

NPIV- Optimizing the storage virtualization in servers

Kiran Anumalasetty

Yadagiri Rajaboina

IBM Hyderabad, India

- Challenges in SAN Data Center
- Why N_Port ID Virtualization (NPIV) ?
- What is NPIV?
- Benefits of NPIV
- NPIV Login Sequence
- Limitations of NPIV

Challenges in SAN Data Center

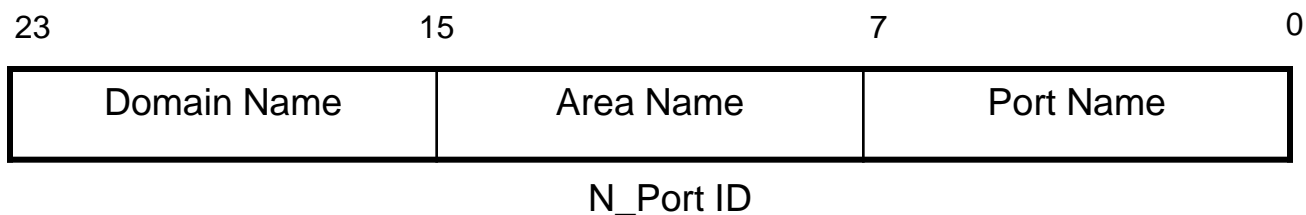
Challenges	Best Practice
Workload consolidation Increasing cost of HBAs and switches. Scalability	Running workloads on Virtual Machines. I/O Virtualization. Dynamic provisioning of VMs & Partition Mobility.
Isolation of storage access	Zoning LUN masking.

Challenges in SAN Data Center

- In a highly scalable SAN, VMs owning a physical HBA is a costly affair and not viable.
- VMs owning physical HBAs lose flexibility.
- Without NPIV, other solutions to virtualize the server side storage are required. Such solutions have below limitations :
 1. VMs have no direct visibility to the SAN : e.g. Applications using common HBA library cannot be run on VMs.
 2. VMs doesn't understand fibre channel language: Zoning and LUN masking cannot be used.
 3. Difficult to exploit storage device specific features.
 4. Not standardized: Interoperability issues.

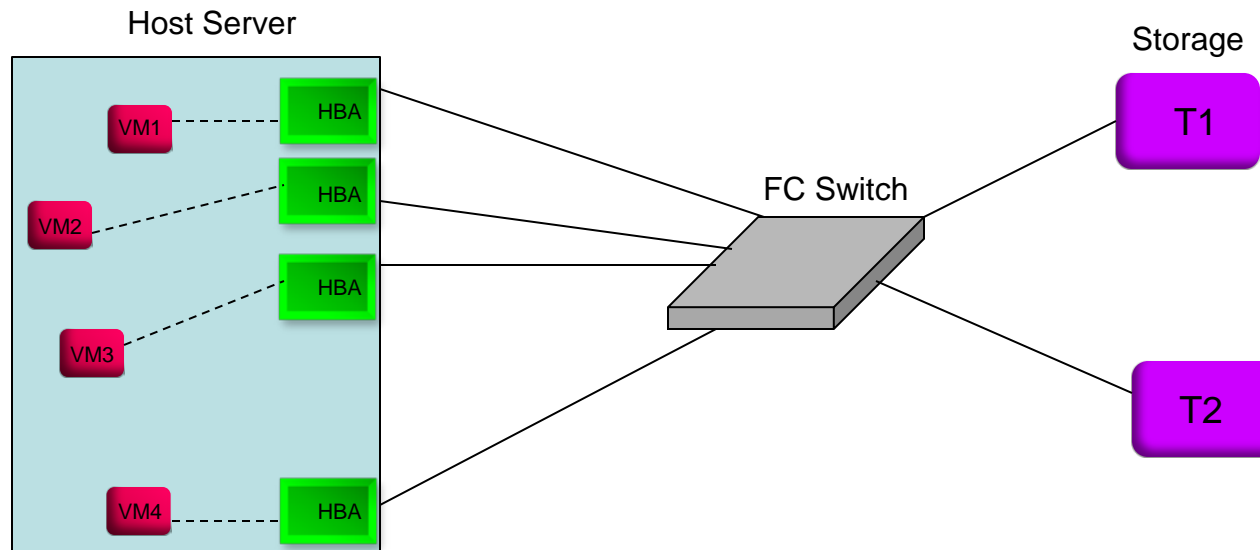
Fibre Channel Basics

- Each Fibre Channel HBA has a unique factory burnt 64 bit identifier called world wide port name (WWPN).
- Using the WWPN a FC HBA port issues Fabric login (FLOGI) request to the switch and acquires a FC address called N_Port ID.
- N_Port ID is used by the FC port for communication with the SAN.
- N_Port ID is a 24 bit address



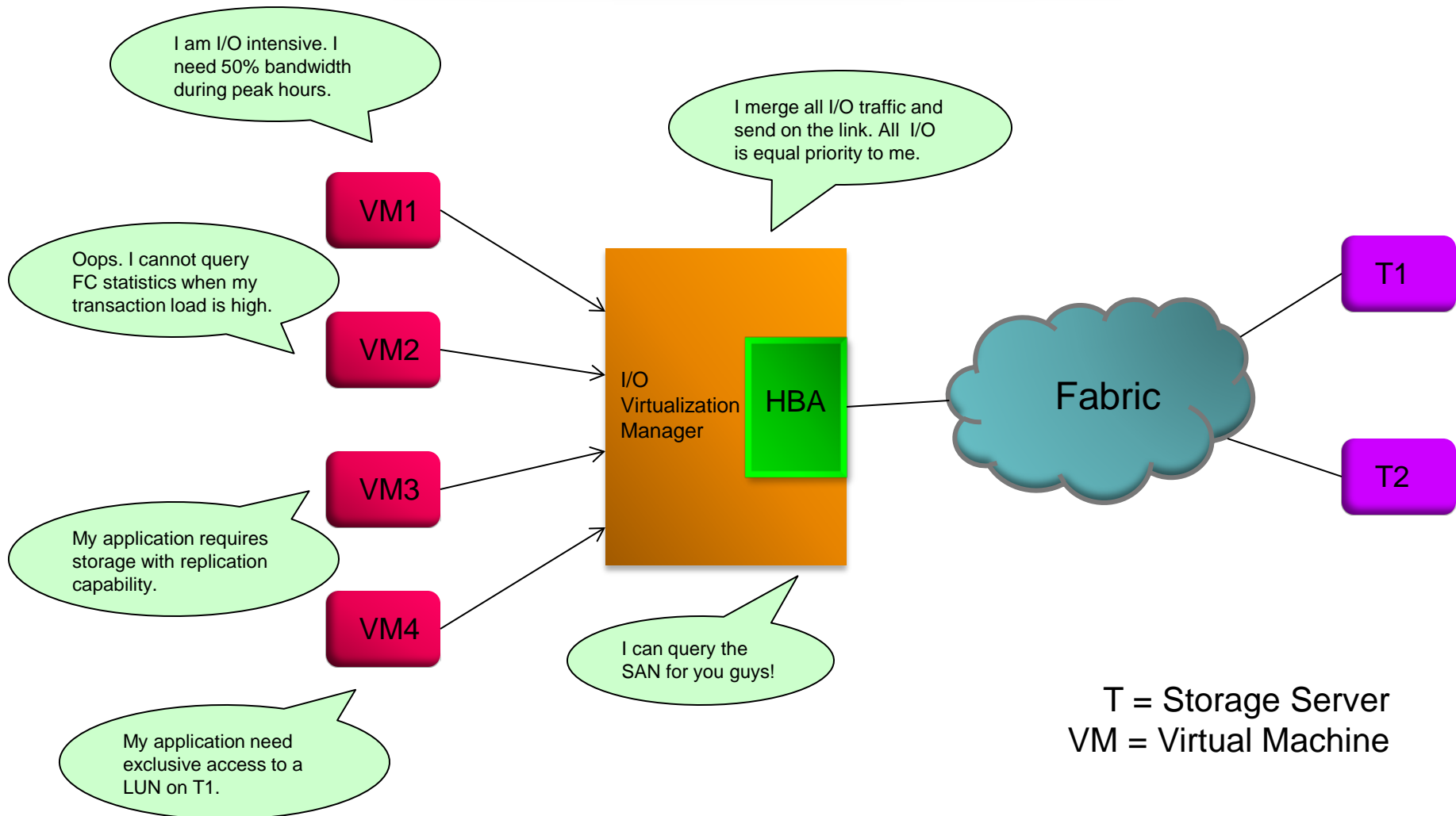
Hosts in a SAN without FC virtualization

- With each VM having a dedicated HBA, the cost of physical HBAs in a server along with the FC switch ports can be very high



T = Storage Server
VM = Virtual Machine

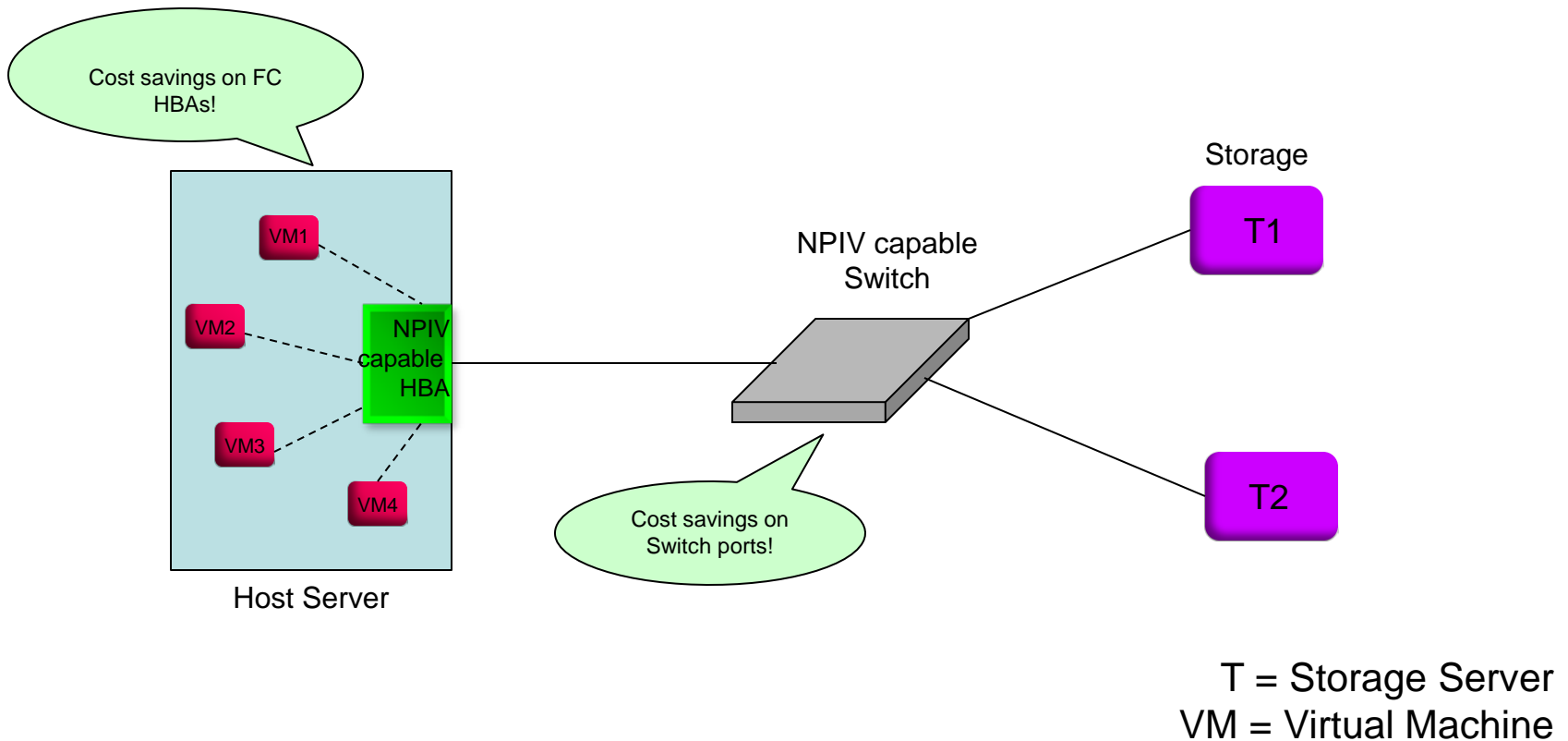
Hosts in a SAN without FC virtualization



What is N_Port ID Virtualization ?

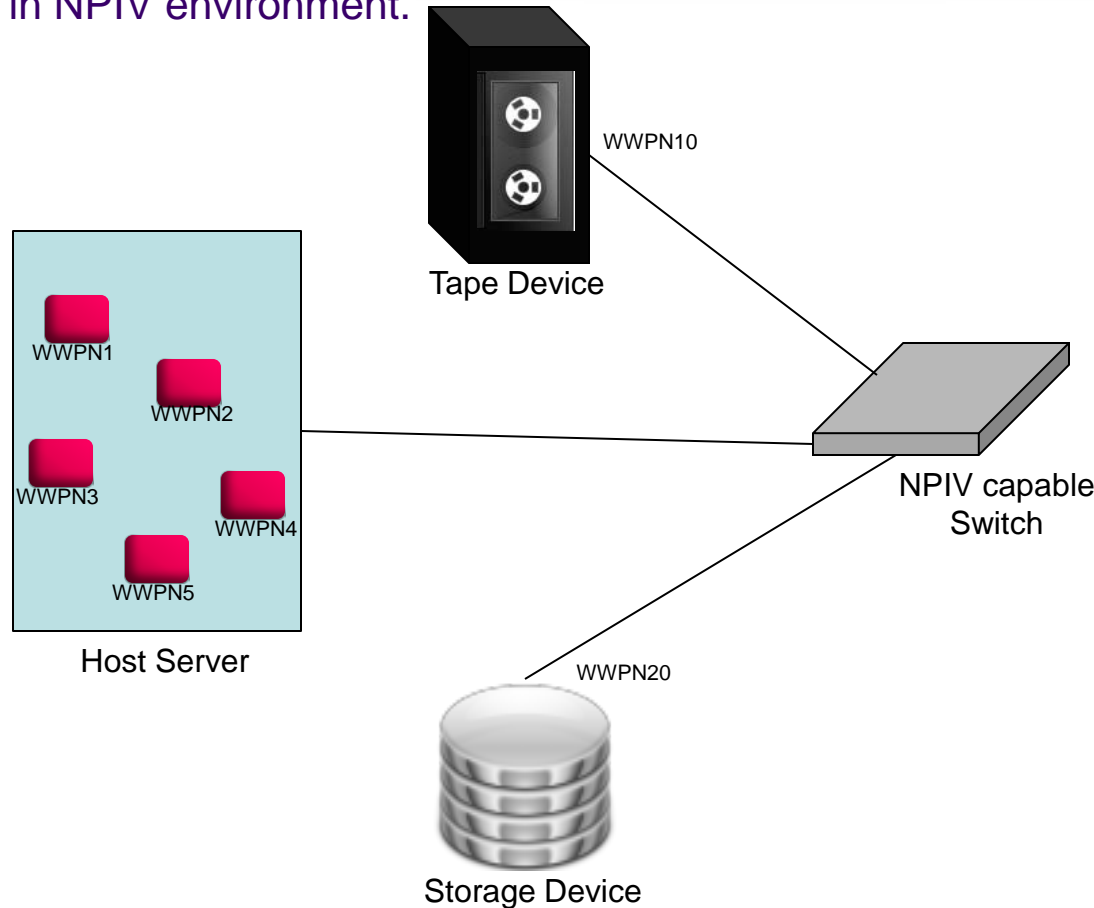
- ◆ N_Port ID Virtualization (NPIV) is ANSI T11 standard.
- ◆ Describes how a single physical FC HBA port can register with FC switch using multiple WWPNs to claim multiple fabric addresses (N_Port IDs) using Discover F_Port (FDISC) ELS command.
- ◆ Each WWPN can be assigned to a VM.
- ◆ Each VM can login to the fabric using its WWPN, via its Virtual N_Port and obtain a N_Port ID.
- ◆ FC HBA and Switch should have support for managing additional N_Port IDs.
- ◆ Each VM acts as an initiator with its WWPN and acquired N_Port ID.

Benefits of NPIV



Benefits of NPIV

➤ Zoning in NPIV environment.



Zoning best practices like Single initiator Zoning, WWPN identification are possible with NPIV.

Zone 0: WWPN1, WWPN20
Zone0_VM1_Storage

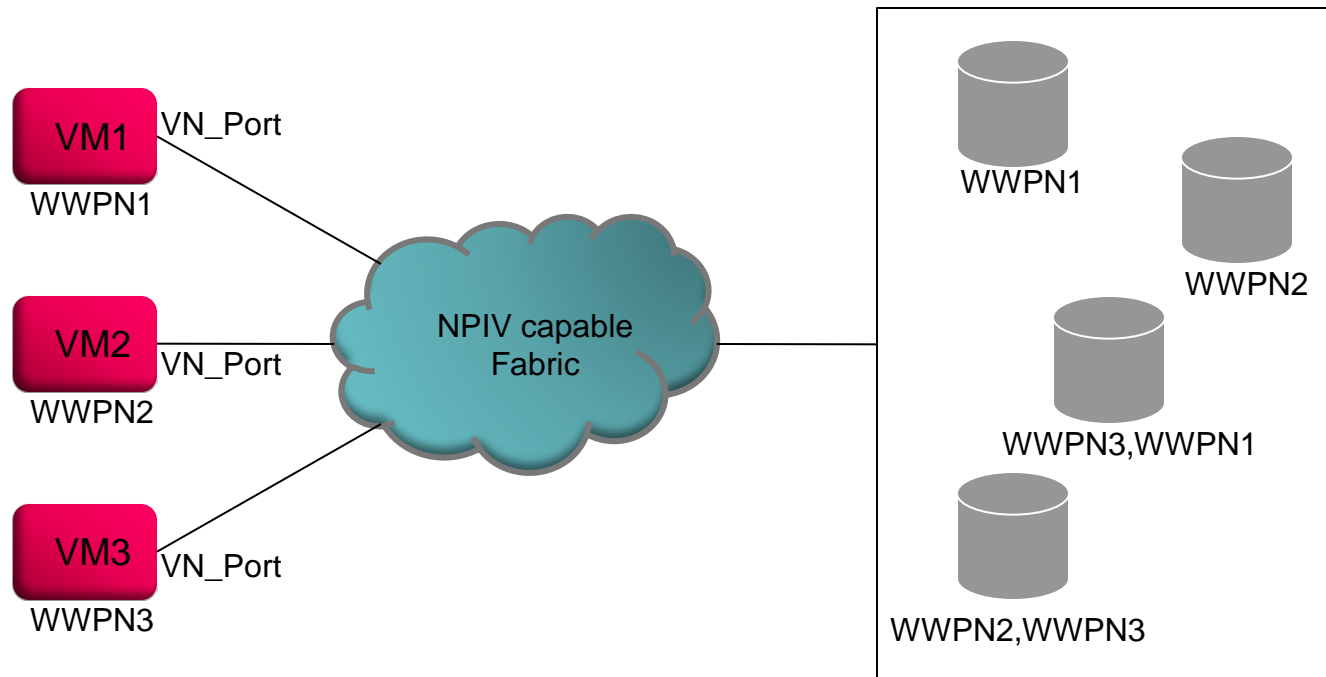
Zone1: WWPN2, WWPN10
Zone1_VM2_Tape

Zone3: WWPN3, WWPN20
Zone3_VM3_Storage

Zone4: WWPN3, WWPN10
Zone4_VM3_Tape

Benefits of NPIV

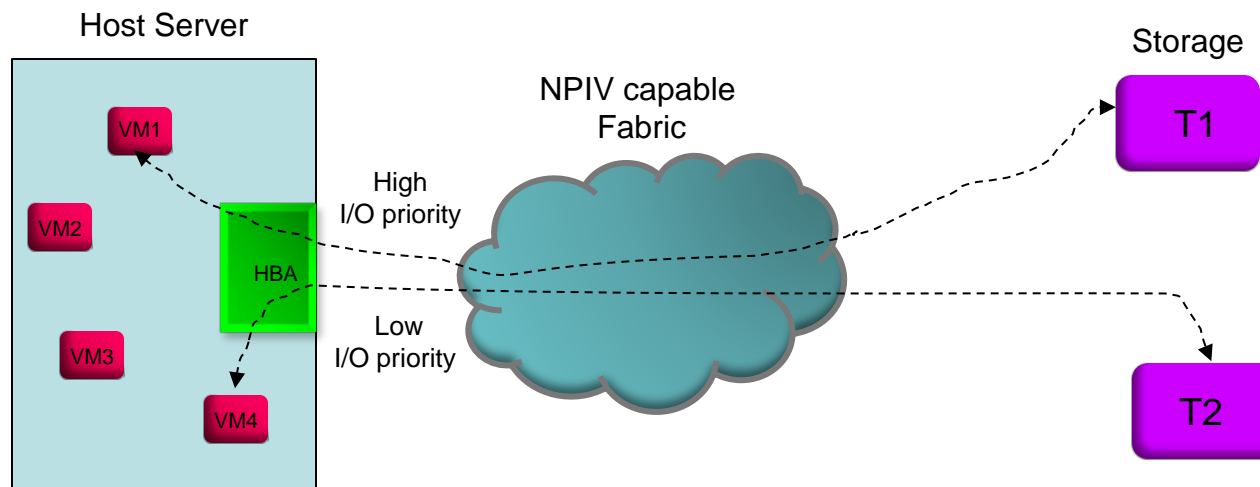
➤ NPIV based LUN Masking.



- Pre-provisioned WWPNs.
- Physical HBA can be replaced without reconfiguring the SAN.

Benefits of NPIV

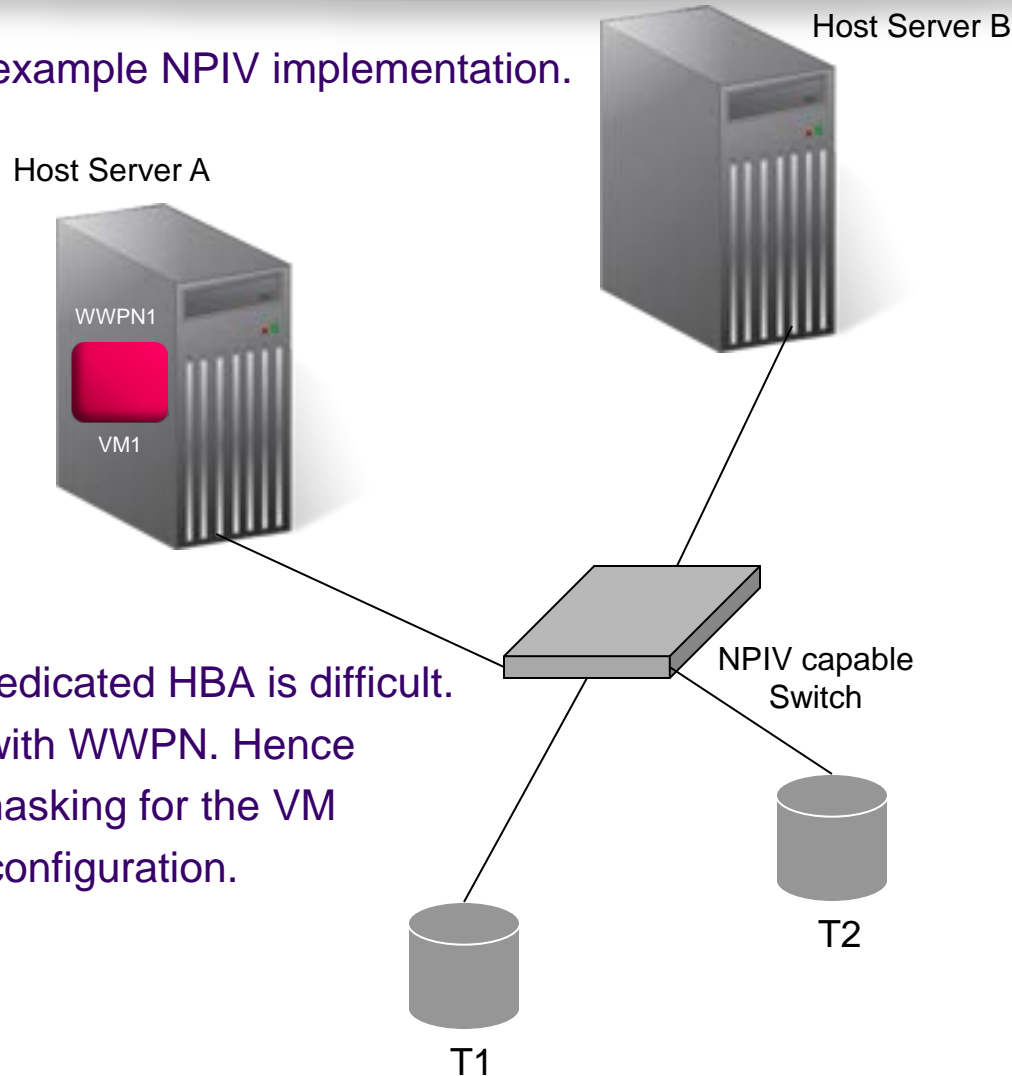
- Quality of Service (QoS) benefits that HBAs and FC switches provide can be extended to VMs using NPIV.



- Allocation of FC Switch /HBA resources to applications based on QoS priority.

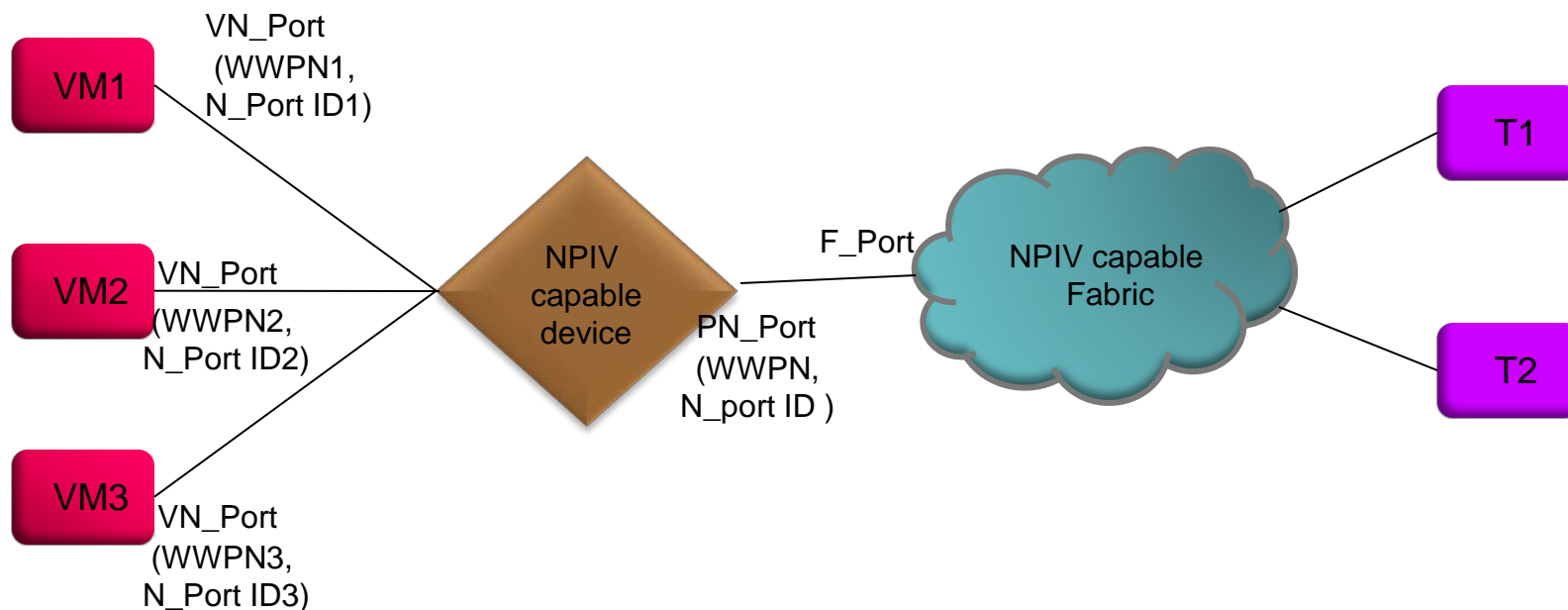
Benefits of NPIV

- VM Mobility in an example NPIV implementation.



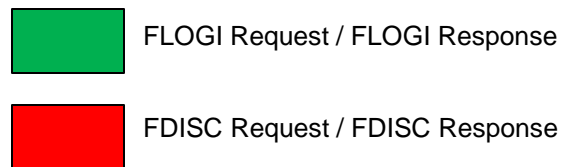
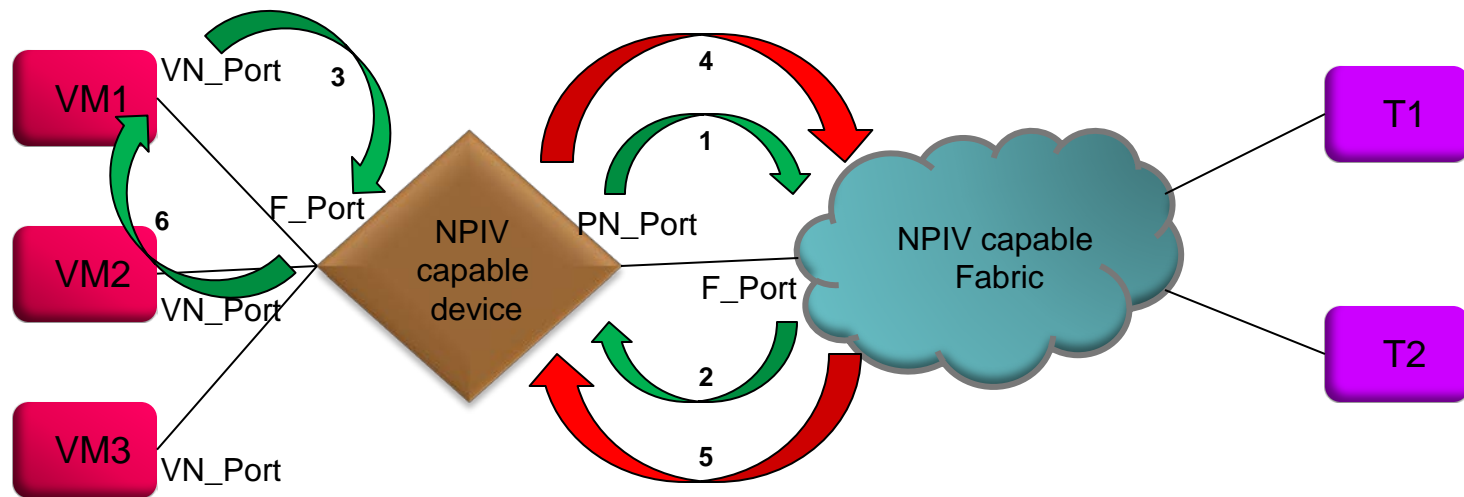
- VM mobility with dedicated HBA is difficult.
- VM moves along with WWPN. Hence zoning and LUN masking for the VM doesn't require reconfiguration.

Example of a NPIV Implementation



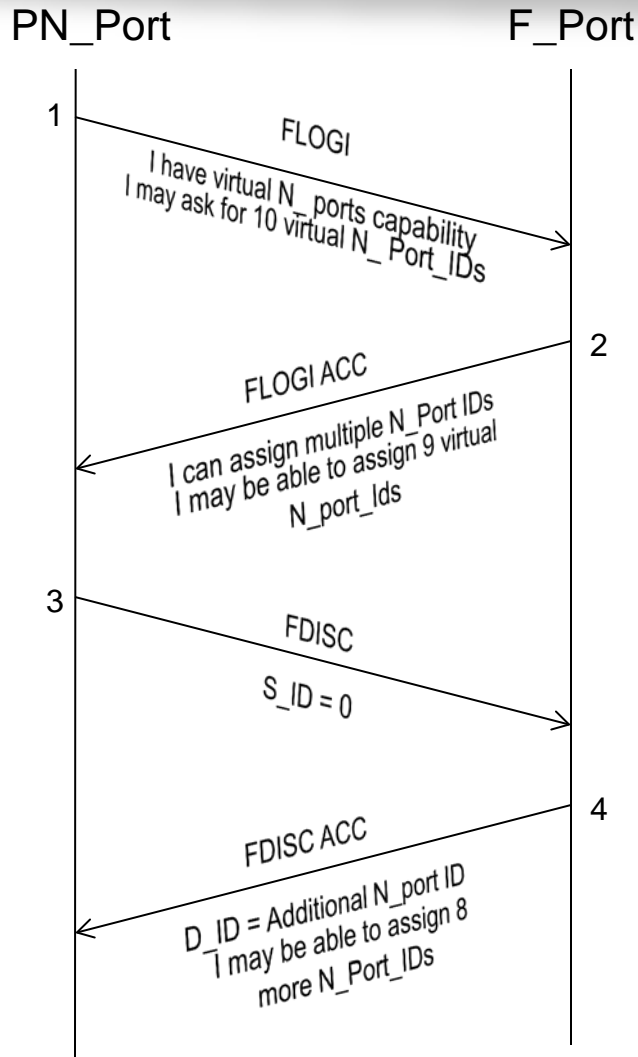
PN_Port – Physical N_Port
VN_Port – Virtual N_Port
F_Port – Fabric Port

NPIV Login sequence



- 1 - FLOGI
- 2 - FLOGI LS_ACC
- 3 - FLOGI
- 4 - FDISC
- 5 - FDISC LS_ACC
- 6 - FLOGI LS_ACC

NPIV Login sequence



- Step 1 : PN_Port issues FLOGI request with NPIV specific parameters set.
- Step 2: F_Port responds to FLOGI request with NPIV specific parameters set in the LS_ACC payload.
- Step 3: PN_Port issues FDISC request with S_ID set to 0 to obtain additional N_Port ID.
- Step 4: F_Port responds to FDISC request with D_ID set to additional N_Port ID.
- Step 3 and Step 4 are repeated for each virtual N_Port served by the PN_Port.

NPIV parameters in FC-LS standards

FLOGI / FDISC / LS_ACC payload (as per FC-LS-3)

Bits Word	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	ELS_Command code			
1	MSB			
..	Common Service Parameters (16 bytes)			
4				LSB
5	MSB			
6	Port_Name			LSB

Common Service Parameters - FLOGI (as per FC-LS-3)

Bits Word	31 .. 24	23 .. 16	15 .. 12	11 .. 08	07 .. 00
0	FC-PH Version - obsolete		Buffer-to-buffer Credit		
1	Common Features		BB_SC_N	Buffer-to-Buffer Receive Data_Field Size	
2	Reserved		Reserved		
3	Reserved		Reserved		

FLOGI Common service parameter applicability

Service Parameter	Word	Bits	Default Login Value	PLOGI and PLOGI LS_ACC Parameter applicability		FLOGI Parameter applicability		FLOGI LS_ACC Parameter applicability	
				Class		Class		Class	
				2	3	2	3	2	3
FC-PH Version - obsolete	0	31-16	2020h	n	n	n	n	n	n
Buffer-to-buffer Credit	0	15-0	0 or 1 ^c	y	y	y	y	y	y
Common Features	1	31-16							
Continuously increasing relative offset	1	31	0	y	y	n	n	n	n
Clean Address	1	31	0	n	n	n	n	y	y
Multiple N_Port_ID Support	1	31	0	n	n	y	y	n	n
Random relative offset	1	30	0	y	y	n	n	n	n
Virtual Fabrics bit	1	30	0	n	n	y	y	y	y
Valid Vendor Version Level	1	29	0	y	y	y	y	n	n
Multiple N_Port_ID Assignment	1	29	0	n	n	n	n	y	y

- The Multiple N_Port ID Support bit (word 1, bit 31) shall be set to one to indicate that the PN_Port supplying this parameter is capable of requesting multiple N_Port IDs using the FDISC ELS.
- This bit is only meaningful in the FLOGI request.

NPIV parameters in FC-LS standards

Common Service Parameters – FLOGI LS_ACC

Bits Word	31 .. 24	23 .. 16	15 .. 12	11 .. 08	07 .. 00
0	FC-PH Version - obsolete		Buffer-to-buffer Credit (Fx_Port)		
1	Common Features		BB_SC_N	Buffer-to-Buffer Receive Data_Field size	
2	R_A_TOV				
3	E_D_TOV				

NPIV parameters in FC-LS standards

FLOGI Common service parameter applicability

Service Parameter	Word	Bits	Default Login Value	PLOGI and PLOGI LS_ACC Parameter applicability		FLOGI Parameter applicability		FLOGI LS_ACC Parameter applicability	
				Class		Class		Class	
				2	3	2	3	2	3
FC-PH Version - obsolete	0	31-16	2020h	n	n	n	n	n	n
Buffer-to-buffer Credit	0	15-0	0 or 1 ^c	y	y	y	y	y	y
Common Features	1	31-16							
Continuously increasing relative offset	1	31	0	y	y	n	n	n	n
Clean Address	1	31	0	n	n	n	n	y	y
Multiple N_Port_ID Support	1	31	0	n	n	y	y	n	n
Random relative offset	1	30	0	y	y	n	n	n	n
Virtual Fabrics bit	1	30	0	n	n	y	y	y	y
Valid Vendor Version Level	1	29	0	y	y	y	y	n	n
Multiple N_Port_ID Assignment	1	29	0	n	n	n	n	y	y

- Multiple N_Port ID Assignment bit (word 1, bit 29) shall be set to one if the F_Port supplying this parameter is capable of assigning multiple N_Port IDs to the attached PN_Port using the FDISC ELS.

- This bit is only meaningful in the FLOGI LS_ACC.

- This bit is set only when the Multiple N_Port ID Support bit in the FLOGI request is one.

NPIV parameters in FC-LS standards

NPIV_CNT significance in

- FLOGI - Number of FDISC with S_ID = 0 that the PN_Port intends to originate.
- FLOGI LS_ACC – Number of FDISC with S_ID = 0 that are likely to be successful.
- FDISC LS_ACC – Number of additional FDISC with S_ID = 0 that are likely to be successful subsequent to this one.

FLOGI / FDISC / LS_ACC payload (as per FC-LS-3)

Bits Word	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	ELS_Command code			
1	MSB			
..	Common Service Parameters (16 bytes)			LSB
4				

⋮

21	MSB			
..	Auxiliary Parameter Data (16 bytes)			LSB
24				

Auxiliary Parameter Data (as per FC-LS-3)

Bits Word	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	Flags	Reserved		
1	Reserved			
2	Reserved			
3	Reserved			NPIV_CNT

NPIV Login - FDISC command

- ◆ Discover F_Port Service Parameters (FDISC) ELS shall transfer Service Parameters from the initiating N_Port to the F_Port at well known F_Port ID (FFFFFFEh).
- ◆ N_Port sends FDISC ELS to the F_Port to obtain additional N_Port ID.
- ◆ The acquired additional N_Port ID is assigned to the Virtual Machine.
- ◆ FDISC provides the means for the exchange of Service Parameters and the assignment of additional N_Port IDs without changing service parameters.

NPIV Login - FDISC command

FDISC Request Sequence

- ◆ When the S_ID of the FDISC ELS is equal to zero, the FDISC ELS shall transfer WWPN and WWNN and request the assignment and login of an additional N_Port ID.
- ◆ The FDISC ELS with an S_ID equal to zero shall only be sent by a PN_Port with at least one N_Port ID that is currently logged in with the Fabric .
- ◆ The interchange of FDISC information shall not modify the operating environment or Service Parameters between the two FC Ports.

FDISC Reply Sequence

- ◆ When the S_ID of the FDISC ELS is zero, the D_ID field of the LS_ACC shall be set to the additional N_Port ID being assigned.
- ◆ If FDISC is received by the F_Port but no N_Port is already logged in then F_RJT shall be returned.

Limitations of NPIV

- NPIV only supported in fabric topology.
- Both FC switch and HBA must support NPIV in a SAN.
- With NPIV VMs may have to install all necessary MPIO device drivers and/or multipath software modules that are aware of specific storage.
- Updates to multipath software will be needed on all the VMs.
- With proliferation of VMs, some targets may not have enough resources to process large number of initiator ports.

Glossary of Terms

- ◆ HBA – Host Bus Adapter is the term used for Fibre Channel adapter on the host.
- ◆ FC – Fibre Channel is a high speed interconnect technology standardized in T11 Technical committee.
- ◆ VM – Virtual Machine is a software emulation of a computer system.
- ◆ SAN – Storage Area Network.
- ◆ N_Port – Node Port in a SAN.
- ◆ F_Port – Port on a switch. Also called a fabric port.
- ◆ PN_Port – Physical N_Port.
- ◆ VN_Port – Virtual N_Port.
- ◆ N_Port ID – Fibre channel N_Port Identifier.
- ◆ WWPN – World Wide Port Name.
- ◆ QoS – Quality of Service.

Thank You