



# **Linux on IBM @server pSeries SAN Overview for Customers**

*January 23, 2004*

## Abstract

This paper is intended to introduce IBM @server® pSeries® customers, IBM Business Partners, sales, marketing, and technical teams to pSeries Storage Area Network (SAN) with Linux. IBM's plans as put forth in this document are subject to change without notice.

## References and Prerequisite Reading

For an overview of Linux on pSeries, Emulex's Linux Fibre Channel device driver, and pSeries SAN solutions, the following background reading is recommended:

- The Linux for pSeries external Web site at <http://www.ibm.com/eserver/pSeries/linux>
- The IBM external Linux Web site at <http://www.ibm.com/linux>
- The IBM Linux Services Web site at <http://www.ibm.com/linux/services>
- The IBM SAN Storage Web site at <http://www.storage.ibm.com/ibmsan/>
- The IBM @server pSeries Server Web site at [http://publib16.boulder.ibm.com/pSeries/en\\_US/infocenter/base/](http://publib16.boulder.ibm.com/pSeries/en_US/infocenter/base/)
- The Emulex FC MultiPulse Device Driver version 4.30o\_M for IBM pSeries OEM-ed adapter (9802 and 9002L) at <http://www.emulex.com/ts/docoem/framibm.htm>
- Subsystem Device Driver (SDD) References
  - IBM Subsystem Device Driver User's Guide
  - Fault Tolerant Storage Multipathing and Clustering Solutions for Open Systems for the IBM TotalStorage® Enterprise Storage Server™ (ESS) is SG24-6295
  - IBM TotalStorage Enterprise Storage Server Implementing ESS Copy Services in Open Environments is SG24-5757
- FASTT References
  - IBM TotalStorage FASTT Storage Manager Concepts Guide, GC26-7560  
[ftp://ftp.software.ibm.com/pc/pccbbs/pc\\_servers/d5bajmst.pdf](ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers/d5bajmst.pdf)
  - IBM TotalStorage FASTT Storage Manager Copy Services User's Guide, GC26-7561  
[ftp://ftp.software.ibm.com/pc/pccbbs/pc\\_servers/fcs0mst.pdf](ftp://ftp.software.ibm.com/pc/pccbbs/pc_servers/fcs0mst.pdf)
  - For hardware reference manuals, check the documentation link from the FASTT Technical support site at <http://www.ibm.com/servers/storage/support/fastt/>
  - IBM TotalStorage: FASTT Best Practices Guide  
<http://publib-b.boulder.ibm.com/Redbooks.nsf/RedbookAbstracts/redp3690.html>

## Overview of Linux on pSeries and Storage Area Network (SAN)

The SAN hardware supported by pSeries servers running the Linux operating system is the same SAN hardware that is supported by pSeries servers running AIX 5L™. The one exception is the IBM 6227 Fibre Channel adapter, as that adapter had been discontinued since the implementation of this document.

There may be differences between AIX 5L SANs and Linux on pSeries SANs. This document is designed to explain how to implement a Linux on pSeries SAN.

**If a function is not explicitly stated to be supported in this document, it should not be assumed that it is supported; even if it is supported in an AIX 5L SAN.** The FAQ section below tries to address the major differences between the two SANs. However, the FAQ may not cover all the differences between the two.

## Overview of IBM pSeries Servers

This document covers all pSeries servers that support the Linux operating system and the 6228 and/or the 6239 Fibre Channel adapters. For server support, please refer to the pSeries documentation found at: [http://publib16.boulder.ibm.com/pSeries/en\\_US/infocenter/base/](http://publib16.boulder.ibm.com/pSeries/en_US/infocenter/base/)

Certain features of pSeries servers that are supported under AIX 5L are not supported under Linux. Hot-swapping of a Fibre Channel adapter is one such example. The Linux kernel does not have the capability to power off/on a PCI slot and integrate newly added hardware as dynamically as AIX 5L.

The term “host”, “server”, and “system” have been used in this document and can be interchanged. The meaning of these words in this document is a single operating system instance. This could be a standalone server, or a single logical partition (LPAR) from a pSeries, which is hosting multiple LPARs. Two independent LPARs on the same physical hardware are considered two hosts (servers, systems) and should be treated as such.

## Applicable SAN Environments for IBM’s pSeries Servers Running Linux

1. The Emulex adapter and driver are used in IBM @server iSeries™/pSeries with 64-bit PowerPC Architecture™ only. There is no support for older IBM RS/6000® or 32-bit PowerPC® machines.
2. Currently, the only supported Linux version is SuSE Linux Enterprise Server (SLES) 8 with SP3 applied.
3. Binary kernel versions:
  - a. 2.4.19-186 and later for Enterprise Storage Subsystem (ESS)
  - b. 2.4.19-219 and later for Fibre Array Storage Technology (FAStT)
4. For adapter placement and limitations please refer to the *PCI Adapter Placement Guide* (SA38-0538).
5. The Emulex device driver, version 4.30o\_M, should be used with either of Host Bus Adapters (HBA).
  - a. IBM FC6228
  - b. IBM FC6239
6. Disk subsystems
  - a. ESS Model 800 and F20
    - i. (R9e.8b030117) LIC EC ID 2.1.1.101 or above  
OR
    - ii. (R9p.8b030606a) LIC EC ID 2.2.0.459 or above
  - b. FAStT Model 200, 500, 600, 700, and 900 with IBM FAStT Storage Manager v8.3

7. Currently, there is no Storage Manager Client Software available for pSeries host running Linux. To configure the FAStT disk subsystems, an AIX 5L, Windows® or PC-based Linux server is necessary.
8. SAN Topologies
  - a. Direct connect, via loop mode only, to ESS and FAStT.
  - b. Switch (Fabric) connect to ESS and FAStT.
  - c. No loop via hub connection is supported.
  - d. No direct connection with multiple paths (LUNs seen by multiple direct connections) is supported.
9. What things can NOT interoperate in the SAN environment?
  - a. No IP support
  - b. No ESS support on HBAs configured for MultiPulse.
10. Device sharing
  - a. A device/LUN can be shared with other HBAs on the same host as long as some type of multipath control is in place and an application is not using disk reservations.
  - b. A device/LUN can be shared with different hosts as long as all those hosts are pSeries hosts with Linux and there are no reservations used.
  - c. A device/LUN can not be shared with different hosts in a heterogeneous environment (one or more servers are NOT pSeries running Linux).

## **Emulex's Linux on pSeries MultiPulse Driver Capabilities**

Emulex's Linux on pSeries MultiPulse driver provides single path and multipath capabilities for the FAStT disk subsystem, when configured accordingly. It also provides only single path capabilities for the IBM Enterprise Storage Server (ESS). Multipath capability for the ESS is handled by the ESS's multipath SDD software. The Emulex MultiPulse driver does not provide multipath capability for the ESS.

Below is a list of capabilities for the Emulex MultiPulse driver in a pSeries host running the Linux operating system. Some of these capabilities are restricted by entities outside of the driver, like the kernel or device restrictions.

- Up to 256 LUNs per server
- Up to a queue depth of eight
- Up to 32 LUNs per HBA
- FAStT RDAC support only
- No support for FAStT Auto Volume Transfer (AVT) mode
- Path Failover Capability for FAStT disk subsystems running with IBM FAStT Storage Manager V8.3 code.
- Multipath capabilities via Switch attach only.
- Manual path recovery (automatic path recovery not yet supported)

- FASSt FlashCopy® supported as long as Autofailback is not enabled
- Up to 16 Emulex HBAs per host. pSeries server limitations may restrict this further.
- Both ESS and FASSt subsystems are supported on a single host, provided each subsystem is accessed via it's own HBA (FASSt and ESS are not accessed via the same HBA)
- IBM OEM-ed HBAs are the only supported Fibre Channel adapter. Emulex Fibre Channel adapters that are not OEM-ed by IBM will not work with the Emulex MultiPulse driver.
- Dynamic FASSt RAID-type changes are supported (e.g. converting a RAID-0 array to a RAID-5, or changing the RAID segment size while I/O is active).
- Dynamic volume expansions are supported. However, the new space may not be available for use, depending on the filesystems used and the Linux version.
- Dynamic LUN/SAN changes, which are not listed above, are not supported.
- I/O must be halted prior to making the following FASSt changes:
  - host type (i.e. AIX 5L to LNXCL)
  - primary and secondary path preference swapping

## Different Driver Versions

Driver packaged with the Linux distribution (just load it and run):

- SLES8 GA: The Emulex Driver version 4.21f comes with the Linux distribution in SLES8 GA. This driver does not support MultiPulse (multi-path) devices.
- SLES8 SP1: The Emulex Driver version 4.22d comes with the Linux distribution in SLES8 Service Pack 1. This driver does not support MultiPulse (multi-path) devices.
- SLES8 SP3: The Emulex Driver in SP3 is the same driver version as SP1 (4.22d).
- MultiPulse: The Emulex Driver version 4.30o\_M is MultiPulse enabled (provides multi-path functionality) and will fix three critical problems for the older drivers such as:
  1. Kernel panic when load or unloading the driver in race condition.
  2. Performance improvement from 64KB to 128KB SCSI buffer size.
  3. MultiPulse multi-path functionality.

The latest Emulex FC MultiPulse Device Driver (version 4.30o\_M) for IBM pSeries OEM-ed adapter (9802 and 9002L) at <http://www.emulex.com/ts/docoem/framibm.htm>

## Before Installing Emulex's Linux FC Adapter Driver

1. Storage: allocate Volumes/LUNs to each HBA on the pSeries host running Linux.
2. Switch: properly zone storage devices and the pSeries host running Linux; DO NOT allow any storage that is not configured for the pSeries host running Linux in its zone.
3. Server: it is recommended that the cable not be connected between the host and storage device unless there is a proper configuration created for the storage, including a matching persistent binding configuration between the storage controllers and server HBAs compiled into the associated MultiPulse Emulex FC driver. You can load the Non-MultiPulse

Emulex FC driver to verify the WWN and note the automatically driver assigned controller and target identifiers (look at “/proc/scsi/lpfc/\*” files) to prepare for binding the intended MultiPulse or Non-MultiPulse storage configuration.

## Adding a new FC Adapter

1. You MUST shutdown the pSeries host running Linux to install new Emulex FC Adapter. Linux does not have support for powering off an individual PCI slot.
2. After installing the new Emulex adapter, you MUST verify the MultiPulse configuration and re-compile the driver before loading the MultiPulse driver.

## Quick Installation

There is no quick method for installation. Please follow the installation instructions below.

## Installing and Configuring the Emulex FC MultiPulse Driver

The directions below must be followed in the exact order to guarantee the driver is built, installed, and loaded correctly. Whenever “/etc/lpfc.conf” is modified manually, or updated by the use of “hbacmd” or “hbanyware”, the lpfcdd driver must be re-compiled with the resulting modified “/etc/lpfc.conf” file before the bindings become effective.

1. Create temporary directory to build Emulex Driver and copy the driver you want to use into that directory.
2. Install the SLES8 SP1 or SP3 kernel source code from SuSE distribution and also install the PowerPC GCC 64-bit compiler on the target pSeries host running Linux.
3. Insure you have the correct kernel source and build environment for Emulex MultiPulse Driver 4.30o\_M as detailed in FAQ.
4. Un-tar the Emulex MultiPulse driver at temporary Emulex directory
5. Install and build the Emulex MultiPulse driver as with no MultiPulse devices (this is default) configured as follow:

Change directory to where the Emulex Driver source code resides

```
#sh Install.sh
#make build
#make install
```

6. Load the driver, making sure that you have the storage and switch configured correctly. Load the Emulex driver with the command “modprobe lpfcdd”
  - a. Confirm that the driver loads correctly with the command “lsmod” and look for the driver name “lpfcdd”
  - b. NOW, you can connect the fibre cable to HBA and see if the Linux on pSeries is able to detect the switch and storage devices with the following commands:

```
#ls /proc/scsi/lpfc/      * check number of HBA installed
#cat /proc/scsi/lpfc/1   * check status of 1st HBA
#cat /proc/scsi/lpfc/2   * check status of 2nd HBA
```

```
#cat /proc/scsi/lpfc/3 * check status of 3rd HBA, ...
```

- c. If the driver is correctly built and loaded, you should be able to see the WWPN and WWNN of all connected storage devices listed in the /proc/scsi/lpfc/\* files. (Make sure that if you are using fabric zoning that the devices are where they are supposed to be.)
- d. Now double-check to see if all the storage subsystem with its WWNNs and WWPNs are seen by the HBA and their devices are configured for the pSeries host running Linux. If they are correct, you can unload and reload the driver quickly and ready to use it.

```
#rmmod lpfcdd * Unload the Emulex driver  
#modprobe lpfcdd * Load the Emulex driver
```

7. Now install and build the Emulex MultiPulse driver with multiple paths to the storage devices. **IMPORTANT: ALWAYS CONFIGURE PERSISTENT BINDINGS BY WWPN AND NOT BY WWNN NOR BY DOMAIN ID AS THE /etc/lpfc.conf DOCUMENTS AS OTHER BINDING METHODS (WWNN and DID binding does not work).**
  - a. After you are done with previous steps to install the Emulex driver as non-MultiPulse devices, you should have obtained all the information needed to rebuild the Emulex driver with MultiPulse device support.
  - b. Note all the WWNN and WWPN associated with the HBAs
  - c. Edit the /etc/lpfc.conf to bind the HBA to the Storage Devices with its WWNN or WWPN; OR
  - d. Use command line interface program /usr/sbin/hbanyware/hbacmd mpMenu <AdapterWWPN> to create the new /etc/lpfc.conf; (see section on hbacmd utility) OR
  - e. Use GUI interface program “/usr/sbin/hbanyware/hbanyware” to create the new /etc/lpfc.conf. (see section on hbanyware utility)
  - f. Rebuild the Emulex driver with make build then make install commands as outlined earlier from the local Emulex driver source code directory.
  - g. Load the Emulex driver via modprobe lpfcdd command.

## Driver Customization

Setting up consistent device numbering (implications of changing configurations and rebooting)

- Consistent device numbering only for Emulex binding parameters. The device instances may not consistent after booting or reconfiguration if SAN or storage changes have been made. However, there are free Linux utilities that may help provide consistent device numbering or consistent device instances, such as scsimap.

Use of persistent bindings feature in the MultiPulse driver is highly recommended

All ESS and FASSt devices may have persistent bindings configured for them. Persistent bindings help to insure that specific WorldWide PortNames are always given the same SCSI target id by the lpfcdd driver, which provides consistency and easier storage management.

It is highly recommended that you use persistent bindings, and also configuring the `/etc/lpfc.conf` file to manage only those storage targets with persistent bindings, by setting the “`lpfc_automap`” variable in `/etc/lpfc.conf` to equal 0 (i.e. `lpfc_automap=0`”).

Configuration of persistent bindings should be done before configuring MultiPulse routes and can be done one of 2 ways.

1. Modify the `/etc/lpfc.conf` file according to the instructions found therein
2. Use `hbacmd` utility’s “Persistent Bindings Menu”.

Choose “Bind Automapped Targets” which instructs the driver to bind the targets to adapter and target instances that it prefers to select. Doing this, each adapter must be automapped this way. Also, you must bind all targets by WWPN, or “By Port Name”. (The other methods will not work, which includes binding by World Wide Node Name or by Domain ID.)

After that is done, you can then choose “Delete Binding(s)” and remove those storage targets which you do not want. By selectively removing bindings, and setting the “`lpfc_automap`” variable in `/etc/lpfc.conf` to equal 0, you can control which storage devices the driver will manage, or in other words, you can control which targets your pSeries host running Linux will “see” and scan for at each boot, and each time the driver is loaded. This gives some control over driver load times and thus, boot times.

#### Persistent Bindings Notes:

- All bindings must be by WWPN.
- When manually editing `/etc/lpfc.conf` for bindings, first compile and load `lpfcdd` with the default `lpfc.conf` file, and use the adapter and target instances found in the `/proc/scsi/lpfc/*` files. During testing it has been observed that using target ids other than those “preferred” by the `lpfcdd` driver will not work.
- When using persistent bindings, the `/etc/lpfc.conf` file must be manually edited and the variables “`lpfc_automap`” and “`lpfc_scandown`” set equal to zero (“0”).
- Remember that each change to the persistent binding configuration requires the `lpfcdd` driver to be re-compiled. Otherwise, the changes will have no effect.

Configuration of MultiPulse routes can be done one of 3 different ways.

1. Manually edit the `/etc/lpfc.conf` file according to the instructions found therein.
2. Use the `hbacmd` utility’s MultiPulse menu, using the command:

```
# /usr/sbin/hbanyware/hbacmd mpMenu <AdapterWWPN>
```

The WWPN being the adapter Portname which can be found in any of the `/proc/scsi/lpfc/*` adapter configuration files.

Within the MultiPulse menu choose “1. Create Route”

select the LUN for which you want multi-path support, for example “1. `lpfc0t1d0`”

select the path to include in the route, for example “1. `lpfc0t1d0`”



select the second path to include in this route, for example "2. lpfc2t4d0"  
 select  
 choose "y" to enable AutoFailBack; (Unless using FlashCopy: see section on FAStT copy services.)  
 choose "n" for InquiryHeartBeat, FailoverAbortTaskSet, FailoverLunReset, and FailoverTgtReset (Use of these features is not supported yet.)  
 select "1. Simple Failover" for the control type  
 select the path you want to be the "primary path", or preferred path for I/O.  
 For example "2. lpfc0t1d0" will cause I/O to run to that device when driver is loaded, regardless of how you have configured the FAStT Preferred Controller Ownership for that volume using the SMclient graphical user interface.  
 (Hint: If you alternate the paths here you will essentially load-balance your I/O.)

answer "Y" to "Are you sure you want to create this route?"  
 After the configuring the first few routes, you can make this shorter work by typing in all of the responses at one time from the main MultiPulse menu. For example, I could have accomplished all of the above steps by doing the following:

Within the MultiPulse menu enter "1" for "Create Route", then after the target choices are displayed, at the "Select a target =>" prompt, enter "1 1 2 0 y n n n n 1 1 y" and it will pick the next route available and configure the paths the same as going through each step individually as outlined above. Be careful using this method!  
 Log it!

3. Use the hbanyware graphical user interface, starting it using the command below, making sure that your DISPLAY environment variable is correctly set:

# /usr/sbin/hbanyware/hbanyware

In resulting GUI:

Find HBA WWN in "Discovered Elements" list

click "+" symbol next to storage icon with WWNN you want to manage, then

click on "LUN 0" -> "MultiPulse" tab

In the MultiPulse tab section, click on the "Avail. Path ID:" for the path you want to be the primary path. This should be the path to the controller which is the preferred owner of the volume mapped to this LUN instance. and then click "Create MP Route" check box.

click "Apply Checked Item" button to finish configuring this route.

Verify the route is configured by clicking on the "+" symbol next to the storage icon which represents the alternate controller which also owns this LUN, and check the "MultiPulse" tab. This route should be represented.

Do this for each LUN the controller owns, making sure that the first path you add by clicking on "Avail. Path ID:" is the desired primary path.

### Commonly customized parameters found in "lpfc.conf"

lpfc\_fcp\_bind\_WWNN[]=" list of Storage WWNN"

```

lpfc_fcp_bind_WWPN[]=" list of Storage WWPN"
lpfc_fcp_DID[]=" list of DID of Storage subsystem"
lpfc_automap=0;                * Set to 0 for persistent binding
lpfc_scandown=0;              * Set to 0 for persistent binding
lpfc_MultiPulse_entries=xx    * MUST match number of entries in
                               * lpfc_MultiPulse_fcp[]
lpfc_MultiPulse_fcp[]=" list of LUNs for binding to WWNN, WWPN"
lpfc_linkdown_tmo=1;
lpfc_nodev_tmo=1;

```

A useful parameter in /etc/lilo.conf that will help to limit number of LUNs scanned at initialization, which for smaller installations of storage will cut device scanning time at driver load time, which reduces the time to boot the Linux kernel.

```
append = "max_scsi_lun=16" * limit SCSI scan up to 16 devices
```

For further advanced configuration, please refer to the Emulex driver documentation.

## Remove or Reinstall the Driver

To remove (unload) the driver, issue `"rmmod lpfcdd"`

To reload the driver, issue `"modprobe lpfcdd"`

To reinstall the driver, follow the section above on how to build and install the Emulex driver.

## Setting up Multiple Paths for FAStT Storage

The Emulex MultiPulse driver only supports multipath for FAStT storage subsystems. There are different classes of FAStT controllers; all of the FAStT Model 200, 500, 600, 700, and 900 are supported. Below is a list of considerations:

- The FAStT firmware required for multi-path support is with IBM FAStT Storage Manager V8.3.
- The Host Type for FAStT Linux on pSeries connections is LNXCLS.
- Connect and configure multiple paths to a FAStT using the IBM FAStT Storage Manger Client.
- Up to two paths to a FAStT are supported, no more than one per controller.
- If you use only FAStT and Multipath Devices, make sure to set `lpfc_automap=0` and `lpfc_scandown=0`.
- Do not configure multipath for ESS devices using the Emulex MultiPulse driver. Make sure to set `lpfc_automap=2` and `lpfc_scandown=0`.

## Path Failures

FAStT Environments:

If a FAStT volume it is configured as a MultiPulse device, and the primary path fails, the MultiPulse driver will automatically failover to the secondary path.

When the primary path has been recovered and is online, manual intervention is required to use “hbanyware” to failback the device to primary path. If all paths are failing, the Emulex driver will report I/O error back to the upper SCSI layer.

Recovery from failed paths and offline paths is the same. There is no difference between failed paths and offline paths.

#### ESS Environments:

There is no multipath support for ESS with the Emulex driver. The ESS device will be treated as single path device. When the path fails, the Emulex driver will report an I/O error to upper SCSI layer.

### **Installing and Booting from a SAN Device**

SAN Installation and booting has not been fully tested and is not currently supported by pSeries servers running Linux.

### **Copy Services**

#### ESS

There are no limitations on ESS copy services beyond the ESS copy services limitations defined for i386 Linux attached servers. Refer to the ESS copy services document [SG245757](#) for information on ESS copy services in a Linux Environment

#### FAStT

Other than the following notes on Path Failover/Failback with FlashCopy there are no limitations on FAStT copy services beyond the FAStT copy services limitations defined for i386 Linux attached servers. Refer to the *FAStT Storage Manager Copy Services User's Guide* for information on FAStT copy services in a Linux Environment

There are two types of failover/failback scenarios for FAStT FlashCopy:

1. Failover/Failback of an entire path

The Failover/Failback of a path is straightforward and works as expected. See the later section in this document titled “Path Failures” for more information on this type of Path Failover/Failback.

2. Failover/Failback of a LUN (LUN ownership)

If LUN Ownership is changed for a FlashCopy Base volume the FlashCopy Logical volume LUN ownership will also change. In order to change the ownership back to the original configuration all I/O to the FlashCopy Logical Volume must be stopped while you manually reconfigure the LUN ownership and put the paths to the FlashCopy Logical Volume online.

Auto Failback should not be used with paths that are involved with FlashCopy Logical Volumes.

## Helpful Commands and Tools

```
cat /proc/scsi/scsi // show SCSI devices
ls /proc/scsi/lpfc/ // show Emulex adapter number
cat /proc/scsi/lpfc/1 // show detail information of Emulex adapter number 1.
cat /proc/scsi/lpfc/2 // show detail information of Emulex adapter number 2.
fdisk // utility to show, create, format disks
yast2 // GUI interface for SLES8 to manage the system
lputil // Emulex command line utility to diagnostic adapters.
hbacmd // Emulex command line utility to diagnostic adapter,
// and configure MultiPulse devices.
HBAnyware // Emulex GUI utility to diagnostic adapter, and
// configure MultiPulse devices.
```

# Frequently Asked Questions

## **Q: Why am I missing some, or all, of the LUNs/devices on my storage subsystem?**

A: If you do not see any working LUNs/devices, you should check the following:

1) You may not have installed and loaded the Emulex driver. Check if the Emulex driver is loaded by using “lsmod” to list all the driver modules that were loaded. If you see the lpfcdd entry on the list, the Emulex driver was loaded. Otherwise, you can use “modprobe lpfcdd” command to load the driver.

2) If you have loaded the driver properly and it is shown as in “lsmod” list, check the following SAN configuration issues:

- fibre cable is properly connected to Emulex HBA, to Switch, and to storage subsystem.
- switch is properly zoned to recognize the Emulex HBA WWNN and storage WWNN
- storage subsystem has configured LUN/devices and associated them with the Emulex HBA WWNN/WWPN
- For **ESS** subsystem: Host type is set to i/Linux on pSeries for the LUNs
- For **FAST** subsystem: Host type is set to “LNXCL” (Linux Cluster).

3) If you do not have LUN0 defined, you need to specify lpfc\_missing\_lun0=1 in lpfc.conf.c before you build and load the Emulex driver. You can also load the parameter at the time of loading the driver as “modprobe lpfcdd lpfc\_missing\_lun0=1”.

4) If you have non-consecutive LUNs defined (i.e. LUN0, LUN2, LUN3 ... and do not have LUN1 defined), you need to specify lpfc\_lun\_skip=1 in lpfc.conf.c before you build and load the Emulex driver. You can also load the parameter at the time of loading the driver as “modprobe lpfcdd lpfc\_lun\_skip=1”.

5) Check to see if the driver can detect the switch and ESS storage subsystem properly by using these following commands.

```
#ls /proc/scsi/lpfc          // check if driver sees any Emulex adapter
#cat /proc/scsi/lpfc/(n)    // list if driver sees any target devices
```

- If it shows that the link is down, it means there is cable problem between Emulex HBA and Switch, or bad cable.
- If it shows that the link is up but did not show any target WWNN/WWPN, it means there is cable problem between Switch and ESS storage subsystem; or bad cable; or incorrect Switch zone setting.
- If it shows that the link is up and target WWNN/WWPN, it means there is incorrect configuration of host type; or Emulex HBA WWNN associated with LUNs/devices; or missing LUN0 defined and the lpfc\_missing\_lun0 was set to 0.

## **Q: Why is there such a long delay when loading the Emulex Driver?**

A: If there is a device connection seen by the pSeries host running Linux, which does not have any LUNs configured, the Emulex driver may experience long time outs as it tries to find a device. To clear up this issue, make sure each connection seen by the HBA has at least one device configured.

**Q: Why do I see more devices than are configured?**

A: If there is a device connection seen by the pSeries host running Linux, which does not have any LUNs configured, the Emulex driver may attempt to configure a LUN for each LUN ID. These are ghost devices and are not valid drives. To clear up this issue, make sure each connection seen by the HBA has at least one device configured.

**Q: Why does the system panic when I try to unload the driver?**

A: If there is a device connection seen by the pSeries host running Linux, which does not have any LUNs configured, the Emulex driver may cause a kernel panic. To clear up this issue, make sure each connection seen by the HBA has at least one device configured.

**Q: How to I assign FASTT devices to a pSeries host running Linux?**

A: Currently, there is no FASTT Storage Manager Client Software available for pSeries host running Linux. To configure the FASTT disk subsystems you will need to use an AIX 5L, Windows or i386 processor-based Linux server, and run a version of the FASTT Storage Manager client that is supported on that server.

**Q: How do I configure MultiPulse devices (multipath devices) for a pSeries host running Linux using the Emulex FC MultiPulse driver?**

A: When you already have the FASTT configured with MultiPulse devices, you can load the Emulex FC MultiPulse driver as single path device driver. This single path driver will be able to detect the MultiPulse device as two independent, separate devices. Using “hbacmd” or “hbanyware”, you can start configure these two independent, separate devices to be a MultiPulse device. The “hbacmd” or “hbanyware” will create a new /etc/lpfc.conf with matching configuration for a MultiPulse device. You now can rebuild the new Emulex FC MultiPulse driver with the new /etc/lpfc.conf. After you build the driver with new /etc/lpfc.conf, you can unload the current single path driver and reload the new driver to use the device as MultiPulse device.

**Q: Why is the MultiPulse driver not managing paths?**

A: If the MultiPulse driver is not managing paths, the devices were probably not configured for MultiPulse operation. Refer to the question above on how to set up MultiPulse operation.

**Q: Can data be shared across adapters and/or servers?**

A:

1. A device/LUN can be shared with other HBAs on the same host as long as some type of multipath control is in place and disk reservations are not used.
2. A device/LUN can be shared with different hosts in a homogeneous environment (ALL servers are Linux on pSeries based) as long as there are no disk reservations.
3. A device/LUN can not be shared with different hosts in a heterogeneous environment (one or more servers is NOT Linux on pSeries based).

**Q: Where do I get the driver?**

A: You can download the latest Emulex FC MultiPulse Device Driver (version 4.30o\_M) for IBM pSeries OEM-ed adapter (9802 and 9002L) at <http://www.emulex.com/ts/docoem/framibm.htm>

**Q: How do I build the driver for SuSE SLES 8 SP1 for ppc64?**

A: You should follow the instructions in the section above on how to build the driver. However, here is a quick list of the necessary operations:

Install the kernel source and ppc64 C++ compiler matching with current binary kernel version. To check for current kernel in use, you can use “uname -a” for active kernel in use; and use “rpm -qa | grep kernel” for current kernel binary and source installed on your system. To install a kernel source rpm, you can use rpm command “rpm -ivh kernel-source-2.4.19-SuSE-xxx.ppc.rpm”. To upgrade a kernel source rpm, you can use rpm command “rpm -Uvh kernel-source-2.4.19-SuSE-xxx.ppc.rpm”.

#### Build kernel dependency for driver build

```
#cd /usr/src/linux
#cp /boot/*.version /usr/src/linux/include/linux/version.h
#sh make_ppc64.sh cloneconfig dep
```

#### Download the driver and un-tar it to local directory.

```
#tar -xvf lpfc-4.30o_M-1_gcc3x.ppc64.tar
```

#### Build driver from this local directory

```
#sh Install.sh
#make build
#make install
```

#### Load the driver

```
#modprobe lpfcdd
```

### **Q: How do I build the driver for SuSE SLES 8 SP3 for ppc64?**

A: You should follow the instructions in the section above on how to build the driver. However, here is a quick list of the necessary operations:

Install the kernel source and ppc64 C++ compiler matching with current binary kernel version. To check for current kernel in use, you can use “uname -a” for active kernel in use; and use “rpm -qa | grep kernel” for current kernel binary and source installed on your system. To install a kernel source rpm, you can use rpm command “rpm -ivh kernel-source-2.4.21-xxx.ppc.rpm”. To upgrade a kernel source rpm, you can use rpm command “rpm -Uvh kernel-source-2.4.21-xxx.ppc.rpm”.

#### Build kernel dependency for driver build

```
#cd /usr/src/linux
#make mrproper
#cp /boot/*.version /usr/src/linux/include/linux/version.h
#cp arch/ppc64/defconfig.pSeries64 .config
#sh make_ppc64.sh oldconfig dep
```

#### Download the driver and un-tar it to local directory.

```
#tar -xvf lpfc-4.30o_M-1_gcc3x.ppc64.tar
```

#### Build driver from this local directory

```
#sh Install.sh
#make build
#make install
```

#### Load the driver

```
#modprobe lpfcdd
```

**Q: I follow the instruction to build the driver but when I load the driver, it gets error messages: “kernel version mismatch” and the driver is not loaded. What should I do?**

A: If you see the kernel version mismatch error, check the following:

- what is the active kernel in use by issue “uname -r”;
- what is the defined version in kernel source by issuing  
`#cat /usr/src/linux/include/linux/version.h`
- If it does not match the “uname -r” version, then make sure that the kernel dependency was built correctly.

After you have matching kernel version from “uname -r” and version.h; you should be able to rebuild, install and load the driver correctly.

**Q: When I load the driver via “modprobe lpfcdd”, it takes a long time and seems like it never loaded. What should I do?**

A: When you don’t have a storage device configured, but you already have connected all of the cables, the driver will take a long time trying to discover devices on the SAN network which are not there or unknown. It may also try to communicate with phantom or unknown devices and will not receive a response, which will cause timeout, and/or retry. You should disconnect the fibre cable at the Emulex HBA, then wait until the driver finishes loading. Then, follow the instructions from previous chapter on how to configure and install the driver.

**Q: When I load the driver via “modprobe lpfcdd”, the system panics. What should I do?**

A: You should make sure that you have a correct kernel installation. Do not use unknown kernel images or those built on a different system. If you have a good Linux kernel, make sure that you have correctly installed the Emulex driver. If you have a good Emulex driver installed, make sure that you have the latest Emulex adapter firmware, and only use IBM OEM-ed adapters.

**Q: When I unload the driver via “rmmod lpfcdd”, the system panics. What should I do?**

A: You should make sure that the driver has been built with these parameter settings in lpfc.conf.c as `lpfc_missing_lun0=1` and `lpfc_lun_skip=1`, especially when you have a storage subsystem connected to the SAN network but is not configured for a pSeries host running Linux.

**Q: I have built everything correctly, but I still cannot get the adapter configured.**

A: Make sure the adapter is an IBM OEM-ed adapter. The Emulex MultiPulse driver has been designed to work with IBM OEM-ed adapters only.

**Q: Shouldn’t the performance be better than I am seeing?**

A: The Linux kernel has some limitations when it comes to SCSI based protocol devices. Mainly those drivers must be configured to operate on a single CPU only and have a buffer limitation of 128K. Tuning general SAN components, like queue depth, may also help performance.

**Q: Can I hot swap a Fibre Channel adapter like I can on AIX 5L?**

A: No. Linux does not support PCI slot power control, which is needed for hot swapping of PCI adapters.

**Q: Are there any known issues and their workarounds with the Emulex driver?**

A: Here is a list of the known Emulex driver issues:

1. (T082) FlashCopy does not work properly with autofailback. Needs to disable autofailback and manual intervention from user with `hbacmd` or `hbanyware` to move secondary copy back.
2. (T068) The Emulex 4.30o\_M MultiPulse Driver only supports up to 16 Emulex HBAs per pSeries host running Linux.



3. (T080) If you build the MultiPulse driver with predefined MultiPulse devices in lpfc.conf and do not have matching storage network devices matching; you will not be able to load the driver correctly. You should build the driver with default lpfc.conf; then load this driver and use “hbacmd” or “hbanyware” to configure MultiPulse devices with this driver; after you have create a new lpfc.conf, you can build and load this new MultiPulse driver.
4. (T078) If you try to unload the driver and receive error: “<3>Attached usage count = -1”; you should rebuild the driver with lpfc\_missing\_lun0=1 and lpfc\_lun\_skip=1 in lpfc.conf.c; or load the driver as “modprobe lpfcdd lpfc\_missing\_lun0=1 lpfc\_lun\_skip=1”.
5. (T077) If you run application and receive kernel panic with “PCI\_DMA: TCE allocation failure in create\_tcsg”; you should build the driver with lpfc\_lun\_queue\_depth with smaller value in lpfc.conf.c, lpfc\_lun\_queue\_depth=8 is recommended; or load the driver as “modprobe lpfcdd lpfc\_lun\_queue\_depth=8”.
6. (T067) (T075) If you are using SuSE SLES8 SP1 kernel and experience system hung unresponsive but you still can ping to the system; you should build the driver with lpfc\_lun\_queue\_depth with smaller value in lpfc.conf.c, lpfc\_lun\_queue\_depth=8 is recommended; or load the driver as “modprobe lpfcdd lpfc\_lun\_queue\_depth=8”. If you have more than 32 LUNs/devices configured for an Emulex HBA, you should only have less than 32 LUNs/devices in used concurrently at a time. It is best to have 32 or less LUNs/devices per Emulex HBA.
7. (T074) If you are unable to configure MultiPulse devices for FASSt storage subsystem with MultiPulse Emulex HBAs, make sure that you have IBM OEM-ed adapters. The Emulex MultiPulse driver only supports IBM OEM-ed adapters on pSeries.
8. (T072) (T073) If you experience an application, using MultiPulse devices, running but no active I/O data transfer, you may run into problem with switch port failing with constant link down/up or RSCN and does not failover to secondary path; you should check the switch port for errors. If you determine that the switch port is indeed having the problem, you can 1) disconnect the fibre cable from the bad port and move to different working port; 2) disconnect the fibre cable from the bad port; 3) disable the primary path via “hbacmd” or “hbanyware” utilities.
9. (T071) If you see lot of messages “!lpfc(x):0756:FPw:Local\_Timeout” in /var/log/messages; you should check to make sure that the driver have lpfc\_log\_verbose set to 0 in lpfc.conf.c; or load the driver as “modprobe lpfcdd lpfc\_log\_verbose=0”.
10. (T069) If you load the driver and receive error “bad virtual addr” in /var/log/messages; you should make sure that you have enough free memory available in the system. You should check if there are too many processes running and you are running out of memory. Use “cat /proc/meminfo” to see how much memory available; you can also use “top” to see how your system is running.
11. (T060) If you experience a long delay after you have closed or killed the application using I/O devices via Emulex adapters; you should check to see if there are any error messages posted in /var/log/messages; these errors may cause long delay in cleaning up the memory and I/O request queues before the application is actually closed.
12. (T062) If you want to configure FASSt with host type Linux and RDAC enable but the default was AVT; you should configure the host type as LNXCL (Linux Cluster); or you can use script to enable RDAC for host type Linux.
13. (T041) If you are unable to build and install Emulex driver on SuSE 2.6 kernel or Red Hat; the current Emulex 4.30o\_M driver only support SuSE 2.4 – SLES8 SP1 or SP3.
14. (T080) If you do not see any path or route after loading a MultiPulse driver; probably the MultiPulse driver does not match the current SAN configuration (i.e. the SAN configuration has changed, a LUN/device has been removed, a path is offline, a fibre cable has been disconnected. You can use “hbanyware” to reconfigure the MultiPulse devices or you can Unload the current driver, Reload the single path driver, and Rebuild a new MultiPulse device configuration and then load this new driver.
15. (T081) The system is panic when you try to unload the driver with 16 HBAs. This problem is still under investigation and will be fixed in next release.
16. (T084) Swapping primary and secondary paths preferences while I/O active does not work. This problem is known and will be fixed in next release. The current solution is to stop all I/O before swapping primary and secondary paths.
17. (T085) No MultiPulse devices entries available on primary HBA is bad. This problem is still under investigation and will be fixed in next release. The current solution is to replace the bad HBA.
18. (T086) When failback occurs, one device or more MultiPulse devices were inaccessible because the primary paths for these MultiPulse devices were offline. This problem is known and under investigation.

The current solution is to stop I/O to the failing devices and use “hbanyware” to enable the path that was offline.

19. (T079) When failback, the system is panic with “kernel panic: NULL devp in flush\_done”. The problem is fixed in 4.30o\_M driver. However, if you are still experience the problem. You may have the FASiT and ESS sharing the same HBA in MultiPulse device configuration. You should move the ESS devices to different HBA.



© IBM Corporation 2004

IBM Corporation  
Marketing Communications  
Systems Group  
Route 100  
Somers, New York 10589

Produced in the United States of America  
January 2004  
All Rights Reserved

This document was developed for products and/or services offered in the United States. IBM may not offer the products, features, or services discussed in this document in other countries. The information may be subject to change without notice. Consult your local IBM business contact for information on the products, features and services available in your area.

All statements regarding IBM's future directions and intent are subject to change or withdrawal without notice and represent goals and objectives only.

IBM, the IBM logo, the e-business logo, @server, AIX 5L, iSeries, PowerPC, PowerPC Architecture, pSeries, and xSeries are trademarks or registered trademarks of International Business Machines Corporation in the United States or other countries or both. A full list of U.S. trademarks owned by IBM may be found at <http://www.ibm.com/legal/copytrade.shtml>.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Windows is a registered trademark of the Microsoft Corporation.

Intel and Itanium are registered trademarks of Intel Corporation in the United States and/or other countries.

Other company, product, and service names may be trademarks or service marks of others.

IBM hardware products are manufactured from new parts, or new and used parts. Regardless, our warranty terms apply.

This equipment is subject to FCC rules. It will comply with the appropriate FCC rules before final delivery to the buyer.

Information concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of the non-IBM products should be addressed with the suppliers.

All performance information was determined in a controlled environment. Actual results may vary. Performance information is provided "AS IS" and no warranties or guarantees are expressed or implied by IBM.

The IBM home page on the Internet can be found at <http://www.ibm.com>.

The pSeries home page on the Internet can be found at <http://www.ibm.com/servers/eserver/pSeries>.