NPIV
FCP Channel Virtualization in a Linux Environment

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

- What is NPIV?
- New Possibilities
- Requirements
- Linux Support
- Getting Started
- Look 'n Feel
Terms and Definitions

- **NPIV**
  - *N-Port ID Virtualization*
  - *IBM System z9 only!*

- **Channel**
  - *Physical FCP adapter*

- **Subchannel**
  - *Logical FCP adapter*
  - *Defined within IOCDS or z/VM*

- **HBA**
  - *Host Bus Adapter (channel or subchannel, context-dependent)*

- **FC SAN**
  - *Fiber Channel Storage Area Network*
Terms and Definitions (cont.)

- **WWPN**
  - Worldwide Port Name
  - Worldwide unique identifier of a FC port
  - „burned in“ during manufacturing
  - 8 byte hexadecimal value
  - E.g. 0x5005076401e06b18

- **D_ID**
  - Destination ID
  - Used internally for active FC connections instead of WWPNs
  - 3 byte hexadecimal value
  - Assigned by FC switch
  - E.g. 0x2a036bc
WWPNs and other Ports

- Many WWPNs in a FC SAN
- Only target WWPN is important
What is NPIV?

- Initiator Port
  - has attributes like:
    - WWPN
    - D_ID
    - PCHID
    - CHPID
    - ...

- Without NPIV: All channel attributes will be duplicated for each subchannel
What is NPIV?

“Think Virtual Adapters”

concept:
virtual HBAs (subchannels) hosted by a physical HBA (channel)
What is NPIV?

**System z9**

**Linux A:**
- **shared ID without NPIV:**
  - WWPN `xx.xx.....xx`
  - D_ID `xx.xx.xx`
- **unique ID with NPIV:**
  - WWPN `aa.aa.....aa`
  - D_ID `aa.aa.aa`

**Linux B:**
- **shared ID without NPIV:**
  - WWPN `xx.xx.....xx`
  - D_ID `xx.xx.xx`
- **unique ID with NPIV:**
  - WWPN `bb.bb.....bb`
  - D_ID `bb.bb.bb`

**Unique SAN Identities!**

**without NPIV:**
- The SAN sees a shared FCP channel as a single initiator.

**with NPIV:**
- Initiators of I/O and their traffic can be distinguished in the SAN through unique WWPNs or D_IDs respectively.
What is NPIV?

Industry-Standard Solution

- NPIV = N-Port Identifier Virtualization
- Standard-based approach being embraced by the industry
- System z9 persistently assigns unique WWPN to each FCP subchannel
- FCP Channel obtains separate D_ID for each subchannel from fabric switch
- result: unique SAN identity for each FCP subchannel

see also Fiber Channel - Framing and Signaling Interface (FC-FS) http://www.t11.org
What is NPIV?

System z9 Implementation

- Connecting a virtual HBA to the SAN

System z9

- virtual Server A
- virtual HBA: subchannel 1
  - WWPN aa.aa......aa
  - D_ID aa.aa.aa

- virtual Server B
- virtual HBA: subchannel 2
  - WWPN bb.bb......bb
  - D_ID bb.bb.bb

SAN

- shared FCP channel
- FC switch
- FCP device
- Zoning
- LUN masking/mapping
What is NPIV?

**System z9 Implementation**

- **Fabric Login**
- **Fabric Logout**

**WWPN a of subchannel A**

- D_ID a'
- D_ID a"

**Fabric Login**

**Fabric Logout**

**WWPN b of subchannel B**

- D_ID b'
- D_ID b"

**IOCDS with A**

**dynamic I/O:**

- B added
- B removed

Rewrote and re-IML-ed IOCDS

A gone, B back

- **Lifetime of virtual WWPN = lifetime of subchannel definition**
- **D_ID lifetime = fabric connection lifetime**
What is NPIV?

System z9 Implementation

- WWPN range assigned to IBM
- System z9 machine identification
- Subchannel identification (derived from LPAR, CHPID, subchannel set, CSS ID)

= WWPN of a virtual HBA in a System z9

- well-defined “ingredients” tie a unique WWPN permanently to a subchannel
What is NPIV?

System z9 Implementation

- Migrating Linux changes WWPNs

**without NPIV**

- Linux
  - WWPN A
  - FCP subchannel a

**with NPIV**

- Linux
  - WWPN A
  - FCP subchannel a

**SAN knows about WWPN A only**

**without NPIV**

- Linux
  - WWPN A
  - FCP subchannel b

**with NPIV**

- Linux
  - WWPN A
  - FCP subchannel b

**SAN knows about WWPN A and WWPN B**
New Possibilities

'Exclusive LUN' policy eliminated

- Many-to-many tape backup solutions
New Possibilities

'Exclusive LUN' policy eliminated

- Shared SAN file systems
New Possibilities

Access Control Done Right

- LUN Masking works for virtual servers!
New Possibilities

- Zoning works as well!

Access Control Done Right

- zone “green”
- zone “red”
- zone “blue”
New Possibilities

Access Control Feature

- NPIV deprecates the FCP Channel Access Control feature

- Access Control feature still available with System z9, (could be used for subchannels operating in backward-compatible non-NPIV mode)

- ACT rules – if defined – are not applied to FCP subchannels in NPIV mode
Requirements

- NPIV is available on System z9 servers.
  - FICON Express 2 adapter running with MCL003 on EC J99658
- z/VM
  - z/VM 5.2
  - z/VM 5.1 with the PTF for APAR VM63744
- Linux Distribution
  - Currently SLES9 SP3 and SLES10 (LPAR mode or z/VM)
- NPIV-Capable Switch
  - only required for switch adjacent to z9
  - Mostly firmware upgrades possible (e.g. McData, Brocade)
NPIV is more or less transparent for operating systems

- Linux uses the new virtual N-Port in the same way as it has used non-virtual N-Ports

- But: some new error codes/messages defined for NPIV-type subchannels (e.g. if you run out of D_IDs)

- Linux code shipped with SLES9 SP3 and SLES10

---

**see also**

- Introducing N_Port Identifier Virtualization for IBM System z9

- Linux on zSeries
Getting Started

“Floods” of WWPNs

- Many new WWPNs to be used by zoning and LUN masking/mapping functions
- Can be exported from SE through FTP
Getting Started

1) Pre-plan SAN with NPIV support
   • see practical limits of components
2) Define FCP subchannels in IOCDS
   • prior to IML, or
   • using dynamic I/O (HCD/HCM)
3) Perform IML, if needed
   • WWPNs for new subchannels get assigned
4) Query WWPNs using SE/HMC panel
   • needed for configuration of SAN functions
   • export function through FTP available
Getting Started

5) Configure switch adjacent to z9
   • ensure NPIV is enabled
   • ensure enough virtual N-Ports per port
   • setup zoning for virtual N-Ports

6) Configure target device
   • setup LUN masking/mapping

7) Enable NPIV-mode for CHPID in LPAR
   • CHPID must be temporarily toggled off

8) Start using FCP subchannel in Linux
   • check for NPIV related error messages
Getting Started

 Implementation Limits

- Switches will presumably allow for 1024 or up to 8192 WWPNs in a SAN.
- Storage device impose limits as well.
- Theoretically up to 255 subchannels per channel connected at the same time.
- \( \leq 510 \) active target port connections for all subchannels of a channel.
- Each System z9 provides a total of 2 million WWPNs for virtual HBAs.

See also: Support of Fiber Channel Protocol for SCSI - FCP channels

[http://www-03.ibm.com/systems/z/connectivity/fcp.html](http://www-03.ibm.com/systems/z/connectivity/fcp.html)
Hints and Tips

- Do not use more than 32 subchannels per physical channel in NPIV mode.
- Zone each NPIV WWPN individually. This can reduce fabric traffic.
- Consider using multipathing (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 subchannel 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.
- Your choice:
  - toggle NPIV on with a per channel-and-LPAR granularity
  (CHPID needs to be toggled off temporarily)
You won't run out of WWPNs, but if ...
Look'n Feel

List Of Virtual N-Ports

<table>
<thead>
<tr>
<th>Partition</th>
<th>CSS</th>
<th>IID</th>
<th>CHPID</th>
<th>SSID</th>
<th>Device Number</th>
<th>WWPN</th>
<th>IOCDS</th>
<th>NPIV Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>T29LP46</td>
<td>03</td>
<td>01</td>
<td>3d</td>
<td>00</td>
<td>5200</td>
<td>c05076ffe803514</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP46</td>
<td>03</td>
<td>01</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe803518</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP46</td>
<td>03</td>
<td>01</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe80351c</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP46</td>
<td>03</td>
<td>01</td>
<td>55</td>
<td>00</td>
<td>1700</td>
<td>c05076ffe8029dc</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP19</td>
<td>01</td>
<td>04</td>
<td>3d</td>
<td>00</td>
<td>5203</td>
<td>c05076ffe8031dc</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP19</td>
<td>01</td>
<td>04</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe8031e0</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP19</td>
<td>01</td>
<td>04</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe8031e4</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP08</td>
<td>00</td>
<td>08</td>
<td>3d</td>
<td>00</td>
<td>5207</td>
<td>c05076ffe803054</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP08</td>
<td>00</td>
<td>08</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe803058</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP08</td>
<td>00</td>
<td>08</td>
<td>3d</td>
<td>00</td>
<td>52fd</td>
<td>c05076ffe80305c</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP08</td>
<td>00</td>
<td>08</td>
<td>55</td>
<td>00</td>
<td>1707</td>
<td>c05076ffe802644</td>
<td>A0 A1</td>
<td>On</td>
</tr>
<tr>
<td>T29LP30</td>
<td>01</td>
<td>0f</td>
<td>3d</td>
<td>00</td>
<td>521d</td>
<td>c05076ffe803288</td>
<td>A0 A1</td>
<td>On</td>
</tr>
</tbody>
</table>
Look'n Feel

Linux Proudly Presents: Its Own SAN Identity

virtual N-Port tied to this FCP subchannel

in non-NPIV mode:
same WWPN

in NPIV mode:
WWPNs differ

physical N-Port tied to FCP channel (CHPID)

# pwd
/sys/class/fc_host/host0

# cat node_name
0x5005076400cd6aad

# cat port_name
0x5005076401008fa8

# cat permanent_port_name
0x5005076401008fa8

in non-NPIV mode:
same WWPN

in NPIV mode:
WWPNs differ

# cat node_name
0x5005076400cd6aad

# cat port_name
0xc05076fffe8031dc

# cat permanent_port_name
0x5005076401008fa8
Keep in Mind

NPIV is NOT about ...

- … reducing SAN maintenance work
  - but it straightens out the way of doing things
- … migrating virtual servers more easily
  - but it enables virtual servers in FC SANs
- … introducing new FC security schemes
  - but it integrates virtual servers into existing ones
- … solving any SAN issue immediately
  - but it provides a foundation to build on
Summary

- NPIV removes the root cause for limited practicability of FCP in virtual server environments
- NPIV allows the use of standard SAN management solutions on IBM System z9
- NPIV gives free rein to sophisticated SAN fantasies

„Ready, steady, ... go!“
NPIV
FCP Channel Virtualization in a Linux Environment
The following are trademarks of the International Business Machines Corporation in the United States and/or other countries:

Enterprise Storage Server, IBM*, IBM logo*, System z9*, IBM eServer, z/VM, zSeries

*Registered trademarks of IBM Corporation

Linux is a registered trademark of Linus Torvalds.

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