

Best Practices for using EMC Storage Arrays with VMware ESX Server



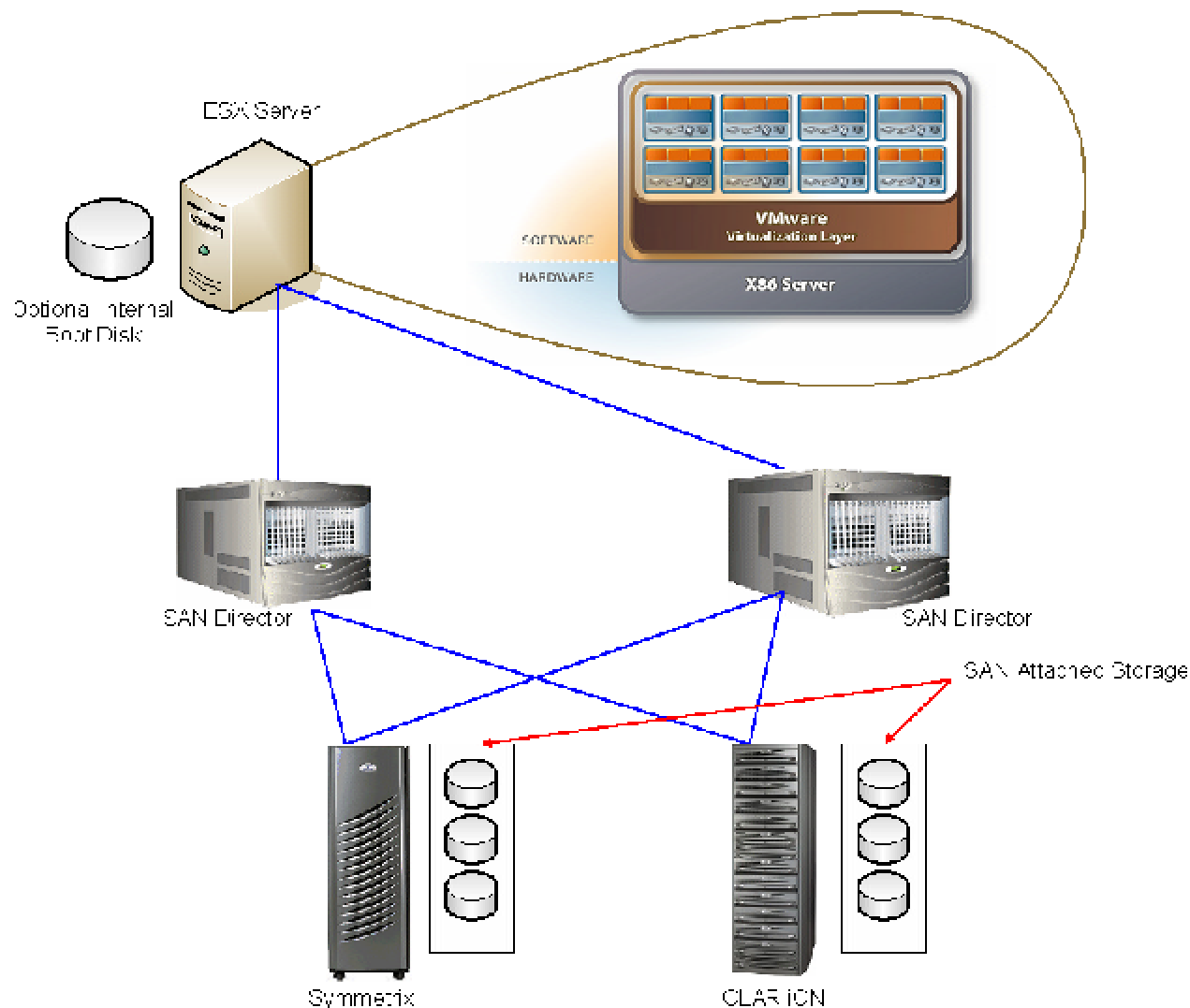
Bala Ganeshan

Eric Horschman

Agenda

- Introduction
- ESX Servers on SAN
- Using ESX Server v2.x on EMC Storage Arrays
 - Using ESX Server v2.x with EMC Symmetrix Arrays
 - Using ESX Server v2.x with EMC CLARiiON Arrays
- Backup and Recovery Considerations for ESX Server v2.x
- VMware ESX Server v3
 - ESX Server 3 New Storage Options
 - ESX Server 3 Consolidated Backup
- Question and Answer Session

SAN Attached ESX Server Configurations



Acronyms and Terms Commonly Used

- ESX Server Farm— A logical grouping of servers that share common pool of storage
- VM— Virtual Machine
- Service Console— Management Interface to ESX server
- MUI— Management User Interface. Web-based tool to manage ESX Server
- VMkernel— The ESX kernel that manages virtual machines
- VMFS— VMware File System. A high-performance clustered file system to store large files
- RDM— Raw Disk Mapping. A construct that maps a virtual disk to a pass-through SCSI device
- MRU— Most recently used. A policy to determine path failover behavior

Using ESX Servers with EMC Storage Arrays

- Recommendations to allow scalability
- Address potential performance limiting characteristics in the future
- Recommendations enable optimal balance between performance, management and functionality
- May not be appropriate for all environments

Storage Layout Considerations – 1

- Do not present SAN storage to ESX server farm as one big SCSI disk
 - Use multiple metavolumes or meta LUNs to present the storage
- Both metavolumes and metaLUNs can be grown non-disruptively and without destroying data
- Use of spanned VMFS with dynamic LUN growth mitigates risk

Storage Layout Considerations – 2

	Storage as single LUN	Storage as multiple LUNs
Management	•Easier management •Storage can be over provisioned •One VMFS to manage	•Slightly harder management •Storage provisioning has to be on demand •One spanned VMFS to manage
Performance	•Can result in poor response time •No manual load balancing	•Multiple queues ensure minimal response times •Static load balancing
Scalability	•Limited # of VMs due to response time issue •Limited # of I/O intensive VMs since one VMFS	•Multiple VMFS allows more VMs per ESX server •Response time of limited concern (can optimize)
Functionality	•All VMs share one LUN •Cannot leverage ALL available storage functionality	•Use VMFS when storage functionality not needed •Judicious use of RDM vs. VMFS

Storage Layout Considerations – 3

- RDMs (raw disk mapping) or VMFS can be used with advanced storage functionality
 - Using VMFS with storage software have restrictions that RDMs do not
- Use physical disk compatibility mode for RDMs
 - Enables vMotion while allowing VMs to recognize storage disk characteristics
- If practical present multiple VMFS per ESX server
 - Ideally 4-5 VMFS per ESX server
 - Separate boot disk and application data
 - Separate log and data volumes
 - Separate classes of applications
- With ESX 2.x, VMFS can span multiple physical disks and/or partitions
 - Can be done using “vmkfstools” or MUI

Storage Layout Considerations – 4

- Span VMFS across SAN volumes if needed
 - Advantages:
 - Allows ESX administrator to manage storage as one entity
 - Allows VMkernel to schedule multiple I/Os to underlying disks
 - Provides flexibility to both ESX server and storage administrator in assigning storage
 - Disadvantages:
 - Limits the number of I/O intensive VMs on a ESX server farm
 - Approximately 32 I/O intensive VMs per VMFS
 - **Potential availability issue. Loss of one member of spanned VMFS will cause disruption and loss of all data on that VMFS**
 - Not a major issue with RAID protected storage accessed using multiple paths
 - Ensure VMFS does not span different storage classes
 - File system is spanned. I/Os not balanced across VMFS “physical extents”

Using ESX Servers with EMC Storage Arrays

Path Management Considerations

Path Management Considerations – 1

- PowerPath is not supported.
 - If applicable, uninstall PowerPath from virtual machine after P2V is run
- ESX Server 2.x does not support active load balancing. Only path failover is available
- Set preferred path on EMC Symmetrix to enable static load balancing across HBAs and FAs
- Set active path on EMC CLARiiON to enable static load balancing across HBAs and SPs
- Statistics will continue to be assigned on the first path discovered by VMkernel
 - However, I/Os will be directed to the active path

Path Management Considerations – 2

```

root@losap152:~
[root@losap152 root]# let i=1; while (( $i <= 3 )); do vmkmultipath -q vmhba1:0:
$i; let i+=1; done
Disk and multipath information follows:

Disk vmhba1:0:1 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:1      on (active, preferred)
    vmhba2:0:1      on
Disk and multipath information follows:

Disk vmhba1:0:2 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:2      on (active, preferred)
    vmhba2:0:2      on
Disk and multipath information follows:

Disk vmhba1:0:3 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:3      on (active, preferred)
    vmhba2:0:3      on
[root@losap152 root]#

```

**All Active (and Preferred)
Path on vmhba1**

Path Management Considerations – 3

```

root@losap152:~
[root@losap152 root]# vmkmultipath -s vmhba1:0:2 -r vmhba2:0:2
VMware::ExtHelpers::System[788]: '/sbin/fdisk' -l /dev/sdo
[root@losap152 root]# r let
let i=1; while (( $i <= 3 )); do vmkmultipath -q vmhba1:0:$i; let i+=1; done
Disk and multipath information follows:

Disk vmhba1:0:1 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:1      on (active, preferred)
    vmhba2:0:1      on
Disk and multipath information follows:

Disk vmhba1:0:2 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:2      on
    vmhba2:0:2      on (active, preferred)
Disk and multipath information follows:

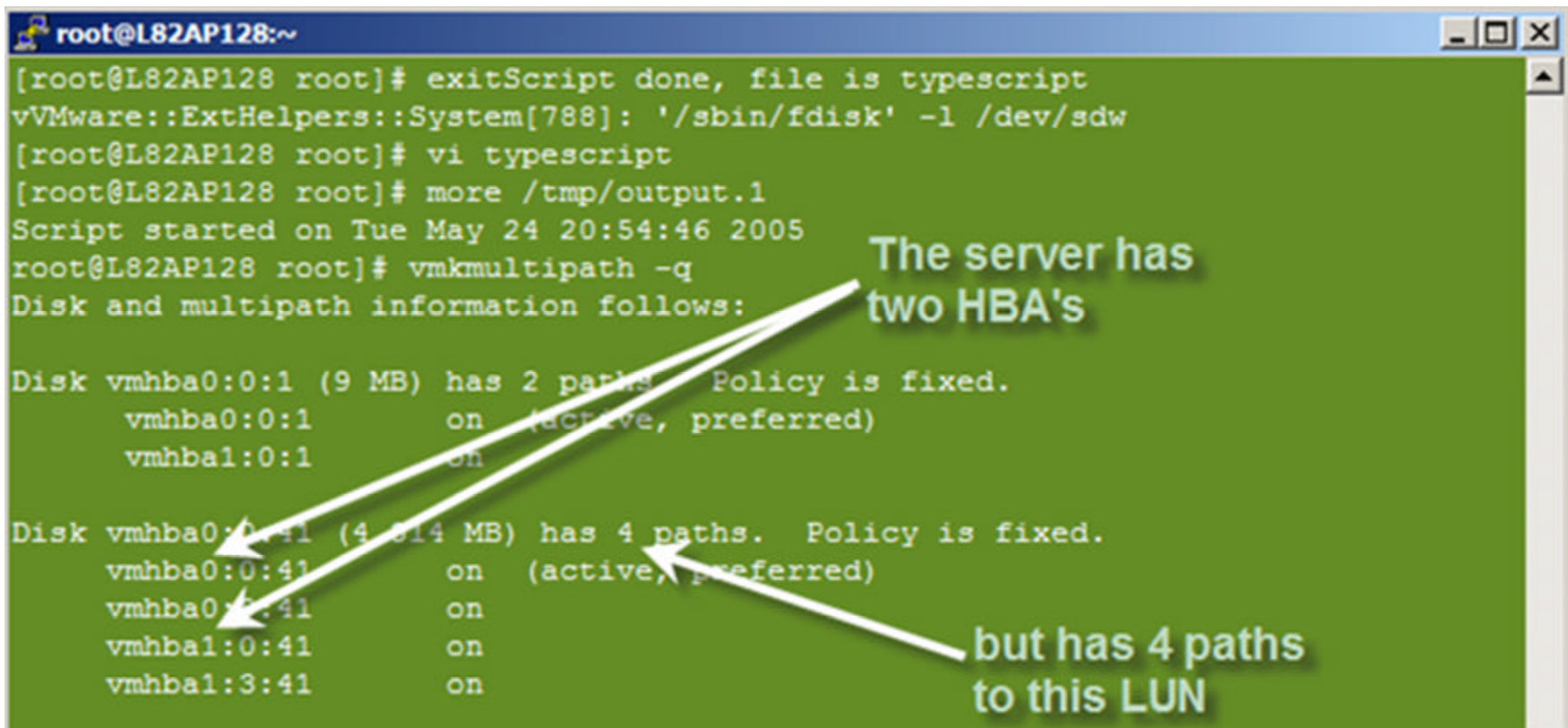
Disk vmhba1:0:3 (8,631 MB) has 2 paths. Policy is fixed.
    vmhba1:0:3      on (active, preferred)
    vmhba2:0:3      on
[root@losap152 root]#

```

Active (and Preferred) Paths distributed between HBA's

Path Management Considerations – 4

- Use multi-target zoning if possible
 - Provides greater resiliency
 - Better opportunity to statically load balance on the back end



The screenshot shows a terminal window with the following text:

```

root@L82AP128:~
[root@L82AP128 root]# exitScript done, file is typescript
vVMware::ExtHelpers::System[788]: '/sbin/fdisk' -l /dev/sdw
[root@L82AP128 root]# vi typescript
[root@L82AP128 root]# more /tmp/output.1
Script started on Tue May 24 20:54:46 2005
root@L82AP128 root]# vmkmultipath -q
Disk and multipath information follows:

Disk vmhba0:0:1 (9 MB) has 2 paths. Policy is fixed.
    vmhba0:0:1      on (active, preferred)
    vmhba1:0:1      on

Disk vmhba0:0:41 (4,314 MB) has 4 paths. Policy is fixed.
    vmhba0:0:41     on (active, preferred)
    vmhba0:1:41     on
    vmhba1:0:41     on
    vmhba1:3:41     on
  
```

Annotations on the screenshot:

- A green text box with the text "The server has two HBA's" has two white arrows pointing to the first two disks (vmhba0:0:1 and vmhba0:0:41).
- Another green text box with the text "but has 4 paths to this LUN" has a white arrow pointing to the vmhba0:0:41 disk entry.

Using ESX Servers with Symmetrix Arrays

Best Practices for Using EMC Symmetrix Arrays – 1

- Use striped metavolumes
 - Follow EMC best practices when creating metavolumes
 - Use the standard split size if it exists
 - Ideally 4, 8 or 16 way metavolumes
 - Metavolumes should not have members on the same physical disks
 - Stripe size should be 960 KB.
- For optimal performance and scalability metavolume should not be much larger than 250 GB
 - Balance between performance versus management
 - Can be grown as appropriate

Best Practices for Using EMC Symmetrix Arrays – 2

- Solutions Enabler 5.4 and later can be used
 - Installation of extraneous software on service console is NOT recommended.
 - Solutions Enabler 6.0 or later cannot be installed on service console
 - Do not install Solutions Enabler in VM unless RDM is used (no value add)
- Use Symmetrix GNS to enable propagation of information
- Use of RDM is recommended if advanced functionality is needed
 - Leverage consistency technology
 - Storage functionality can be coordinated across multiple OS and applications
- Gatekeepers need to be provided and mapped to VM to leverage storage functionality

Using ESX Servers with CLARiiON Arrays

Best Practices for Using EMC CLARiiON Arrays – 1

- Refer to *VMware ESX Server using EMC CLARiiON Storage Systems* solution guide for details
- Use of just one LUN per ESX server is not recommended
- Use striped metaLUNs
 - Follow EMC best practices when creating metaLUNs
 - See [“EMC CLARiiON Best Practices for Fibre Channel Storage”](#)
- For optimal performance and scalability metaLUN should not be much greater than 250 GB
 - Balance between performance versus management
 - Can be grown as appropriate

Best Practices for Using EMC CLARiiON Arrays – 2

- Currently, there is no I/O load balancing with VMware
 - Use native failover with policy set to MRU for CLARiiON
- Use multi-target zoning if possible
 - Provides greater resiliency
 - On the CLARiiON reduces LUN trespass events
 - Better opportunity to manually load balance on the back end

Best Practices for Using EMC CLARiiON Arrays – 3

- Navisphere agent can be installed on the ESX service console. However, it cannot be installed in virtual machines
- Navisphere CLI to manage CLARiiON storage array can be installed on either the service console or virtual machines
- CLARiiON Layered Application Support is available with FLARE R19 and ESX server version 2.5.1 or later

Backup and Recovery Considerations

Backup and Recovery Considerations – 1

- ESX Server and Virtual Machines/Applications
 - Native tools provided by VMware
 - vmsnap, vmsnap_all, vmres etc.
 - Symmetrix Replication Software
 - TimeFinder/Mirror, TimeFinder/Clones, TimeFinder/Snap, SRDF/S, SRDF/A and Open Replicator
 - CLARiiON Replication Software
 - SnapView, MirrorView, SAN Copy
 - CLARiiON Disk Library

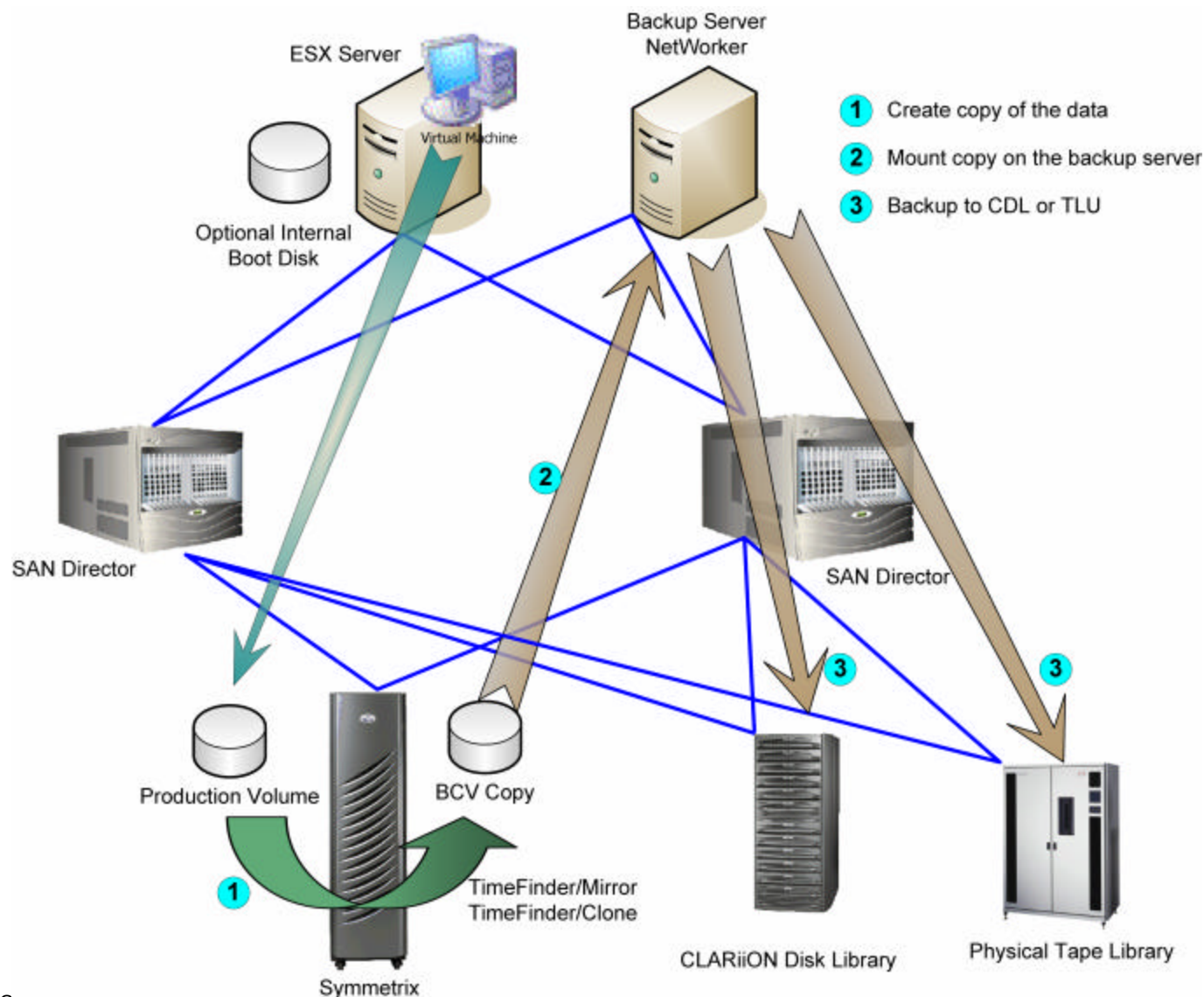
Backup and Recovery Considerations

- Backup Software
 - VMware tests major backup software. Refer to matrix http://www.vmware.com/pdf/esx2_backup_guide.pdf
 - Legato NetWorker has been tested
 - “Storage Node” in Virtual Machines (Linux and Windows)
 - “NetWorker Client” in Virtual Machines
 - “Storage Node” or “NetWorker Client” for Linux on Service Console
 - EMC recommends configuring “Storage Node” or “NetWorker Client” in Virtual Machines
 - Provides granular backup and restores in this configuration
 - Add on modules (Oracle, Exchange, etc.) can be used
- VMware add-on tools (vmware-mount.pl and vmware-loop) maybe needed for restores of individual files
- VMware ESX server 3.0 also offers “Consolidated Backups”

Backup and Recovery Considerations

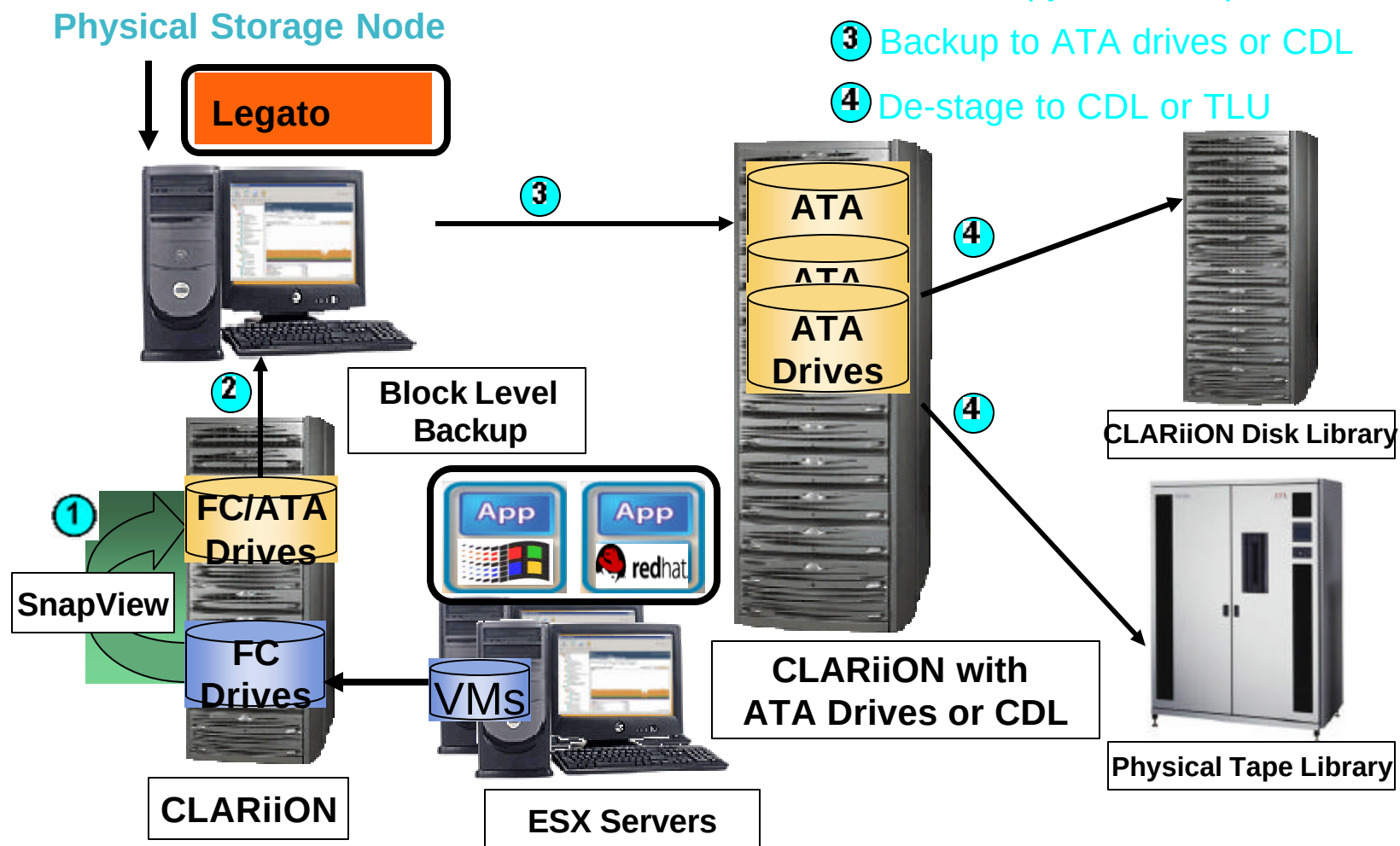
- On the Symmetrix side, extensive testing has been done with TimeFinder and SRDF
 - Refer to solution guide, “VMware ESX Server using EMC Symmetrix Storage Systems” for further details
- All CLARiiON layered applications listed earlier are fully supported
 - Refer to the solution guide *VMware ESX Server using EMC CLARiiON Storage Systems* for further details

Backup and Recovery Considerations



Backup and Recovery Considerations

- ① Create copy of data
- ② Mount copy on backup server
- ③ Backup to ATA drives or CDL
- ④ De-stage to CDL or TLU



ESX Server 3

New Storage and Backup Options

What's New in ESX Server 3.0 Storage?

- Support for iSCSI and NAS (NFS) storage targets
- VMFS3
- SNIA-compliant CIM provider
- Consolidated backup

ESX Server 3 New Features: NAS and iSCSI Details

Lower Cost Access to the Benefits of Virtual Infrastructure

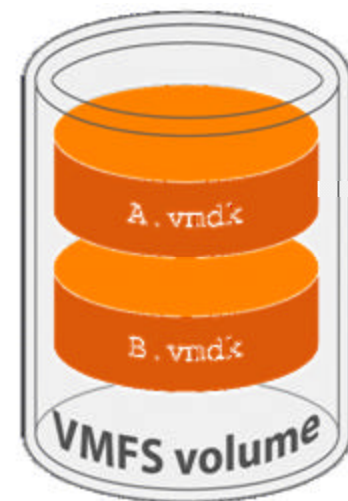
- NAS
 - NFS client built into ESX Server
 - NAS can be used in place of VMFS as the file system for virtual disks
- iSCSI
 - Use hardware iSCSI card or built-in software iSCSI initiator
 - ESX Server boot from iSCSI (for hardware iSCSI only)
- Both
 - VMotion, distributed resource scheduling, and distributed availability services all extend to non-Fibre Channel shared storage
 - Storage option transparent to guests

ESX Server 3 NAS and iSCSI Hardware Support

- Storage arrays
 - Refer to *ESX Server 3 SAN Compatibility Guide* at http://www.vmware.com/products/beta/esx-vc/esx3_san_guide.pdf
 - CLARiiON, Celerra and NetApp arrays
- Software NFS servers (Linux/Windows) not yet supported

VMFS3

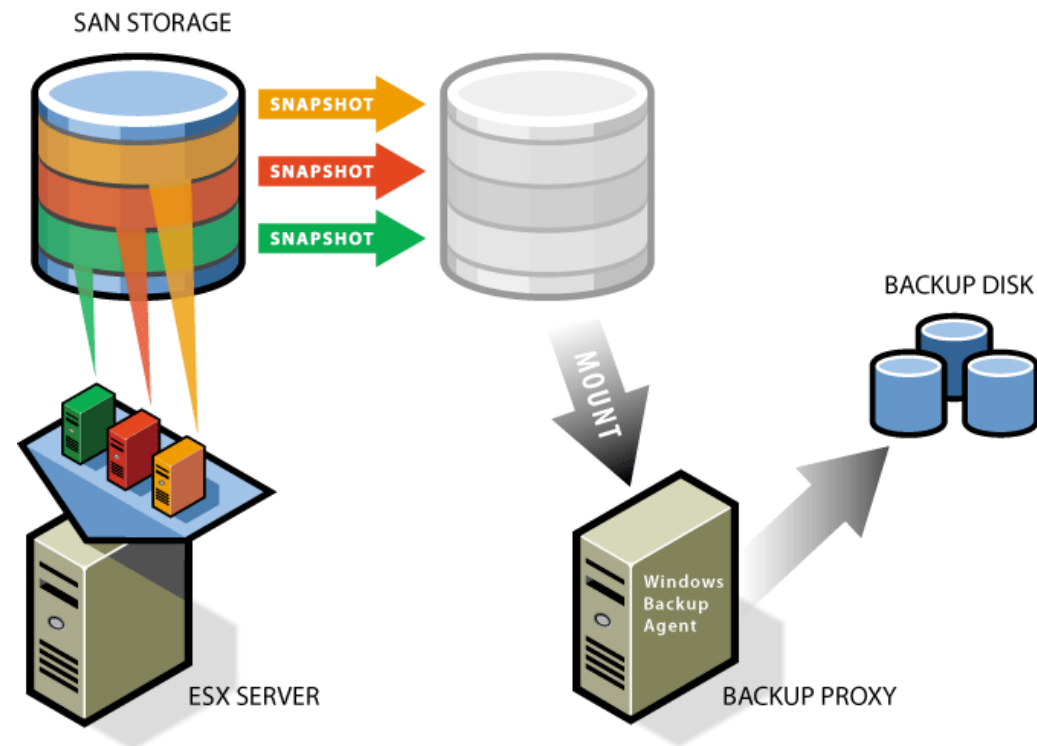
- Revamped disk locking
 - Better scaling for access by large numbers of hosts simultaneously
 - Enables large clusters for distributed resource scheduling and distributed availability services
- Greater reliability and flexibility
 - Distributed journaling for faster recovery
 - Logical volume manager
 - Resize LUNs, add LUNs on the fly
 - Volume availability not compromised due to spanning
- Store more than virtual disks
 - Exclusive repository for virtual machines and VM state
 - Virtual disks, configuration files, snapshots
 - Directories to organize files
 - Optimized for large and small files
 - Optimized for a large number of files
- Virtual disk performance remains close to native



ESX Server 3 Consolidated Backup

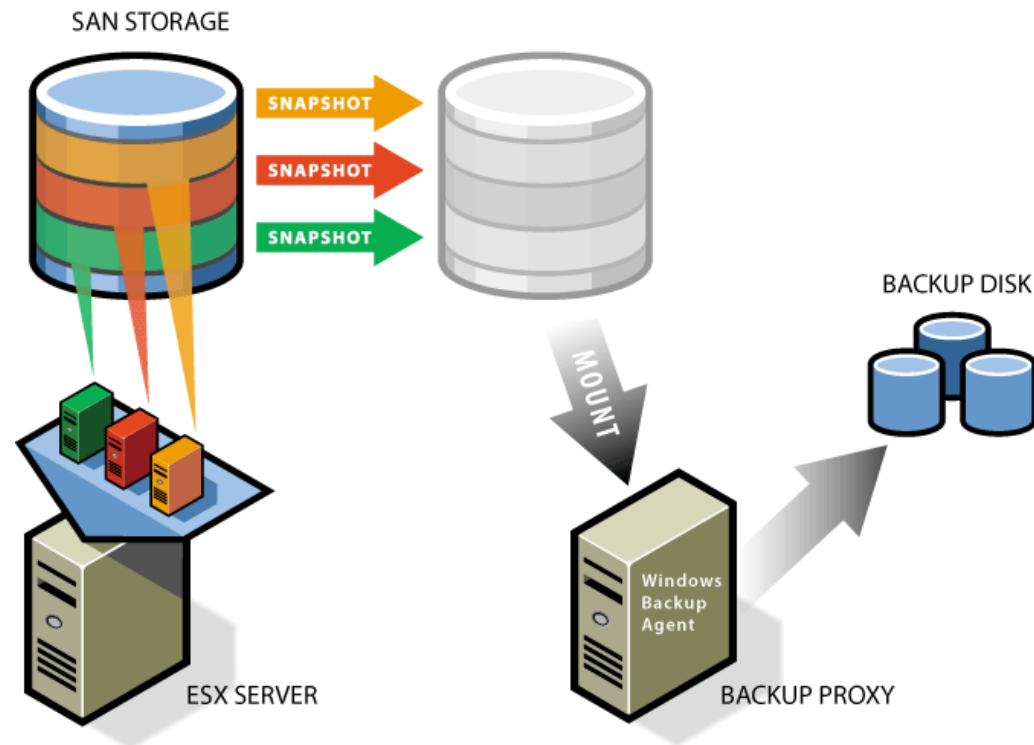
VMware Consolidated Backup

- Centralized File-Level Backups with No Agents
- Move backup out of the virtual machine and into the infrastructure
- Manage centralized backup service instead of agents
- File system-consistent live snapshots
- Remove backup load from ESX Server hosts and virtual machines
- Pre-integrated with major third-party backup products



How Does it Work? – Overview

- Backup SW schedules backup job
 - Job represents single or group of virtual machines
- Consolidated Backup's pre-backup script runs
 - Create file system quiesced virtual machine snapshot(s)
 - Make data available on proxy
- Backup SW performs backup
- Consolidated Backup's post-backup script runs
 - Remove mount from proxy
 - Commit backup snapshot(s)



How Does it Work? – Quiesceing

1. Consolidated Backup triggers virtual machine snapshot through VirtualCenter
2. VMware Tools freeze I/O in virtual machine:
 - In-virtual machine pre-freeze scripting hook
 - Filesystem quiesceing (Windows only)
3. Disk Snapshots are created
4. VMware Tools thaw I/O in virtual machine:
 - In-virtual machine post-thaw scripting hook

Related Presentations...

- New Storage Options with VMware ESX Server 3
- VMware ESX with CLARiiON – Best Practices
- Using VMware for Business Continuity and DR
- VMware Remote Disaster Restart using CLARiiON Replication Software
- Business Continuity in a CLARiiON and VMware Environment

EMC²
where information lives[®]