



# Advanced POWER6 Virtual I/O Configurations Overview

Speaker Name: Janel Barfield

**PowerHA Development**  
email: [jgbarfie@us.ibm.com](mailto:jgbarfie@us.ibm.com)

# Agenda

- Discuss criteria and grading scale used to rate virtual I/O configurations
- Define and discuss some typical virtual I/O configuration scenarios
- Answer questions as time permits
  - Email me with any unanswered questions  
[jgbarfie@us.ibm.com](mailto:jgbarfie@us.ibm.com)

## Complex VIO Architectural Considerations

- Complex VIO configurations have these architectural considerations:
  - **Resilience**: sustain operations during unplanned outage
  - **Serviceability**: sustain operations during planned outages
  - **Scalability/throughput**: can the configuration scale and increase throughput?
  - **Flexibility**: how easy is it to add virtual clients and devices without extensive administration operations
- Scenarios in this course will be graded on how well they meet these considerations
  - Grades:  
Standard (default), Medium, High
  - Example:

### Dual VIOS with client disk mirroring

<b>Resilience</b>	<b>High</b>
<b>Serviceability</b>	<b>High</b>
<b>Scalability</b>	<b>Standard</b>
<b>Flexibility</b>	<b>High</b>

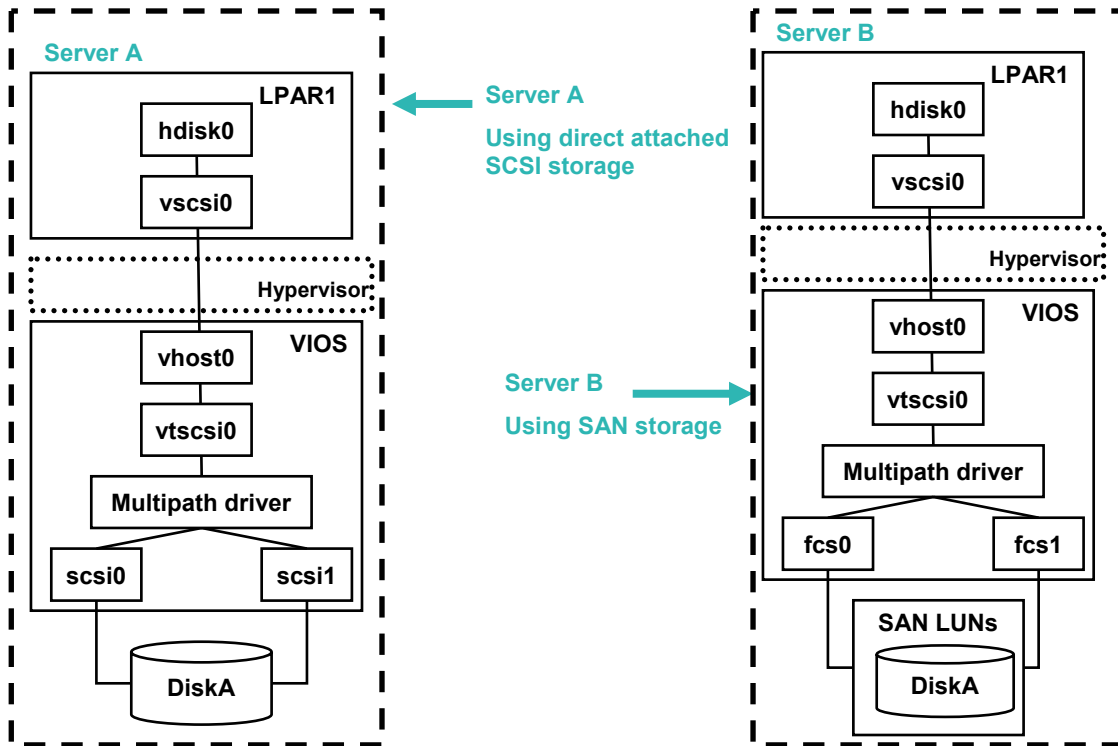
## Overview of Scenarios

- Virtual SCSI:
  - VIOS with multi-path I/O (MPIO) to storage
  - Dual VIOS with client disk mirroring
  - Dual VIOS with MPIO *and* client-partition with MPIO
- Virtual Ethernet:
  - Client-partition backup link feature
  - Link aggregation and backup feature
  - Dual VIOS with client LA backup channel
  - Dual VIOS with Shared Ethernet adapter failover

# Highly Available Virtual Storage Configurations

- We'll start with storage configurations:
  - VIOS with multi-path I/O (MPIO) to storage
  - Dual VIOS with client disk mirroring
  - Dual VIOS with MPIO *and* client-partition with MPIO

# Storage: VIOS with Multiple Paths to Storage



## Storage: VIOS with Multiple Paths to Storage Configuration Rating

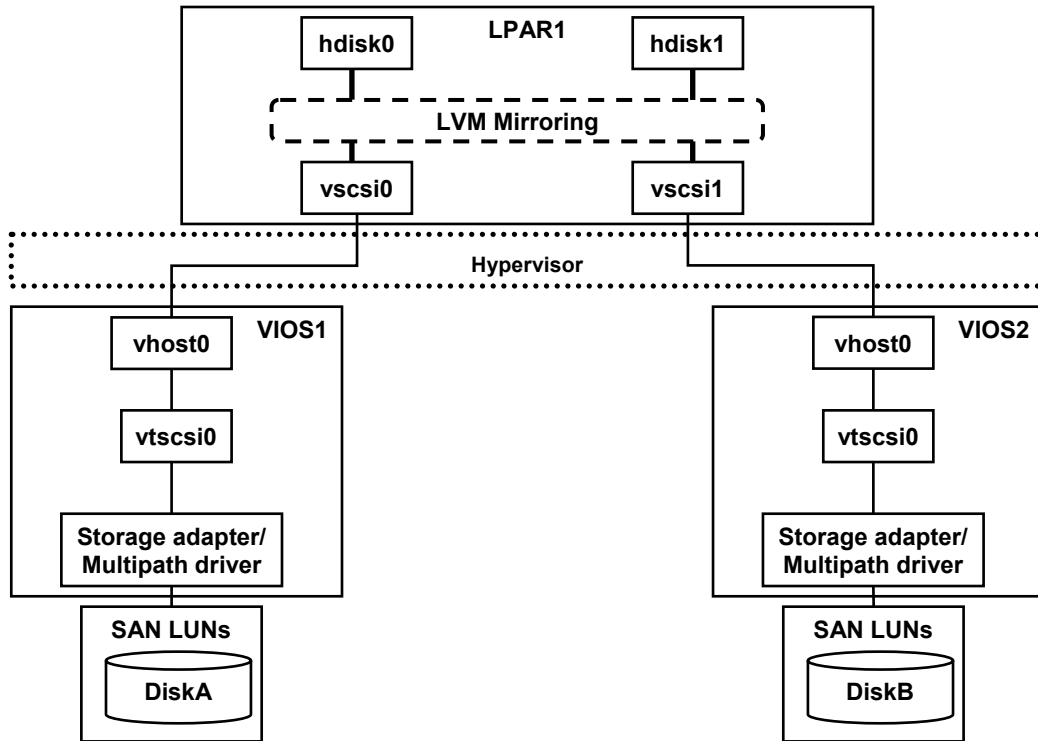
- Clients have a single virtual adapter to access virtual disk(s) through the server adapter
  - Simple and flexible
  - Can increase resilience with mirroring
- Disk associated with the virtual server adapter is reached via multiple paths using a multipath driver
  - Scalable and resilient
  - Requires MPIO software (AIX MPIO used by default)

<b>Resilience</b>	Medium
<b>Serviceability</b>	Standard
<b>Scalability</b>	High
<b>Flexibility</b>	High

## Storage: VIOS with Multiple Paths to Storage Configuration Considerations

- **Resilience is Medium**
  - Sustain failure of storage adapter or SAN path (and SAN switch when appropriately configured), but not VIOS failure
  - Resilience is improved if mirroring is used
- **Serviceability is Standard**
  - VIOS maintenance requires client outage
- **Scalability is High**
  - With SAN storage, load balancing is performed when multiple paths are available, improving throughput
- **Flexibility is High**
  - Clients can be added easily without additional hardware as long as storage is available on the server by creating new virtual adapters

# Storage: Dual VIOS with Client Disk Mirroring



## Storage: Dual VIOS with Client Disk Mirroring Rating

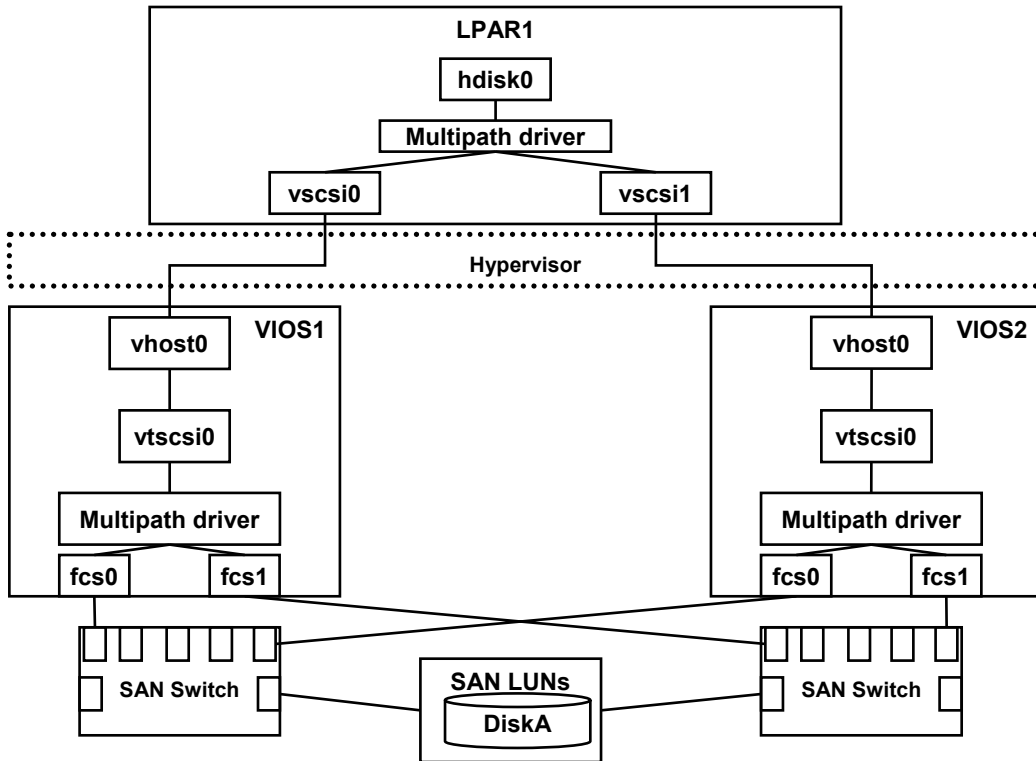
- Client has two virtual adapters
  - One assigned to each VIOS
  - Provide access to different disks, often in different storage subsystems for redundancy
  - Disks provided from the VIO servers can come from the same SAN
- Mirror client data or rootvg across them
  - You should mirror virtual disks from the client, not the VIO server

Resilience	High
Serviceability	High
Scalability	High
Flexibility	High

## Storage: Dual VIOS with Client Disk Mirroring Considerations

- **Resilience is High**
  - Sustain failure of any single infrastructure component
- **Serviceability is High**
  - Allows maintenance of VIOS without client outage
- **Scalability is High**
  - Use MPIO on VIOS to make scalability/throughput **high**
- **Flexibility is High**
  - Easily add new clients as long as storage is available from both VIOS partitions (may be limited by storage adapters if all clients use single path to storage)

# Storage: Dual VIOS with MPIO on Server and Client



## Considerations For Using Multiple Virtual Storage Paths from Dual VIO Servers

- Only supported when whole physical volume is backing device (cannot belong to a volume group on the server)
- Set `no_reserve` policy from each host (VIO Server) when paths to a disk are provided from multiple hosts (from VIO Server CLI)
  - Do this **before** you associate the disk with a server adapter using `mkvdev`
    - `lsdev -dev hdiskX -attr reserve_policy`
    - `chdev -dev hdiskX -attr reserve_policy=no_reserve`
- Since only one disk is being provided to the client, important that the disk is protected by RAID in the storage or mirroring on the client to maintain high resilience

## Storage: Dual VIOS with MPIO in Server and Client Rating

- Client has two virtual adapters that point to the same disk through multiple paths and multiple servers
- Client adapters attach to different VIOS partitions
- Uses MPIO on server and client

<b>Resilience</b>	High
<b>Serviceability</b>	High
<b>Scalability</b>	High
<b>Flexibility</b>	High

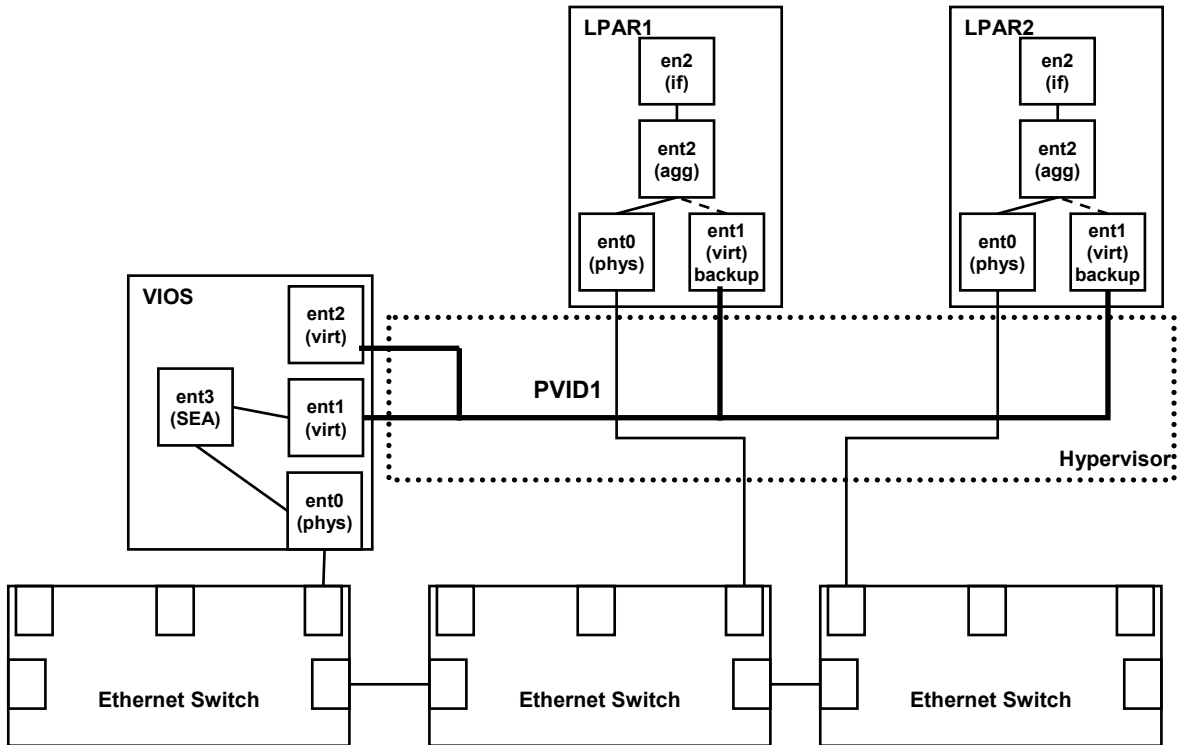
## Storage: Dual VIOS with MPIO in Server and Client Considerations

- **Resilience is High**
  - Sustain failure of any single infrastructure component
- **Serviceability is High**
  - Allows maintenance of any component without client outage, including the VIO server
- **Scalability is High**
  - MPIO on VIO servers provides scalability and increased throughput for all clients
- **Flexibility is High**
  - Easily add new clients without adding any new hardware

## Highly Available Virtual Networking

- Next are some highly available virtual networking solutions:
  - Client-partition backup link feature
  - Link aggregation and backup feature
  - Dual VIOS with client LA backup channel
  - Dual VIOS with Shared Ethernet adapter failover

# Networking: Client Backup Link Feature



## Networking: Client Backup Link Feature Rating

- Each client LPAR has a physical and virtual Ethernet adapter assigned
  - The physical adapter is connected to the external network and is the primary adapter in the link aggregation
  - The virtual adapter has access to the external network provided via an SEA on the Virtual I/O Server and is the backup adapter in the aggregation

<b>Resilience</b>	<b>Medium</b>
<b>Serviceability</b>	<b>High</b>
<b>Scalability</b>	<b>Standard</b>
<b>Flexibility</b>	<b>Medium</b>

## Networking: Client Backup Link Feature Considerations

- **Resilience** is **Medium**

- Can withstand failure of either client or server physical adapter but not both

- **Serviceability** is **High**

- Allows maintenance of any component without client outage, including the VIO server

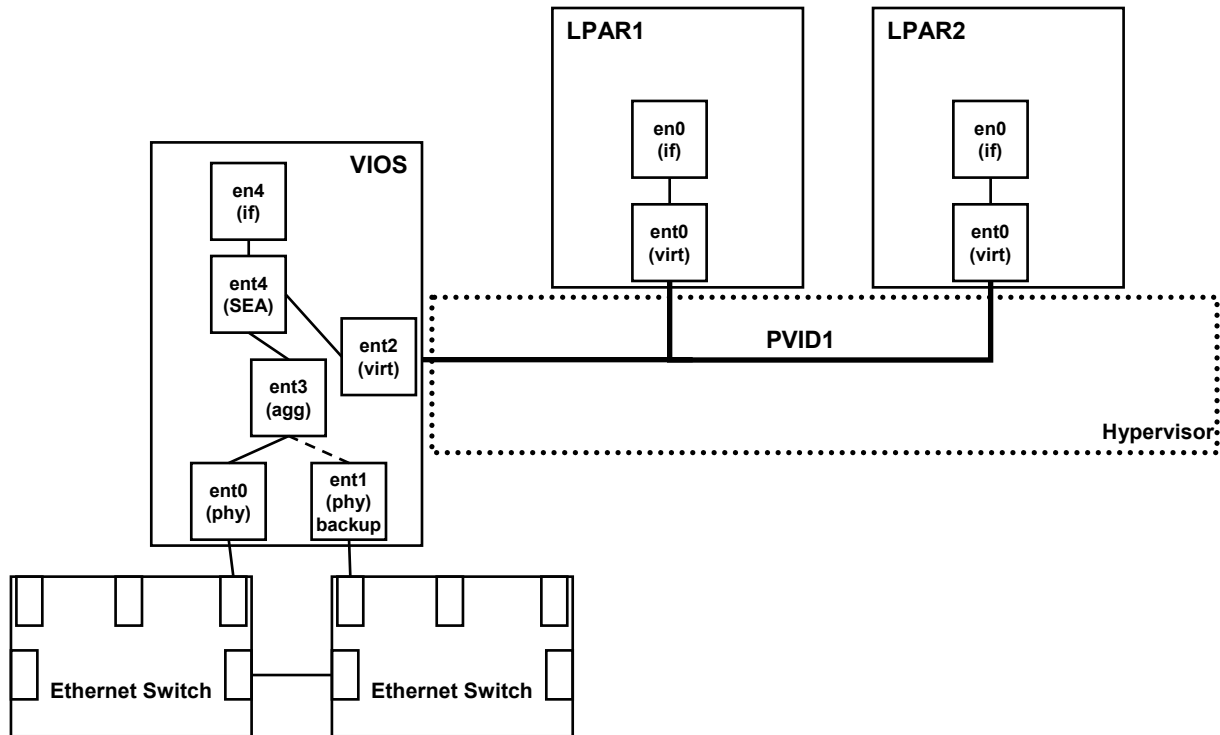
- **Scalability** is **Standard**

- If there is more than one unplanned client outage, additional bandwidth may be needed on the server to support multiple clients

- **Flexibility** is **Medium**

- May have to add more physical adapters to support additional clients

# Networking: Link Aggregation and Backup Feature



## Networking: Link Aggregation and Backup Feature Rating

- Each client LPAR has one virtual adapter and no physical adapters
- Redundancy provided in the Virtual I/O Server using two physical adapters in a backup link aggregation configuration

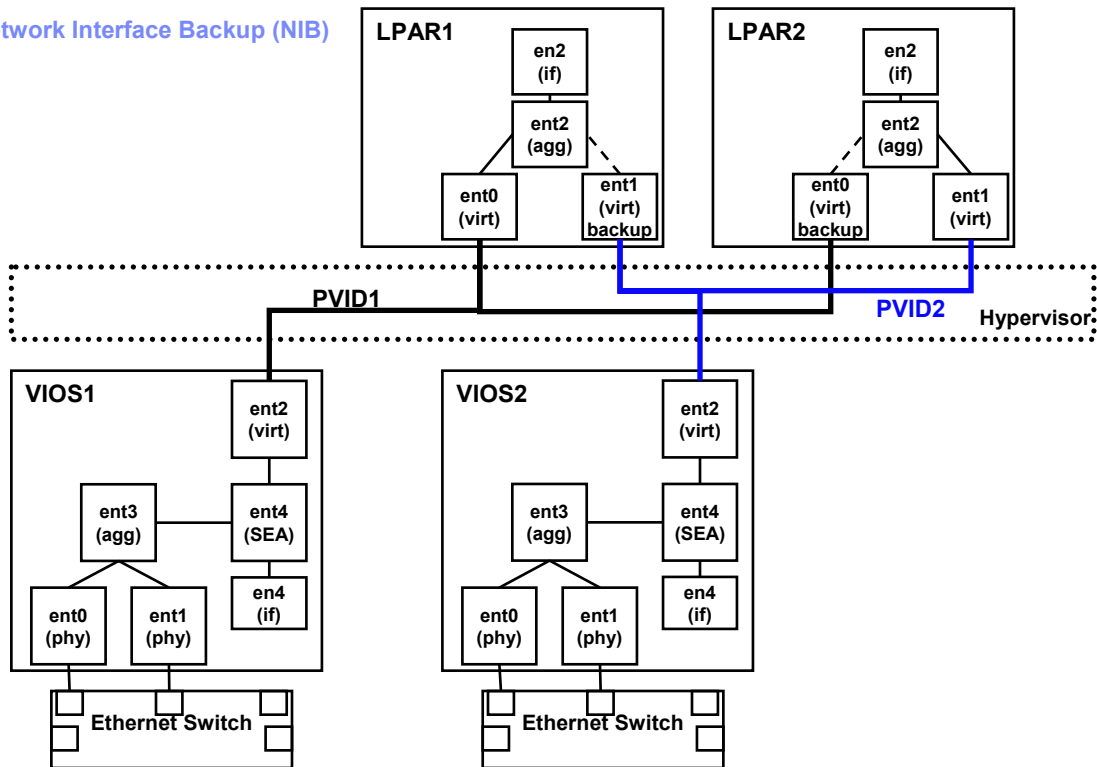
<b>Resilience</b>	Medium
<b>Serviceability</b>	Standard
<b>Scalability</b>	Standard
<b>Flexibility</b>	High

## Networking: Link Aggregation and Backup Feature Considerations

- **Resilience** is **Medium**
  - Can withstand failure physical adapter but not VIO server
- **Serviceability** is **Standard**
  - Maintenance of VIO server requires client outage
- **Scalability** is **Standard**
  - All clients are supported by bandwidth in single VIOS
- **Flexibility** is **High**
  - Clients are easily added with only virtual adapters

# Networking: Dual VIOS with Client LA Backup Channel

a.k.a. Network Interface Backup (NIB)



## Networking: Dual VIOS with Client LA Backup Channel Rating

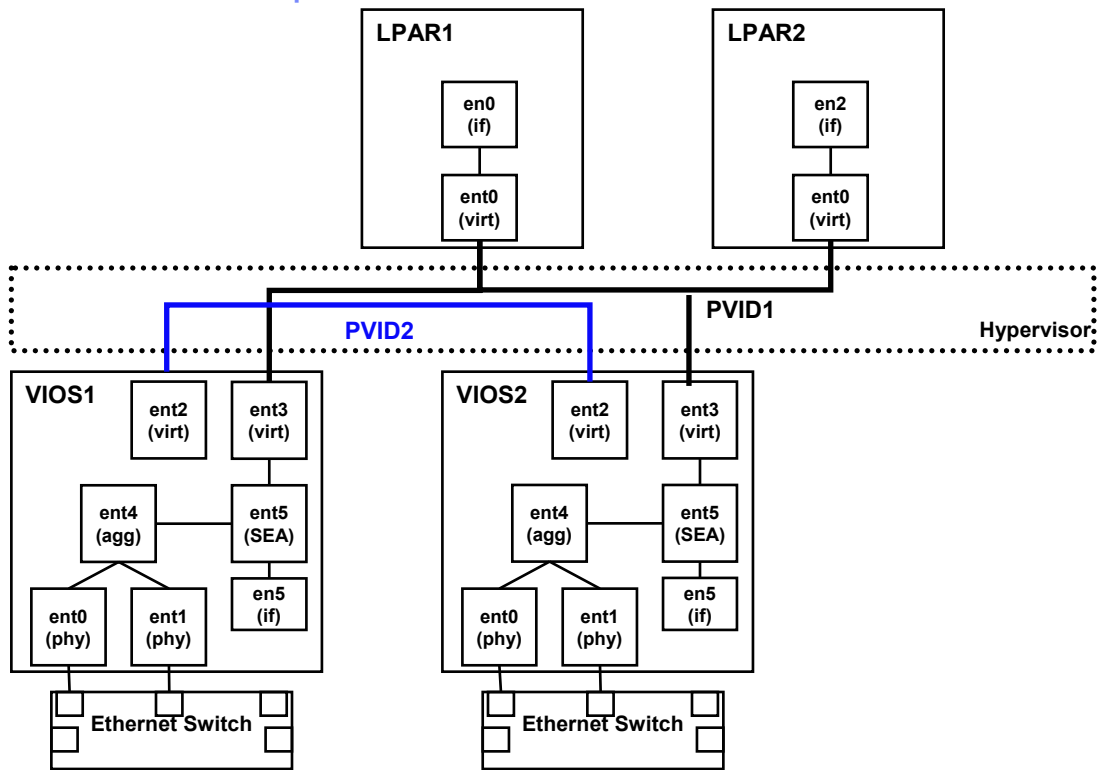
- Each client has two virtual adapters in a link aggregation backup configuration
  - Each uses a different VLAN for its primary network path
- Total bandwidth of 4 physical adapters is available for clients
  - Clients can be split evenly over 2 VLANs
  - If one VIOS fails, all traffic switches to available VLAN

<b>Resilience</b>	<b>High</b>
<b>Serviceability</b>	<b>High</b>
<b>Scalability</b>	<b>High</b>
<b>Flexibility</b>	<b>Medium</b>

## Networking: Dual VIOS with Client LA Backup Channel Considerations

- **Resilience is High**
  - Protects against VIOS, switch, and adapter failure
- **Serviceability is High**
  - Maintenance of VIOS, switch, or adapters can be performed without client outage
- **Scalability is High**
  - More physical adapters can be added to each VIOS to increase physical bandwidth available
- **Flexibility is Medium**
  - Clients can be added without installing additional hardware, but configuration is more complex

# Networking: Dual VIOS with Shared Ethernet Adapter Failover



## Networking: Dual VIOS with Shared Ethernet Adapter Failover Rating

- Client configuration is simplified – each only requires one virtual Ethernet adapter
- Complexity moves to the servers (easy to add new clients), but pay in available bandwidth
  - Operates in failover mode only
- Care must be taken to avoid a broadcast storm when configured

<b>Resilience</b>	<b>High</b>
<b>Serviceability</b>	<b>High</b>
<b>Scalability</b>	<b>High</b>
<b>Flexibility</b>	<b>High</b>

## Networking: Dual VIOS with Shared Ethernet Adapter Failover Considerations

- **Resilience is High**
  - Like NIB, protects against VIOS, switch, and adapter failure
- **Serviceability is High**
  - Maintenance of VIOS, switch, or adapters can be performed without client outage
- **Scalability is High**
  - More physical adapters can be added to each VIOS to increase physical bandwidth available
    - Only one SEA is used at a time – no load balancing
- **Flexibility is High**
  - Can support VLAN tagging through SEA adapters (not allowed with NIB) – means you can easily add new clients on different VLANs

## Conclusion

- Virtual I/O technology for POWER systems is highly configurable and flexible
- Use multiple paths, multiple Virtual I/O Servers, and redundant hardware (SAN and network switches) for highly available configurations
- Benefits of virtualized I/O configurations
  - Resiliency with less hardware
    - Multiple clients can share hardware and paths
  - Improved serviceability
    - Components can be maintained without disrupting clients
  - Scalability and flexibility
    - Add new clients without adding new hardware or disrupting other clients operation

# Trademarks

## The following are trademarks of the International Business Machines Corporation in the United States, other countries, or both.

Not all common law marks used by IBM are listed on this page. Failure of a mark to appear does not mean that IBM does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Those trademarks followed by ® are registered trademarks of IBM in the United States; all others are trademarks or common law marks of IBM in the United States.

For a complete list of IBM Trademarks, see [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml):

\*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

## The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

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

### Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprocessing in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

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.

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

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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.