Hewlett Packard Enterprise

NonStop Technical Boot Camp 2023
TBC23-TB57 HPE Virtualized NonStop Storage

Marcelo de Azevedo and Lars Plum

September 2023

Forward-looking statements

This is a rolling (up to three year) Roadmap and is subject to change without notice

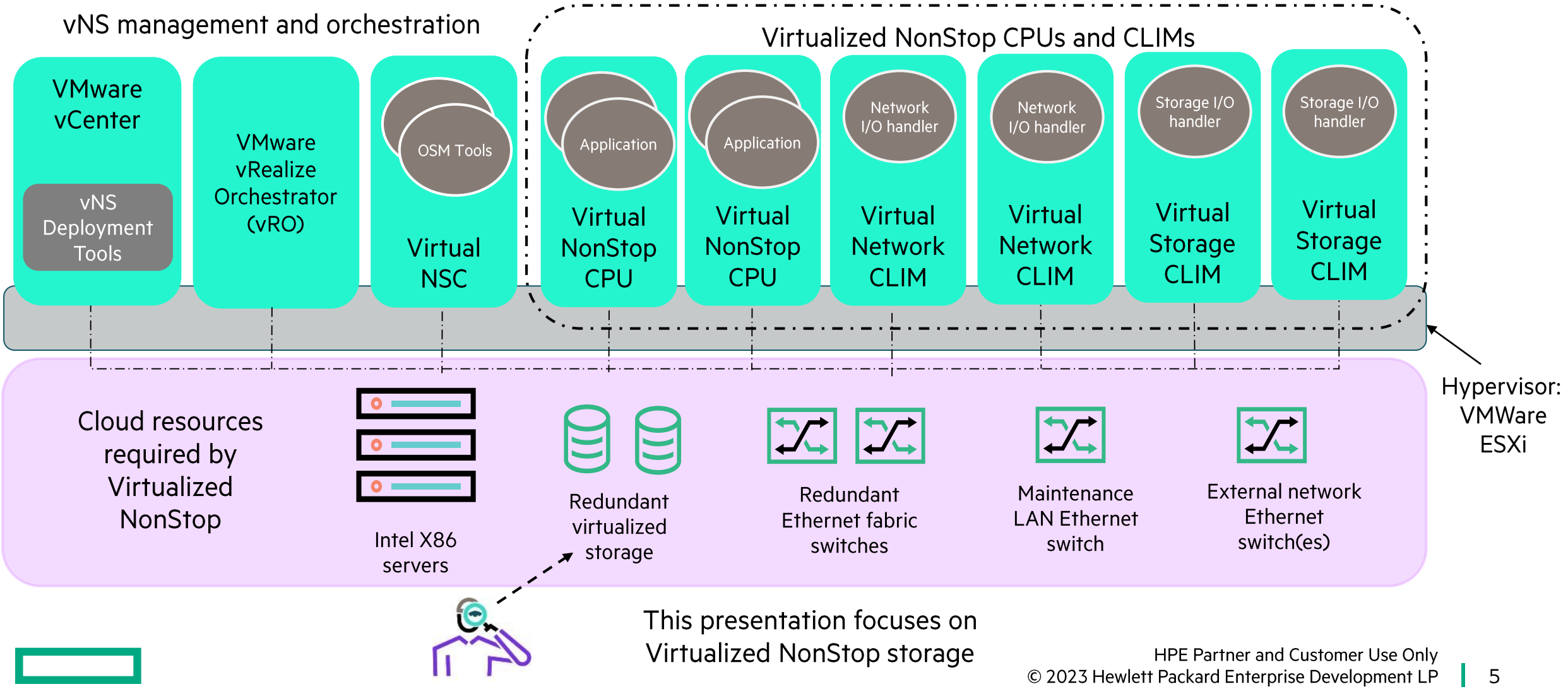
This document contains forward looking statements regarding future operations, product development, product capabilities and availability dates. This information is subject to substantial uncertainties and is subject to change at any time without prior notification. Statements contained in this document concerning these matters only reflect Hewlett Packard Enterprise's predictions and / or expectations as of the date of this document and actual results and future plans of Hewlett Packard Enterprise may differ significantly as a result of, among other things, changes in product strategy resulting from technological, internal corporate, market and other changes. This is not a commitment to deliver any material, code or functionality and should not be relied upon in making purchasing decisions.

Agenda

- HPE Virtualized NonStop storage overview
- HPE Virtualized NonStop storage options
- HPE Virtualized NonStop LUNs (Logical Unit Numbers)
- New HPE Virtualized NonStop storage features
- Related TBC talks and resources
- Backup slides: Examples of failure handling scenarios

HPE Virtualized NonStop storage overview

Virtualized NonStop on VMware
Architectural diagram



This presentation focuses on Virtualized NonStop storage

VMware storage virtualization overview

- VMWare supports multiple storage technologies and storage connectivity options
- vNS requires block storage devices that are supported by VMware
  - Refer to the *VMware Storage/SAN Compatibility Guide* for more details (https://www.vmware.com/resources/compatibility/pdf/vi_san_guide.pdf)
- Examples tested with vNS:
  - Storage arrays:
    - HPE MSA 2050 and MSA 2060 Modular Smart Arrays
    - HPE Nimble AF20 All Flash Storage
    - HPE 3PAR StoreServ 7000
    - HPE XP7 Storage
  - Internal drives in ESXi hosts (with and without VMware vSAN)
- Regardless of the storage and connectivity options used, the ESXi hypervisor presents storage to the VMs as virtual disks connected via a virtual SCSI interface
- ESXi provides multiple software controller options for the virtual SCSI interface to storage
- For vNS, the software SCSI controller must be PVSCSI (paravirtual SCSI)

Virtualized NonStop Storage vCLIMs

- Storage is provisioned as block devices by the orchestrator and presented to Storage vCLIMs through the PVSCSI interface
- NonStop volumes are subsequently configured for vNS access through SCF (Subsystem Control Facility)
- Storage vCLIMs are deployed with dedicated processor cores (4 or 8 cores) and memory (8 GB) and are subsequently configured via OSM
  - 8 cores is the default configuration and is required if Volume Level Encryption (VLE) is in use
- Storage vCLIMs belonging to the same failover pair must be deployed on different physical servers
- Number of storage vCLIMs per vNS system:
  - High end: 2 to 54
  - Entry class: 2 to 4

Virtualized NonStop Storage vCLIMs (continued)

- If VLE is used, storage vCLIMs require connectivity to a pair of Enterprise Secure Key Managers (ESKMs)
  - These are IP connections, and can be provisioned over a VMXNET3 interface of the storage vCLIM
- Storage CLIMs require storage I/O cards supported by VMware
  - Refer to the *VMware I/O Compatibility Guide* for more details (https://www.vmware.com/resources/compatibility/pdf/vi_io_guide.pdf)
- For connecting to external SAN storage, it is recommended to have one 2-port storage NIC in the physical server for each Storage vCLIM deployed on it
- External storage connectivity options: iSCSI (Ethernet) or Fibre Channel (FC) networks
- For backup needs, HPE vNS supports these options:
  - Virtual BackBox (vBB) using iSCSI connectivity
  - Physical BackBox connected using FC passthrough
- Multiple HPE vNS or converged HPE NonStop systems can be connected to a virtual BackBox or to a physical BackBox

HPE Virtualized NonStop storage requirements

Volume	Size[1]	Remarks
$SYSTEM	100 to 1600 GB	1600 GB is a new limit as of L22.09
$AUDIT (or possibly multiple audit volumes)	100 to 1600 GB	
$DSMSCM	100 to 1600 GB	
$OSS (or possibly multiple OSS volumes)	100 to 1600 GB	
$SWAP (or possibly multiple swap volumes)	100 to 1600 GB	Use the formula 1/2 x memory per CPU x number of CPUs
$DATA volumes	1 to 1600 GB	Based on user requirements
Storage vCLIM OS (first pair of storage vCLIMs)	300 GB	600 GB can also be selected on high-end systems
Storage vCLIM OS (additional storage vCLIMs)	100 GB	
IP and Telco vCLIM OS	100 or 300 GB	Use larger size to support longer TCP/IP monitor dumps
HPE vNSC	250 GB	
HPE Virtual BackBox	300 GB	

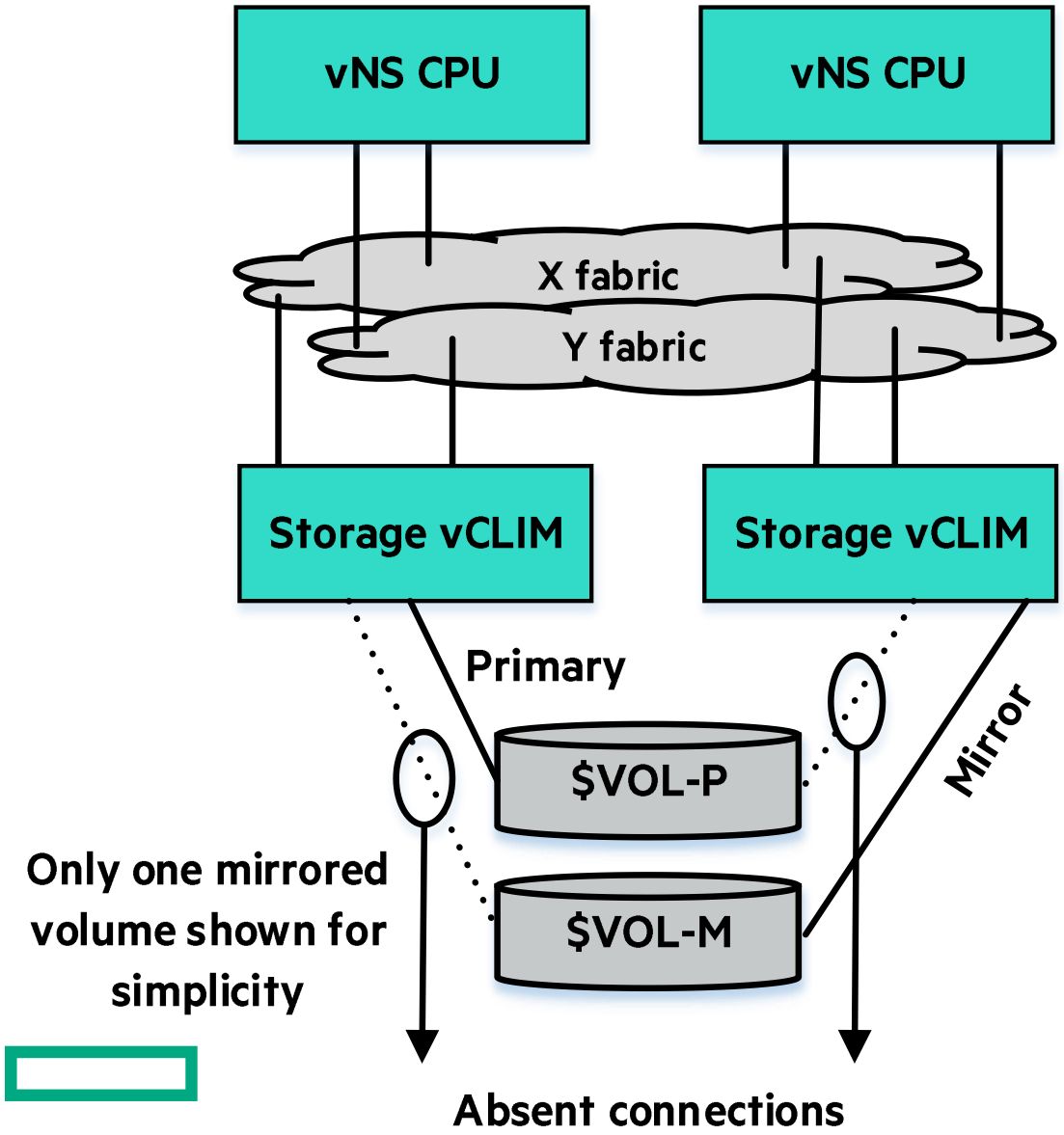
[1] All sizes must be specified in 1 GB increments

HPE Virtualized NonStop storage requirements

- Separate physical storage resources for fault tolerance (see more details later)
- All disks must be thick provisioned (“eagerzeroedthick”)
  - This means that the entire provisioned space for a disk is committed to the virtual disk, and
  - The disk is written with zeroes upon provisioning
- Reasons for “eagerzeroedthick” provisioning requirement:
  - Ensure low latency on disk I/Os by avoiding storage provisioning during system operations
  - Required by VMware to support multi-writes
    - Multi-writes: configurations with a virtual drive attached to two VMs such as a storage vCLIM pair (see more details later)
- Storage requirements for VMware vCenter and vRealize Orchestrator also apply
  - For information on storage requirements for VMware products go to https://docs.vmware.com/

HPE Virtualized NonStop storage options

2c-2d configuration with 2 paths to storage



Storage technology

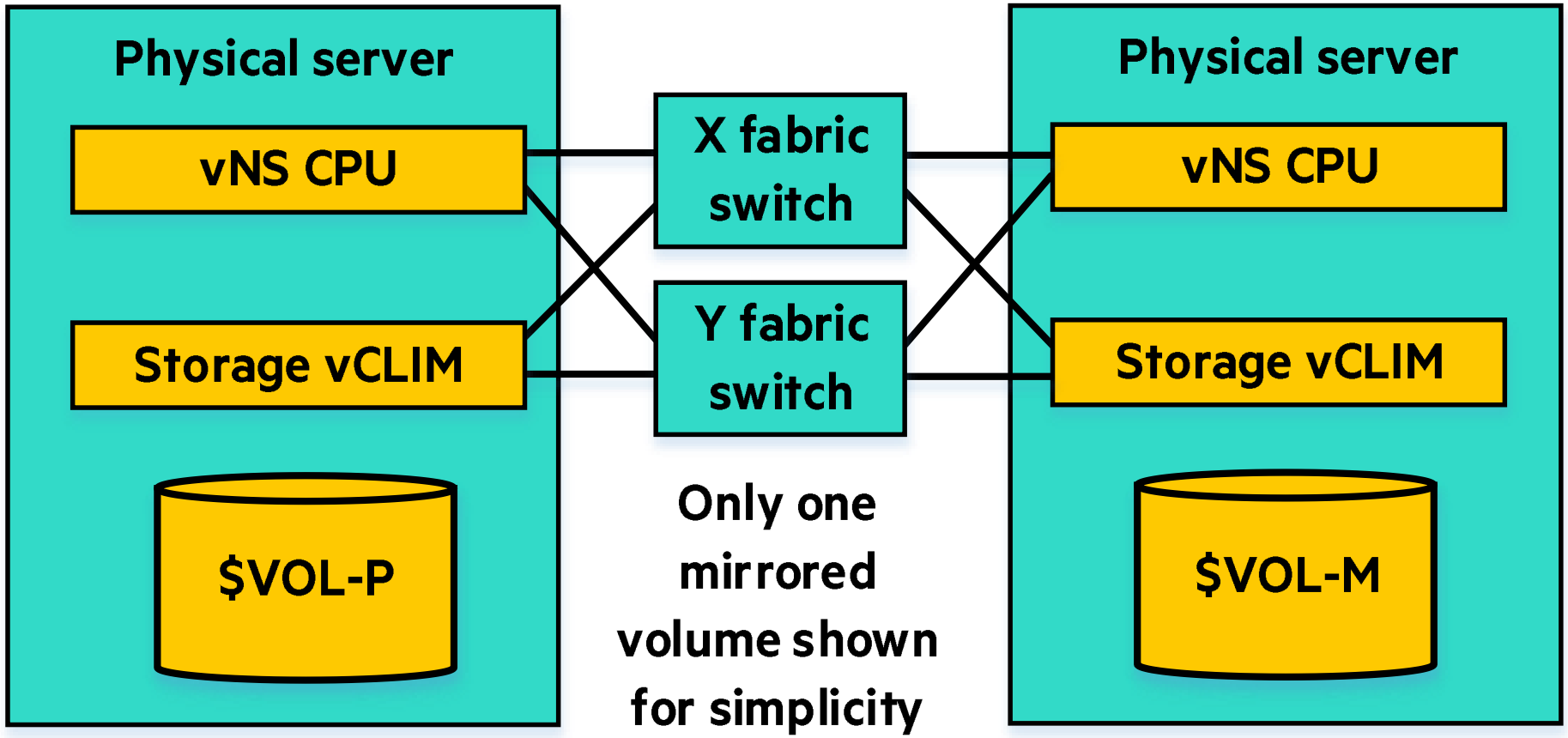
- Internal drives in ESXi hosts without vSAN
- SSD drives recommended for performance
- Requires redundant datastores

Key considerations

- DP2 disk revives required after loss of an SCLIM or ESXi host due to planned maintenance or failure
- Rolling upgrades of ESXi hosts not supported[1]

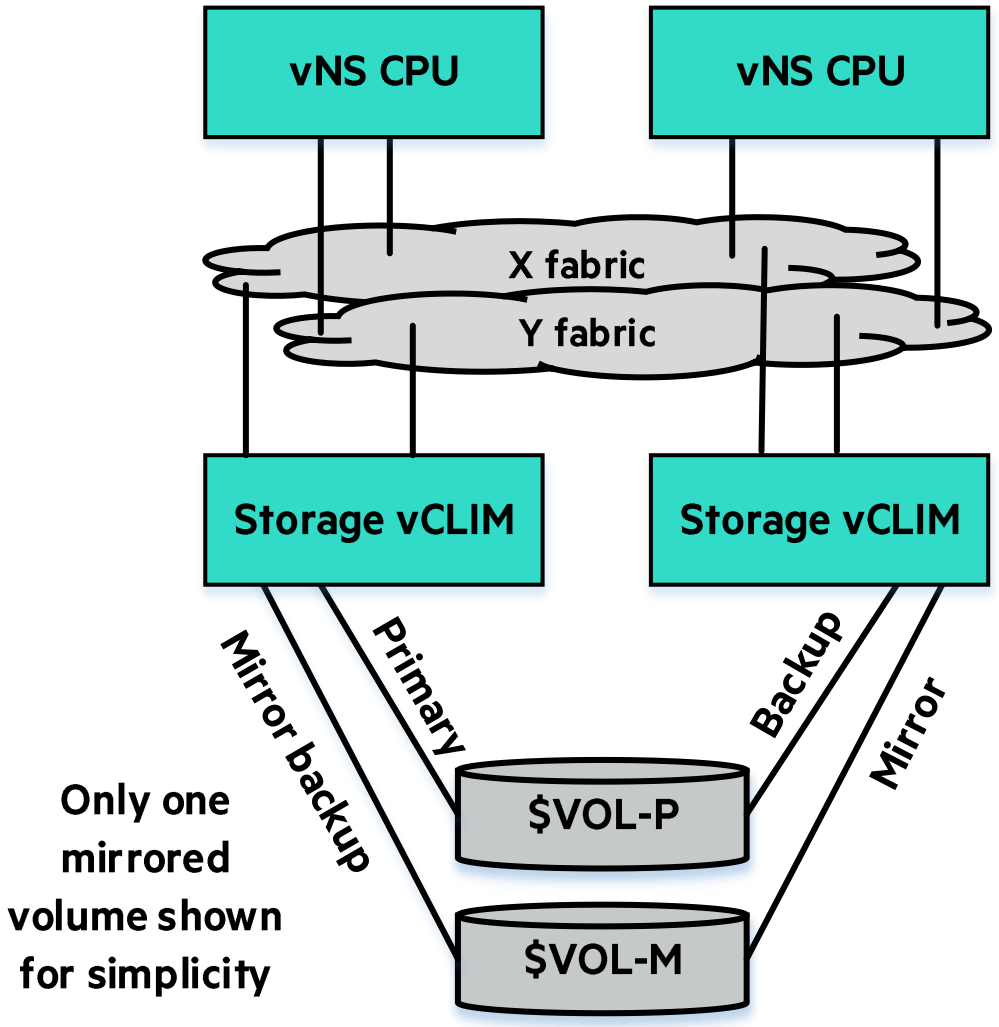
[1] Rolling upgrades would be time consuming since it would require migrating Storage vCLIMs and associated drives.

2c-2d 2-path configuration with internal drives in ESXi hosts



This example shows: RAID 0 (within ESXi host) + RAID 1 (NonStop volume mirroring)

2c-2d configuration with 4 paths to storage



2c-2d configuration

Storage technology

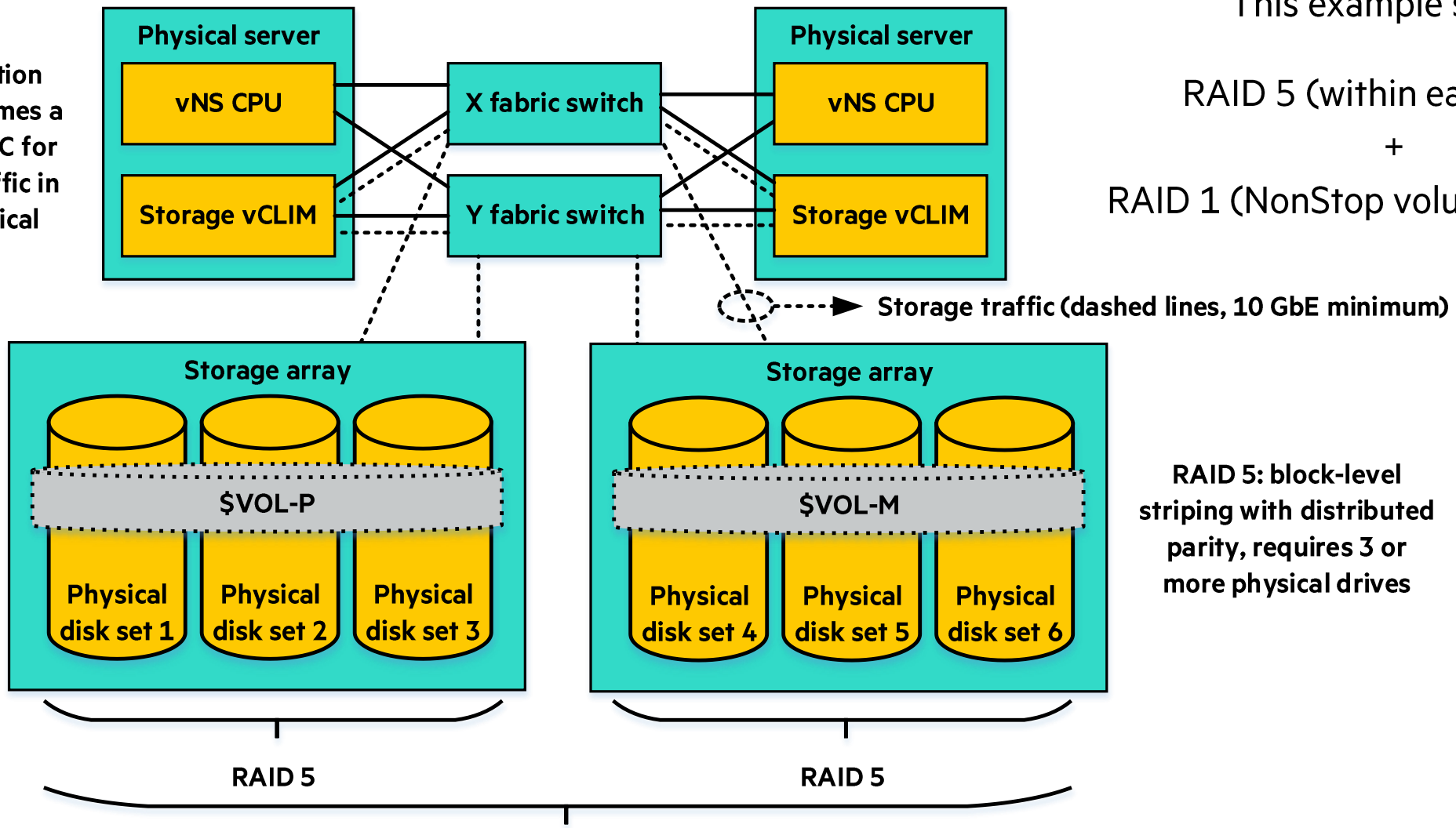
- External storage arrays or VMware vSAN
- Requires redundant datastores

Key considerations

- DP2 disk revives not required after loss of an SCLIM or ESXi host
- However, if an ESXi host in a vSAN cluster is lost and replaced the vSAN RAID array must be rebuilt
- Rolling upgrades of ESXi hosts ✓

2c-2d 4-path configuration with external storage arrays

Configuration shown assumes a separate NIC for storage traffic in each physical server



Only one mirrored volume shown for simplicity

This example shows:

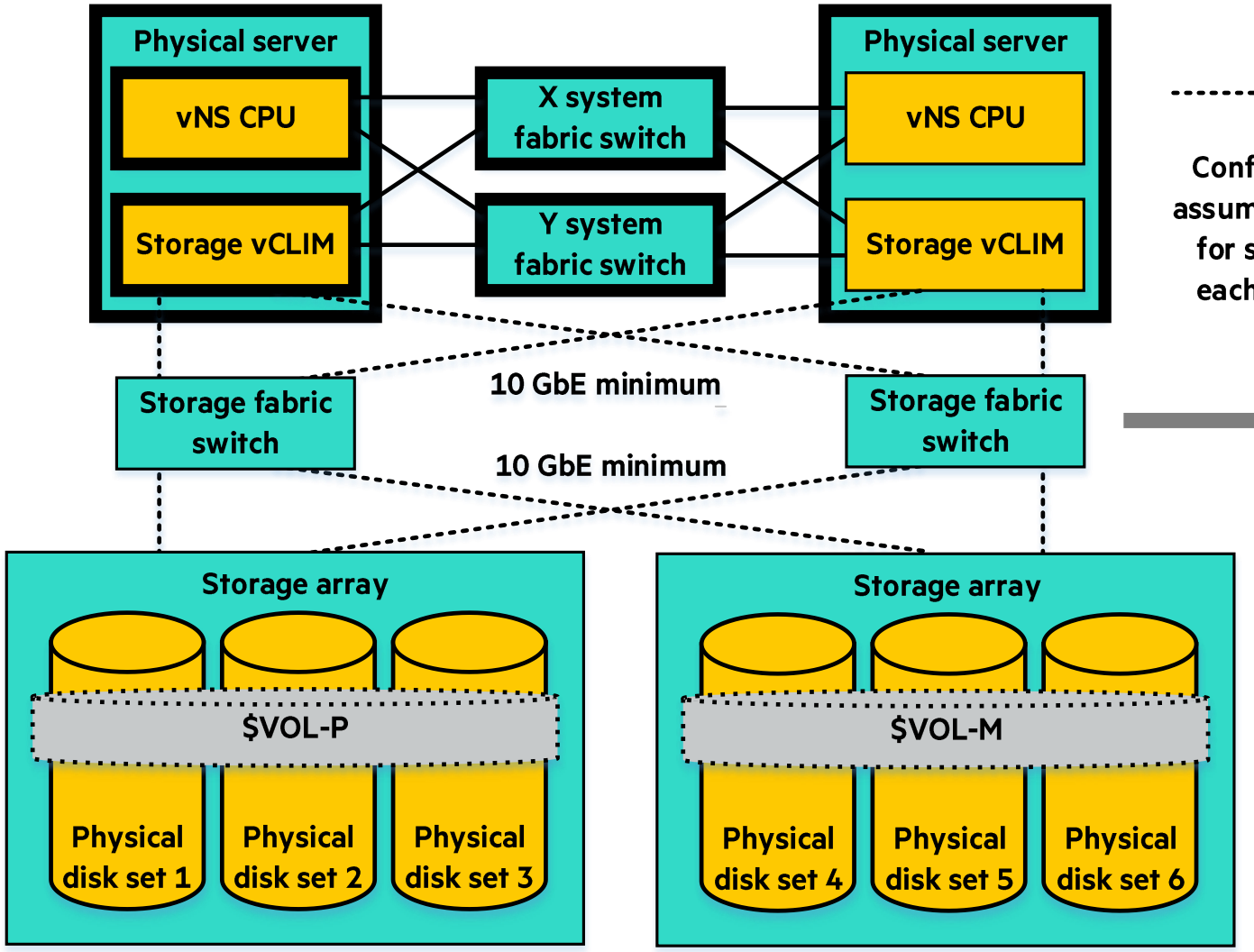
RAID 5 (within each array)

+

RAID 1 (NonStop volume mirroring)

RAID 5: block-level striping with distributed parity, requires 3 or more physical drives

2c-2d 4-path configuration with storage arrays and storage fabric switches



Legend:

-------- Storage traffic

Configuration shown assumes a separate NIC for storage traffic in each physical server

Storage data traffic flows through storage fabric switches as opposed to through the system fabric switches

This example shows:

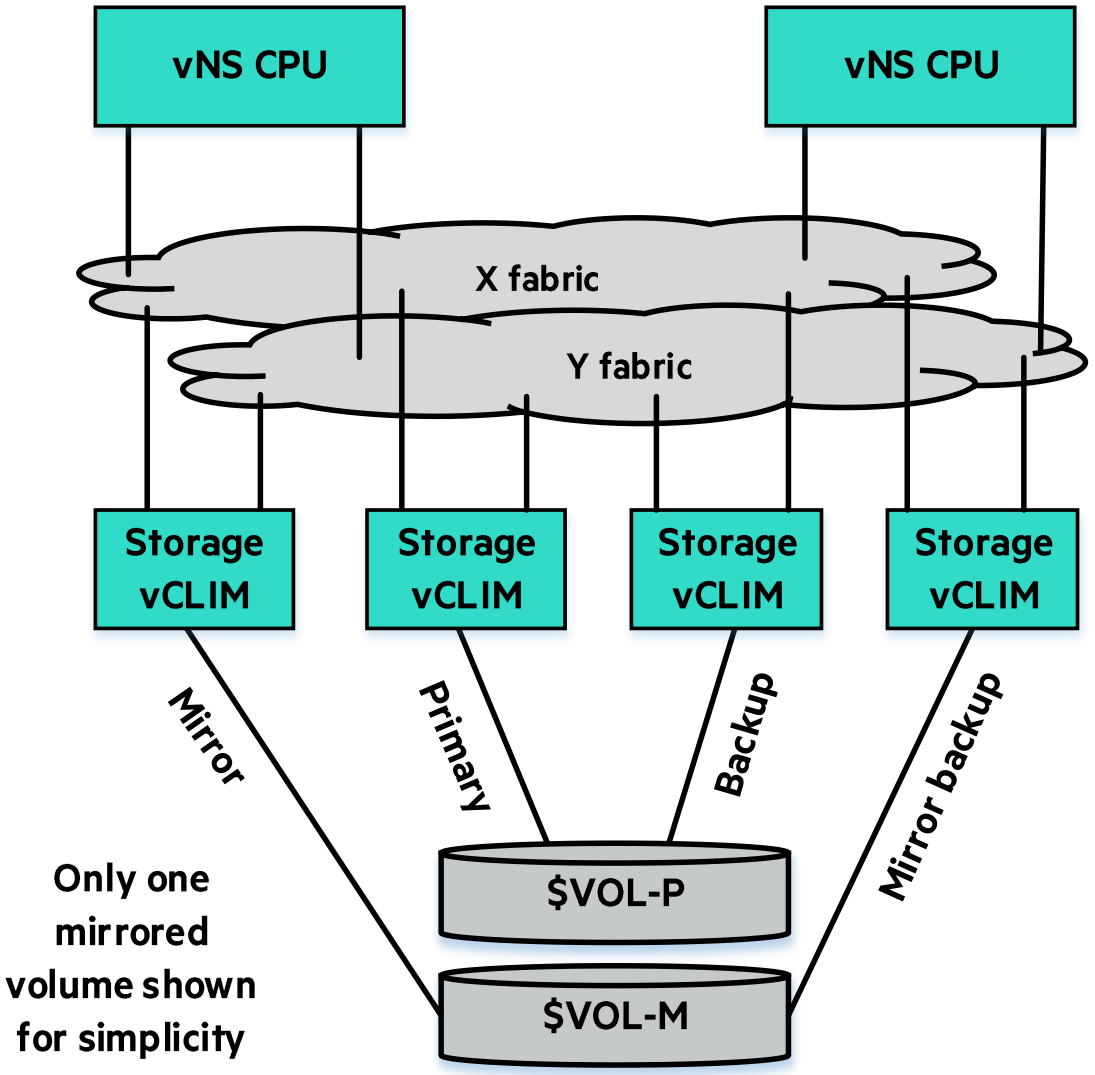
RAID 5 (within each array)

+

RAID 1 (NonStop volume mirroring)

Only one mirrored volume shown for simplicity

4c-2d configuration with 4 paths to storage



4c-2d configuration

Storage technology

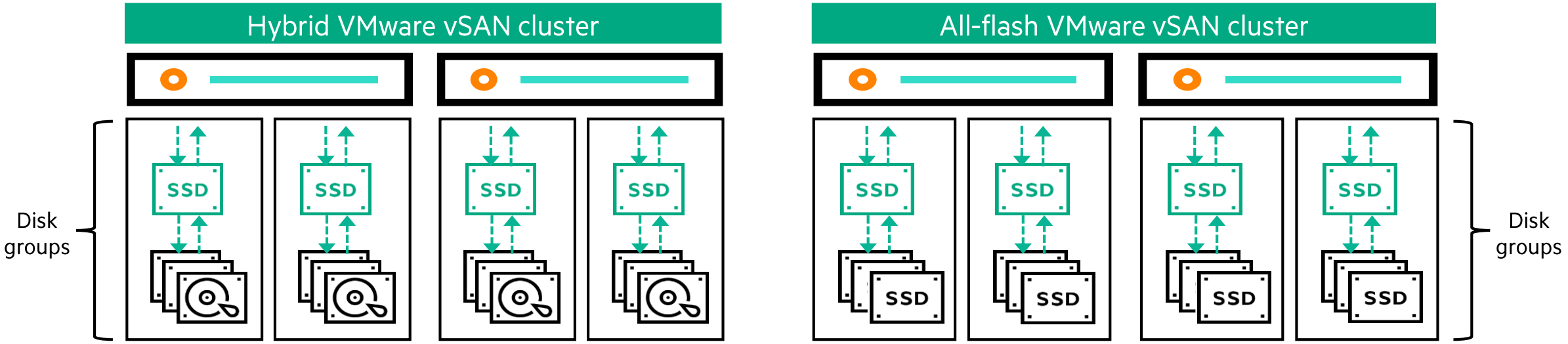
- External storage arrays or VMware vSAN
- Requires redundant datastores

Key considerations

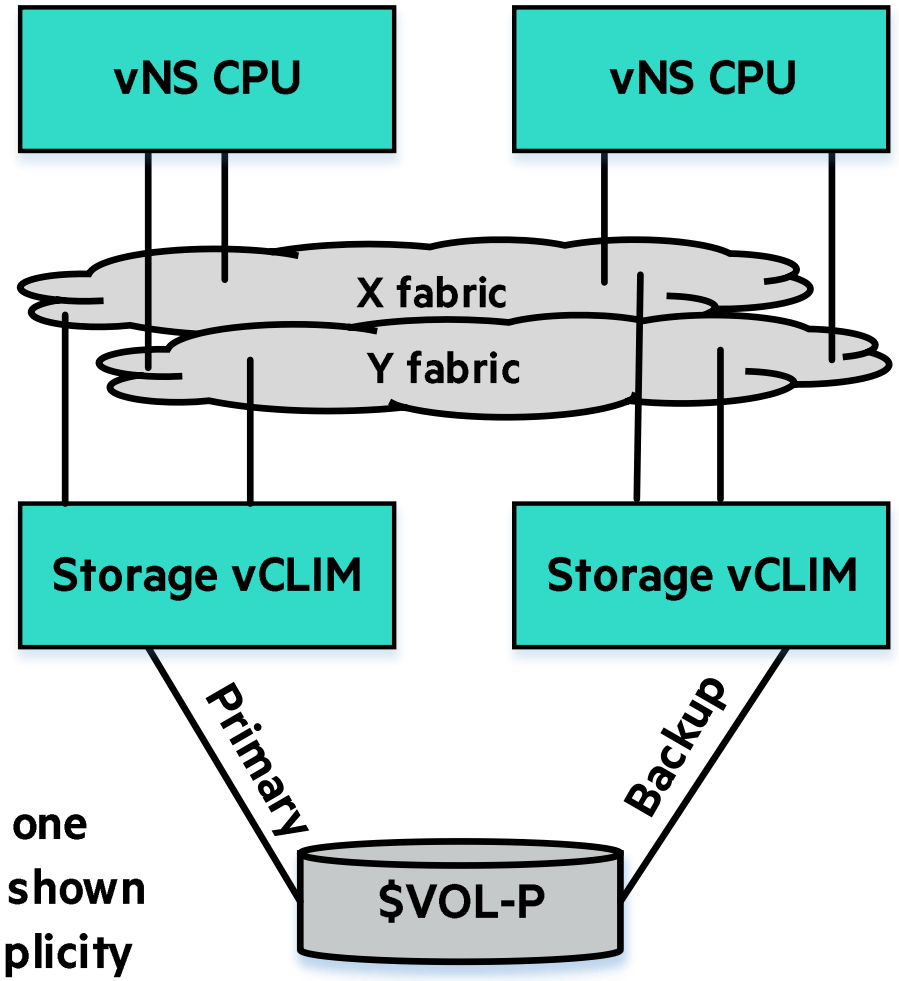
- DP2 disk revives not required after loss of an SCLIM or ESXi host
- However, if an ESXi host in a vSAN cluster is lost and replaced the vSAN RAID array will need to be rebuilt
- Rolling upgrades of ESXi hosts ✓
- Better load balancing upon loss of a CLIM

vNS on Hyperconverged Storage: VMware vSAN

- VMware vSAN is a storage virtualization software that abstracts and pools physical storage belonging to a cluster of servers and presents them as datastores to the VMs above
- Disk groups reside within the ESXi hosts that provide storage capacity to a vSAN cluster
- In the vSAN Original Storage Architecture (OSA), each disk group must have one flash cache device and one or multiple capacity devices
  - Hybrid vSAN clusters uses flash devices for the cache layer and HDDs for the capacity layer
  - All-flash vSAN clusters use flash devices for both cache and capacity



2c-1d configuration with 2 paths to storage



Only one volume shown for simplicity

2c-1d configuration

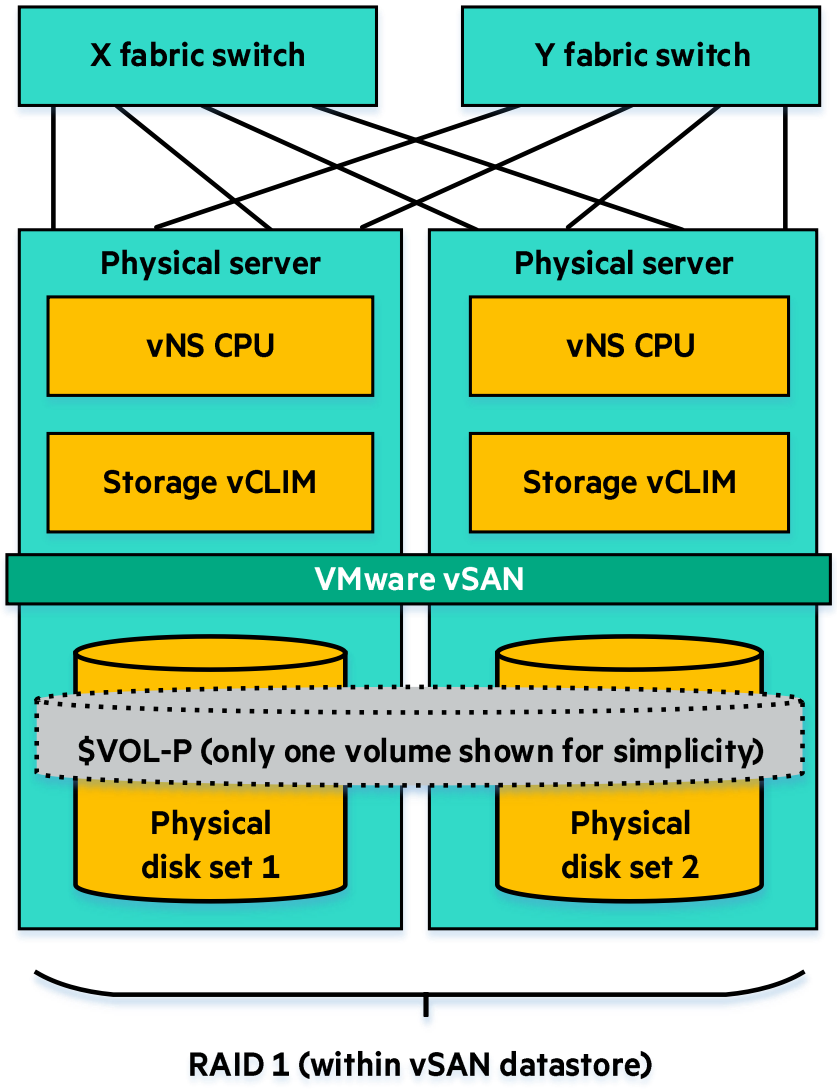
Storage technology

- Leverages built-in VMware vSAN redundancy to achieve fault tolerance with a single RAID-1 datastore
- Does not use NonStop DP2 disk mirroring

Key considerations

- Rolling upgrades of ESXi hosts ✓
- NSVLE (NonStop Volume Level Encryption) not supported

2c-1d 2-path configuration with single vSAN cluster



- This configuration relies on RAID 1 fault tolerance provided by VMware vSAN for storage availability
- This configuration does not offer redundancy through NonStop DP2 mirroring
- NonStop Volume Level Encryption (NSVLE) is not supported in this configuration
  - NSVLE key rotation is not possible without NonStop DP2 mirroring

vSAN OSA versus ESA

- vSAN 8 introduces a new architecture referred to as ESA (Express Storage Architecture)
- ESA uses a single tier of NVMe devices for storage instead of separate drives for cache and for capacity
- ESA requires a minimum of 512GB in host server memory vs a minimum of 32GB in host server memory for OSA
- HPE Virtualized NonStop has not been tested with ESA yet

Overview of supported storage configurations

Configuration	2c-2d	4c-2d	2c-1d
Number of CLIMs with access to each NonStop logical volume	2	4	2
Virtual drives per NonStop logical volume	2	2	1
Redundancy through NonStop DP2 mirroring?	Yes	Yes	No
Number of paths between CLIMs and NonStop logical volume	2 or 4	4	2
Supported with internal drives without VMware vSAN?	Yes, but with 2 paths only	No	No
Supported with external storage arrays?	Yes	Yes	No
Supported with VMware vSAN?	Yes	Yes	Yes
NonStop Volume Level Encryption (NSVLE) supported?	Yes	Yes	No

HPE Virtualized NonStop LUNs (Logical Unit Numbers)

HPE Virtualized NonStop LUNs (Logical Unit Numbers)

- Each storage vCLIM can support up to 100 primary LUNs and 100 mirror LUNs (i.e., a total of 200 LUNs)
- Number of mirrored volumes that can be configured with a storage vCLIM pair: from 1 to 100
- Examples of LUNs displayed by `SCF INFO DISK $<volume_name>, DETAIL`:

SCF INFO DISK $SYSTEM, DETAIL (2c2d, 4 paths)

```
Primary Path Information:
 CLIM..................................... SCLIM000
 LUN...................................... 10001

Backup Path Information:
 CLIM..................................... SCLIM001
 LUN...................................... 10001

Mirror Path Information:
 CLIM..................................... SCLIM001
 LUN...................................... 10201

Mirror Backup Path Information:
 CLIM..................................... SCLIM000
 LUN...................................... 10201
```

SCF INFO DISK $OSS, DETAIL (2c2d, 2 paths)

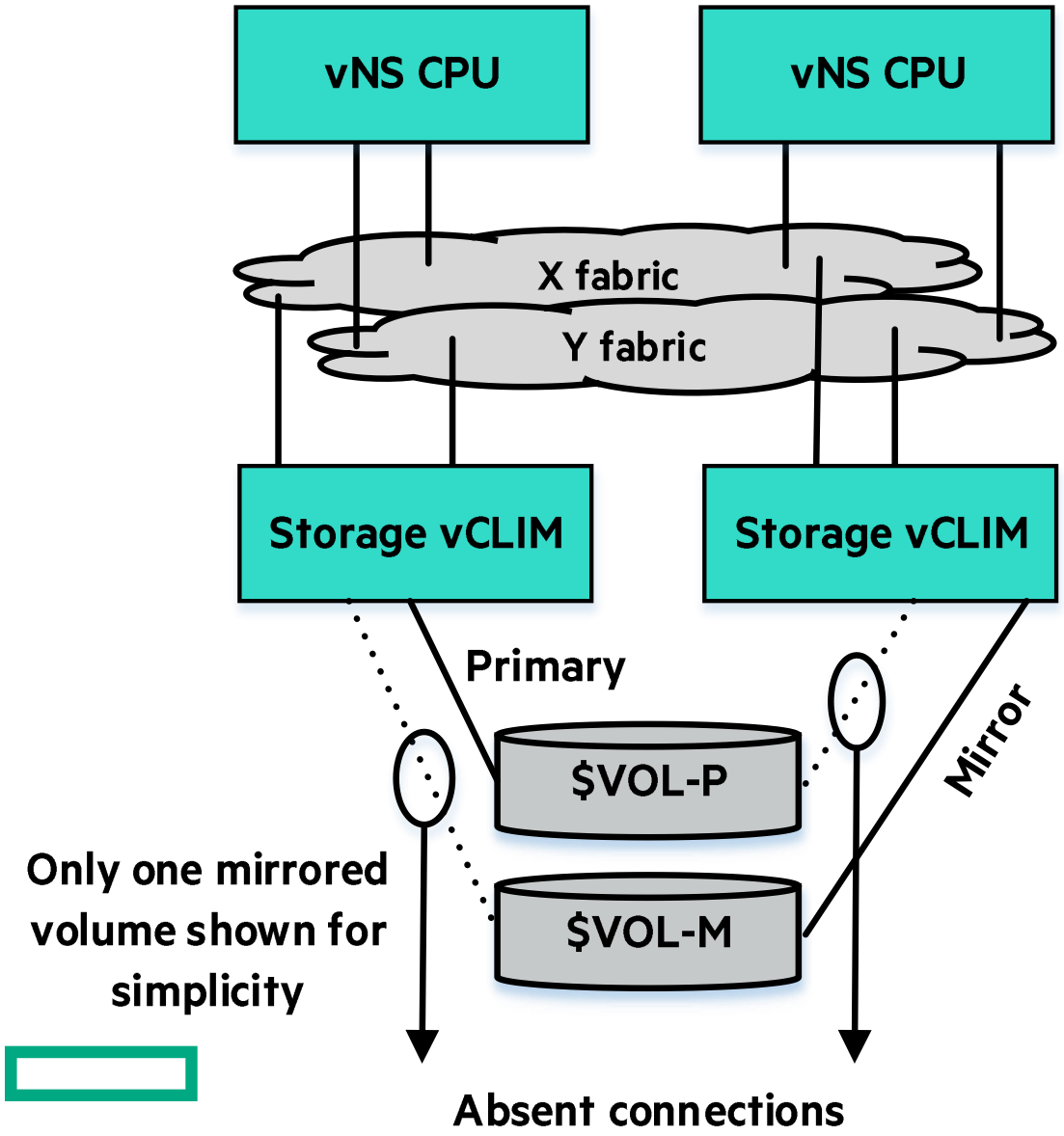
```
Primary Path Information:
 CLIM..................................... SCLIM000
 LUN...................................... 10004

Mirror Path Information:
 CLIM..................................... SCLIM001
 LUN...................................... 10004
```

HPE Virtualized NonStop LUNs (Logical Unit Numbers) – Continued

- More examples of LUNs displayed by `SCF INFO DISK $<volume_name>, DETAIL`:

SCF INFO DISK $SYSTEM, DETAIL (4c2d, 4 paths)	**SCF INFO DISK $SYSTEM, DETAIL (2c1d, 2 paths)**
Primary Path Information: CLIM..... SCLIM000 LUN..... 10001 Backup Path Information: CLIM..... SCLIM001 LUN..... 10001 Mirror Path Information: CLIM..... SCLIM002 LUN..... 10201 Mirror Backup Path Information: CLIM..... SCLIM003 LUN..... 10201	Primary Path Information: CLIM..... SCLIM000 LUN..... 10001 Backup Path Information: CLIM..... SCLIM001 LUN..... 10001

2c-2d configuration with 2 paths to storage

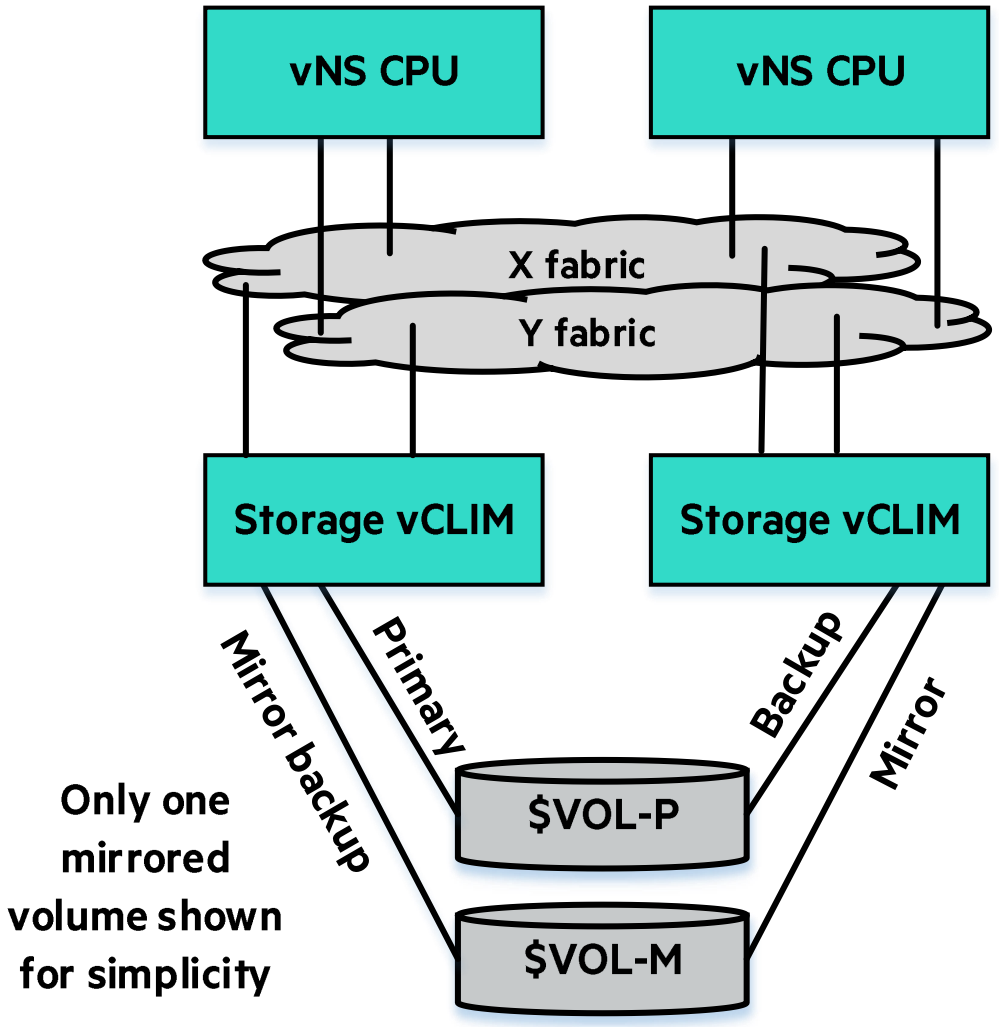


- Redundant datastores configured in internal drives in ESXi hosts without vSAN

Path	LUN number range
Primary	10001 – 10200
Backup	N/A
Mirror	10001 – 10200
Mirror backup	N/A

Example	LUN number
$VOL primary	10060
$VOL backup	N/A
$VOL mirror	10060
$VOL mirror backup	N/A

2c-2d configuration with 4 paths to storage



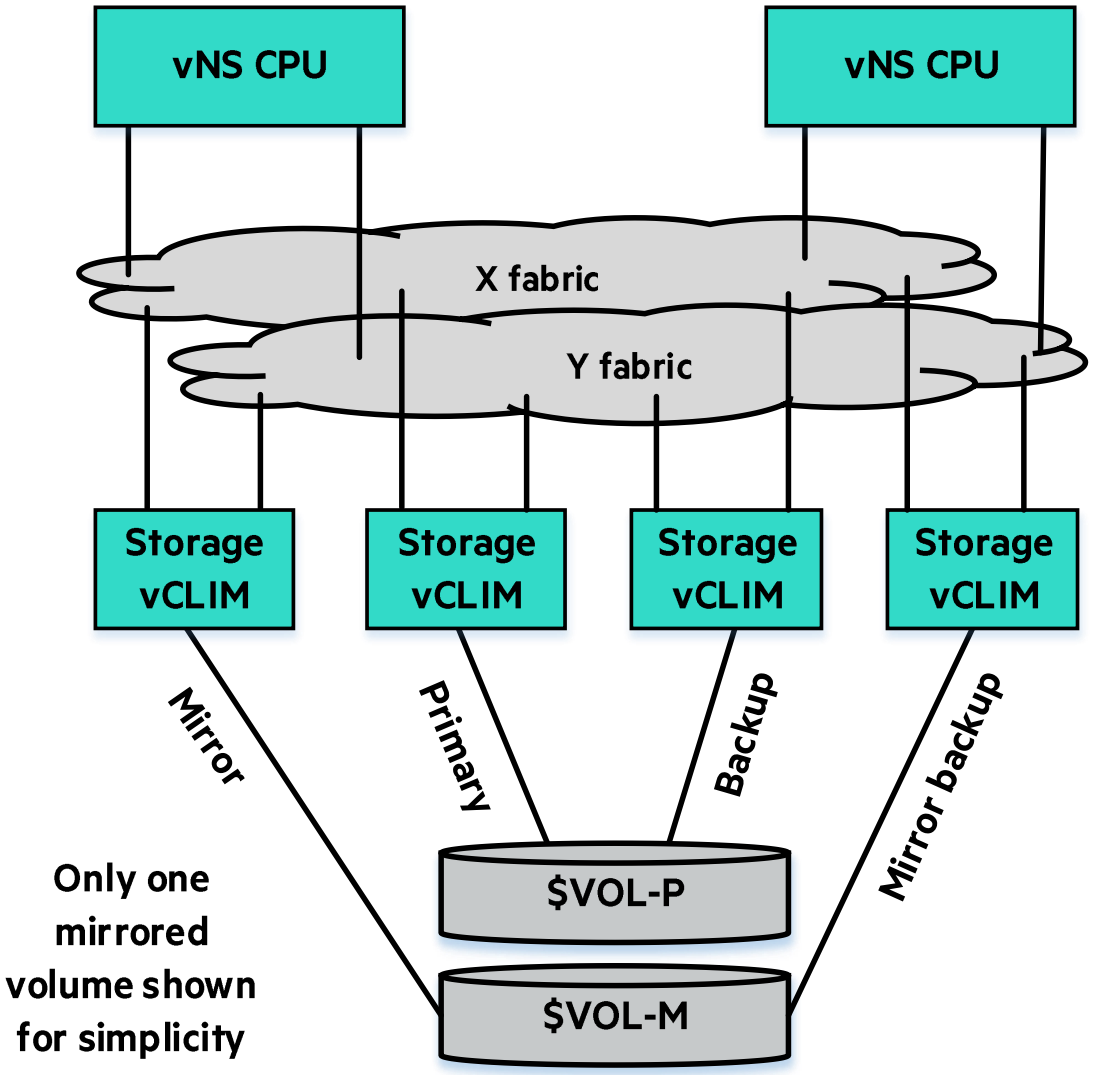
2c-2d configuration

- Redundant datastores configured on external storage arrays or VMware vSAN

Path	LUN number range
Primary	10001 – 10200
Backup	10001 – 10200
Mirror	10201 – 10400
Mirror backup	10201 – 10400

Example	LUN number
$VOL primary	10060
$VOL backup	10060
$VOL mirror	10260
$VOL mirror backup	10260

4c-2d configuration with 4 paths to storage



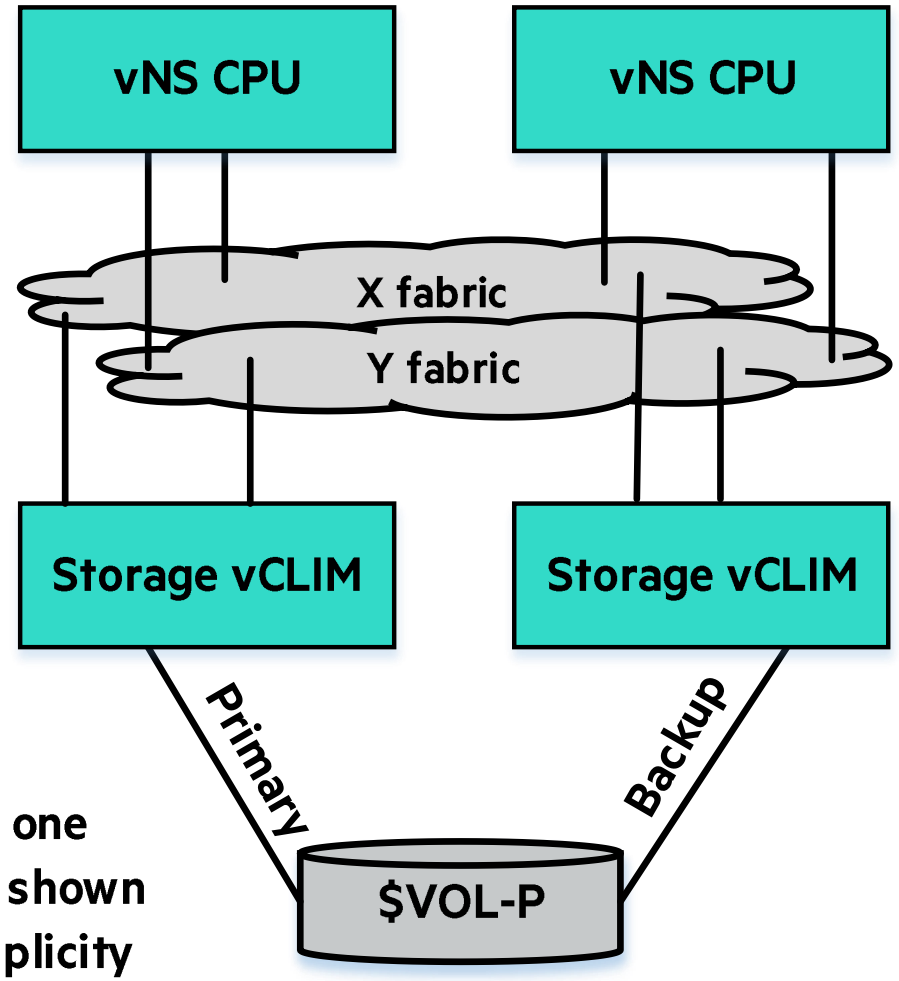
4c-2d configuration

- Redundant datastores configured on external storage arrays or VMware vSAN

Path	LUN number range
Primary	10001 – 10200
Backup	10001 – 10200
Mirror	10201 – 10400
Mirror backup	10201 – 10400

Example	LUN number
$VOL primary	10060
$VOL backup	10060
$VOL mirror	10260
$VOL mirror backup	10260

2c-1d configuration with 2 paths to storage



2c-1d configuration

- Configuration without NonStop DP2 disk mirroring
- Leverages VMware vSAN redundancy to achieve fault tolerance with a single RAID-1 datastore

Path	LUN number range
Primary	10001 – 10200
Backup	10001 – 10200
Mirror	N/A
Mirror backup	N/A

Example	LUN number
$VOL primary	10060
$VOL backup	10060
$VOL mirror	N/A
$VOL mirror backup	N/A

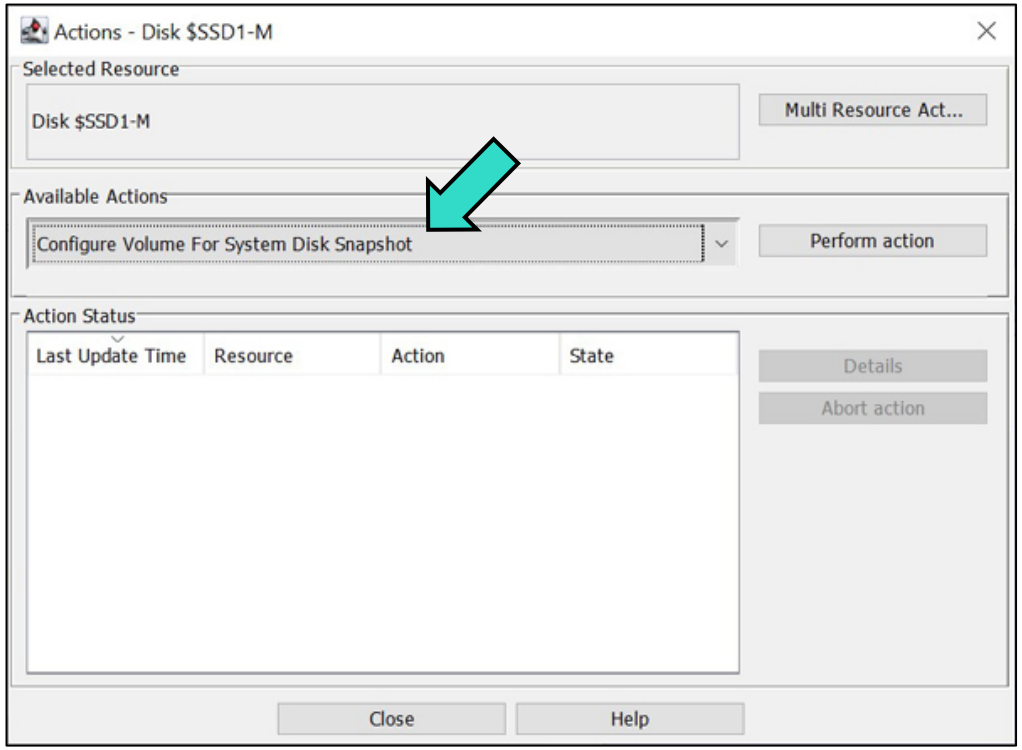
New HPE Virtualized NonStop storage features

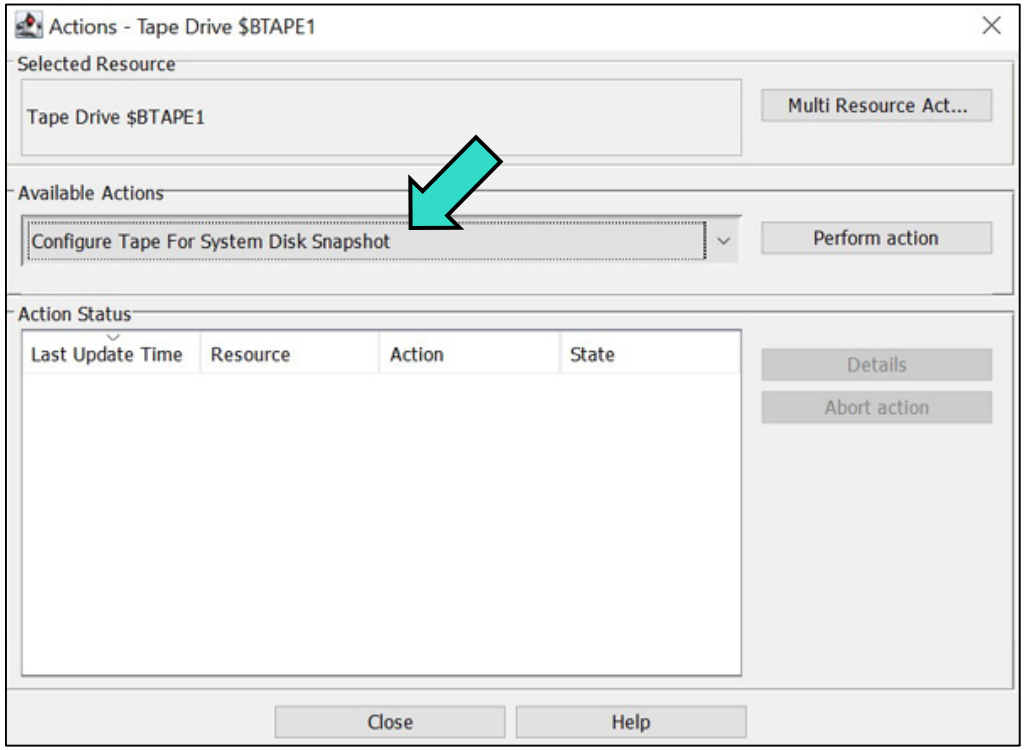
New L23.08 HPE Virtualized NonStop storage features

- No change in T0876 (vNS Deployment Tools for VMware) in L23.08
  - The latest T0876 version is T0876L02^AAR (released in conjunction with the L22.10 RVU)
- New NonStop $SYSTEM disk snapshot and restore feature
  - New feature released with L23.08
  - Relevant SPRs:
    - T0853L03^DCR — CLIM DVD Software
    - T0682L02^BBN — OSM Service Connection Suite
    - T0634H06^ACY — OSM Console Tools
    - T0954V04^AAV — ETI-NET BackBox H4.11 (required only if backup destination is tape)
    - T0964V01^AAJ — VTR R1.05 (required only if backup destination is tape)
  - Minimum required RVU: L21.06 plus the SPRs above
  - Supported HPE NonStop systems:
    - HPE Virtualized NonStop
    - Converged HPE NonStop — NS4 X4, NS8 X4, NS7 X2 & X3, and NS3 X2 & X3
    - Virtualized Converged NonStop — NS2 X2 & X3

New L23.08 NonStop $SYSTEM disk snapshot and restore feature

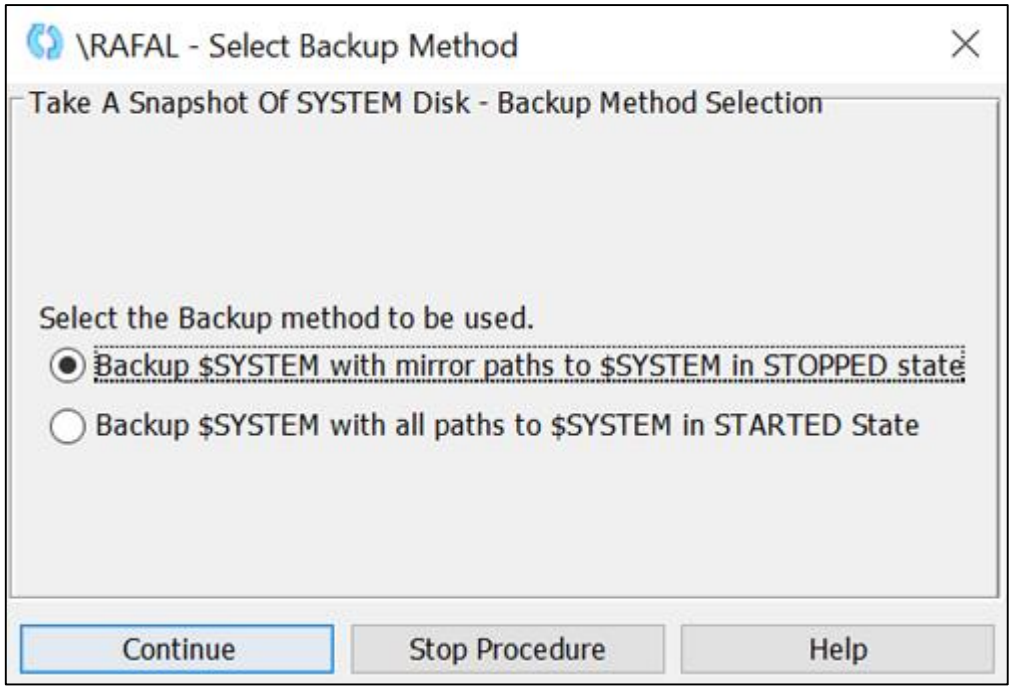
- Either disk or tape devices can be configured as the destination for $SYSTEM disk snapshots
- The destination disks or tapes used must be connected to the same storage CLIM pair as $SYSTEM is connected
- The size of the destination disks or tapes must be equal to or greater than those of the $SYSTEM volume



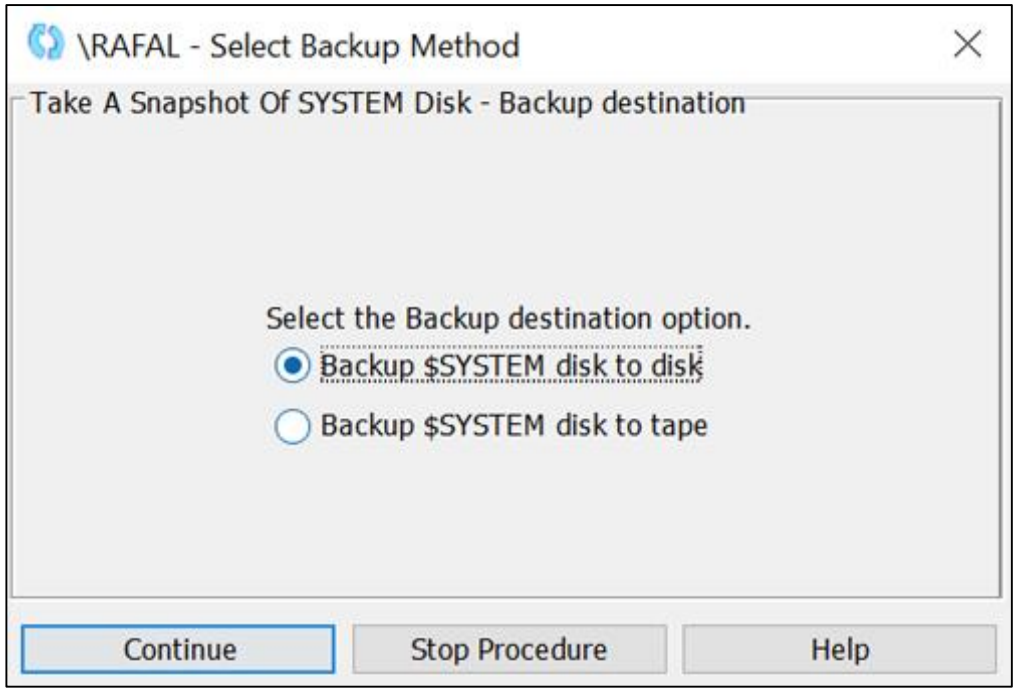


$SYSTEM disk snapshot methods and destinations

- Two types of backup methods:
  - Backup $SYSTEM with mirror paths to $SYSTEM in stopped state
  - Backup the $SYSTEM volume when both the primary and mirror disks are in the STARTED state

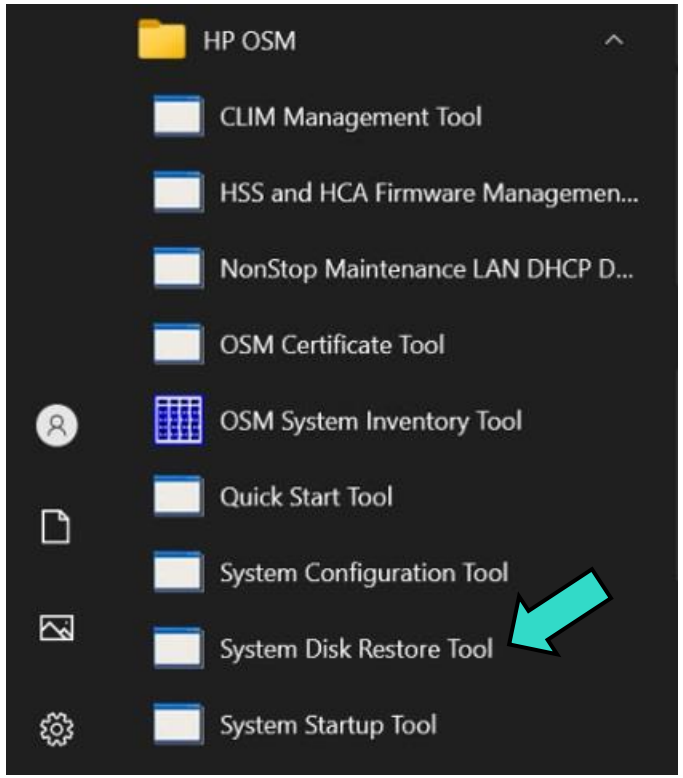


- Both backup methods support these snapshot destinations:
  - $SYSTEM disk to disk
  - $SYSTEM disk to ETI-NET Backbox tape



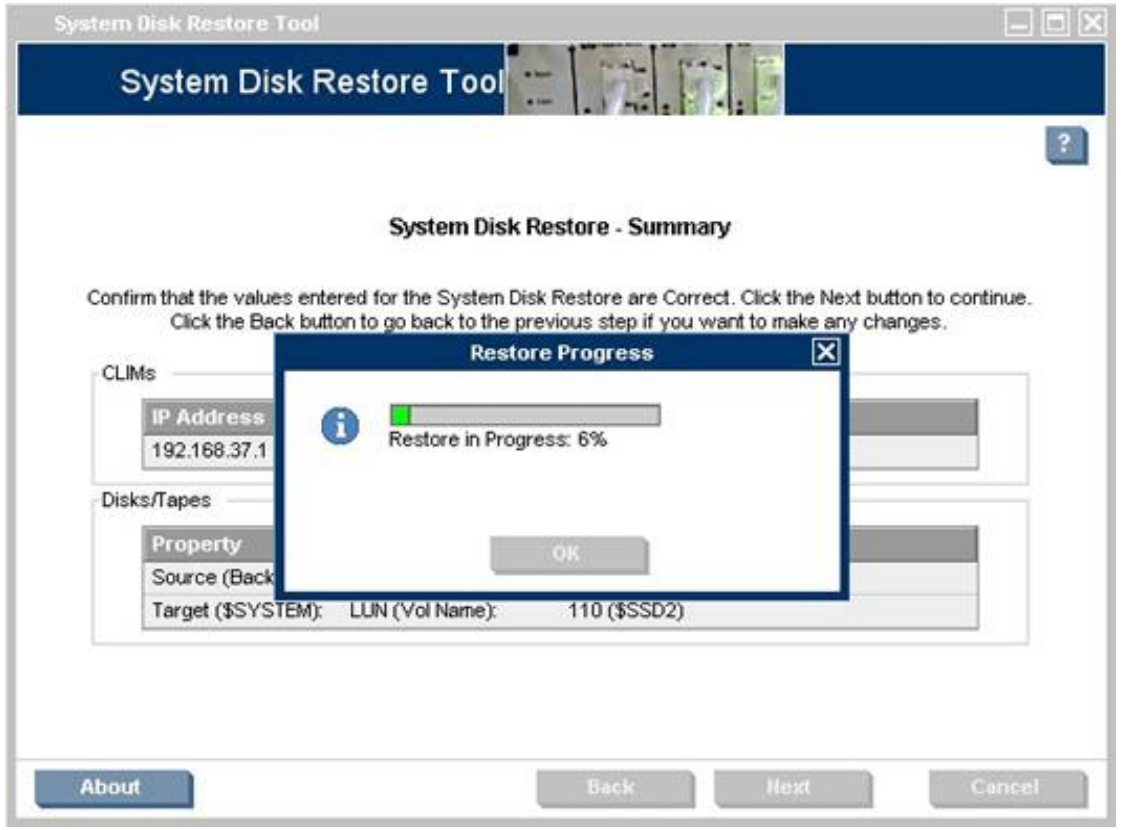
System disk restore tool

- The System Disk Restore Tool is installed as part of OSM Console tools
- The tool runs from the NSC, and needs maintenance LAN connection to SCLIMs
- The NonStop OS must not be running when the System Disk Restore tool runs
- The tool assumes that the destination disk for restore must be equal to or greater than the disk or tape storing the snapshot of $SYSTEM

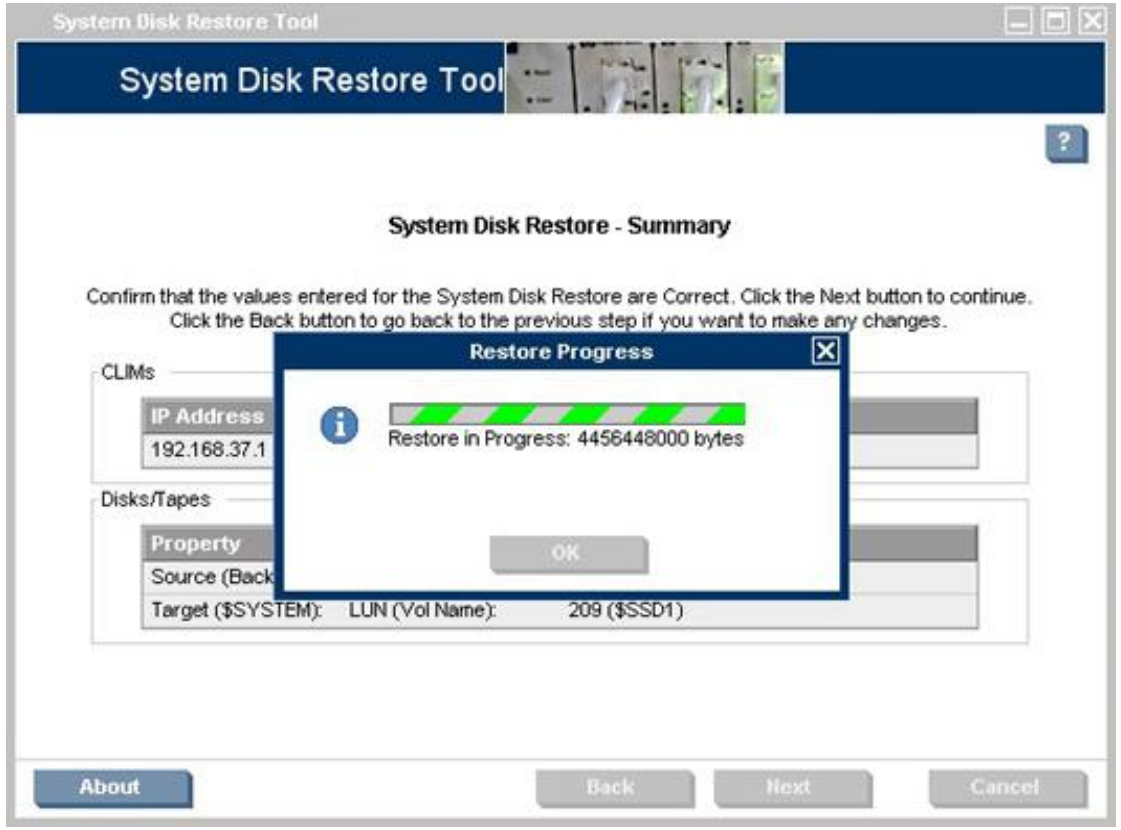


Tracking system disk restore progress

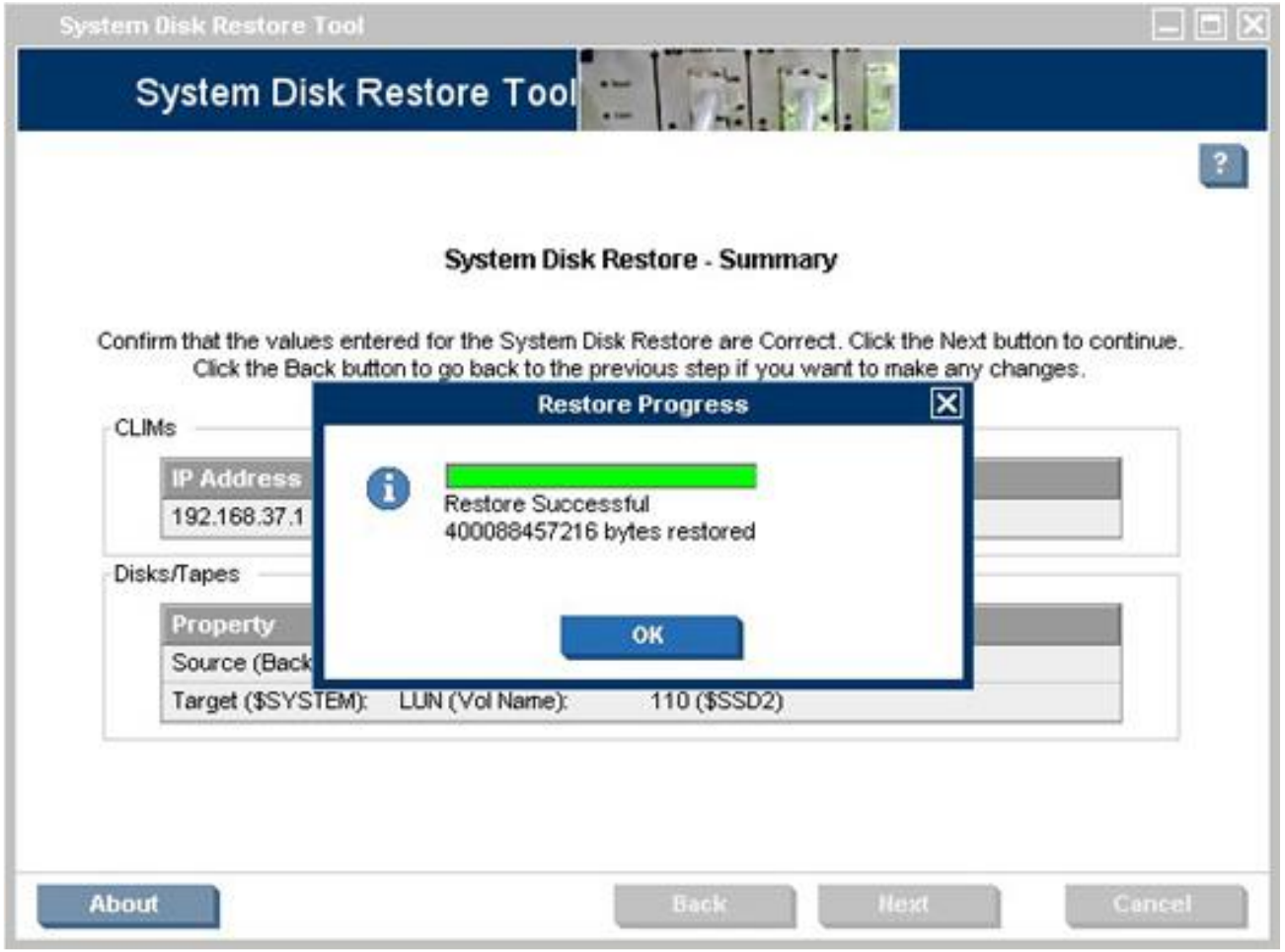
- Disk to disk restore



- Tape to disk restore

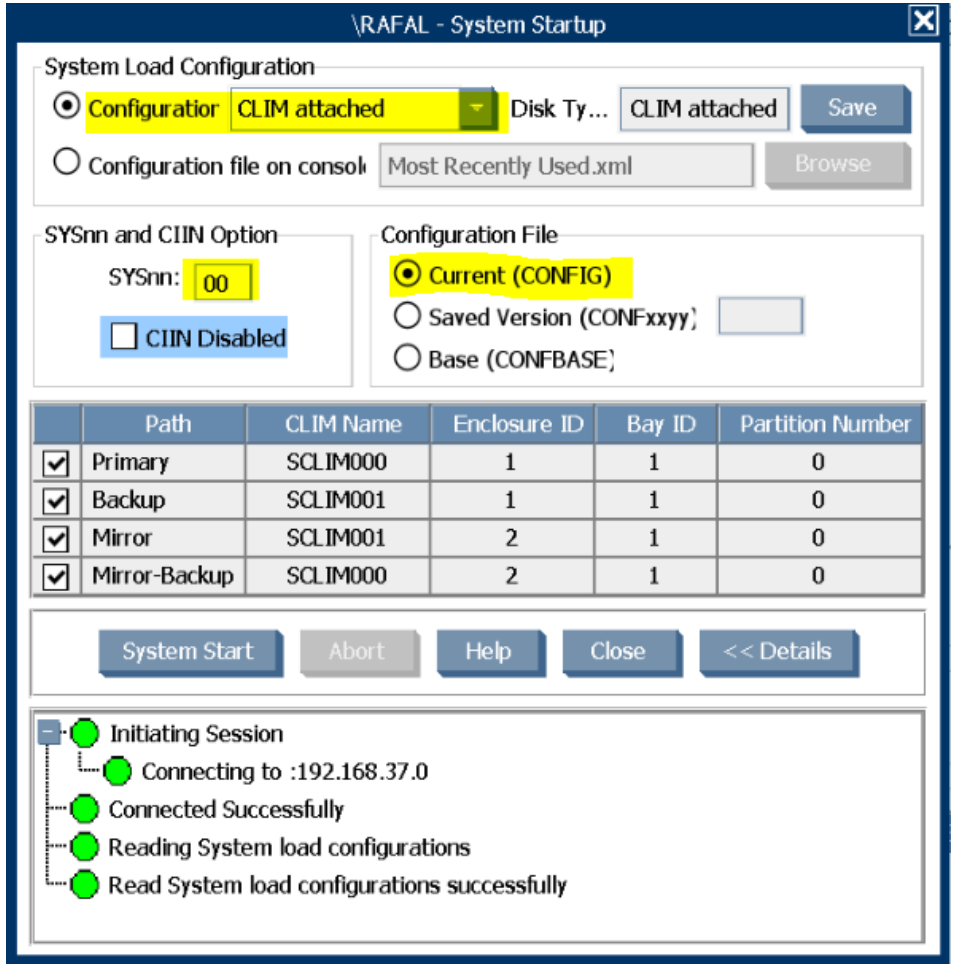


System disk restore success



System load with restored system disk

- In the System Startup Tool, select **Operations** > **Start System**
- Enter the boot variables noted when the system disk backup was taken:
  - **Configuration**: Select *CLIM-attached*
  - **SYSnn**: Enter *SYSnn* that was used while taking backup
  - Under **Configuration File**, select *Current (CONFIG)*
- Verify all parameters and click on **System Start** to start the System load



Related TBC talks and resources

Related or recommended talks on the subject

Talk ID	Title	Presenter	Date & Time
TBC23-TB55	HPE Virtualized NonStop Continues the Journey to the Cloud (repeats on Wednesday)	Marcelo de Azevedo Ken James Bryce Kosinski Spencer Kropp Lars Plum	Tuesday, Sep 12, 2023 11:00 AM - 12:00 PM Wednesday, Sep 13, 2023 10:30 AM - 11:30 AM
TBC23-TB58	Best Practices for Configuring and Managing HPE Virtualized NonStop Systems for Mission Critical Workloads	Marcelo de Azevedo Lars Plum Bryce Kosinski Mark Thompson John Zimsky	Tuesday, Sep 12, 2023 2:15 PM – 3:15 PM
TBC23-TB63	Evolving Your vNS Environment to Keep Your NonStop Business Running	Mark Thompson	Wednesday, Sep 13, 2023 11:30 AM - 12:30 PM
TBC23-TB56	Provide the mission-critical advantage to workloads in your private cloud with HPE Virtualized NonStop	Marcelo de Azevedo Bryce Kosinski	Wednesday, Sep 13, 2023 2:45 PM – 3:45 PM
TBC23-DEV3	HPE Virtualized NonStop developer chat room	Marcelo de Azevedo Lars Plum Bryce Kosinski	Thursday, Sep 14, 2023 9:00 AM – 10:00 AM

Resources

- Product Website: **https://www.hpe.com/us/en/servers/nonstop.html**
- Technical Manuals (available at **www.hpe.com/info/nonstop-ldocs**)
  - *HPE Virtualized NonStop deployment and configuration guide for VMware*
  - *Hardware architecture guide for HPE Virtualized NonStop on VMware*
  - Relevant documents for new NonStop $SYSTEM disk snapshot and restore feature:
    - *OSM Configuration Guide*
    - *OSM Service Connection User Guide*
    - *NonStop System Console Installer and Management Guide*
    - *HPE NonStop OSM SYSTEM Disk Snapshot and Restore User Guide*
    - *BackBox User Guide*
    - *VTR User Manual*

NonStop Partnership– It's a Beautiful Thing!

Thank you for attending this talk
TBC23-TB57 HPE Virtualized NonStop Storage

For more information send email to:

P.Kamath@hpe.com
Marcelo.DeAzevedo@hpe.com
Lars.Plum@hpe.com

Backup slides: Examples of failure handling scenarios

Examples of failure handling scenarios

Internal drives without vSAN (2c-2d with 2 paths)

Sample scenario	Effect	Recovery
ESXi host OS update without a spare host	• Temporary loss of all VMs in host • Temporary loss of access to virtual drives attached to Storage vCLIMs running on host	Revive virtual drives after the ESXi host reboots and VMs are restarted
ESXi host OS update with a spare host	Rolling upgrade not recommended in this configuration. This would be time consuming since it would require migrating Storage vCLIMs and associated virtual drives	
Storage vCLIM software update	Temporary loss of access to virtual drives attached to Storage vCLIM	Revive virtual drives after Storage vCLIM update completes
ESXI host failure or ESXi host hardware upgrade	• Loss of all VMs in host • Loss of access to virtual drives attached to Storage vCLIMs running on host	Revive virtual drives after the ESXi host is replaced and VMs are restarted
Failure of an internal drive in an ESXI host	Loss of access to virtual drives deployed on the failed internal drive	Revive virtual drives after the failed internal drive is replaced

Examples of failure handling scenarios

2c-2d with 4 paths to external storage arrays or vSAN

Sample scenarios	Effect	Recovery
ESXi host OS update without a spare host	• Temporary loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to both primary and mirror volumes, taking double the I/O load	No need to revive virtual drives after the ESXi host reboots and VMs are restarted
ESXi host OS update with a spare host	Rolling upgrade leveraging VMware vMotion supported in this configuration. Similar effect as described above, but with shorter duration.	
Storage vCLIM software update	• No loss of access to virtual drives attached to Storage vCLIM being updated • Surviving storage vCLIM writes to both primary and mirror volumes, taking double the I/O load	No need to revive virtual drives after Storage vCLIM update completes
ESXI host failure or ESXi host hardware upgrade	• Loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to both primary and mirror volumes, taking double the I/O load	No need to revive virtual drives after the ESXi host is replaced and VMs are restarted

Note: there is a difference in recovery between external storage arrays and vSAN in an ESXi host failure in that there will be drive rebuild activity to recover fault tolerance after the host is replaced.

Examples of failure handling scenarios

4c-2d with 4 paths to external storage arrays or vSAN

Sample scenarios	Effect	Recovery
ESXi host OS update without a spare host	• Temporary loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to both primary and mirror volumes, without having to take double the I/O load	No need to revive virtual drives after the ESXi host reboots and VMs are restarted
ESXi host OS update with a spare host	Rolling upgrade leveraging VMware vMotion supported in this configuration. Similar effect as described above, but with shorter duration.	
Storage vCLIM software update	• No loss of access to virtual drives attached to the Storage vCLIM being updated • Surviving storage vCLIMs write to both primary and mirror volumes, without having to take double the I/O load	No need to revive virtual drives after Storage vCLIM update completes
ESXI host failure or ESXi host hardware upgrade	• Loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to both primary and mirror volumes, without having to take double the I/O load	No need to revive virtual drives after the ESXi host is replaced and VMs are restarted

Note: there is a difference in recovery between external storage arrays and vSAN in an ESXi host failure in that there will be drive rebuild activity to recover fault tolerance after the host is replaced.

Examples of failure handling scenarios

Single vSAN cluster (2c-1d with 2 paths)

Sample scenarios	Effect	Recovery
ESXi host OS update without a spare host	• Temporary loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to the drives, taking any I/O load migrated over from downed Storage vCLIMs	No need to revive virtual drives after the ESXi host reboots and VMs are restarted (no DP2 mirroring)
ESXi host OS update with a spare host	Rolling upgrade leveraging VMware vMotion supported in this configuration. Similar effect as described above, but with shorter duration.	
Storage vCLIM software update	• No loss of access to virtual drives attached to Storage vCLIM being updated • Surviving storage vCLIM writes to the drives, taking any I/O load migrated over from the Storage vCLIM being updated	No need to revive virtual drives after Storage vCLIM update completes (no DP2 mirroring)
ESXI host failure or ESXi host hardware upgrade	• Loss of all VMs in host • No loss of access to virtual drives attached to Storage vCLIMs running on host • Surviving storage vCLIMs write to the drives, taking any I/O load migrated over from downed Storage vCLIMs	No need to revive virtual drives after the ESXi host is replaced and VMs are restarted (no DP2 mirroring)

HPE Slides and Materials Usage

This content is protected

This presentation is the property of Hewlett Packard Enterprise and protected by copyright laws of the United States. The material in this presentation is provided to attendees of the NonStop Technical Boot Camp 2023 as part of their registration and attendance at the event. Attendees are free to use this material and share it with others within their own company.

This material may not be quoted, copied, communicated or shared with third parties or mutual customers without permission from HPE. To request permission to share material in this presentation outside of your company, send an email to mark.pollans@hpe.com explaining the usage you are intending and your request will be considered.