
  - Much more flexible for future enhancements e.g. CD, DVD
New I/O Devices

- New I/O devices which can be used during IPL
  - SCSI over Fibre Channel I/O devices
- Different access method compared to CCW I/O devices
- More addressing parameters
- No ECKD emulation overhead
- No disk size restrictions
SAN Addressing

Device Number (devno)  e.g. 0x6000

Worldwide Port Name (WWPN)  e.g. 0x5005076300ce93a7

Logical Unit Number (LUN)  e.g. 0x1234000000000000
IOCDS - FCP Configuration

CHPID PATH=(CSS(0),50),SHARED,*
   PARTITION=((LP01,LP02,LP03,LP04,LP05,LP06,LP07,LP08,LP09*,
   LP10,LP11,LP12,LP13,LP14,LP15),(=)),PCHID=160,TYPE=FCP
CHPID PATH=(CSS(1),50),SHARED,*
   PARTITION=((LP16,LP17,LP18,LP19,LP20,LP21,LP22,LP23,LP24*,
   LP25,LP26,LP27,LP28,LP29,LP30),(=)),PCHID=161,TYPE=FCP

CNTLUNIT CUNUMBR=5402,PATH=((CSS(0),50),(CSS(1),50)),UNIT=FCP

...  

IODEVICE ADDRESS=(5400,002),CUNUMBR=(5402),*
   PARTITION=((CSS(0),LP01),(CSS(1),LP16)),UNIT=FCP
IODEVICE ADDRESS=(5402,002),CUNUMBR=(5402),*
   PARTITION=((CSS(0),LP02),(CSS(1),LP17)),UNIT=FCP

...  

IODEVICE ADDRESS=(5460,144),CUNUMBR=(5402),*
   PARTITION=((CSS(0),LP15),(CSS(1),LP30)),UNIT=FCP
SCSI IPL - LPAR

- Enhanced load panel
  - Only visible when SCSI IPL is enabled
  - SE remembers last set of parameters

- Enhanced activation profile
  - SCSI IPL possible when activating an LPAR
### SCSI IPL - LPAR - Load Panel

**CPC:** P000F12B  
**Image:** ZFCP4  
**Load type:**  
- Normal  
- Clear  
- SCSI  
- SCSI dump  

- **Store status**  
- **Load address:** 5C00  
- **Load parameter:**  
- **Time-out value:** 060 (60 to 600 seconds)  
- **World wide port name:** 5005076300CE93A7  
- **Logical unit number:** 5732000000000000  
- **Boot program selector:** 0  
- **Boot record logical block address:** 0000000000000000  
- **OS specific load parameters:**

![Image of the Load Panel]
SCSI IPL - LPAR - OS Messages

Message Text

MLQEVLO121: Machine loader up and running (version 0.12).
MLQPDMD0031: Machine loader finished, moving data to final storage location.
Linux version 2.4.20-06.0-s390xdebug (root@gsusr04) () #1 SMP Thu Jun 5 13:21:32
CEST 2003
We are running native (64 bit mode)
On node 0 totalpages: 16384
zone(0): 16384 pages.
zone(1): 0 pages.
zone(2): 0 pages.
Kernel command line: zfcp_map="0x9100 0x0:0x5005076300ce93a7 0x0:0x5733000000000000"
root=/dev/sdal ro nominitrd
Highest subchannel number detected (hex): 0419
SNID - Device 1304 on Subchannel 00E5, Lpm 80, became 'not operational'
SNID - Device 1305 on Subchannel 00E6, Lpm 80, became 'not operational'
SNID - Device 1306 on Subchannel 00E7, Lpm 80, became 'not operational'
SNID - Device 1307 on Subchannel 00E8, Lpm 80, became 'not operational'
SNID - Device 1308 on Subchannel 00E9, Lpm 80, became 'not operational'
SNID - Device 1309 on Subchannel 00EA, Lpm 80, became 'not operational'
SNID - Device 130A on Subchannel 00EB, Lpm 80, became 'not operational'
SNID - Device 130B on Subchannel 00EC, Lpm 80, became 'not operational'
SNID - Device 130C on Subchannel 00ED, Lpm 80, became 'not operational'
SNID - Device 130D on Subchannel 00EE, Lpm 80, became 'not operational'
SCSI IPL - z/VM

- SCSI IPL of a z/VM guest
- Two new CP commands
  - SET LOADDEV
  - QUERY LOADDEV
- LOADDEV directory statement
- Enhanced CP IPL command
  - IPL <fcp_adapter_devno>
SCSI IPL – z/VM – Example

att 50aa *
00: FCP 50AA ATTACHED TO LINUX18 50AA
Ready; T=0.01/0.01 13:16:20

q v fcp
00: FCP 50AA ON FCP 50AA CHPID 40 SUBCHANNEL = 000E
00: 50AA QDIO-ELIGIBLE QIOASSIST-ELIGIBLE
Ready; T=0.01/0.01 13:16:24

set loaddev portname 50050763 00c20b8e lun 52410000 00000000
Ready; T=0.01/0.01 13:16:33

q loaddev
PORTNAME 50050763 00C20B8E LUN 52410000 00000000
BOOTPROG 0 BR_LBA 00000000 00000000
Ready; T=0.01/0.01 13:16:38
SCSI IPL - z/VM - Example

i 50aa

00: HCPLDI2816I Acquiring the machine loader from the processor controller.
00: HCPLDI2817I Load completed from the processor controller.
00: HCPLDI2817I Now starting machine loader version 0001.
00: MLOEVL012I: Machine loader up and running (version 0.12).
00: MLOPDM003I: Machine loader finished, moving data to final storage location.

Linux version 2.4.21 (root@tel15v18)(gcc version 3.3 (Red Hat Linux 8.0 3.3-5bb9)) #3 SMP Mon Sep 15 15:28:42 CEST 2003

We are running under VM (64 bit mode)

On node 0 total pages: 32768
SCSI IPL - z/VM - Profile Exec

/* PROGRAM: PROFILE EXEC A */
CALL DIAG 8,'TERM MORE 0 50' /* ACCELERATE WAIT ON MORE... */
CALL DIAG 8,'SET RETR 50' /* SET RETRIEVE BUFFER */
CALL DIAG 8,'TERM CHARDEL OFF' /* TO USE @ IN INTERNET ADDRESSES */
CALL DIAG 8,'SET RUN ON' /* AVOID CP-READ AT RECONNECT */

... ACC 592 T /* ACCESS TCP/IP DISK */
SCREEN CPOUT YEL
SCREEN INREDISP BLUE
ATT 5480 *
SET LOADDEV PORT 50050763 00CB93CB LUN 51220000 00000000
VSPF /* INVOKE PF-KEY SETTINGS */
VMFCLEAR /* VMFCLEARS SCREEN */
Q V FCP
Q LOADDEV
Q V DASD
SCSI IPL Parameters

- Load type
  - Conventional load types
    - Normal
    - Clear
  - Two new load types
    - SCSI load
    - SCSI dump
SCSI IPL Parameters

- **Load address**
  - 2-byte hexadecimal number
  - Device number of the FCP adapter
  - Not associated with an I/O device
  - The only SCSI IPL parameter defined in IOCDS
  - Required parameter
SCSI IPL Parameters

- WWPN
  - Worldwide port name
  - 8-Byte hexadecimal number
  - Identifier of the FCP adapter port of the SCSI target device
  - Worldwide unique
  - Required parameter
SCSI IPL Parameters

- **LUN**
  - Logical unit number
  - 8-Byte hexadecimal number
  - Identifier of the logical unit
  - Representing the IPL device
  - Required parameter
**SCSI IPL Parameters**

- **Boot program selector**
  - Used to select a boot configuration
  - Up to 31 different configurations possible (decimal 0 - 30)
  - Simple Boot Loader
  - Prepared with Linux zipl tool
  - Partition independent
  - 0 is default
  - Optional
SCSI IPL Parameters

- **Boot record LBA**
  - Used to locate an OS or OS loader on an IPL disk
  - IPL entry on disk
  - Specifies the block number, containing the boot record
  - Normally located at LBA 0 (default)
  - Optional
**SCSI IPL Parameters**

- **OS specific load parameter**
  - Intended to hand over parameters to the operating system or dump program
  - Only passed through
  - Currently restricted to
    - 256 Bytes (SE)
    - 4096 Bytes (z/VM)
  - Optional
SCSI IPL Parameters

- Unsupported IPL parameters
  - Store status
  - Time-out value
- SCSI IPL independent IPL parameters
  - Load parameter
# Terms and Definitions

<table>
<thead>
<tr>
<th>LPAR</th>
<th>z/VM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load Type</td>
<td>-</td>
</tr>
<tr>
<td>Load Address</td>
<td>&lt;fcp_vdev&gt;</td>
</tr>
<tr>
<td>Load parameter</td>
<td>LOADParm &lt;value&gt;</td>
</tr>
<tr>
<td>World wide port name</td>
<td>PORTname &lt;value&gt;</td>
</tr>
<tr>
<td>Logical unit number</td>
<td>LUN &lt;value&gt;</td>
</tr>
<tr>
<td>Boot program selector</td>
<td>BOOTprog &lt;value&gt;</td>
</tr>
<tr>
<td>Boot record logical block address</td>
<td>BR_LBA &lt;value&gt;</td>
</tr>
<tr>
<td>OS specific load parameters</td>
<td>SCPdata &lt;value&gt;</td>
</tr>
</tbody>
</table>
SCSI Disk Installation

- Direct installation to SCSI disk possible with SLES9
- Migration guide available for SLES8 and RHEL3
  - Migration from existing ECKD installation to SCSI disk installation
SCSI Installation - SLES9

```
s already in use.
Sep 14 09:32:52 suse sshd[504]: fatal: Cannot bind any address.

/sbin/ifconfig eth0
eth0   Link encap:Ethernet  HWaddr 0E:06:29:29:29:29
       inet addr:8.152.84.228  Bcast:9.152.87.255  Mask:255.255.248.0

*** sshd has been started ***

*** login using 'ssh -X root@53v30g21.boeblingen.de.ibm.com' ***
*** run 'yast' to start the installation ***
```
login as: root
Password:

>>> >>> SUSE Linux Enterprise Server 9 for S/390 and zSeries <<< <<<

Check with the file README on CD1 for any preparational steps
before starting the installer.

Then start installation with the ‘yast’ command.

have a lot of fun...

...your SUSE team
Welcome to the inst-sys...
Linux 53v30g21 2.6.5-7.69-s390x #1 SMP Wed Jun 2 22:53:38 UTC 2004 s390x s390x s90x GNU/Linux
/root
Probing connected terminal...

Initializing virtual console...

Found a PuTTY terminal on /dev/pts/0 (80 columns x 24 lines).

run yast to start the installation

inst-sys:~ # yast
SCSI Installation - SLES9
SCSI Installation - SLES9

Add New ZFCP Disk

Enter the identifier of the disk to add. Enter the Channel Number of the ZFCP controller, the worldwide port number (WWPN), and the FCP-LUN number.

Channel Number: 0.0.5586

WWPN: 0x500507630e6936cb

FCP-LUN: 0x51250e00000000000
SCSI Installation - SLES9
SCSI Installation - SLES9

Create a primary partition on /dev/sda

- **Format**
  - ( ) Do not format
  - (x) Format
    - File system: Ext3
    - Options

- **Size**
  - Cylinder size: 1.87 M
  - Start cylinder: 0
  - End: (9 or +9M or +3.2GB) 1015

- **Mount Point**
  - [ ] For fs options

- [OK] [Cancel]
SCSI Installation - SLES9

Use accept to perform a new installation with the values displayed.
Change the values by clicking on the respective headline or by using the Change... menu.
To update an existing SUSE Linux system instead of doing a new install, click the Mode headline or select Mode in the Change... menu.

Your hard disk

Installation Settings
Click any headline to make changes or use the "Change..." menu.

Mode
* New installation

Partitioning
* Create root partition 1.8 GB (/dev/sdal with ext3)

Software
* Default system
* + Authentication Server (NIS, LDAP, Kerberos)
* + Linux Tools
* + Print Server (CUFS)
* + KDE Desktop Environment
* + Basic Sound Libraries and Tools
* + Help & Support Documentation
* + Basic Runtime System
* + YaST

[change...]
[Abort]
[Accept]
SCSI Installation - SLES9
SCSI Installation - SLES9

Finishing Basic Installation

- Update configuration
- Copy files to installed system
- Install boot manager
- Prepare system for initial boot

Initializing fonts 37%
SCSI IPL for IBM zSeries Server

Ready(06704); T=0.01/0.01 12:48:19
set loaddev portname 50050763 00c693cb lun 51250000 00000000
Ready; T=0.01/0.01 12:49:25
q loaddev
PORTNAME 50050763 00c693CB LUN 51250000 00000000 BOOTPROG 0
BR_LBA 00000000 00000000
Ready; T=0.01/0.01 12:49:30
i 5588
HCPLDI2816I Acquiring the machine loader from the processor controller.
HCPLDI2817I Load completed from the processor controller.
HCPLDI2817I Now starting machine loader version 0001.
MLOEVL0121: Machine loader up and running (version 0.13).
MLODM0031: Machine loader finished, moving data to final storage location.
Linux version 2.6.5-7.97-390x (geeko@buildhost) (gcc version 3.3.3 (SuSE Linux)
) #1 SMP Fri Jul 2 14:21:59 UTC 2004
We are running under VM (64 bit mode)
On node 0 totalpages: 65536
   DMA zone: 65536 pages, LIFO batch:16
   Normal zone: 0 pages, LIFO batch:1
   HighMem zone: 0 pages, LIFO batch:1
Built 1 zonelists
Kernel command line: root=/dev/sda1 selinux=0 TERM=dumb elevator=cfq

Hold BOEVIM530
Linux SCSI Disk Preparation (manually)

- Linux disk preparation tool „zipl“
  - Boot loader for IBM S/390 and zSeries architectures
  - Command line versus configuration file
  - Makes SCSI disks IPL’able as well as ECKD DASDs
  - Boot menu (multi-boot option)
    - For more than one boot configuration
    - Boot configuration is kernel, parmline and ramdisk
  - Prepares disk for SCSI IPL and SCSI dump
  - IPL and dump programs can be on the same disk
  - More information on zipl and zipl.conf man pages
SCSI Disk Preparation - Example

/etc/zipl.conf

```
[defaultboot]
default = scsi-ipl-1
[scsi-ipl-1]
target   = "/boot"
image    = "/boot/kernel-image-1"
parmfile = "/boot/parmfile-1"
[scsi-ipl-2]
target   = "/boot"
image    = "/boot/kernel-image-2"
parmfile = "/boot/parmfile-2"
ramdisk  = "/boot/initrd-2"

:menu1
target   = "/boot"
1=scsi-ipl-1
2=scsi-ipl-2
default=2
```

[root@host /]# zipl -m menu1
Using config file '/etc/zipl.conf'
Building bootmap '/boot/bootmap'
Building menu 'menu1'
Adding #1: IPL section 'scsi-ipl-1'
Adding #2: IPL section 'scsi-ipl-2'
(default)
Preparing boot device: 08:00
Done.
[root@host /]#
SCSI Dump

- Stand-alone dump to a SCSI disk
- IPL of an OS dependent dump program
- LPAR only
- Automatic store status
- Reset normal instead of reset clear
- Machine loader and system dump program run in same LPAR memory, which has to be dumped.
- Lower-address area of the LPAR memory will be copied into a reserved area (HSA).
- Serial access, one save area for all LPARs.
SCSI Dump With Linux on zSeries

- **zfcpdump** - Linux SCSI dump program
- Part of **s390-tools**
- Prepared with **zipl tool**
- Independent Linux
  - Kernel 2.4.19
  - Ramdisk with busybox
- The dump program determines where to put the dump.
  - Currently the same SCSI disk
  - Maybe in the future: using “OS specific parameter” field to pass additional target parameters
- The dump program retrieves two parts of the dump
  - From machine dependent storage area (HSA space)
  - From main storage
- Machine depending storage area can be released after the first part
SCSI Dump With Linux on zSeries

- Dump disk
  - contains dump program
  - contains file system
  - is mountable

- Dumps are files

- Several dumps on one disk possible

- Readable with lcrash
Summary

- New IPL method for IBM zSeries server
  - Available for LPAR and z/VM
- Expands the set of IPL devices
- Enhanced set of parameters
  - Three required parameters
  - Several optional parameters
References

  *SCSI initial program loading for zSeries*


- IBM Corporation, *z/Architecture Principles of Operation*, Order No. SA22-7832; available through IBM branch offices.


- IBM Corporation, *IBM eServer zSeries z990 System Overview*, Order No. SA22-1032; available through IBM branch offices.


SCSI IPL for IBM zSeries Server

Questions?
Trademarks

- The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

  * Registered trademarks of IBM Corporation

- The following are trademarks or registered trademarks of other companies.
  Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
  LINUX is a registered trademark of Linus Torvalds
  Penguin (Tux) complements of Larry Ewing
  Tivoli is a trademark of Tivoli Systems Inc.
  UNIX is a registered trademark of The Open Group in the United States and other countries.

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

Notes:
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.

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.