IBM Power Systems Technical University

October 18–22, 2010 — Las Vegas, NV

VIOS NextGen: Server & Storage Integration

Session ID: VM19

PowerVM™: *Virtualization for IBM Power Systems*

Speaker Name: Bob Kovacs

Authorized IBM Training

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

VIOS Basics

- **Power LPAR based I/O virtualization appliance**
  - Facilitates sharing physical I/O resources amongst LPARs
  - Power5, Power6, Power7, pBlade
  - VIOS serves AIX, Linux, and i operating systems
  - Multiple VIOS's per CEC, typically deployed in pairs
  - Packaged with PowerVM editions (optional feature)
  - Introduced in 2004
  - Significant attachment rate, particularly for Mid-High End client environments
- **Virtual I/O**
  - Storage
    - Storage Virtualizer (vSCSI)
    - Pass-through (NPIV adapter sharing)
  - Virtual Networking
    - Ethernet Bridging
- **Advanced Virtualization**
  - LPAR Mobility
  - AMS (Active Memory Sharing)

vSCSI specifics

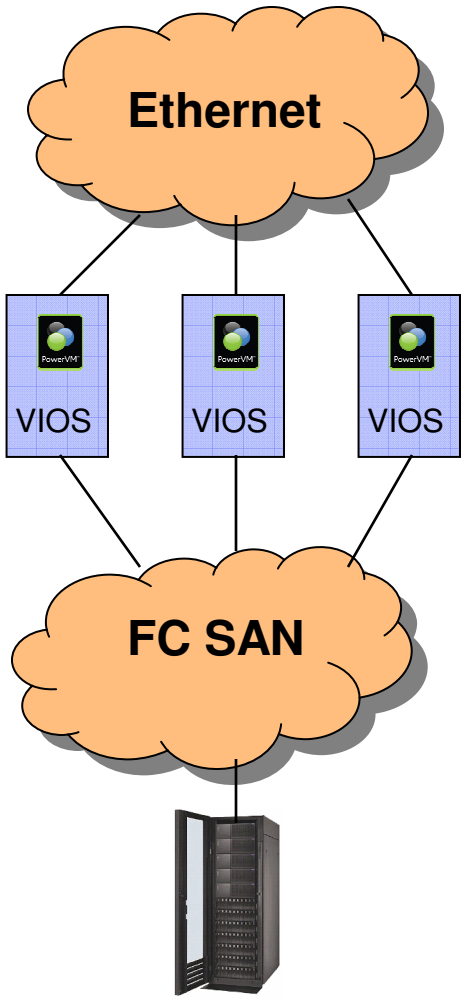
- Storage Virtualizer:
  - pools heterogeneous storage
  - FC, FCoE, SCSI, iSCSI, SAS, SATA, USB
  - SCSI Target
  - SCSI peripheral device types:
    - Disk (backed by physical volume, logical volume, or file)
    - Optical (backed by physical optical, or file)
    - Tape (backed by physical tape)
  - Adapter and device level sharing
- Introduced in 4Q/2004
- PowerVM Express, Standard, and Enterprise
- client OS support: AIX, IBM i, and Linux
- Power 5, Power 6, Power 7, and Blade
- Compatible with LPM (live partition mobility)

NextGen vSCSI - Server and Storage Integration

- **NextGen vSCSI is a server based storage virtualizer that is clustered across multiple Power servers**
  - It's an extension of VIOS's existing storage virtualization (vSCSI)
  - Combines existing SCSI emulation with clustering technology and a distributed data object repository
  - VIOS provides the same standard SCSI Target interface to host
- **Distributed data object repository**
  - Developed for VIOS storage virtualization
  - Object-backed vSCSI device, client devices are encapsulated in objects in the repository
  - Object repository provides advanced features
  - Linked clones, snapshot / rollback, thin provisioning / reclamation, pooling and tiering
  - Admin manages pools, tiers, capacity
  - Distributed object abstraction and namespace
- **Much broader pooling & aggregation that spans servers**
  - Pooled SAN storage with a distributed data object repository and namespace

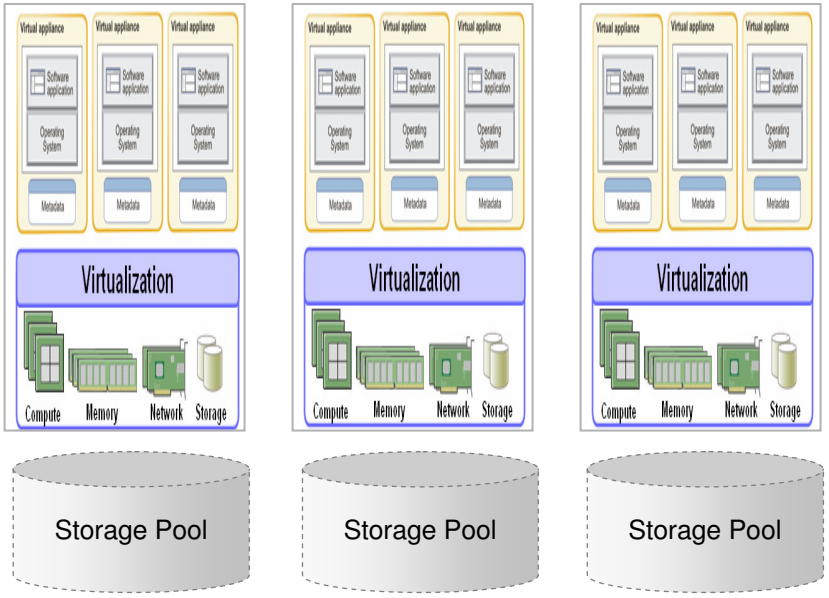
VIOS Storage Clustering Model

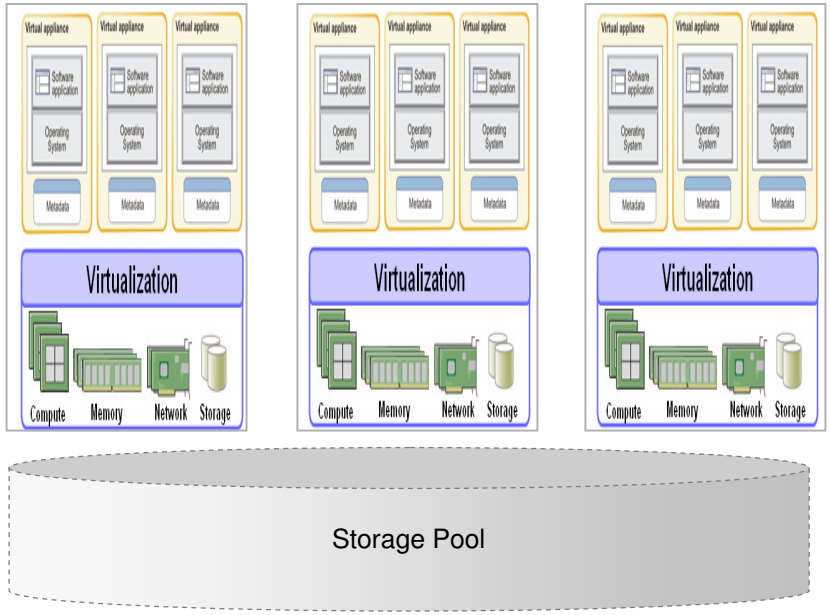
- VIOS nodes are joined together to form a cluster
- All VIOS nodes see all disks for storage pools they access
- All VIOS nodes are eligible to read and write user data via the SAN
- One or more VIOS nodes are meta-data managers for sub-units of each storage pool
- Non managers communicate meta-data access and update via the ethernet
- Cluster-wide concurrent access to namespace and data available from any node



PowerVM – VIOS Next Generation

Extending Storage Virtualization Beyond a Single System





vSCSI Classic – storage virtualization

- Storage pooled at VIOS for a single system
- Enables dynamic storage allocation
- Supports Local and SAN Storage, IBM and non-IBM Storage

vSCSI NextGen – clustered storage virtualization

- Storage pool spans multiple VIOS's and servers
- Enabler for federated management
- Location transparency
- Advanced capabilities
- Supports SAN and NAS, IBM and non-IBM Storage

Power your planet.

vSCSI NextGen Objectives

1. **Reduce complexity**

2. **Facilitate key capabilities needed for emerging technologies and the Power datacenter**

vSCSI NextGen

1. Reduce complexity

- Centralized server and storage domains in one view (Director)
  - Server and storage integration
- Automated storage provisioning
- Location transparency (simplified server and storage mobility)
- Reduce the frequency of server and SAN team interaction
- End-To-End provisioning, compliance, security, usage and performance
- Control point that maps virtual machine to its storage
  - Device operations orchestrated at the virtual machine level (LPAR)

vSCSI NextGen

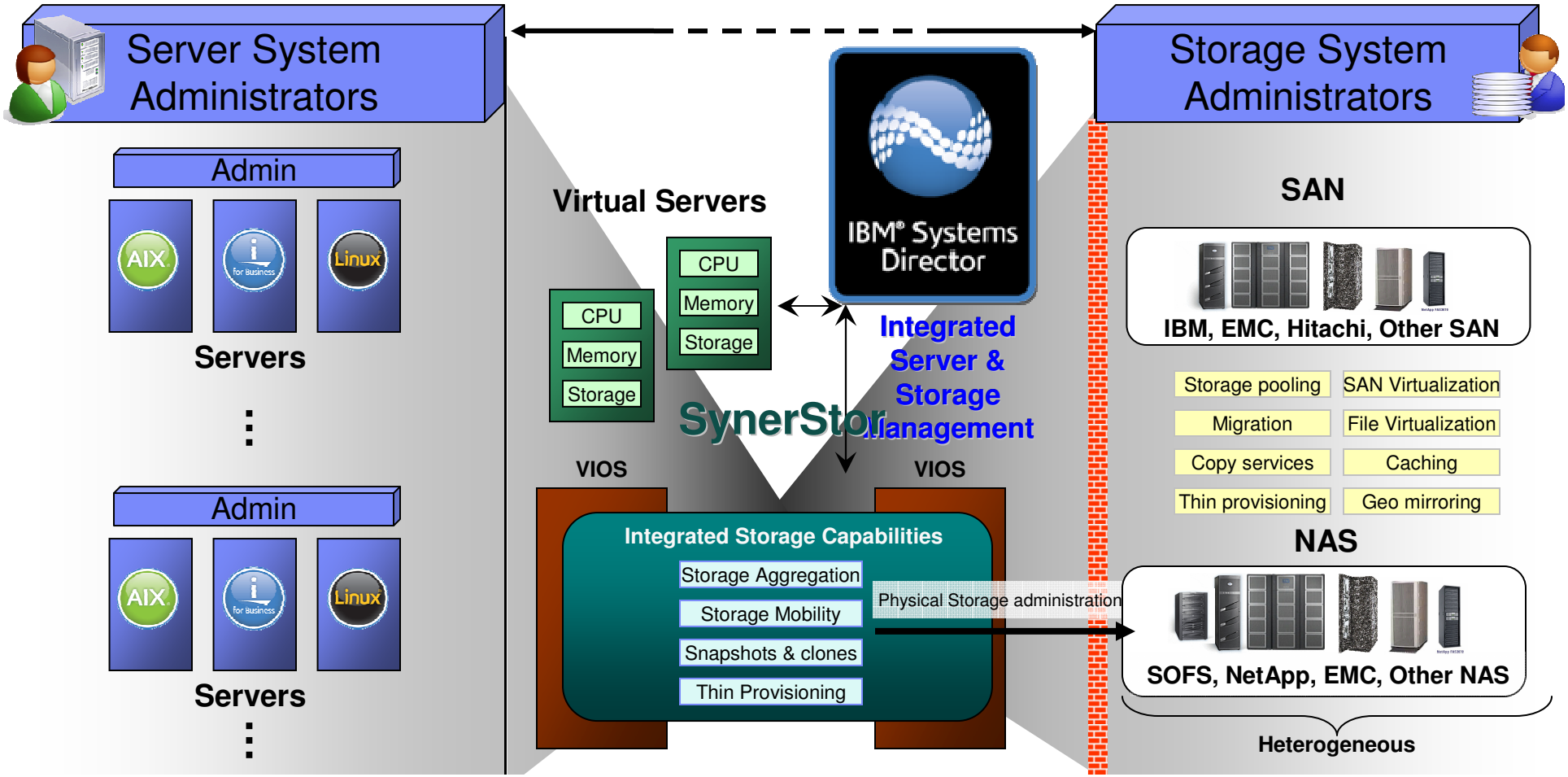
Integrated Storage Virtualization increases Platform Value

- Automated storage provisioning
- Simplified, integrated Director Mgmt
- Advanced image management

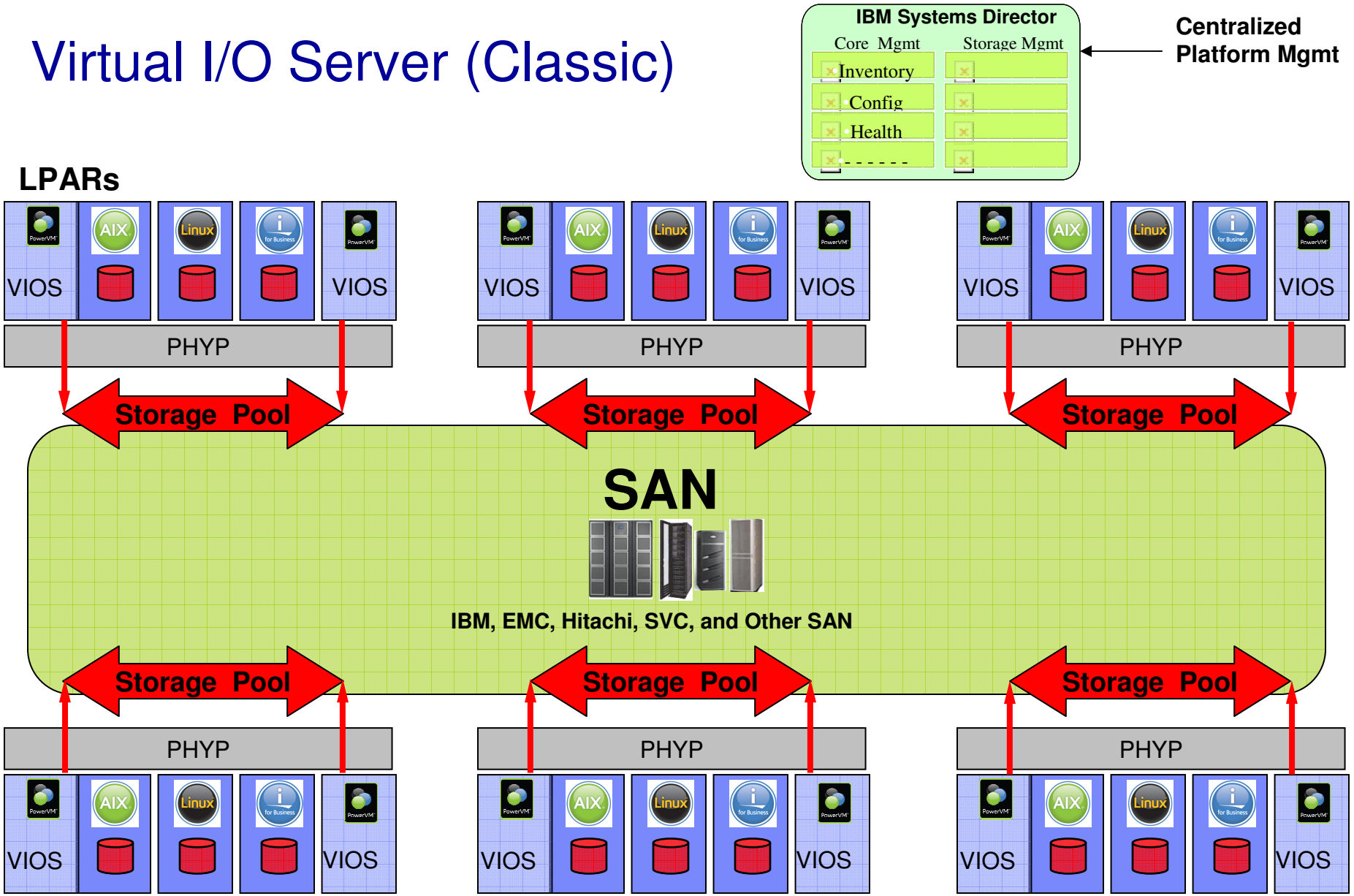
Client Benefits

- Few interactions between mgmt domains
- Consolidated on-line backup
- Consistent capabilities with different storage

= Decreased complexity and cost



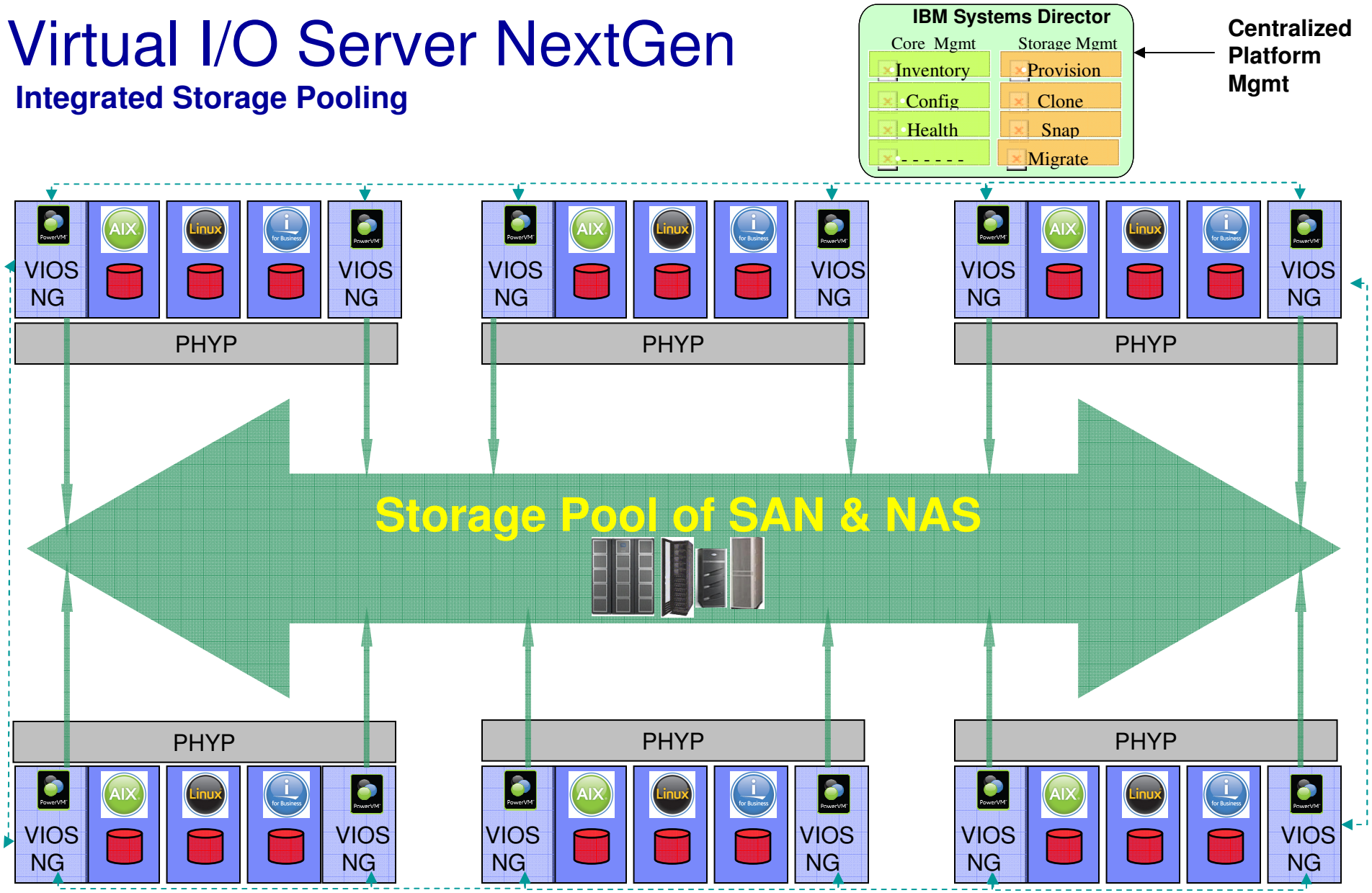
Virtual I/O Server (Classic)



Power your planet.

Virtual I/O Server NextGen

Integrated Storage Pooling



Power your planet.

vSCSI NextGen Capabilities

2. Base Capabilities

- Linked clones (space efficient virtual disks)
- Snapshot & rollback (virtual machine & device level)
- Thin provisioning & reclamation
- Storage tiering & multiple storage pools
- Live Storage Mobility
- Location transparency (server and storage mobility)

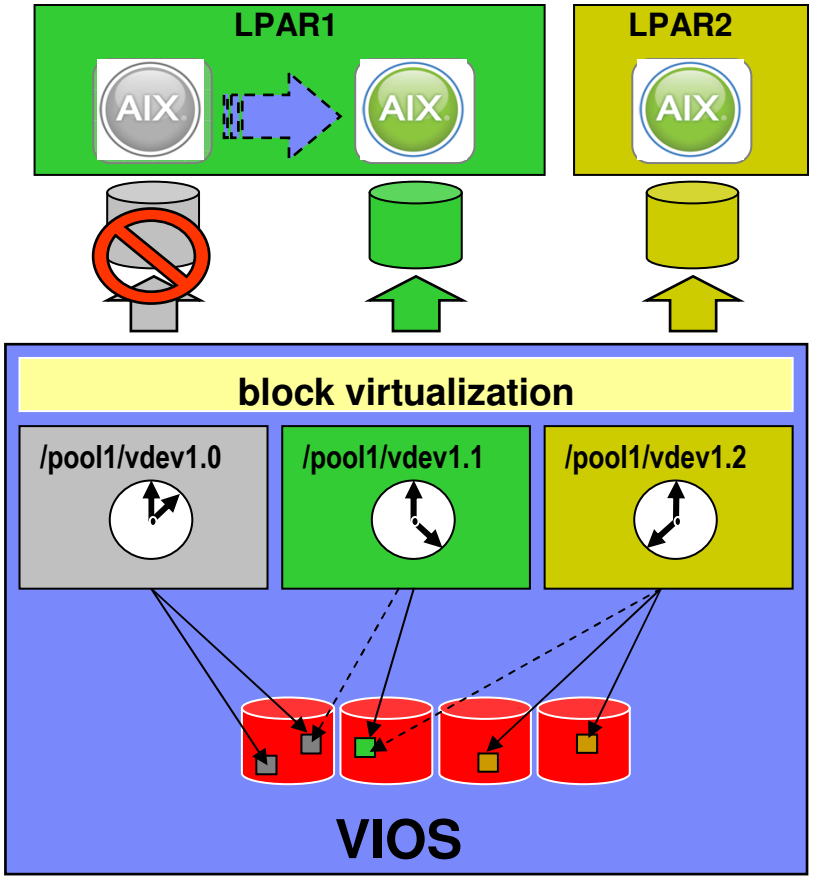
Advanced Capabilities

- Automation & Agility
  - quick & simple provisioning: storage, AMS, hibernate
- Image Management (capture / deploy, linked-clones)
- Consolidated backup / restore
- Life Cycle Management
- Security: isolation framework for multi-tenancy
- Disaster Recovery

vSCSI NextGen Capabilities

- **Snapshot & rollback**
  - Persistently saved snapshot of a point in time, rollback to prior snapshot
  - Consistency group snapshots
  - Sub-second creation
  - No practical limit on the number of snapshots that can be persisted
  - Basis for:
    - Capture / deploy
    - Disaster Recovery
    - Consolidated Backup / Restore

- **Linked Clones**
  - Blocks are all shared between logical copies initially, blocks diverge as they are modified
  - Basis for:
    - Storage utilization
    - Quick provisioning
    - Life Cycle Management



vSCSI NextGen Capabilities

- **Pools**
  - group of disks managed as a block of common storage
  - simplify the aggregation of large numbers of disks
  - better storage utilization
  - simplified administration
- **Tiers**
  - Virtually unlimited LUN segregation into storage tiers
  - Can leverage easy tier
- **Location transparency**
  - Virtual machines and their data
  - Enabler for mobility automation
- **Storage utilization**
  - Thin Provisioning (monitored)
  - Space efficient Linked Clones
  - Storage aggregation via pooling
- **Reliability**
  - Support for up to 10 data mirrors
  - De-clustered RAID layout and rebuild
- **Security**
  - Security domains extend from VM to storage pool
- **Scalability**
  - No practical limit on the number of virtual server objects, snapshots, clones

vSCSI NextGen

- **SAN Integration**
  - Geo and Metro mirror usage for Disaster Recovery
  - Server/storage integration for offload optimizations
- **Data Mobility and Consolidated Backup**
  - Online disk entry and replacement in the pool (Storage Mobility)
  - Import of pre-existing LUN with data in-place
  - Consolidated backup of Virtual Server data
- **Datacenter Operations**
  - Extends existing Power virtualization capabilities (VIOS) (no new layers)
  - Does not require changes in data center roles
  - Use existing data center resources including SAN and NAS storage

VIOS 2.2 transforming the datacenter

- **Simplification**
  - **Server & Storage integration (server and storage mgmt domains centralized)**
- **Provisioning time**
- **LifeCycle Management (first time power is applied to the server's retirement)**
  - **Image Management: capture, deploy**
  - **Agility (linked-clones)**
- **Location transparency**
- **Availability**
- **Backup / Restore & Disaster Recovery**
- **Compliance**

vSCSI NextGen

Overview

- **Power Server based Storage Virtualizer – clustered VIOS's**
  - Storage integration with the Power Server
  - Server and storage mgmt domains in one view
  - Control point that maps virtual machine to its storage
  - Distributed object repository and global namespace
  - Serves AIX, Linux, i, PHYP
  - Proven model
  - Staged delivery

Capabilities

- **Provides the base capabilities for Power servers…….**
  - Thin provisioned devices & reclamation
  - Linked clones
  - Snapshot / rollback
  - Live storage mobility
  - Storage pooling and tiering
- **… that are the basis for key Power services & capabilities**
  - Automation & Agility
    - quick & simple provisioning: storage, AMS, hibernate
  - Image Management (capture, deploy, linked-clones)
  - Consolidated backup / restore
  - Utilization (thin devices, reclamation, linked clones)
  - Location transparency (server and storage mobility)
  - Security: isolation framework for multi-tenancy
  - Disaster Recovery
- **Server / Storage collaboration**
  - Offload, optimize, accelerate data ops

Value

- **Simplification**
- **Agility (needed for emerging workloads such as Cloud)**
- **Location transparency (VM and storage)**
- **Improved utilization & TCO reduction**
  - cost savings through sharing; efficient utilization of phys I/O
  - enables more LPARs facilitating server consolidation (densit
  - reduce datacenter footprint (ie power, space)
  - reduce SAN infrastructure and SAN management costs
- **Facilitate and Simplifies Advanced PowerVM capabilities**
  - Platform Integration (ie AMS, Hibernation, etc)
  - Widespread exploitation of distributed storage (LPM,HA,DR)
- **Highly Scalable**
- **Reliability**
- **Uses existing DC structure (P6 P7 srvrs, VIOS, SAN / NAS)**

Supports all previous VIOS capabilities

- **Enterprise PowerVM**
  - LPAR mobility, AMS (active memory sharing)
- **, pBlade)**
  - Virtual SCSI, NPIV, SEA

Datacenter Operations

- **Extends existing Power virtualization capabilities (VIOS)**
- **Does not require changes in data center roles**
- **Use existing data center resources including SAN**

Power your planet.

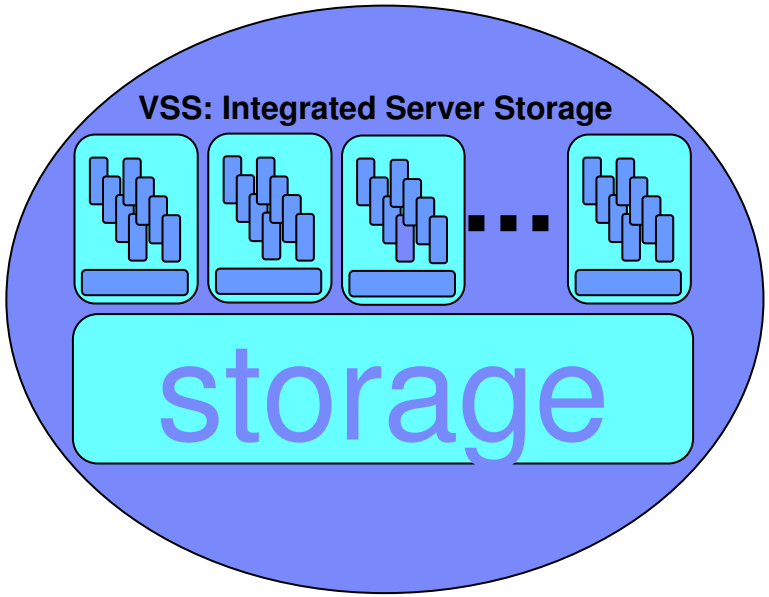
VIOS NG Base Capabilities - Staged

- vSCSI (standard vSCSI Target, including Persistent Reserve)
- Storage aggregation / pooling
- Thin provisioning (including notification framework)
- Thick provisioning
- Snapshot / rollback
- Consistency groups
- Linked-clones (space-efficient clones)
- Storage tiering
- Multiple storage pools
- Structured / distributed namespace
- CLI from any node in the cluster

Advanced Capabilities - Staged

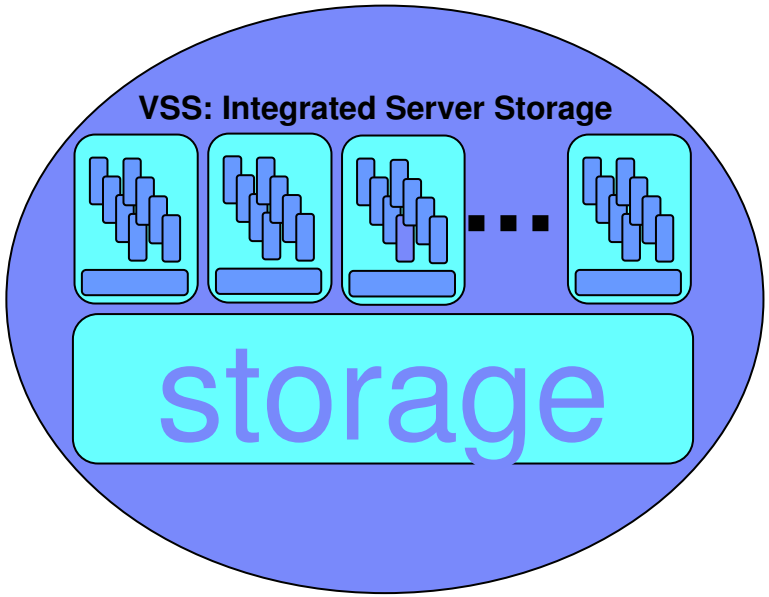
- Import existing storage to VIOS NextGen
- Automated provisioning (storage, AMS, Hibernation)
- Live Storage Mobility
- Application consistent snapshot framework
- Consolidated backup / restore framework
- Virtual optical
- Pool Mirroring
- Storage isolation infrastructure for multi-tenancy
- Server / storage integration (accelerate/offload data ops to SAN)
- NAS support (NAS filer on the back-end)
- vSCSI device data encryption, compression, de-dup
- Centralized management console (GUI)

Q & A



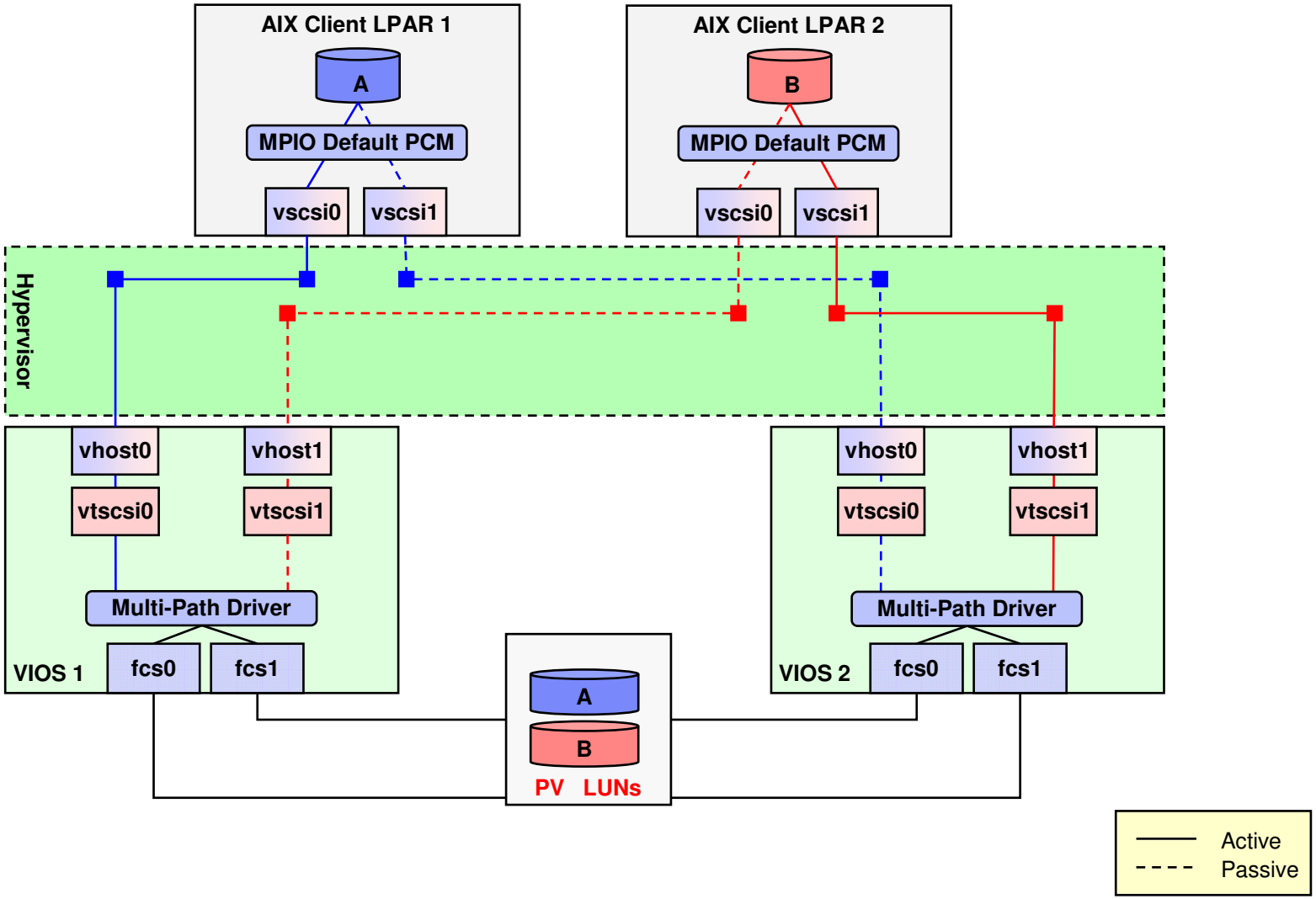
Power your planet.
Smarter systems for a Smarter Planet.

Backup Charts



Power your planet.
Smarter systems for a Smarter Planet.

Typical vSCSI redundant configuration



Power your planet.

(Banchy, 2010)

Capabilities: NPIV, vSCSI Classic, and vSCSI NextGen

	NPIV	vSCSI Classic	vSCSI NextGen
Server Based Storage Virtualization	×	✓	✓
Adapter Level Sharing	✓	✓	✓
Device Level Sharing (partition a device)	×	✓	✓
Server Based Storage Aggregation / Pooling	×	✓	✓
Automated provisioning	–	–	✓
Virtual server services (snapshot, clone, storage migration, IM, consolidated bkup.	?	×	✓
Platform Simplification	–	–	✓
Live Partition Mobility (LPM) capable	✓	✓	✓
Existing Storage Management Tools	✓	–	–
SCSI-3 Compliant (Persistent Reserve)	✓	×	✓
Generic Device Interface	×	✓	✓
Physical <-> Virtual device compatibility	✓	×	×
Peripheral Device Types			
Disk	✓	✓	✓
vOptical	×	✓	✓
vTape	×	✓	✓
Tape Libraries	✓	×	×

Power your planet.

Considerations when choosing NPIV / vSCSI

Both models offer:

- Efficient utilization of physical I/O; cost savings through sharing; VM density, etc.
- Reduce datacenter footprint (ie power, space) & SAN infrastructure

- **NPIV:**
  - Simplified provisioning (compared to vSCSI Classic)
  - Simplified Management (including LPM)
  - Clear view of virtual machine allocation and usage
  - Shared physical resources without breaking existing tools / solutions
  - Tape libraries
  - Load balancing (active/active) and heterogeneous multipathing options
  - SCSI-3 Persistent Reserve
  - Physical ←→ Virtual device compatibility

- **vSCSI Classic**
  - Standard / generic interface
  - Storage Virtualization
    - Disk / Optical / Tape
    - Device sharing
    - Virtualization of non-SAN storage (iSCSI, SAS, parallel SCSI, etc)
  - Server based storage pooling
  - Agility: quick LPAR deployment

Considerations cont'd

- VIOS NextGen vSCSI will take time to mature
  - Capabilities will be staged over time
  - Integration with platform management & enterprise management solutions
- VIOS NextGen being developed to meet the needs of emerging technologies and the datacenter of the future

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:

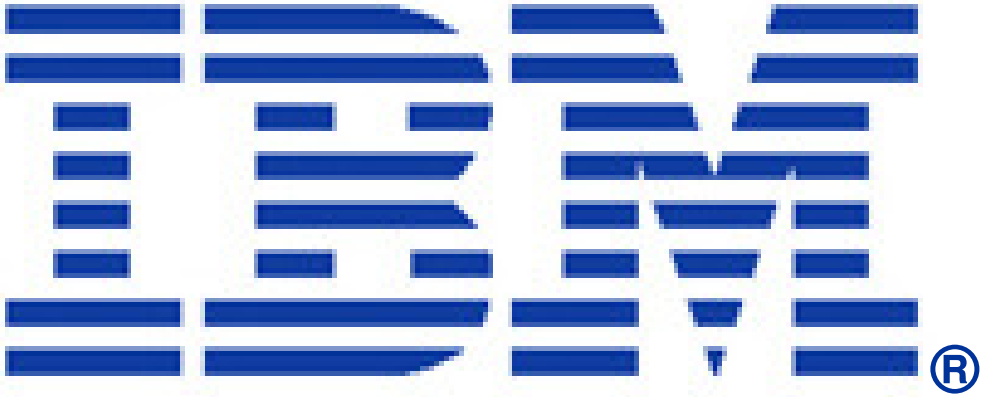
*, AS/400®, e business(logo)®, DBE, ESCO, eServer, FICON, IBM®, IBM (logo)®, iSeries®, MVS, OS/390®, pSeries®, RS/6000®, S/30, VM/ESA®, VSE/ESA, WebSphere®, xSeries®, z/OS®, zSeries®, z/VM®, System i, System i5, System p, System p5, System x, System z, System z9®, BladeCenter®

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 multiprogramming 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.



Power your planet.