IA B29: Best Practices for Backup and Recovery of VMware

Jim Olson
Technical Director

George Winter   Twitter: @vmxgeorge
Technical Product Manager, NetBackup

Disclaimer

This information is about pre-release software. Any unreleased update to the product or other planned modification is subject to ongoing evaluation by Symantec and therefore subject to change.

This information is provided without warranty of any kind, express or implied. Customers who purchase Symantec products should make their purchase decision based upon features that are currently available.

Symantec.

NetBackup 7.6 Beta Sign Up

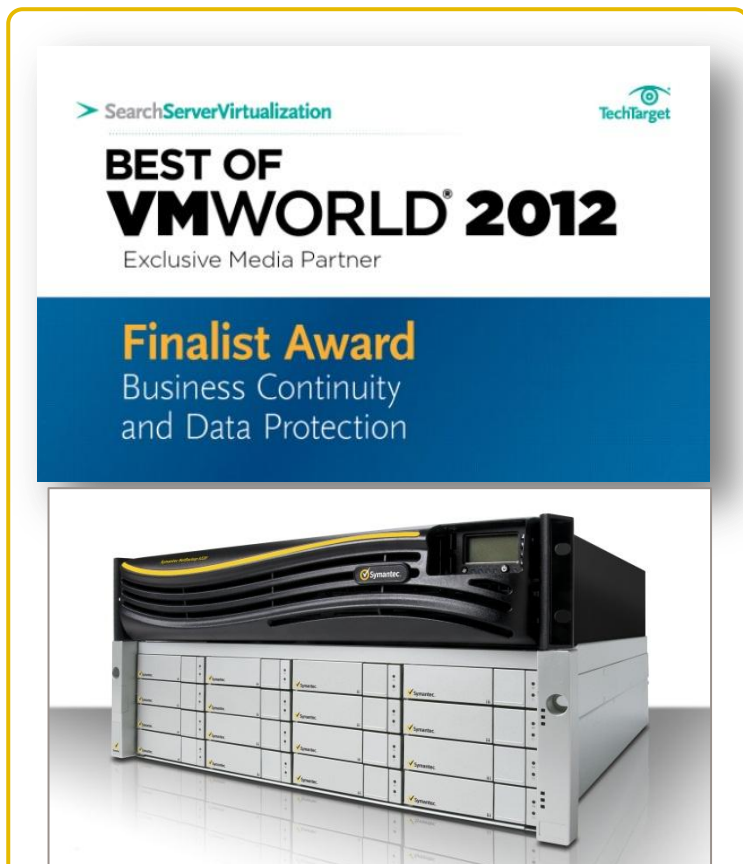**Norton™ Snap**
Protecting you from dangerous QR codes
www.norton.mobi

http://tinyurl.com/NB76beta

NetBackup Has Won More Best of VMworld Awards Than Any Other Vendor

Business Continuity and Data Protection Category

Best of VMWORLD 2012
Exclusive Media Partner

Finalist Award
Business Continuity and Data Protection

Symantec NetBackup 5220
One Appliance, end-to-end protection for enterprise

Previous NetBackup Best of VMworld Awards

GOLD

SearchServerVirtualization.com
BEST OF vmworld 2010

FINALIST

SearchServerVirtualization.com
BEST OF vmworld 2008

GOLD

SearchServerVirtualization.com
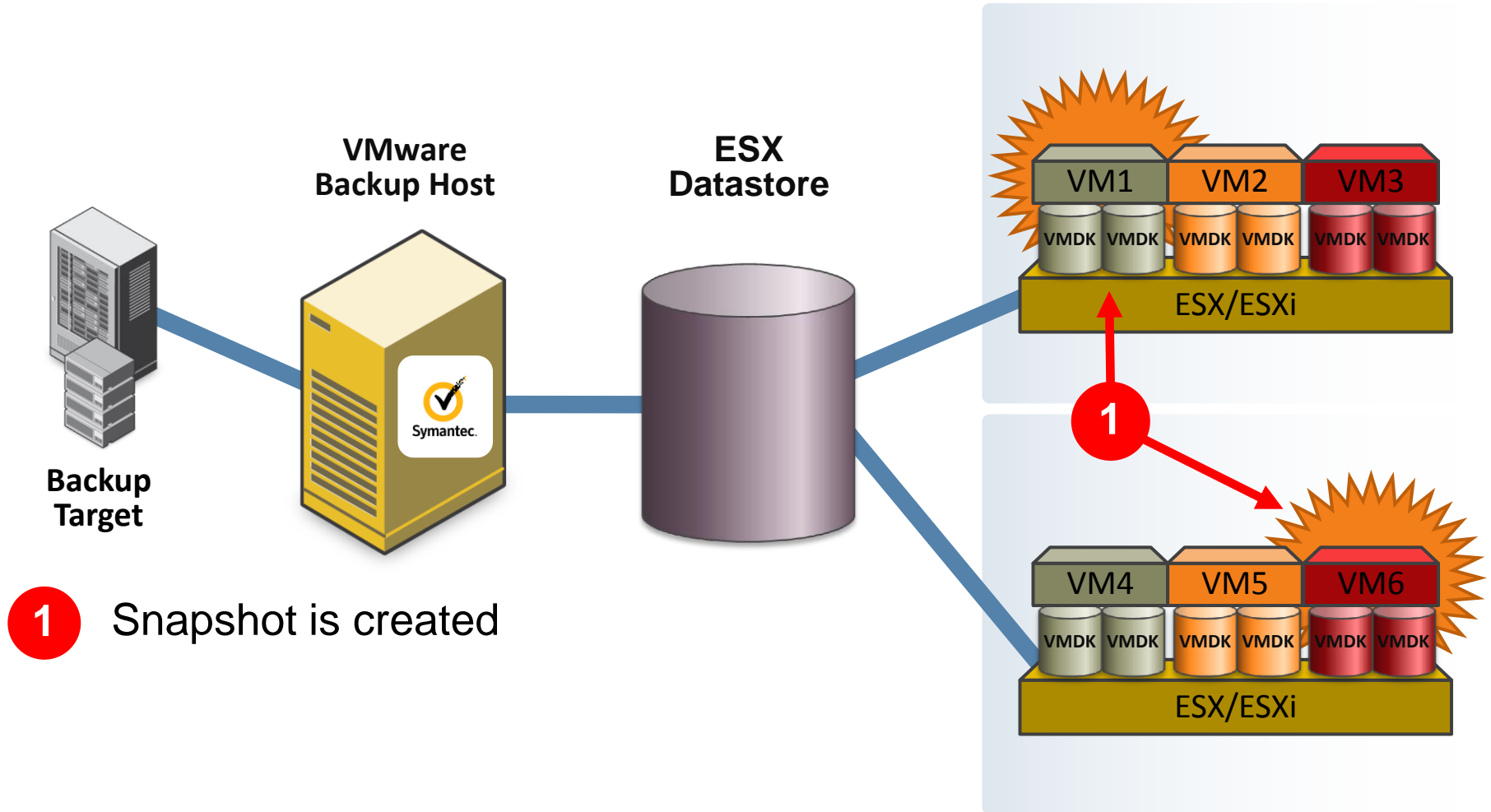BEST OF VMWORLD 2007

Agenda

- VADP – Backup Process

- Backups And VMDK Files

- NetBackup Application & VADP

- Virtual Machine Restore Considerations

- VMware Intelligent Policy

- VMware Intelligent Policy Use Cases
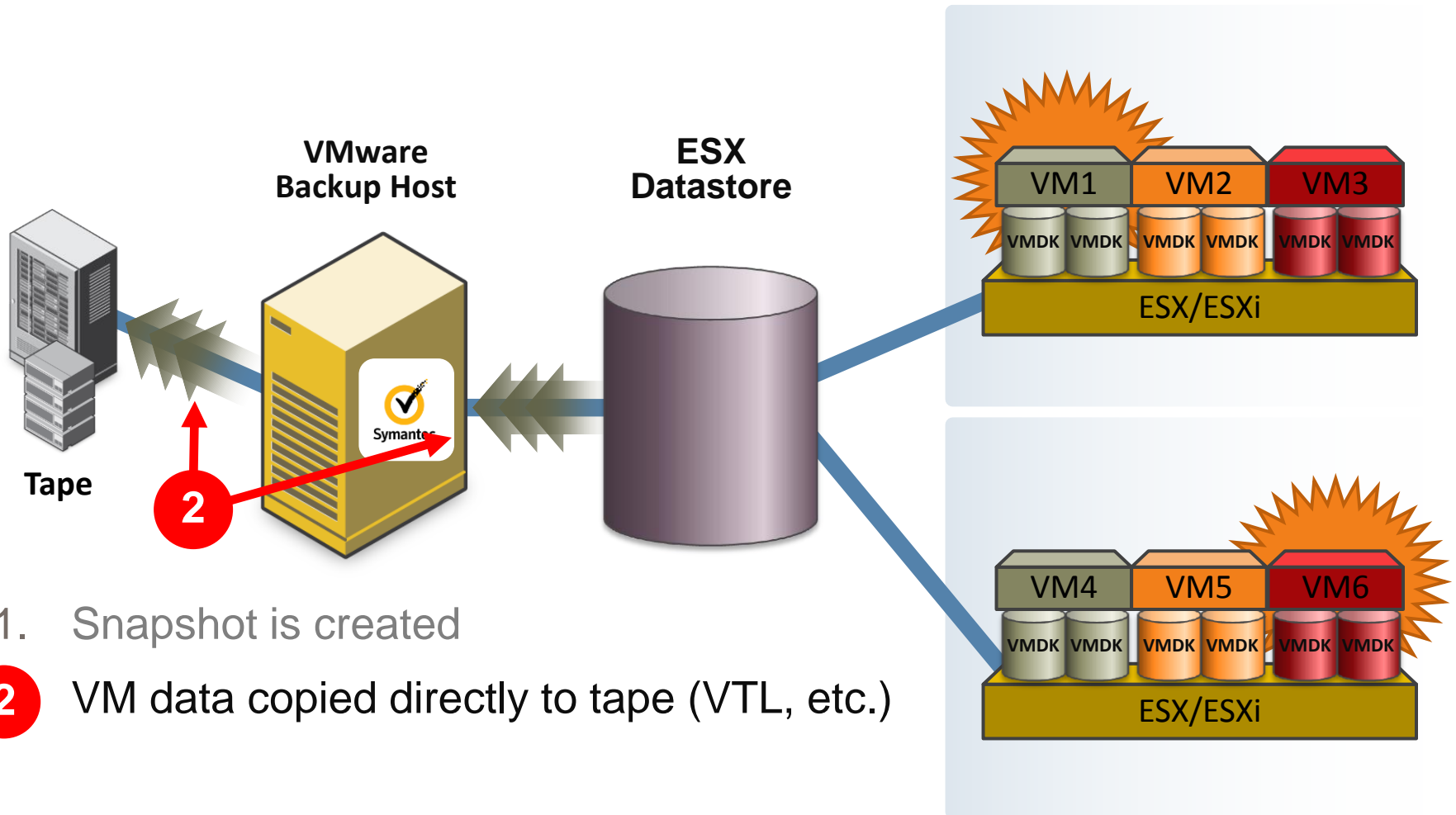
- VIP Configuration Exercise

- Q&A

vStorage API for Data Protection (VADP)

Symantec.

vStorage API Backup Process

1 Snapshot is created

NetBackup 7 for VMware Backup Process

1. Snapshot is created

2 VM data copied directly to tape (VTL, etc.)

Symantec.

NetBackup 7 for VMware Backup Process

1. Snapshot is created
2. VM data copied directly to tape (VTL, etc.)
3 VM Snapshot released

✓Symantec.
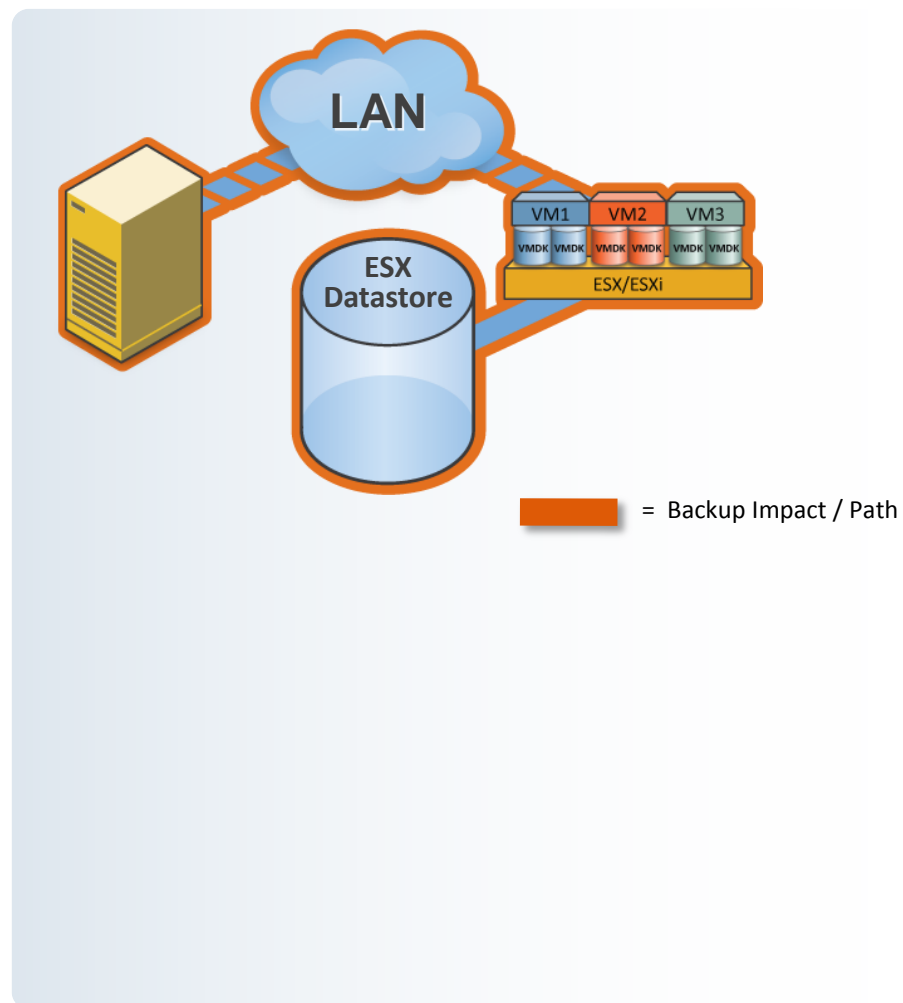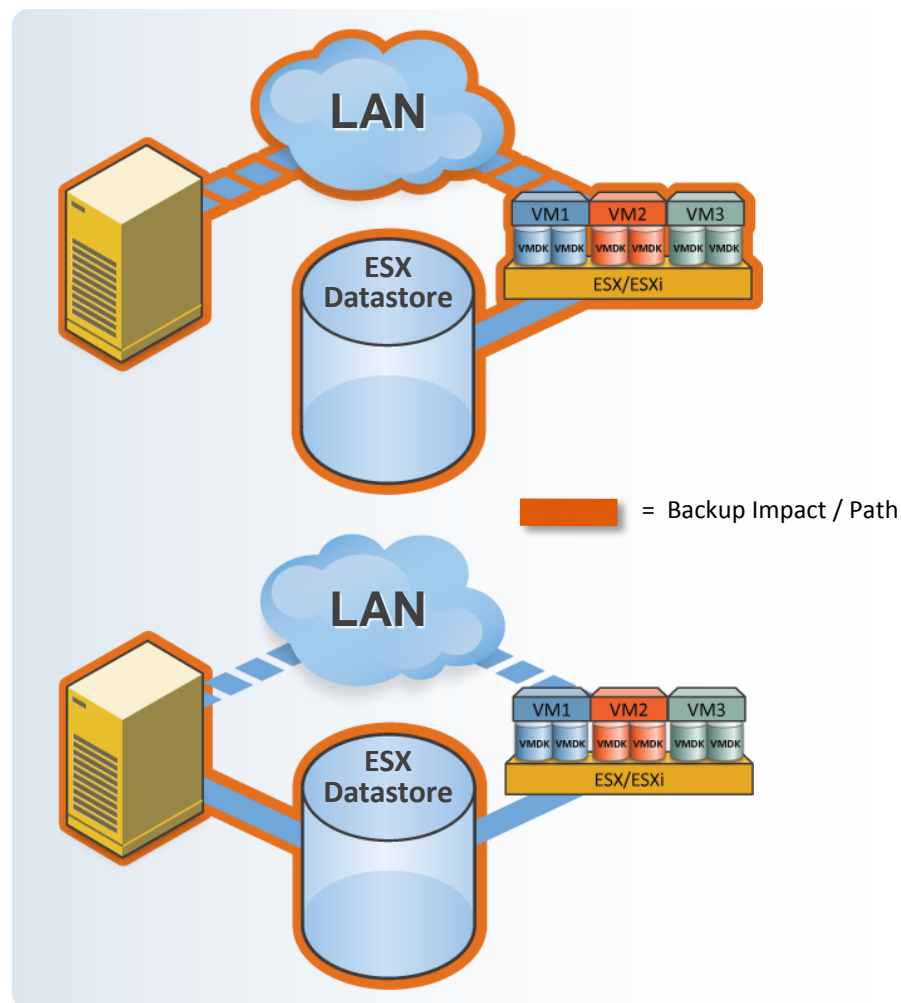
NetBackup 7 for VMware Backup Data Path

- Network Backups (NBD)

    - NFS or DAS fully supported

    - No loss in backup or restore functionality

    - Direct communication with ESX server required (DNS, etc.)

    - ESX server directly impacted

LAN

VM1 VM2 VM3
VMDK VMDK | VMDK VMDK | VMDK VMDK
ESX/ESXi

ESX Datastore

= Backup Impact / Path

NetBackup 7 for VMware Backup Data Path

- Network Backups (NBD)
  - NFS or DAS fully supported
  - No loss in backup or restore functionality
  - Direct communication with ESX server required (DNS, etc.)
  - ESX server directly impacted

- Shared storage configuration
  - Fibre or iSCSI
  - Near zero impact on ESX
  - No loss in backup or restore functionality
 - No communication with ESX required
- Hotadd

LAN

ESX Datastore

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX/ESXi

= Backup Impact / Path

LAN

ESX Datastore

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX/ESXi

Symantec.

Backups And VMDK Files

Symantec.

VMDK Creation Process

- VMware rules for creation of new VMDK

  1. Empty VMDK is created first

  2. Space within the VMDK must be zeroed (mandatory)

  3. Data is now written to zeroed space within VMDK

- Redo log creation follows these same rules

  - Expanded in 16 MB extension – each extension must be "zeroed"

  - Three steps: create, zero, write

  - NOTE: similar process for thin provisioned VMDK creation

- Redo is a temporary VMDK

  - At end of snapshot redo must be reapplied to original VMDK

✔Symantec.
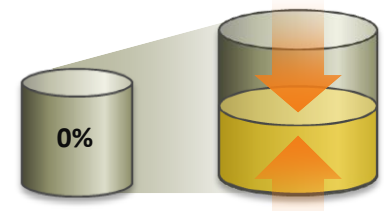
# A Closer Look - What Happens During a Backup?

A Detailed Look at the VADP **Snapshot** Process

- During VM backups, VMware creates temp snapshot of VM. Process is as follows:

  - VSS provider flushes OS buffers in VM

  - Snapshot of VM is taken (vmdk(s) are frozen) (SCSI reservation of LUN)

  - Redo log created – all writes redirected to redo log (Redo expanded in 16 MB chunks)

  - VM is backed up

  - Redo log data applied to original vmdk(s)

  - Snapshot released - backup completed

- Why does this matter?

  - Every one of these steps involves significant amount of I/O

  - Reducing **number of snapshots per Datastore** improves backup perf & reliability

  - Incrementals are quick and can reduce snapshot impact

Symantec
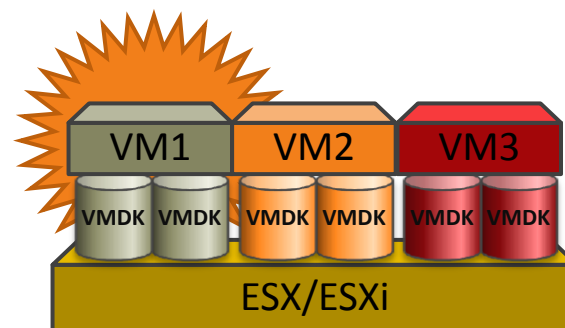
Reducing Backup I/O Impact – Improve B/U Reliability

- Shorten time backup snapshot is open
  - Backup during periods of low VM activity
  - Use incremental backups (CBT) liberally
  - Limit simultaneous backups per ESX / Datastore (VIP)

- Configure NetBackup for optimal performance
  - Design backup policies to evenly **balance load across ESX / Datastores**
  - Faster backups = shorter time snapshots are open
  - Tune backup application buffers for optimal performance

- Result:
  - Snapshots more reliable
  - Overall backup processing faster
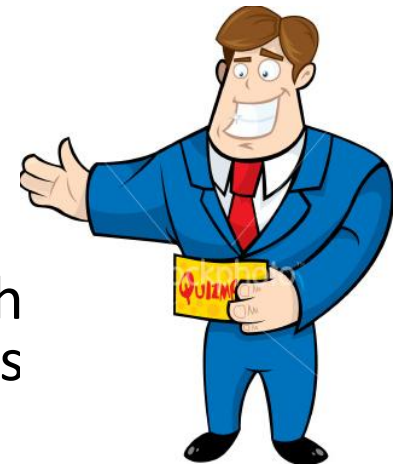 - Higher level of backup success
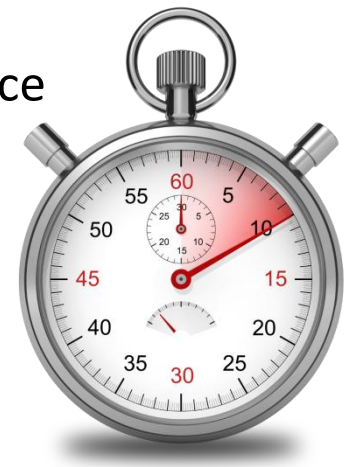
Symantec.

POP QUIZ

Q: How many additional IOPS are required for each write I/O that occurs while the VADP snapshot is active?

Hint: Redo logs are created like thin provisioned VMDKs

Symantec.

POP QUIZ

Q: How many additional IOPS are required for each write I/O that occurs while the VADP snapshot is active?

A: Actual number depends on what is happening during backup but it can be a lot:

Snapshot creation:

- Create REDO log

- Expand 16MB chunks

- Zero-out 16 MB chunks

  (repeat as necessary)

Snapshot Deletion:

- Read data from REDO

- Apply data to original VMDK

- Remove REDO

  (Thin VMDK can increase this)

Symantec.

POP QUIZ

Notes:

- None of the IOPS listed below occur during standard VM I/O activity

- The operations listed occur for **every** actively backed up VM

Snapshot creation:

- Create REDO log

- Expand 16MB chunks

- Zero-out 16 MB chunks

  (repeat as necessary)

Snapshot Deletion:

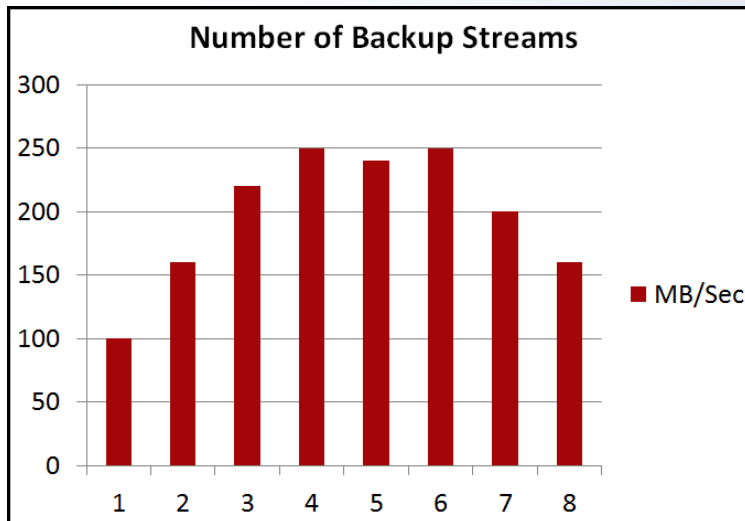- Read data from REDO

- Apply data to original VMDK

- Remove REDO

NetBackup Application &
VADP Performance Configuration

Ensure Optimal Backup Performance

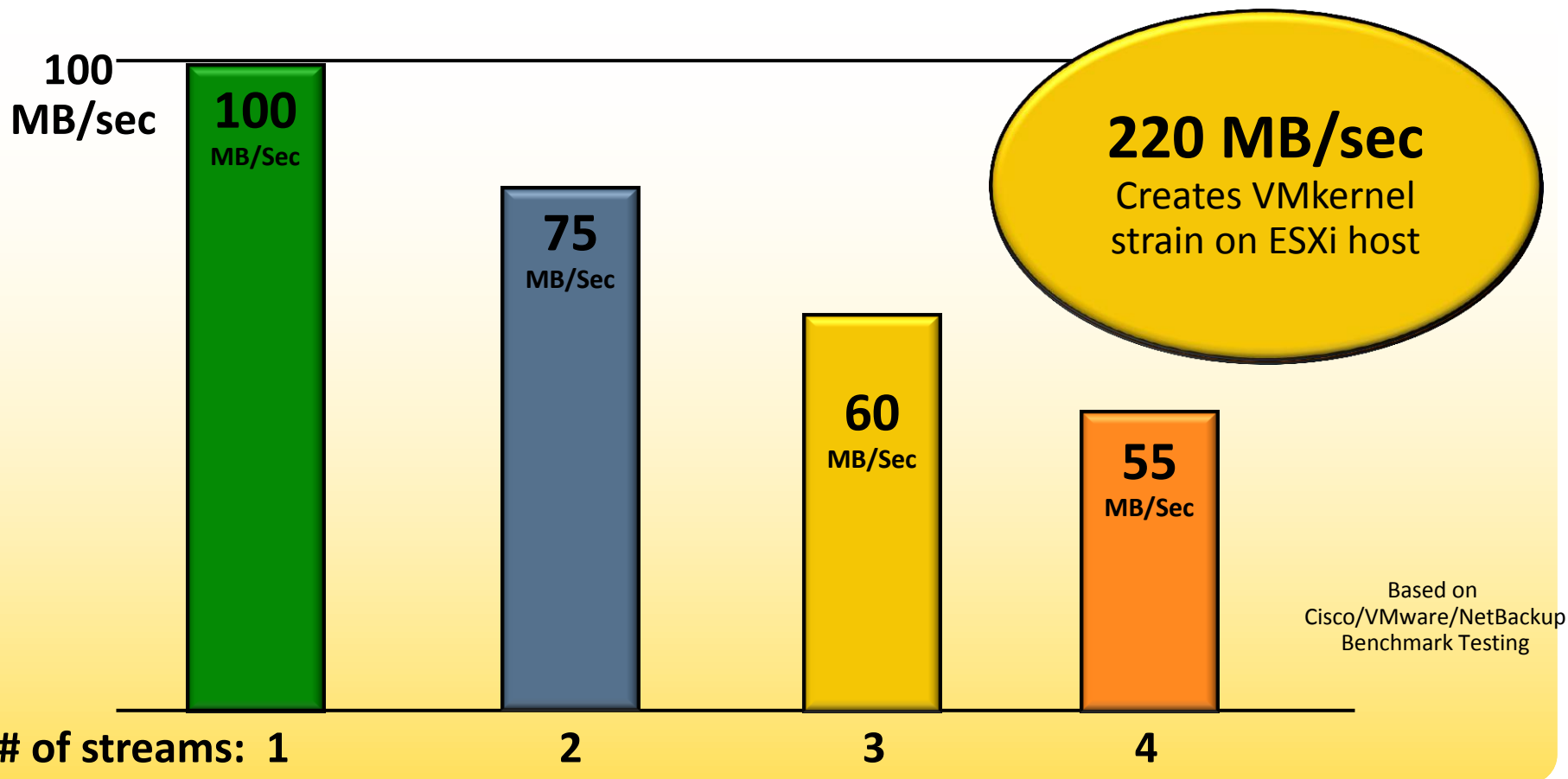- Number of simultaneous backups?

Number of Backup Streams

Aggregate Backup Speed

- Single VADP backup stream won't saturate backup path

- **VMware throttles VADP throughput to reserve VMkernel port bandwidth**

- Max backup performance achieved by creating simultaneous backup streams

- Design backups so that data is streamed across multiple VMware components (ESXi/DS)
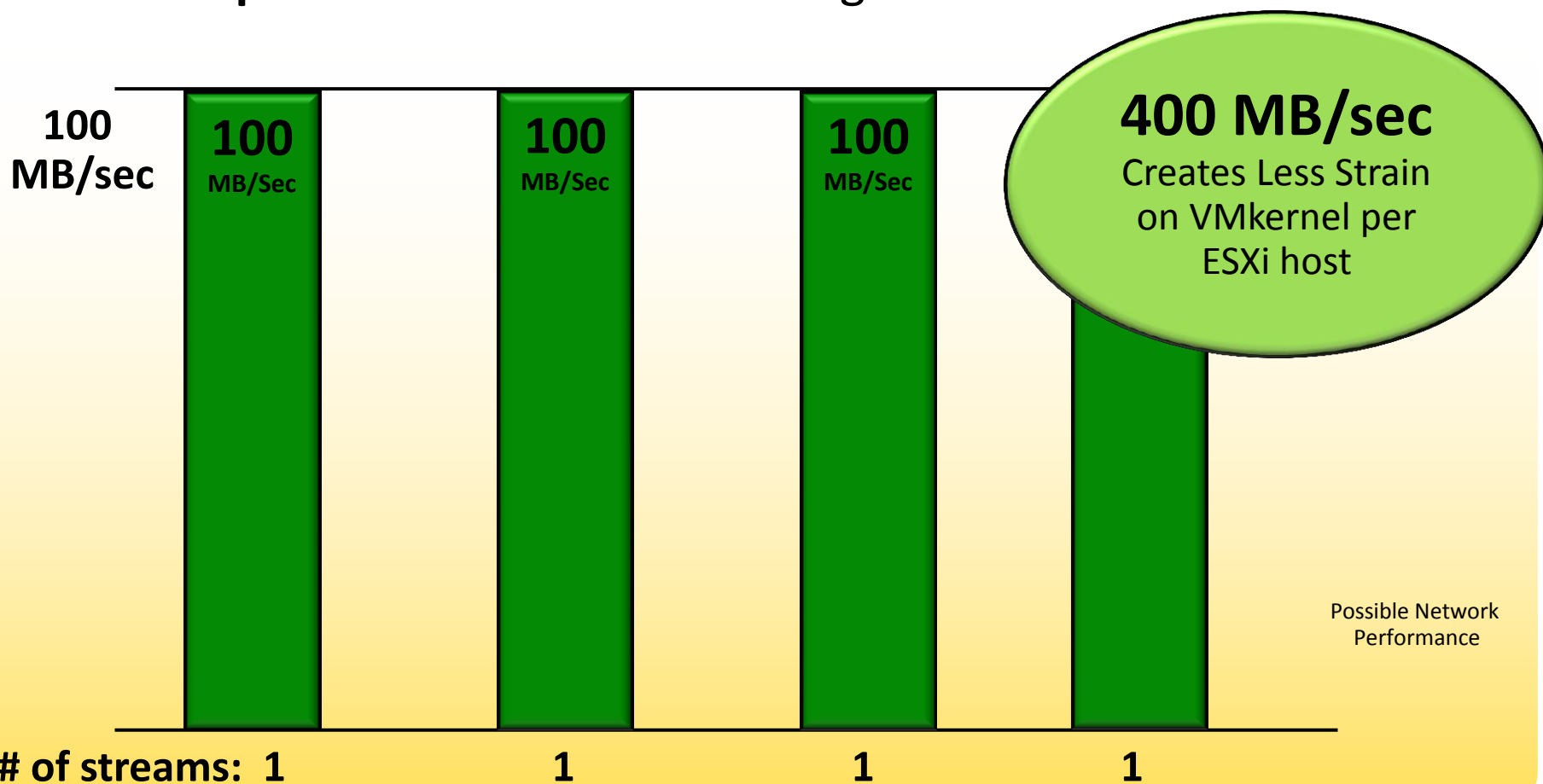
- Improve backup speeds, shorten backup window

✔Symantec™

VADP Performance Characteristics

- **Single ESXi host** – NBD Performance **Per Stream**

# of streams:	1	2	3	4
	100 MB/Sec	75 MB/Sec	60 MB/Sec	55 MB/Sec

100 MB/sec

220 MB/sec
Creates VMkernel strain on ESXi host

Based on Cisco/VMware/NetBackup Benchmark Testing

✓ Symantec.

VADP Performance Characteristics

- **Four Separate ESXi hosts** – Creating One Stream Per ESXi

100 MB/sec

100 MB/Sec	**100** MB/Sec	**100** MB/Sec

400 MB/sec
Creates Less Strain on VMkernel per ESXi host

Possible Network Performance

# of streams: 1   1   1   1

✔Symantec.

Virtual Machine Restore Considerations

Symantec.
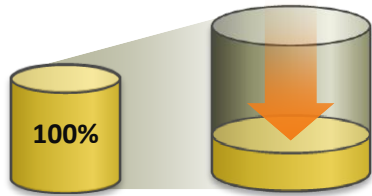
General Restore Considerations

- Restore process involves more I/O than backup process
  - Disks (vmdk's) must be first created as target of restore
  - Type of vmdk can impact restore speed and I/O required
- Single restore typically won't saturate restore path
  - DR – create simultaneous restore jobs
  - As with backup – balance restores across ESX or Datastore
- Slow vCenter can also cause restore perf issues
  - Optional: bypass vCenter by restoring directly to ESX(i) server
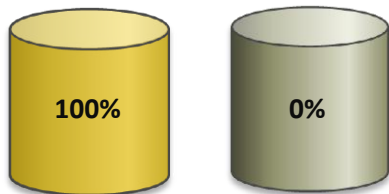 - Known to significantly improve restore perf in some cases

✔Symantec.

VMDK Type Can Impact Restore Speds
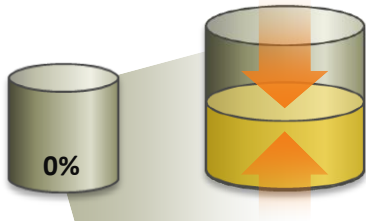
- thin
  - Space not allocated during creation
  - Space is supplied then zeroed out – on demand
  - Creation slow if vmdk turns out to be full
- zeroedthick
  - 100% of space allocated during creation
  - Zeroed out on demand
  - Can be faster than "thin" – especially if vmdk is nearly full
- eagerzeroedthick
  - 100% of space allocated during creation
  - 100% of disk zeroed out during creation
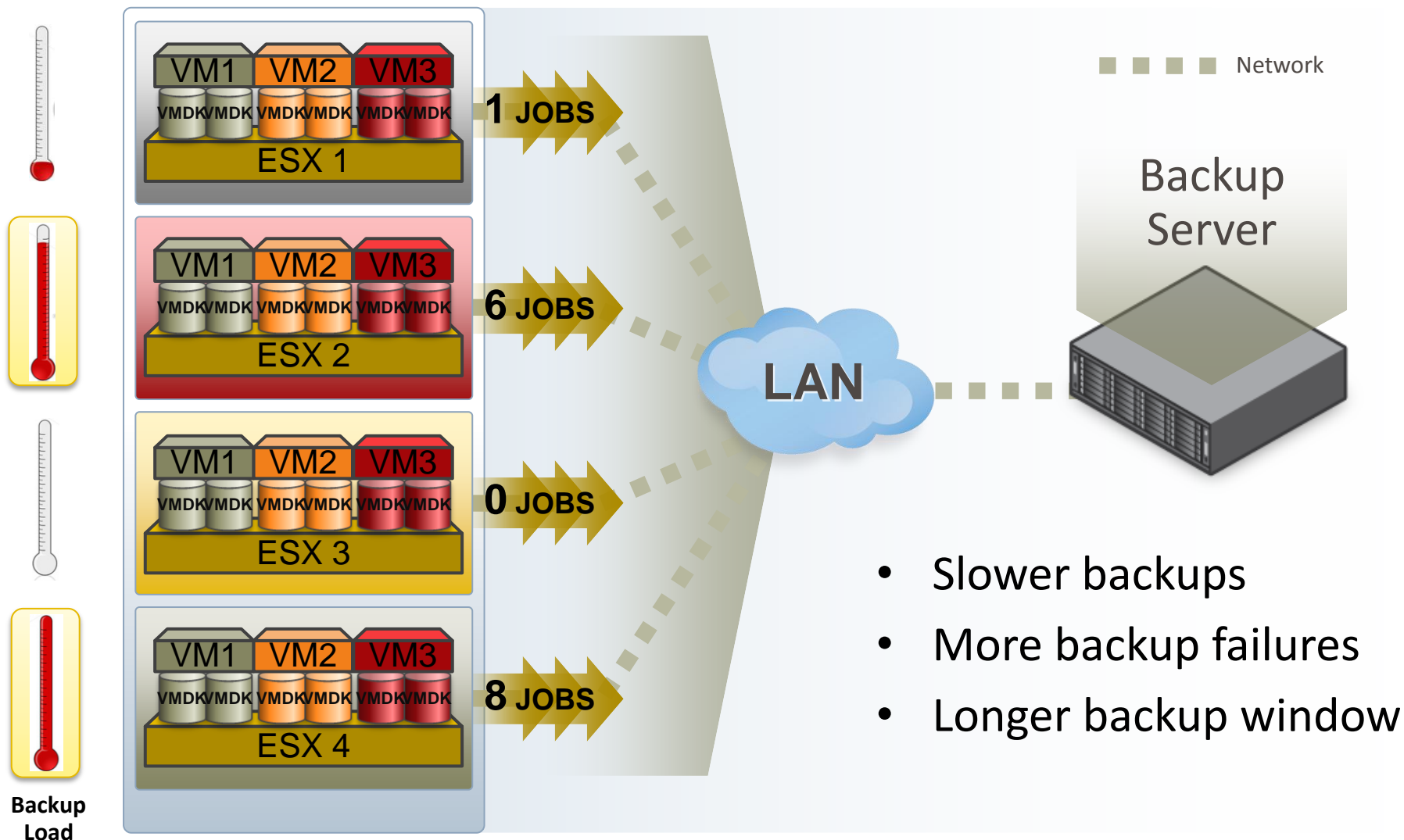 - Could take long time (and create lots of I/O) to complete entire process

✓Symantec.

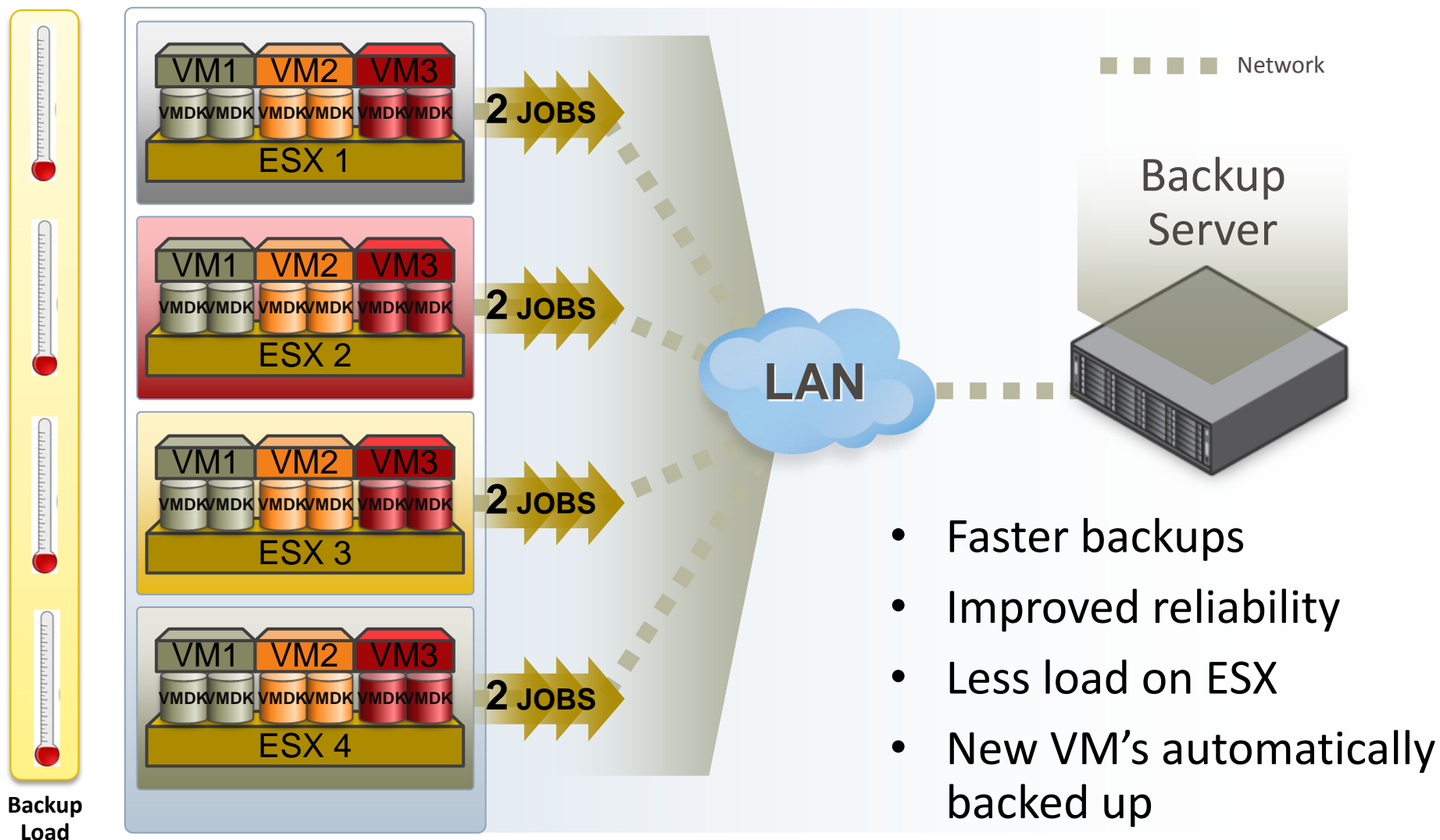# VMware Intelligent Policy Overview

VIP Designed for Two Major Tasks

1) Automatically add and backup new and moved VM's

2) Automatically balance backups across entire vSphere environment (Fibre or network)

- VMs protected based on physical location
  - ESX server
  - ESX Datastore

- VMs protected based on logical attributes
  - vCenter folder
  - Resource pool

- Backup performance is maximized with VIP by using physical location of VM to define backup

# Without VIP – Backup Activity Unbalanced

- Slower backups
- More backup failures
- Longer backup window

With VIP – Automatic Load Balancing

Backup Load

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 1

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 2

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 3

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 4

2 JOBS

2 JOBS

2 JOBS

2 JOBS

Network

LAN

Backup Server

- Faster backups
- Improved reliability
- Less load on ESX
- New VM's automatically backed up

✓Symantec.

VIP Query Generator Overview – Auto Select VMs

- VIP is selected via standard NetBackup policy client tab

- Traditional static or automatic (VIP) VM selection can be used

- VIP cache automatic update can be defined to minimize impact on vCenter server
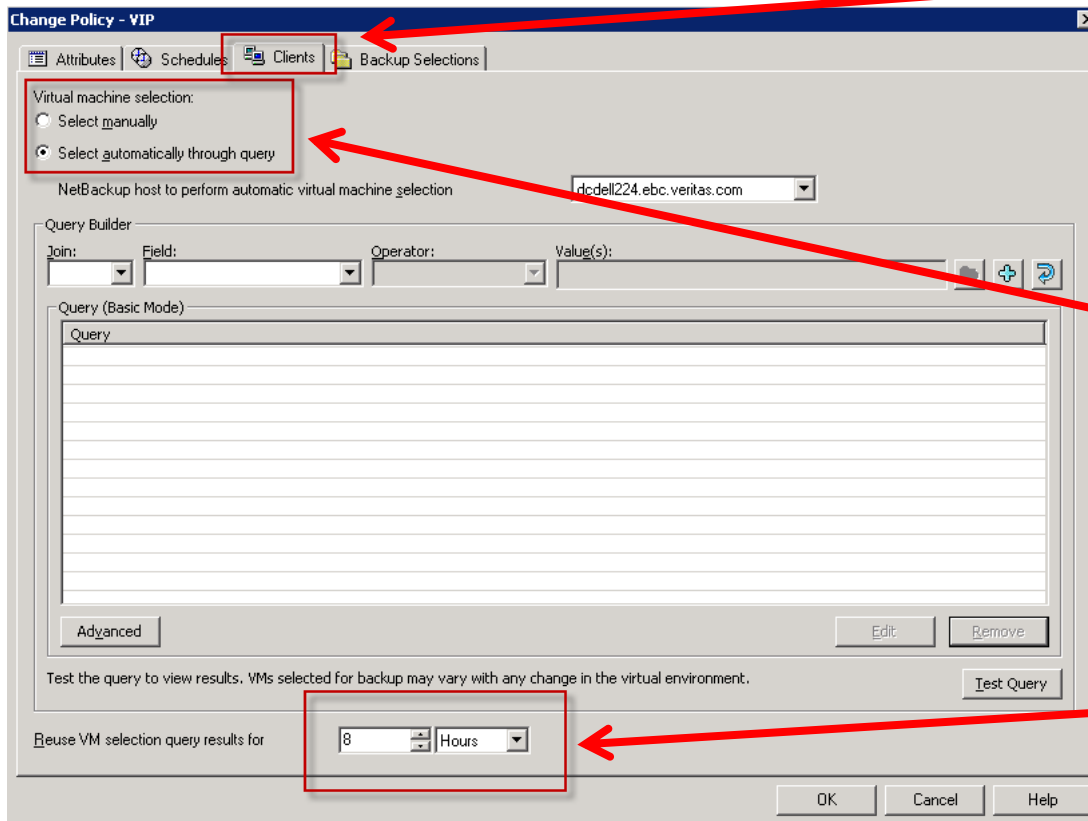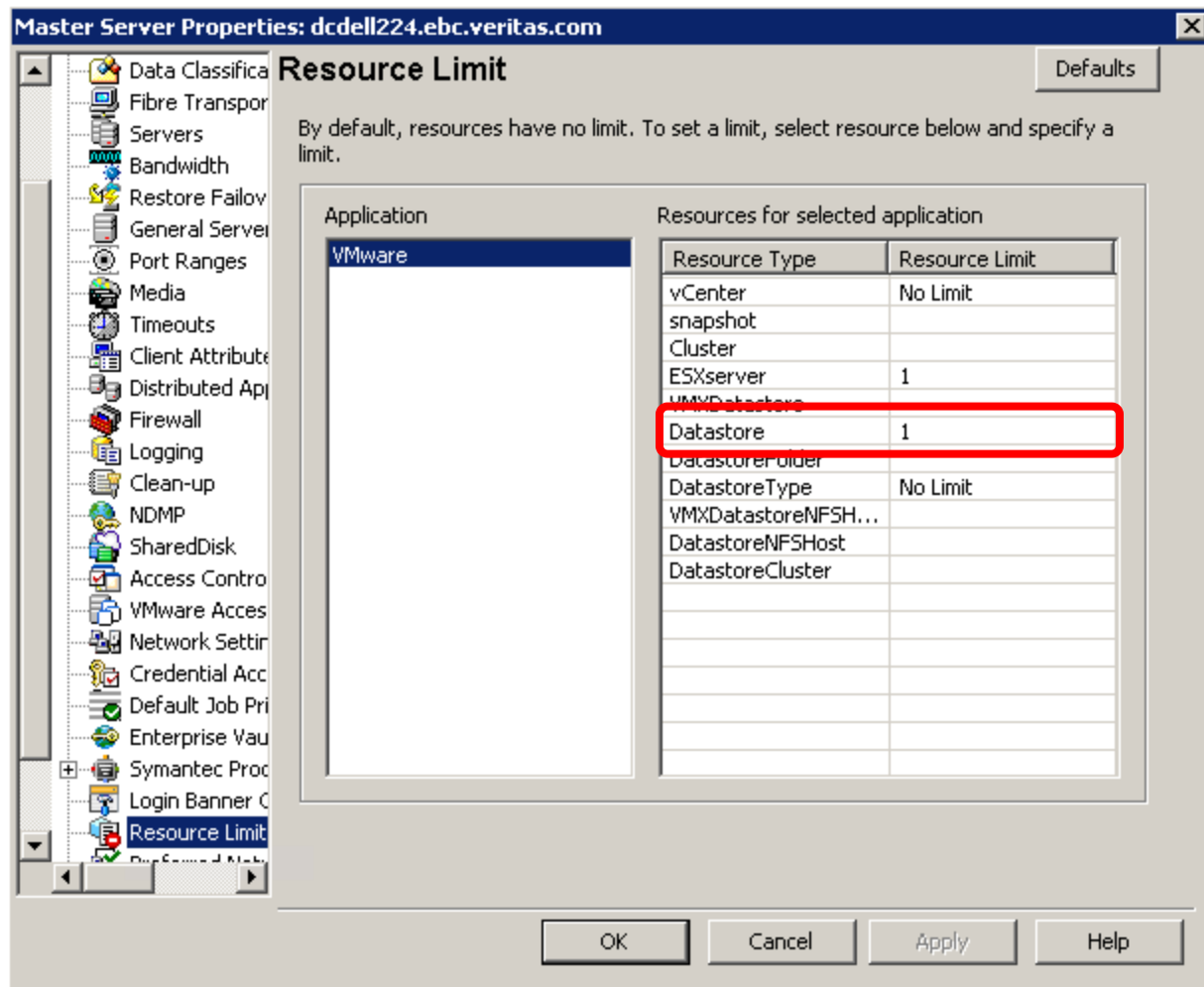
VIP Resource Limits – Limit Jobs Per ESXi or Datastore

- These are global settings

- Two most commonly used settings:

  1. ESXserver (NBD transport)

✔ Symantec.
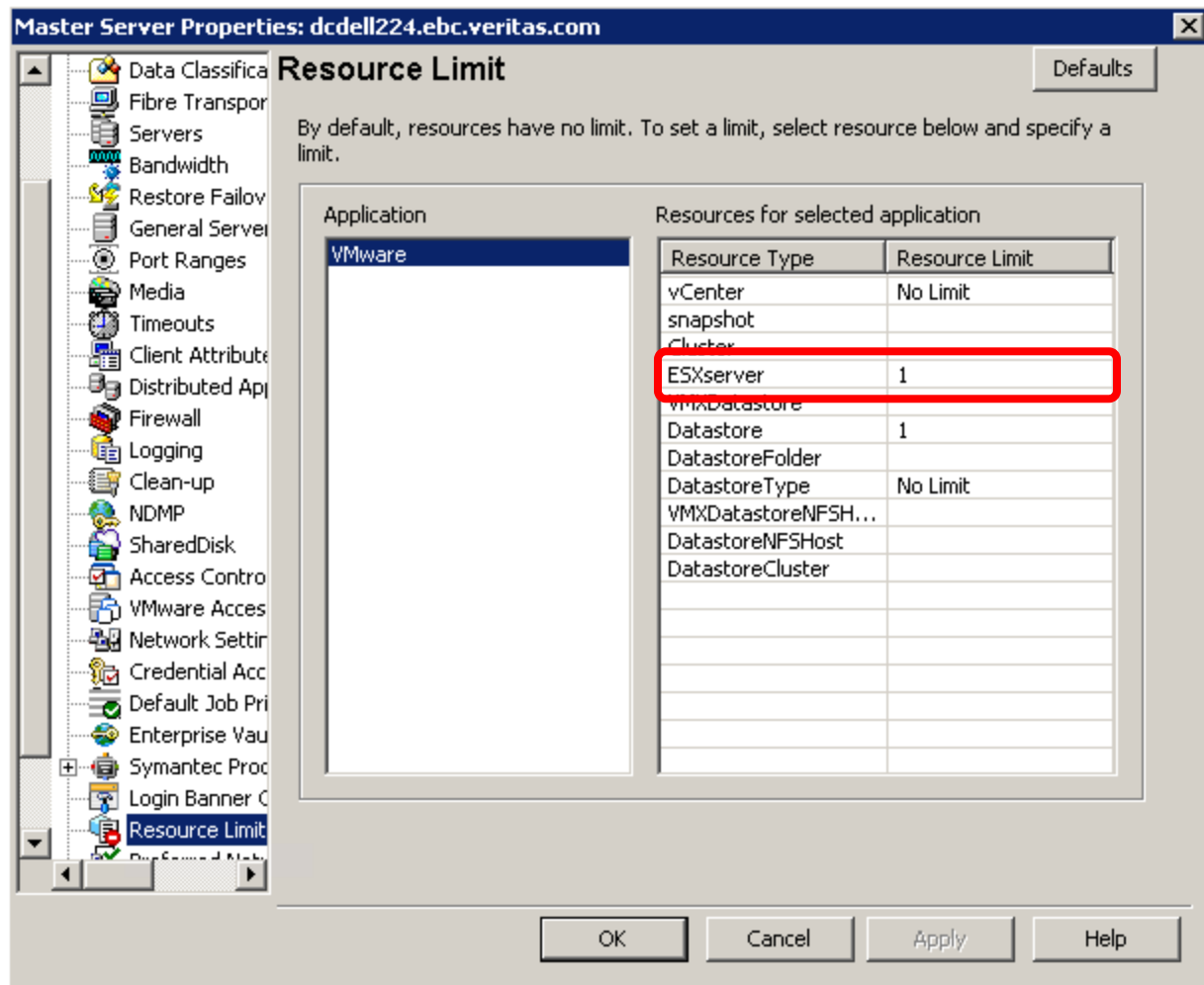
VIP Resource Limits – Limit Jobs Per ESXi or Datastore

- These are global settings

- Two most commonly used settings:

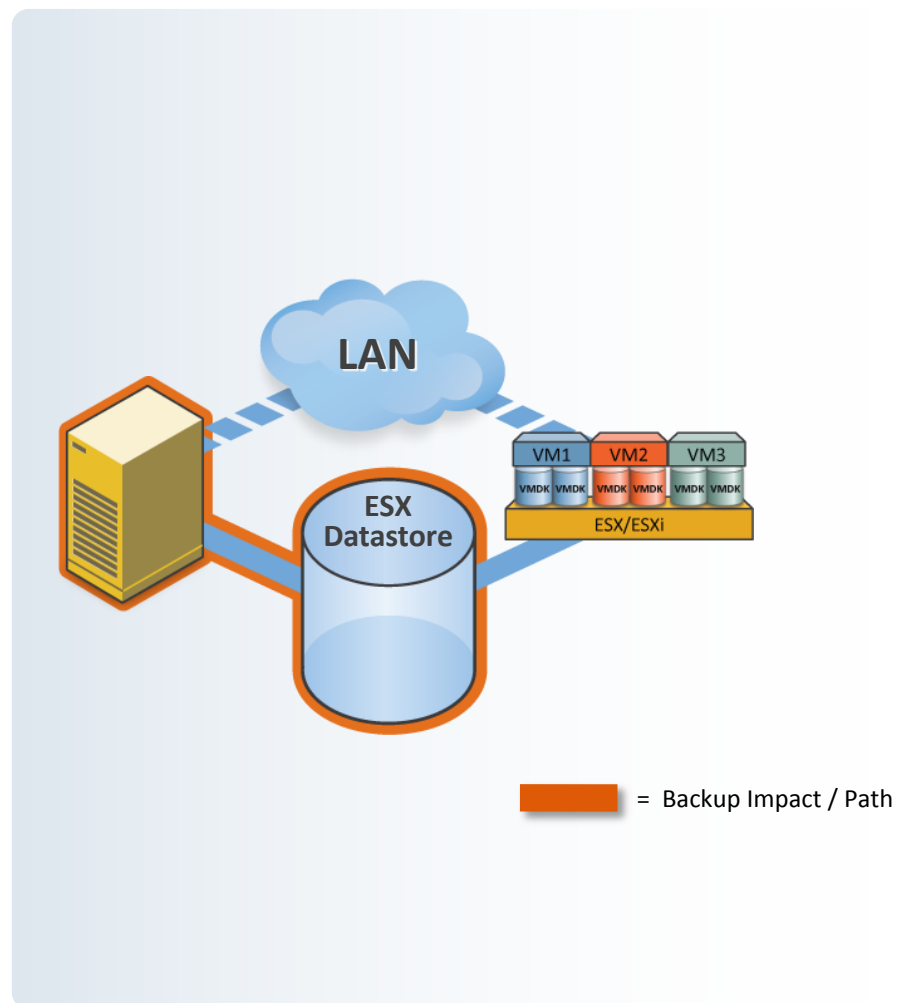 1. ESXserver (NBD transport)

  2. Datastore (SAN transport)

# VMware Intelligent Policy Use Cases

Problem: Equalize Backup Load in SAN Environment

- Shared storage configuration

  - Fibre or iSCSI

  - Near zero impact on ESX

  - No loss in backup or restore functionality

  - No communication with ESX required

LAN

VM1   VM2   VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX/ESXi

ESX Datastore

= Backup Impact / Path

Symantec.

Problem: Equalize Backup Load in SAN Environment

- Manage SAN (iSCSI - shared storage) backups at Datastore level

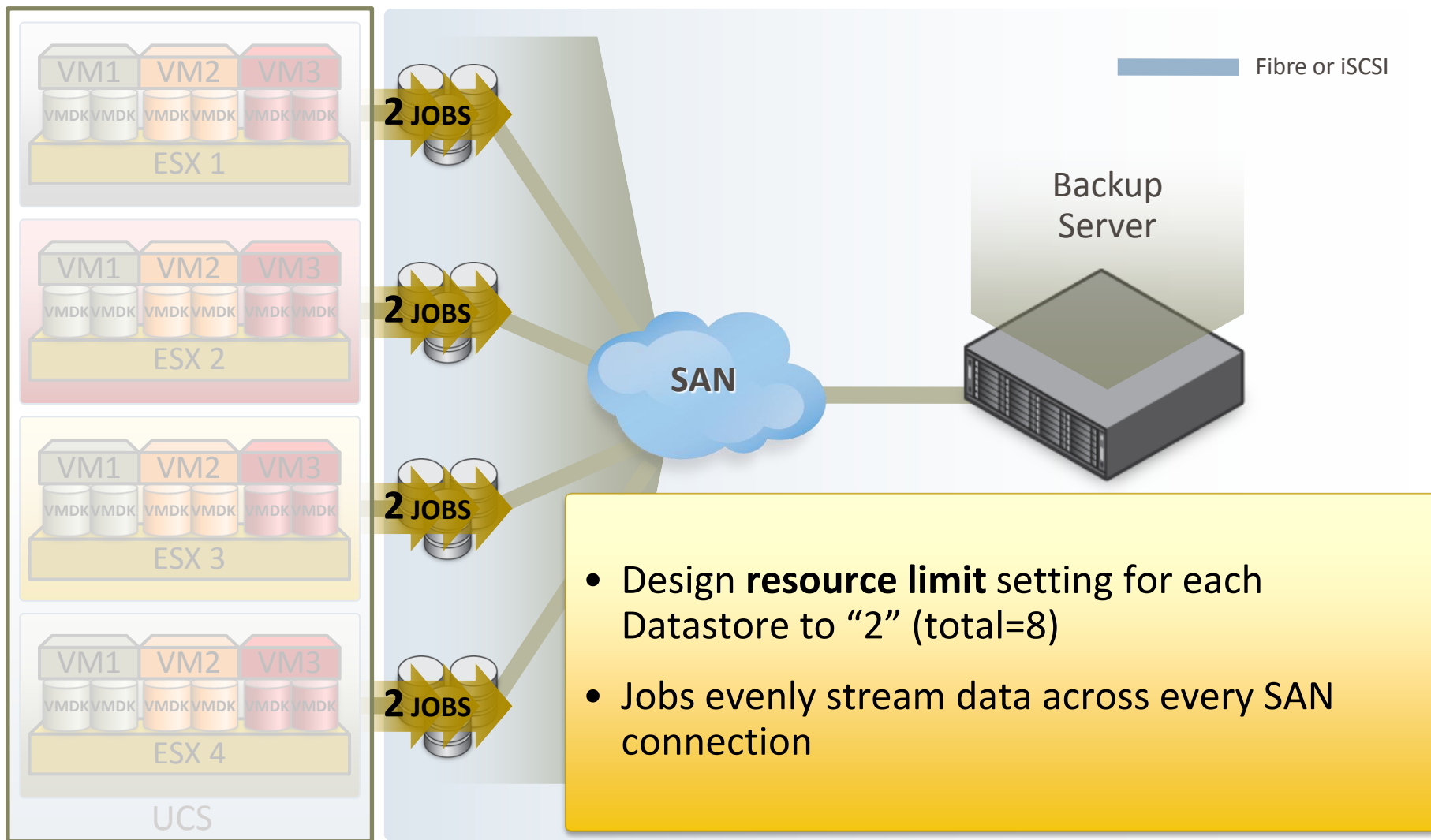- Managing backups at ESX server does not balance load at storage level

Query Builder

Join: | Field:

Query (Basic Mode)

Query
Datastore Contains "PROD"
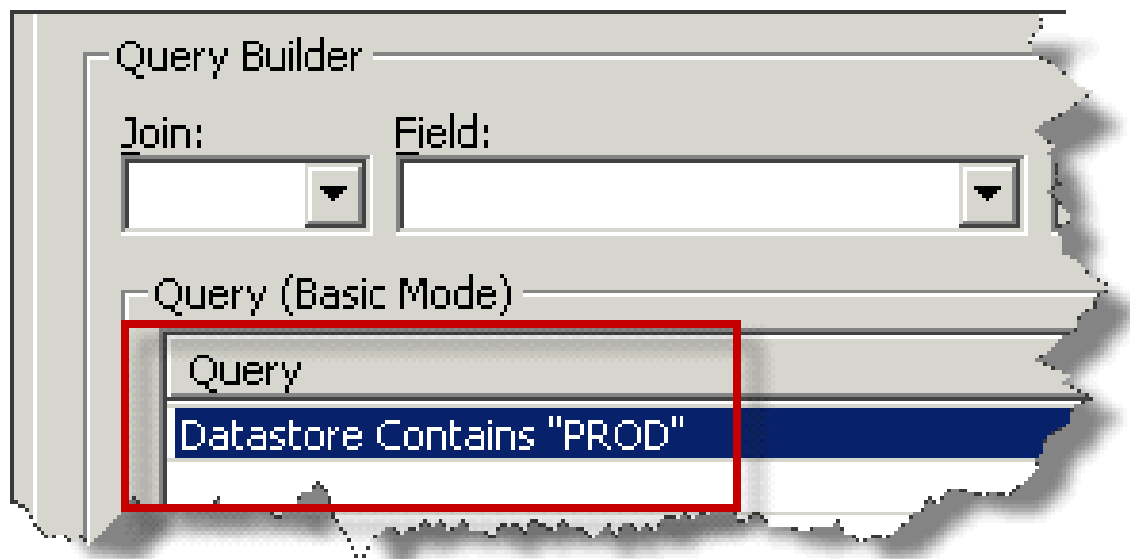
Solution: Will protect every existing and new VM on any Datastore with "PROD" in name

Fully compliant with Storage vMotion

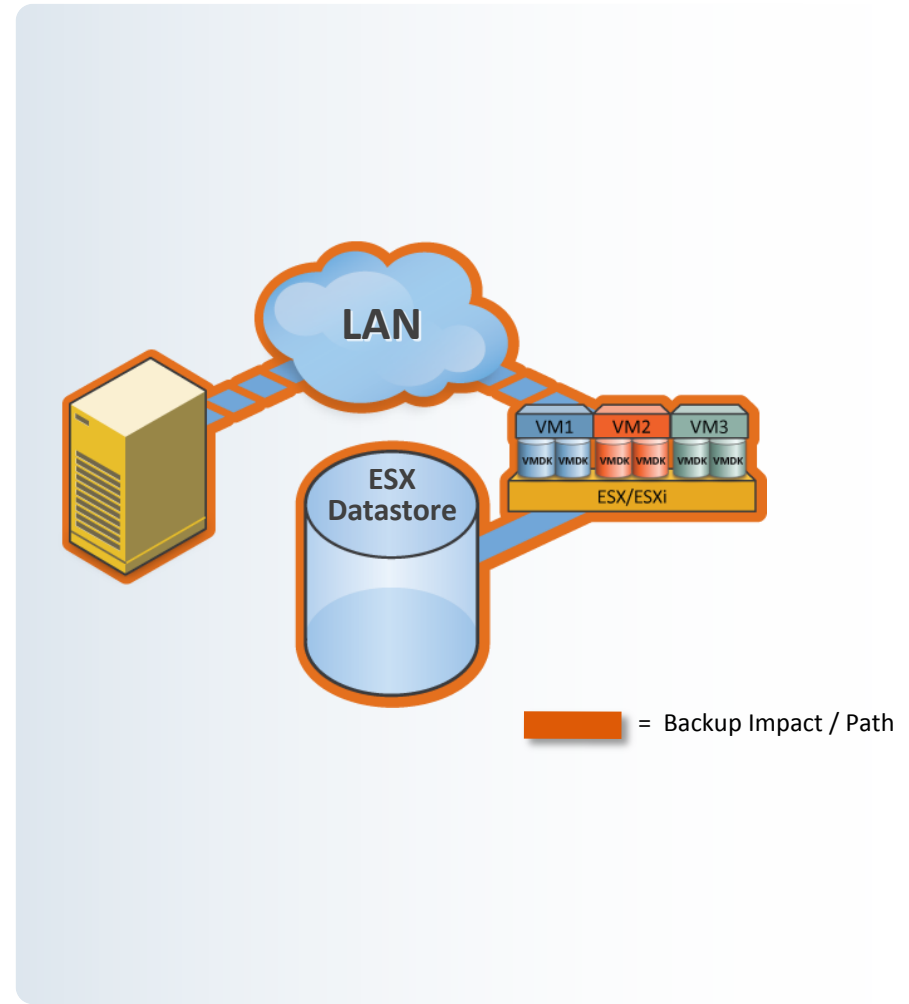Set "Datastore" resource limit to balance backup load across all Datastores

SAN Backups - Equalize Load at Datastore

Fibre or iSCSI

2 JOBS

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 1

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 2

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 3

VM1 VM2 VM3
VMDK VMDK VMDK VMDK VMDK VMDK
ESX 4

UCS

SAN

Backup Server

- Design **resource limit** setting for each Datastore to "2" (total=8)

- Jobs evenly stream data across every SAN connection

✓Symantec.

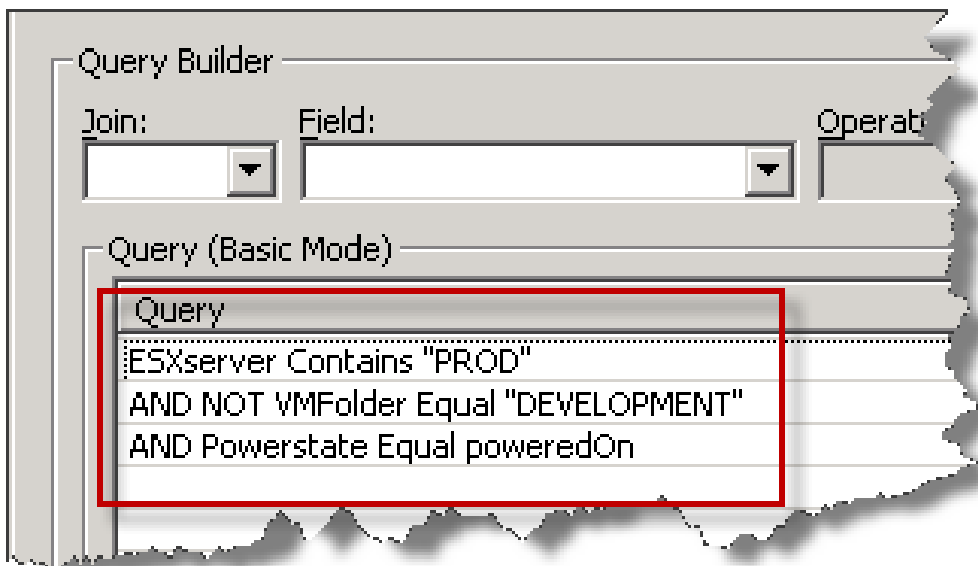# Problem: Network backups – Equalize Load at ESX NIC

- Network Backups (NBD)

  - NFS or DAS fully supported

  - No loss in backup or restore functionality

  - Direct communication with ESX server required (DNS, etc.)

  - ESX server directly impacted

LAN

VM1 VM2 VM3

VMDK VMDK VMDK VMDK VMDK VMDK

ESX/ESXi

ESX Datastore

= Backup Impact / Path

✓Symantec™

Problem: Network backups – Equalize Load at ESX NIC

- All backup traffic over network – no shared storage available
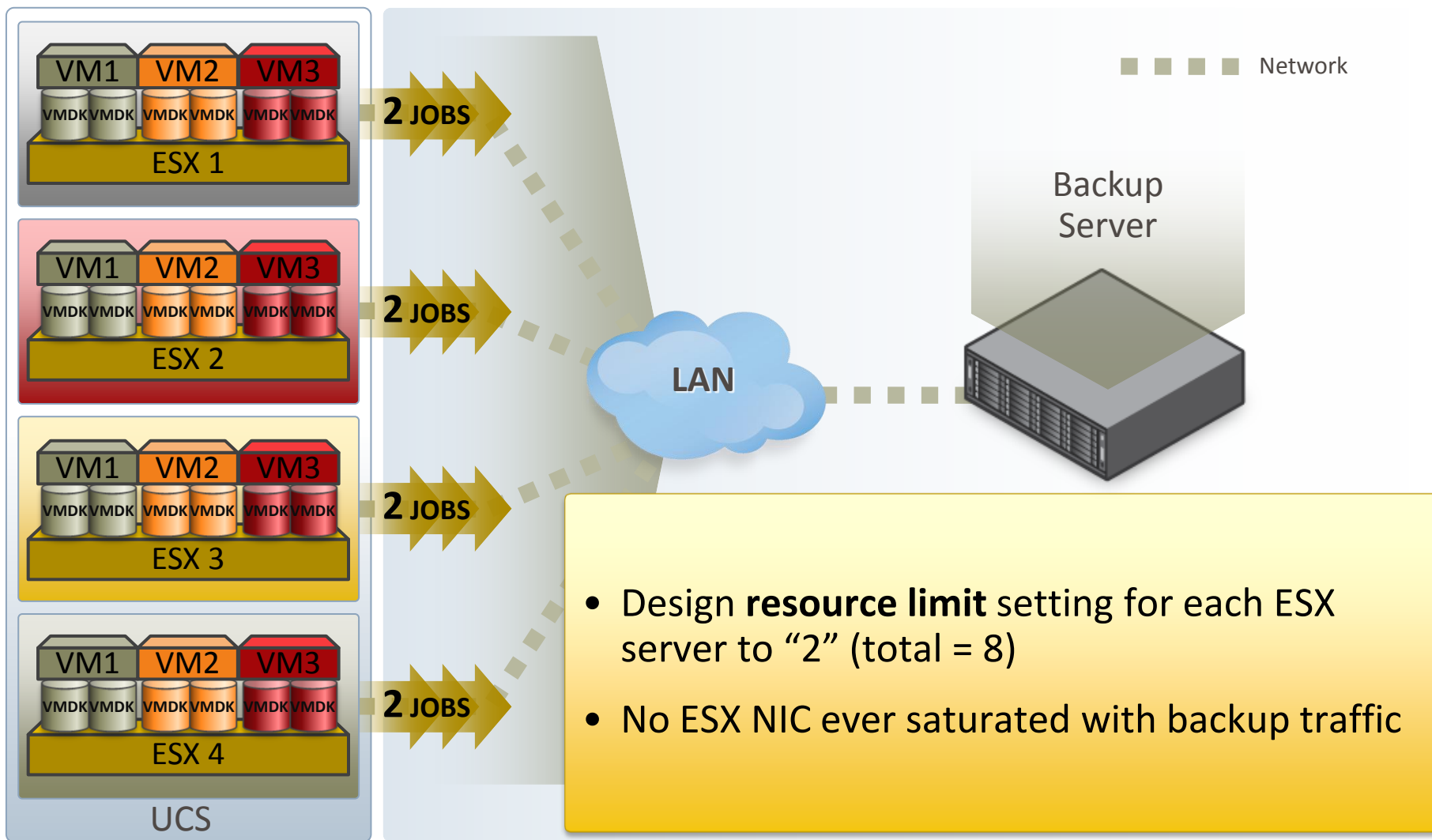
- Don't want to saturate ESX network interface

Solution: All powered on VMs on any ESX server in "Production" will be protected.

Unimportant VMs in the "DEVELOPMENT" folder (only used for development) will be skipped

Set "ESX" resource limit to balance backup load across all ESX NICs
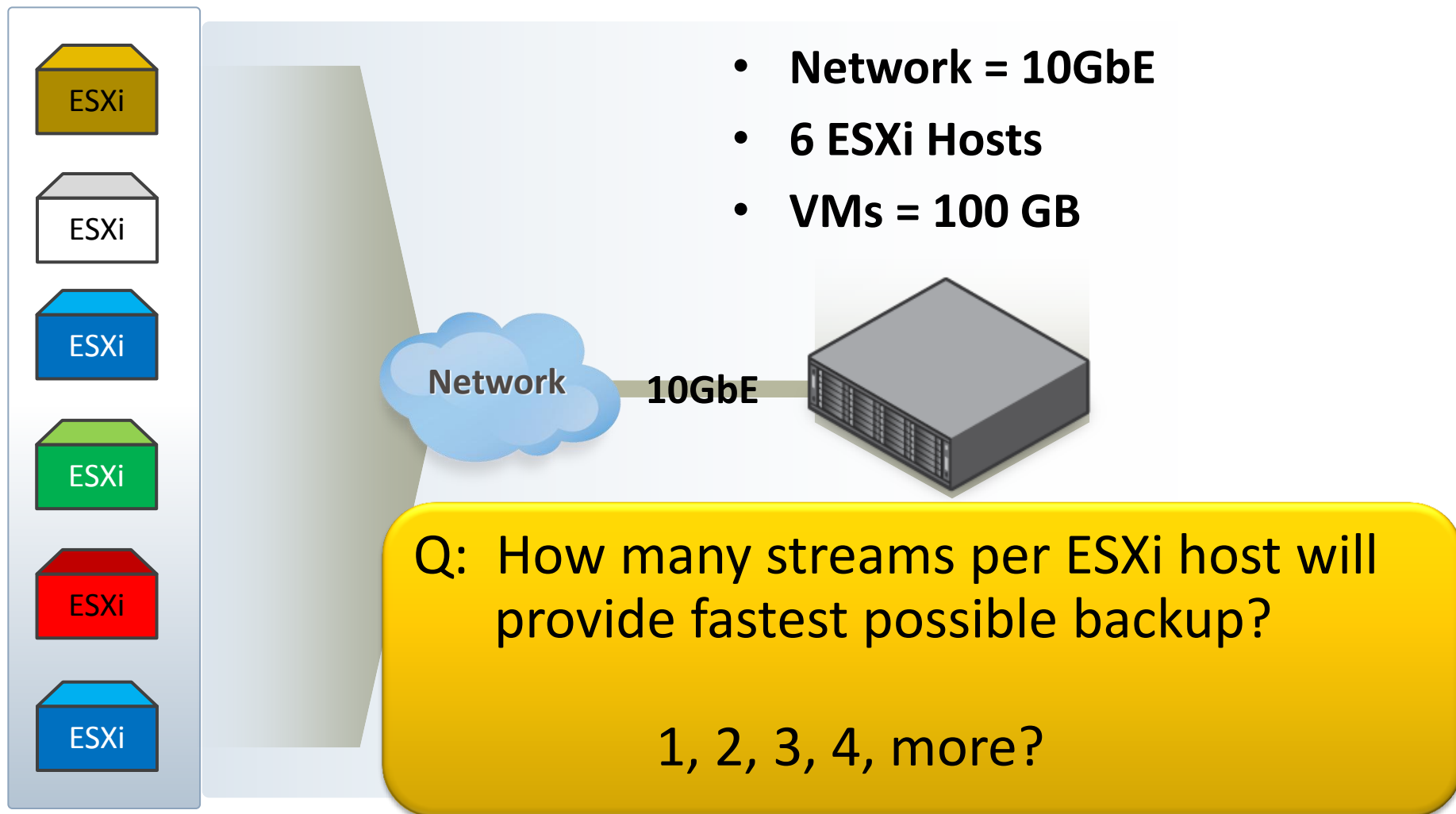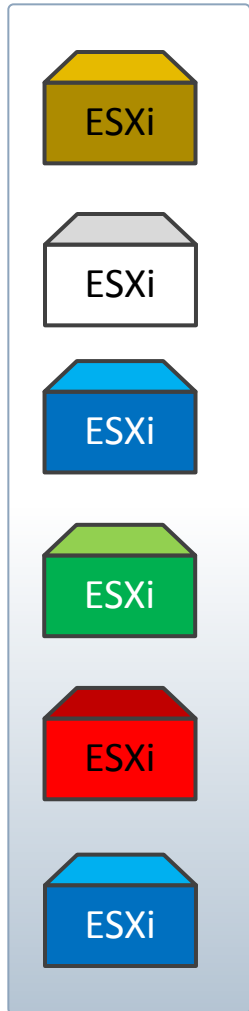
Network backups – Equalize Load at ESX NIC

VM1 VM2 VM3 — VMDK VMDK VMDK VMDK VMDK VMDK — ESX 1 — **2 JOBS**

VM1 VM2 VM3 — VMDK VMDK VMDK VMDK VMDK VMDK — ESX 2 — **2 JOBS**

VM1 VM2 VM3 — VMDK VMDK VMDK VMDK VMDK VMDK — ESX 3 — **2 JOBS**

VM1 VM2 VM3 — VMDK VMDK VMDK VMDK VMDK VMDK — ESX 4 — **2 JOBS**

UCS

Network

LAN

Backup Server

- Design **resource limit** setting for each ESX server to "2" (total = 8)
- No ESX NIC ever saturated with backup traffic

✔Symantec.

VIP Configuration Exercise

Symantec.

VIP Configuration Exercise – NBD Transport

- **Network = 10GbE**
- **6 ESXi Hosts**
- **VMs = 100 GB**

Network — **10GbE**

Q: How many streams per ESXi host will provide fastest possible backup?

1, 2, 3, 4, more?

NBD Transport – How Many Streams?

	1 Stream / ESXi Host (6 streams total)	**3** Streams / ESXi Host (18 streams total)
	Network = 10 GbE	Network = 10GbE
	100 MB/sec per ESXi	100 MB/sec per ESXi
	Each VM = 100 MB/sec	**Each VM = 33 MB/sec**
	Backup 21 VMs / Hour	Backup 21 VMs / Hour

ESXi
ESXi
ESXi
ESXi
ESXi
ESXi

VIP Configuration Example – NBD Transport

- **Network = 10GbE**
- **6 ESXi Hosts**
- **VMs = 100 GB**

Network

10GbE

Backup
Server

Symantec.
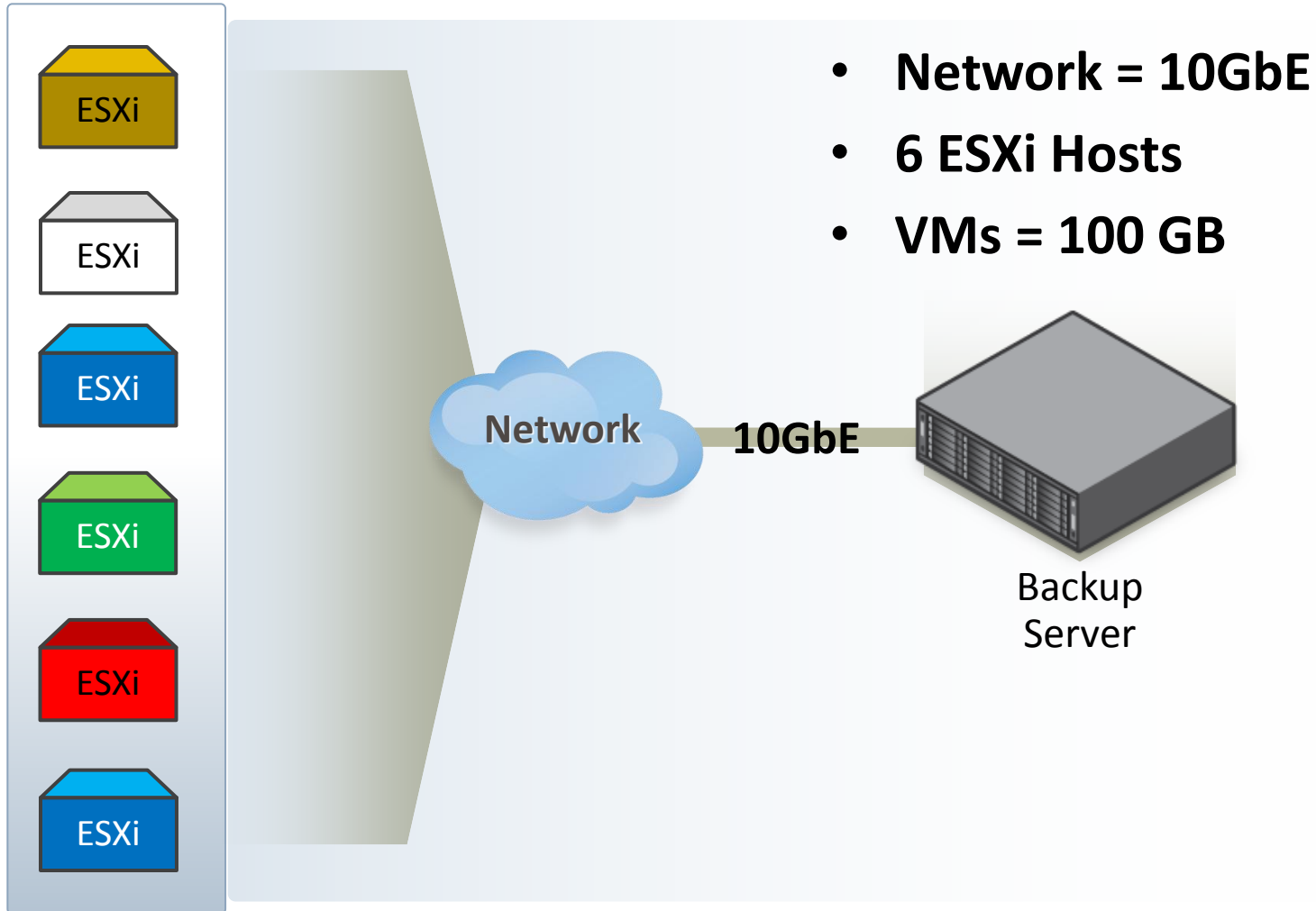
NBD Transport – Impact Of Snapshots Per Datastore

ESXi
ESXi
ESXi
ESXi
ESXi
ESXi

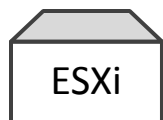Why 45 GB for 3 simultaneous backups?

- Each VM backs up at slower rate

- Snapshot for each VM is open 3 times as long

- Time to capture data change in redo is longer

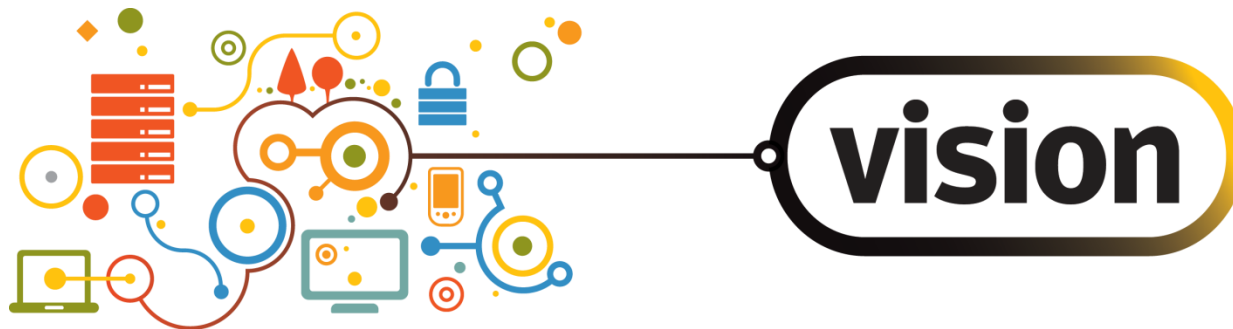- Amount of data to re-apply to VMDK larger

5 GB

Per Datastore

???

Per Datastore

Symantec.

Thank you!

Jim Olson

George Winter