SCSI on Linux for zSeries – Early Experiences

Neale Ferguson
Software AG
August 2002
• Software AG began participation in the ESP in July, 2002
• IBM have authorized Software AG to present this material
• Thanks for their assistance during the ESP goes to:
  – Ingolf Salm – IBM Germany
  – Pam Hares – IBM UK
  – Christoph Arenz – IBM Germany (& US)
Acknowledgements

• Uli Kuhna – for SAN and Shark configuration
• Wolfgang Buettner – for z/VM & IOCDS configuration
• Dr. Gerhard Banzhaf (IBM) – material from his presentation “FCP Channel for z800 and z900”
• Material adapted from “Device Drivers and Installation Commands”.
• Material sourced from “Getting Started with zSeries Fibre Channel Protocol”.

Agenda

- Background
- Environment
- Configuration
- 31-bit Experiences
- 64-bit Experiences
- Further work
• Allows Linux for S/390 (31-bit mode) and Linux for zSeries (64 bit mode) to access:
  – Distributed storage devices
    • With FCP interfaces (via switches)
    • With parallel SCSI interfaces (via additional bridges)
  – With Linux for zSeries (S/390) running
    • In a partition
    • Under z/VM (requires z/VM 4.3 RSU001)
• Provides access to distributed (open) storage and SAN world
• Based on existing FICON and FICON Express hardware
Background

- New CHPID type: FCP
- Uses 2-port Fibre Channel cards FICON and FICON Express
  - Optical only
  - Short wave and long wave
  - 1 Gbit/s today: 2 Gbit/s has been announced
  - Currently 232KB buffer in card: 2MB proposed
- Different Firmware Load
  - Selected via definition of CHPID type in IOCP (HCD)
- QDIO protocol for communication between Processor/Memory and Channel
  - Based on scheme introduced with OSA Express
  - Continuously running channel programs
    - Reduces I/O path lengths
    - Reduces number of interrupts
• The zfcp driver is a low-level or host-bus adapter driver supplementing the Linux SCSI I/O subsystem (SCSI stack).
• zfcp driver is open source.
• Linux for zSeries and S/390 can make use of all SCSI device types currently supported by Linux on other platforms including
  – SCSI disks,
  – Tapes,
  – CD-ROMs, and,
  – DVDs.
Background

- **Filesystem**
  - block devices: DASD, CD-ROM, DVD
  - character devices: CD Writer, Scanner

- **Common SCSI Stack**
  - SR: SCSI device driver
    - CD RIM
  - SD: SCSI device driver
    - disk
  - ST: SCSI device driver
    - Tape
  - SG: SCSI Generic Device driver

- **zFCP**:
  - zSeries FCP adapter device driver

Linux

- SCSI Stack unchanged

- Replaces adapter specific device drivers
SAN Topologies

- **Point-to-point**
  - [Diagram showing a point-to-point topology with a server and a controller connected directly.]

- **Arbitrated Loop**
  - [Diagram showing an arbitrated loop topology with two servers and two controllers connected in a loop.]

**point-to-point**

**arbitrated loop**
SAN Topologies

- Switched Fabric
## Environment

### Hardware

<table>
<thead>
<tr>
<th>Component</th>
<th>Model</th>
<th>Microcode Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM z900</td>
<td>2064-103</td>
<td>J11233.003 FCP OFCP15 Level 0.1A</td>
</tr>
<tr>
<td>McData SAN Switch</td>
<td>Connectrix ED-64M</td>
<td>01.04.002</td>
</tr>
<tr>
<td>Shark ESS</td>
<td>2105-F20</td>
<td>1.5.0.107</td>
</tr>
</tbody>
</table>

### Software

<table>
<thead>
<tr>
<th>Component</th>
<th>Function</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>z/VM</td>
<td>Hypervisor</td>
<td>4.3 RSU 001</td>
</tr>
<tr>
<td>Linux for S/390</td>
<td>31-bit kernel</td>
<td>2.4.7+</td>
</tr>
<tr>
<td>Linux for zSeries</td>
<td>64-bit kernel</td>
<td>2.4.17+</td>
</tr>
<tr>
<td>ESM Manager</td>
<td>SAN Manager</td>
<td>2.0</td>
</tr>
</tbody>
</table>
Environment

• Linux details
  – CONFIG_MSDOS_PARTITION=y
  – CONFIG_ZFCP=m
  – New utils-linux – fdisk required
  – Will (eventually) need new devs.rpm but not mandatory
  – Manuals geared towards devfs but we were using SLES7
Other LPARs

<table>
<thead>
<tr>
<th></th>
<th>DAEV005</th>
<th>DAEV006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linux</td>
<td>Linux</td>
</tr>
<tr>
<td></td>
<td></td>
<td>z/VM 4.3</td>
</tr>
</tbody>
</table>

SAN

Shark

0 1
**Configuration**

- **z/VM User Directory**

  **** PROFILE LNXGST for LINUX Guests
  
  PROFILE LNXGST
  
  MACHINE ESA
  
  OPTION QUICKDSP
  
  SHARE REL 1000
  
  IPL CMS PARM AUTOCHR
  
  CONSOLE 009 3215 T OPERATOR
  
  SPOOL 00C 2540 READER *
  
  SPOOL 00D 2540 PUNCH A
  
  SPOOL 00E 1403 A
  
  LINK MAINT 190 190 RR
  
  LINK MAINT 19D 19D RR
  
  LINK MAINT 19E 19E RR
  
  USER DAEV005 ******** 256M 512M G
  
  INCLUDE LNXGST
  
  ACCOUNT LINUX R.WAITE
  
  MDISK 0191 3390 2676 0025 VVSYS1 MR
  
  MDISK 0192 3390 3839 2838 VVTL99 MR
  
  MDISK 0193 3390 3339 0500 VVTL99 MR
  
  SPECIAL E00C CTCA TCPIP
  
  SPECIAL E00D CTCA TCPIP
  
  * HyperSockets with MTU size of 8k
  
  DEDICATE D04 D04
  
  DEDICATE D05 D05
  
  DEDICATE D06 D06

  Note: The SCSI device ‘E000’ was ATTACHed to each user as required:

  FCP  E000 ATTACHED TO DAEV006  E000 CHPID E0
• IOCDS

```plaintext
CHPID PATH=(E0), SHARED,
    PARTITION=((DAEV, DAEX, DALI), (DAEV, DAEX, DALI)), TYPE=FCP
CNTLUNIT CUNUMBR=E0FC, PATH=(E0), UNIT=FCP
IODEVICE ADDRESS=(E000, 016), CUNUMBR=(E0FC), UNIT=FCP
```
Configuration

- CHPID/Device/SCSI Device
  - S390 device $\neq$ SCSI device
  - S390 device is conduit to SCSI
  - May be $1, \ldots, n$ SCSI devices at end of conduit
  - Each SCSI device may be partitioned to produce multiple targets
A new addressing scheme was developed for Fibre Channel Protocol (FCP) usage, built around World Wide Names (WWN) that are eight bytes long.

Part of the name represents an address type, part is a number that identifies the manufacturer, and part is a unique number assigned (by the manufacturer) for each port or node.

A node is typically a box that contains information. Nodes have one or more ports (and only FC ports are relevant here). A given box may have several addresses; one for the node itself, and one for each FC port contained in the node. Abbreviations are:

- WWNN is a World Wide Node Name
- WWPN is a World Wide Port Name
- WWN is any World Wide Name (WWNN or WWNP)
Linux provides extensive support for SCSI devices, based on the more advanced forms of traditional SCSI addressing.

A SCSI device is expected to have a number (target address, typically a single digit) and this target may have multiple LUNs.

The full SCSI addressing scheme can be used:

- Device number (from your IOCDS) (coded as a “host number”)
- A bus number on this adapter (always zero in current zSeries implementations)
- A SCSI target number on this bus
- A LUN number on this SCSI target
- A partition or device within this LUN
FCP Mapping

1 2 3 4 5
0x0e00 0x01:0xppppppppppppppppp 0x00:0xnnnnnnnnnnnnnnnnn

- LUN name within node
- LUN number used by Linux
- WWPN of SAN node
- SCSI target number
- IOCDS device number
1. Device number
   - This device number must be defined in the IOCDS and be assigned to the FCP channel attached to the FCP switch.
     - You can use the same device number for all your FCP connections (by a given Linux image), or,
     - You can elect to use multiple device numbers (all assigned to the FCP channel, of course).
2. Target number

- The standard Linux SCSI support understands traditional SCSI addressing, with adapter numbers, bus numbers, target ("SCSI address") numbers, and LUNs.
- You assign this number.
- Usable target addresses range from 1 to any positive 31-bit number.
- Traditional SCSI target addresses ranged from 0-7 or 0-15. More recent SCSI architecture allows a much larger number.
3. WWPN

- The World Wide Port Name of the device containing the LUN.
- The WWPN *as seen by the FC switches*. WWPNs reported by devices may be slightly different.
- You do not assign this number; it is built into the FCP devices and you must determine the proper number by querying your FCP elements.
4. Linux LUN
   – The number *to be used by* Linux.
   – You assign this number.
   – Normal usage is to start with 0 and to increment by one for each LUN.

5. SAN Device LUN
   – The number used by the remote device.
   – You do not assign this number.
   – You must determine the number assigned by the node controller.
• There are different parameters that can or must be supplied by the user to allow for proper zfcp operation:
  – Address mappings between Linux SCSI and FCP schemes (optional for each SCSI target)
  – Logging level to determine the verbosity of the zfcp device driver (optional, default value is used if not supplied)
• The zfcp driver provides different means of configuration:
  – Kernel parameters
  – Module parameters (such as for use in modules.conf)
  – Various proc file system entries in /proc/scsi/zfcp
## Configuration

### Module Parameters and proc File System entries

<table>
<thead>
<tr>
<th>Function</th>
<th>Module Parameter</th>
<th>Kernel Parameter</th>
<th>proc-fs entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set logging level</td>
<td>loglevel zfcp_</td>
<td>loglevel</td>
<td>/proc/scsi/zfcp/mod_parm</td>
</tr>
<tr>
<td>Get logging level (and other global module information)</td>
<td>N/A</td>
<td>N/A</td>
<td>/proc/scsi/zfcp/mod_parm</td>
</tr>
<tr>
<td>Add address mapping(s)</td>
<td>map</td>
<td>zfcp_map</td>
<td>/proc/scsi/zfcp/add_map</td>
</tr>
<tr>
<td>Get all existing address mappings</td>
<td>N/A</td>
<td>N/A</td>
<td>/proc/scsi/zfcp/map</td>
</tr>
</tbody>
</table>
Configuration

- Additional device nodes:

  mknod /dev/sda  b 8 0
  mknod /dev/sda1 b 8 1
  mknod /dev/sda2 b 8 2
  mknod /dev/sda3 b 8 3
  mknod /dev/sda4 b 8 4
  mknod /dev/sda5 b 8 5
  mknod /dev/sda6 b 8 6
  mknod /dev/sda7 b 8 7
  mknod /dev/sda8 b 8 8
  mknod /dev/sdb  b 8 16
  mknod /dev/sdb1 b 8 17
  mknod /dev/sdb2 b 8 18
  mknod /dev/sdb3 b 8 19
  mknod /dev/sdb4 b 8 20
## Devices

<table>
<thead>
<tr>
<th>Target</th>
<th>Logical Unit</th>
<th>Size</th>
<th>Device Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>25GB</td>
<td>/dev/sda1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>50GB</td>
<td>/dev/sda2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>25GB</td>
<td>/dev/sda3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>25GB</td>
<td>/dev/sda4</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>5GB</td>
<td>/dev/sdb1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>5GB</td>
<td>/dev/sdb2</td>
</tr>
</tbody>
</table>
• **Driver load script**
  
  - `ln -s /etc/init.d/scsi_load /etc/init.d/rc3.d/S12scsild`

```bash
#!/bin/sh
rmmod zfcp
modprobe qdio
modprobe scsi_mod
insmod zfcp loglevel=0 map="\0xe000 0x01:0x5005076300c38c6d 0x00:0x5200000000000000;\n0xe000 0x01:0x5005076300c38c6d 0x01:0x5201000000000000"
modprobe sd_mod
modprobe st
```
• During scsi_load:

```bash
scsi0 : zfcp
    Vendor: IBM      Model: 2105F20      Rev: .107
    Type: Direct-Access  ANSI SCSI revision: 03
Attached scsi disk sda at scsi0, channel 0, id 1, lun 0
zfcp: FSF: zfcp_fsf_send_fcp_command_task_handler: status for SCSI Command:
    00000000 0000
zfcp: FSF: zfcp_fsf_send_fcp_command_task_handler: SCSI status code 0x2
00000000 00000000 0000202 00000000 00000020 00000000
70000600 00000018 00000000 29000000 00000000 00000000 00000000 00000000

SCSI device sda: 195312512 512-byte hdwr sectors (100000 MB)
dasdbm:(nonl)/       : dasdbm1
```
• Following scsi_load:

<table>
<thead>
<tr>
<th>Module</th>
<th>Size</th>
<th>Used by</th>
</tr>
</thead>
<tbody>
<tr>
<td>st</td>
<td>27696</td>
<td>0 (unused)</td>
</tr>
<tr>
<td>sd_mod</td>
<td>12048</td>
<td>0 (unused)</td>
</tr>
<tr>
<td>zfcp</td>
<td>345312</td>
<td>0</td>
</tr>
<tr>
<td>scsi_mod</td>
<td>61344</td>
<td>3 [st sd_mod zfcp]</td>
</tr>
<tr>
<td>nfsd</td>
<td>69648</td>
<td>4 (autoclean)</td>
</tr>
<tr>
<td>qeth</td>
<td>153072</td>
<td>1 (autoclean)</td>
</tr>
<tr>
<td>qdio</td>
<td>33968</td>
<td>2 (autoclean) [zfcp qeth]</td>
</tr>
<tr>
<td>ipv6</td>
<td>247472</td>
<td>-1 (autoclean) [qeth]</td>
</tr>
<tr>
<td>8021q</td>
<td>12928</td>
<td>0 (autoclean) [qeth]</td>
</tr>
<tr>
<td>ctc</td>
<td>49840</td>
<td>1 (autoclean)</td>
</tr>
<tr>
<td>fsm</td>
<td>1920</td>
<td>0 (autoclean) [ctc]</td>
</tr>
</tbody>
</table>
Partitioning using `fdisk`:

```
Disk /dev/sda: 255 heads, 63 sectors, 12157 cylinders
Units = cylinders of 16065 * 512 bytes

<table>
<thead>
<tr>
<th>Device</th>
<th>Boot</th>
<th>Start</th>
<th>End</th>
<th>Blocks</th>
<th>Id</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dev/sda1</td>
<td></td>
<td>1</td>
<td>3188</td>
<td>25607578+</td>
<td>83</td>
<td>Linux</td>
</tr>
<tr>
<td>/dev/sda2</td>
<td></td>
<td>3189</td>
<td>6376</td>
<td>25607610</td>
<td>83</td>
<td>Linux</td>
</tr>
<tr>
<td>/dev/sda3</td>
<td></td>
<td>6377</td>
<td>12157</td>
<td>46435882+</td>
<td>83</td>
<td>Linux</td>
</tr>
</tbody>
</table>

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```
Configuration

• Mounting devices
  - `ln -s /etc/init.d/scsi_mount /etc/init.d/rc3.d/S12scsimn`

```bash
#!/bin/sh
mount /dev/sda1 /FS/scsi01
mount /dev/sda2 /FS/scsi02
mount /dev/sda3 /FS/scsi03
mount /dev/sdb1 /FS/scsi04
mount /dev/sdb2 /FS/scsi05
```
• Unmounting devices
  – ln -s scsi_unmount rc3.d/K13scsium

#!/bin/sh
umount /FS/scsi01
umount /FS/scsi02
umount /FS/scsi03
umount /FS/scsi04
umount /FS/scsi05
**Configuration**

- Contents of `/proc/partitions`:

<table>
<thead>
<tr>
<th>major</th>
<th>minor</th>
<th>#blocks</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>97656256</td>
<td>dasdbm</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>25607578</td>
<td>dasdbm1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>25607610</td>
<td>dasdbm2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>46435882</td>
<td>dasdbm3</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>9765632</td>
<td>dasdbn</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>5121008</td>
<td>dasdbn1</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>4643840</td>
<td>dasdbn2</td>
</tr>
<tr>
<td>94</td>
<td>0</td>
<td>360000</td>
<td>dasda</td>
</tr>
<tr>
<td>94</td>
<td>1</td>
<td>359988</td>
<td>dasdal</td>
</tr>
<tr>
<td>94</td>
<td>4</td>
<td>2043360</td>
<td>dasdb</td>
</tr>
<tr>
<td>94</td>
<td>5</td>
<td>2043348</td>
<td>dasdl</td>
</tr>
</tbody>
</table>

➤ Suspect device naming due to IBM Partitioning being enabled.
• proc Filesystem has lots of SCSI information:

```
proc/scsi
  ├── scsi
  │    └── zfcp
  │         └── addmap
  │                 └── map
  │                           └── mod_parm
  │                               └── dev0xnnnn
  │       └── id0x00
  │           └── status
  │               └── id0x01
  │                   └── status
  │                       └── lun0x0
  │                                   └── status
  │                                           └── lun0x1
  │                                               └── status
  └── status
```
**Configuration**

- Contents of `/proc/scsi/scsi`:

<table>
<thead>
<tr>
<th>Attached devices:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Host</strong>: scsi0 <strong>Channel</strong>: 00 <strong>Id</strong>: 01 <strong>Lun</strong>: 00</td>
<td></td>
</tr>
<tr>
<td><strong>Vendor</strong>: IBM  <strong>Model</strong>: 2105F20  <strong>Rev</strong>: .107</td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong>: Direct-Access  <strong>ANSI SCSI revision</strong>: 03</td>
<td></td>
</tr>
<tr>
<td><strong>Host</strong>: scsi0 <strong>Channel</strong>: 00 <strong>Id</strong>: 01 <strong>Lun</strong>: 01</td>
<td></td>
</tr>
<tr>
<td><strong>Vendor</strong>: IBM  <strong>Model</strong>: 2105F20  <strong>Rev</strong>: .107</td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong>: Direct-Access  <strong>ANSI SCSI revision</strong>: 03</td>
<td></td>
</tr>
</tbody>
</table>
Contents of /proc/scsi/zfcp/dev0xe000/status:

FCP adapter

FCP driver $Revision: 3.60.4.1$ $ (or for cryptography's sake 0x0003003c)$

device number: 0xe000 registered on irq: 0x0013
WWNN: 0x5005076400c67100
WWPN: 0x50050764012006f6 S_ID: 0x611413
HW version: 0x0002 LIC version: 0x0000001a
FC link speed: 1 Gb/s FC service class: 3
FC topology: fabric
SCSI host number: 0x00000000

Attached ports: 2 QTCB size (bytes): 1696
Max SCSI ID of ports: 0x00000001 Max SCSI LUN of ports: 0x00000001
FSF req seq. no: 0x003d03c8 FSF reqs active: 16
Scatter-gather table-size: 57 Max no of queued commands: 4096
Uses clustering: 1 Uses New Error-Handling Code: 1
ERP counter: 0x00000000 Adapter Status: 0x5400006f

Adapter Structure information:
Common Magic: 0xfcfcfcfc Specific Magic: 0xaaaaaaaa
Adapter struct at: 0x0de25000 List head at: 0x0de25008
Next list head: 0x109d82b0 Previous list head: 0x109d82b0

Scsi_Host struct at: 0x0d868400
Port list head at: 0x0de250a0
Next list head: 0x0d869808 Previous list head: 0x0d868808
List lock: 0x00000000 List lock owner PC: 0x00000000
Configuration

• `/proc/scsi/zfcp/dev0xe000/status` (cont.):

```
O-FCP req list head: 0x0de250d0
Next list head: 0x0d89c608 Previous list head: 0x0d87a608
List lock: 0x00000000 List lock owner PC: 0x00000000

Request queue at: 0x0de250f8
Free index: 002 Free count: 128
List lock: 0x00000000 List lock owner PC: 0x00000000
current TOD: 13263832012677040160
time lock held: 35007809781

Response queue at: 0x0de25550
Free index: 072 Free count: 000
List lock: 0x00000000 List lock owner PC: 0x00000000

DEVICE INFORMATION (devinfo):
Status: "OK"
Control Unit Type: 0x1731 Control Unit Model: 0x03
Device Type: 0x1732 Device Model: 0x03
CIWs: 0x40720080 0x41830004 0x42820040 0x431b1000
0x441f0000 0x00000000 0x00000000

DEVICE INFORMATION (devstat):
Interrupt Parameter: 0x00000002 Last path used mask: 0x00
Channel Status: 0x80 Device Status: 0x00
Flag: 0x00000204 CCW address (from irb): 0x0d863a40
Response count: 0x00000000 Sense Count: 0x00000000
IRB: 0x00c0c0c9 0x0d863a40 0x00800000 0x00000000
Sense Data: 0x00000000 0x00000000 0x00000000 0x00000000
```
Configuration

- Contents of .../dev0xe000/id0x01/status:

<table>
<thead>
<tr>
<th>Port Information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWNN: 0x0000000000000000</td>
</tr>
<tr>
<td>WWPN: 0x5005076300c38c6d</td>
</tr>
<tr>
<td>SCSI-ID: 0x00000001</td>
</tr>
<tr>
<td>Max SCSI lun: 0x00000001</td>
</tr>
<tr>
<td>D-ID: 0x00614013</td>
</tr>
<tr>
<td>Handle: 0x00000025</td>
</tr>
<tr>
<td>Attached units: 2</td>
</tr>
<tr>
<td>ERP counter: 0x00000000</td>
</tr>
<tr>
<td>Port Status: 0x54000000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Port Structure information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Magic: 0xfcfcfcfcfc</td>
</tr>
<tr>
<td>Specific Magic: 0xbbbbbbbbbb</td>
</tr>
<tr>
<td>Port struct at: 0x0d869800</td>
</tr>
<tr>
<td>List head at: 0x0d869808</td>
</tr>
<tr>
<td>Next list head: 0x0d868808</td>
</tr>
<tr>
<td>Previous list head: 0x0de250a0</td>
</tr>
<tr>
<td>Unit list head at: 0x0d869820</td>
</tr>
<tr>
<td>Next list head: 0x0d869508</td>
</tr>
<tr>
<td>Previous list head: 0x0d869208</td>
</tr>
<tr>
<td>List lock: 0x00000000</td>
</tr>
<tr>
<td>List lock owner PC: 0x00000000</td>
</tr>
<tr>
<td>Parent adapter at: 0x0de25000</td>
</tr>
</tbody>
</table>
### Contents of \.../id0x01/lun0x0/status:

<table>
<thead>
<tr>
<th><strong>Unit Information:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SCSI lun:</strong></td>
</tr>
<tr>
<td><strong>FCP lun:</strong></td>
</tr>
<tr>
<td><strong>Handle:</strong></td>
</tr>
<tr>
<td><strong>ERP counter:</strong></td>
</tr>
<tr>
<td><strong>Unit Status:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Unit Structure information:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common Magic:</strong></td>
</tr>
<tr>
<td><strong>Specific Magic:</strong></td>
</tr>
<tr>
<td><strong>Unit struct at:</strong></td>
</tr>
<tr>
<td><strong>List head at:</strong></td>
</tr>
<tr>
<td><strong>Next list head:</strong></td>
</tr>
<tr>
<td><strong>Previous list head:</strong></td>
</tr>
<tr>
<td><strong>Parent port at:</strong></td>
</tr>
<tr>
<td><strong>SCSI dev struct at:</strong></td>
</tr>
</tbody>
</table>
31-bit Experiences

- Vanilla kernel not sufficient:
  - fdisk partitioning “forgotten”
  - Problems with mke2fs:
    
    ```
    daev005:/usr/src/linux/drivers/s390/scsi # mke2fs
    /dev/sda2
    mke2fs 1.19, 13-Jul-2000 for EXT2 FS 0.5b, 95/08/09
    mke2fs: No such device or address while trying to
determine filesystem size
    ```

- Required reconfiguration:
  - File Systems->Partition Types->MSDOS
31-bit Experiences

• Product testing:
  – Adabas databases
    • 100MB data files
  – AQA tables (SQL access to Adabas)
  – All worked as expected
31-bit Experiences

- Simplistic benchmarking
  - Used `dd` to exercise disk writes
  - Changed the underlying file system
    - `ext2`
    - `ext3`
    - `reiser`
  - Uncontrolled environment & unscientific methodology
    - CPUs shared with other LPARs
    - LPAR has low weighting
  - Results are indicative only
31-bit Experiences

Write Performance

![Bar chart showing write performance for ext2, ext3, and reiser file systems. The x-axis represents the file system types, and the y-axis represents MB/s or Time/record.]

```
dd if=/dev/zero of=/FS/scsi03/dummy.file bs=1024 count=43352000
```
31-bit Experiences

Erase Performance

<table>
<thead>
<tr>
<th>File System</th>
<th>Time in Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>ext2</td>
<td>50</td>
</tr>
<tr>
<td>ext3</td>
<td>70</td>
</tr>
<tr>
<td>reiser</td>
<td>60</td>
</tr>
</tbody>
</table>

Graph showing the performance of different file systems (ext2, ext3, reiser) over time (in seconds) for Eased, User, System.
64-bit Experiences

- Product Testing
  - Adabas & AQA as per 31-bit
  - Included Natural
  - All worked as expected
- Similar transfer rate found when doing simplistic "dd" test
64-bit Experiences

Ext3 Write Performance

Block Size

MB/s or Time/Record

512 1024 2048 4096 8192 16384 32768 65536

SysTime/s*10E+4
User/s*10E+6
MB/s
64-bit Experiences

Ext3 Erase Performance

<table>
<thead>
<tr>
<th>Block Size</th>
<th>Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>80</td>
</tr>
<tr>
<td>1024</td>
<td>80</td>
</tr>
<tr>
<td>2048</td>
<td>80</td>
</tr>
<tr>
<td>4096</td>
<td>80</td>
</tr>
<tr>
<td>8192</td>
<td>80</td>
</tr>
<tr>
<td>16384</td>
<td>80</td>
</tr>
<tr>
<td>32768</td>
<td>80</td>
</tr>
<tr>
<td>65536</td>
<td>80</td>
</tr>
</tbody>
</table>
Future Work

- Connection to EMC devices
- Exercising multi-pathing
- Testing copy services
  - Can’t be done from Linux
  - Use the web interface to ESS