vSphere 4.1 Technical Overview

May 2010



Agenda

- Introduction vSphere 4.1 Features and Themes
- Network Network I/O Control, Load-Based Teaming, IPv6, Performance
- Storage Storage I/O Control, VAAI, I/O Metrics, iSCSI Offload enhancements
- ESXi Deployment Methods, Tech Support Mode enhancements
- HA & FT HA Diagnostics & Reliability, FT enhancements, vMotion enhancements
- DRS & DPM DRS Host Affinity, DPM enhancements
- Management vCenter Server & vSphere Platform enhancements

VMware vSphereTM 4.1

The Platform for Cloud Infrastructure



Confidential

VMware vSphere 4.1: What's New?



Network



Confidential

vSphere 4.1—Network Feature Summary

Network Performance & Scale Improvements

- vmkernel TCP/IP stack—vMotion, NFS, FT logging performance gains
- UDP and intra-host VM to VM performance improvements
- vDS scaling to ~350 hosts/vDS (from current 64 hosts/vDS)—final number is TBD!

Traffic Management (*vDS only* features)

- NetIOC (Network I/O Control)
 - Software scheduler to guarantee service levels for specified traffic types
- LBT (Load Based Teaming)
 - Avoid congestion by dynamic adjustment to NIC team based upon pNIC load

IPv6—NIST Compliance

• Compliance with NIST "Host" Profile

Nexus 1000V Enhancements

Additional Nexus 1000V features—Nexus 1000V V1.4 or 2.0

Network Performance Improvements

vMotion

- Throughput improved significantly for single vMotion
 - ESX 3.5 ~1.0Gbps
 - ESX 4.0 ~2.6Gbps
 - ESX 4.1 max 8 Gbps
- Elapsed reduced by 50%+ on 10GigE tests.

Tx Worldlet

- VM VM throughput improved by 2X, to up to 19 Gbps
- VM Native throughput improved by 10%

LRO (Large Receive Offload)

- Receive tests indicate 5-30% improvement in throughput
- 40 60% decrease in CPU cost

NFS & HW iSCSI Support (Storage) in vSphere 4.1

Improved NFS performance

- Up to 15% reduction in CPU cost for both read & write
- Up to 15% improvement in Throughput cost for both read & write

Broadcom iSCSI HW Offload Support

- 89% improvement in CPU read cost !
- 83% improvement in CPU write cost !



Plotted 18-Mar-2010 15:48:25 (Thu)

- + D=2010-03-09.T=0818.CS=993875.P=swiscsi.V=rdm.H=82598.H=1500.Gos=H2008EE.A=x64.PCpu=8.VH=1.Read%=100.Rand%= 0
- * D=2010-03-08.T=1629.CS=993875.P=____nfs.V=nfs.H=82598.H=1500.Gos=H2008EE.A=x64.PCpu=8.VH=1.Read%=100.Rand%= 0
- * D=2010-03-02.T=1932.CS=993875.P=hwiscsi.V=rdm.H=57711.H=1500.Gos=W2008EE.A=x64.PCpu=8.VH=1.Read%=100.Rand%= 0

Network Traffic Management—Emergence of 10 GigE

1GigE pNICs



- NICs dedicated for some traffic types e.g. vMotion, IP Storage
 - Bandwidth assured by dedicated physical NICs

10 GigE pNICs



- Traffic typically converged to two 10 GigE NICs
- Some traffic types & flows could dominate others through oversubscription

Network I/O Control Goals

- Isolation
 - One flow should not dominate others
- Flexible Partitioning
 - Allow isolation and over commitment
 - Guarantee Service Levels when flows compete

Note: NetIOC feature is only available with vDS (Enterprise Plus)

Limits and Shares

- Limits specify the absolute maximum bandwidth for a flow over a Team
 - Specified in Mbps
 - Traffic from a given flow will never exceed its specified limit
 - Egress from ESX host
- <u>Shares</u> specify the <u>relative importance</u> of an egress flow on a <u>vmnic</u> i.e. <u>guaranteed minimum</u>
 - Specified in abstract units, from 1-100
 - Presets for Low (25 shares), Normal (50 shares), High (100 shares), plus Custom
 - Bandwidth divided between flows based on their relative shares
- Controls apply to <u>output</u> from ESX host
- <u>Shares</u> apply to a given <u>vmnic</u>
- <u>Limits</u> apply across the <u>team</u>

Network I/O Control—Configuration from vSphere Client

🛃 tm-vcenter01.vmworld.com - vSphe	re Client			
File Edit View Inventory Administration	n Plug-ins Help			
💽 💽 🏠 Home 🕨 🛃 Invent	ory 🕨 🧕 Networking			
			2	Limits
E 🛃 tm-vcenter01.vmworld.com	dySwitch			Maximum bandwidth
DHCP-VM02	Getting Started Summary	t defines how different network traffic types are propagat	Virtual Machines H	for traffic class/type
Production01 Static-VM01 dvSwitch	To exclude a physical network a	adapter from the network resource management, go to So	ftware > Advanced Sc	
Wowitch-DVUplinks-48 August DHCP-VM02				Shares
	Summary			Guaranteed minimum
vDS only	Total number of physical adapt Total network bandwidth capac	ers: 2 :ity: 2000 Mbit/s		service level
feature!	Network resource management	: Enabled 📀		
Storage01	Network Resource Pool	Host Limit - Mbit/s Physical Adapter Shares	Shares Value	
taxaad	FT Traffic	Unlimited High	100	
	iSCSI Trattic	Unlimited High	100	
Preconfigured	Management Traffic	Unlimited Normal	50 Retwork Re	source Pool Settings
Traffic Classes	NFS Traffic	Unlimited Normal	50 Network Reso	urce Pool: Virtual Machine Traffic
	Virtual Machine Traffic	500 Normal	50 Resource Alloc	ation
	Edit Set	tings	Bhyrical ada	
Anton another addition	line and	and the second		
	and the second	A standard second and the second	Host limit:	300 Mbit/s
e.g. VM traffic in this exa	ample:		Help	OK Cancel

- limited to max of 500 Mbps (aggregate of all VMs)
- with minimum of 50/400 of pNIC bandwidth (50/(100+100+50+50+50))

Network I/O Control – Demonstration

Demonstration Scenario

- IGigE pnics (vmnics)
- Two traffic types competing for link (link oversubscribed)
 - iSCSI set with "100 shares" (generated by iometer)
 - VM traffic set with "25 shares" (generated by iperf)
 - Link capable of ~890 Mbps
 - iSCSI gets 80% → ~710Mbps
 - VM gets $20\% \rightarrow \sim 180$ Mbps



📸 🚯 🚱 🐲 🏨 Chi Alt 📴 🖳 🗙						
gagan-w2k8-vm.vmware.com - vSphere Client						
File Edit View Inventory Administration Plug-ins Help						
🖸 🔯 🥎 Home 🕨 👸 Inv	entory 👂 👰 Networking		🖅 - Se	arch Inventory		
at at at a						
gagan-w2k8-vm.vmware.com Datacenter_1 VM Network Vmnic0 Vmnic0 Vmnic1 Vmnic4 Vmnic5 Vmnic5 Vmnic5 Vmnic6 Vmnic7 Vmnic7 Vmnic9 C Vc-dvs-vmnic3 Vc-dvs-vmnic3-DVU Vc-dvs-vmnic3 Vc-dvs-vmnic4 Vc-dvs-	vc-dvs-vmnic3 Getting Started Summary Network Resource Management de congested physical network adapte To exclude a physical network adapte Settings on the host Configuration Summary Total number of physical adapters Total network bandwidth capacity: Network resource management:	The show different network the show different network the show different network the show of the network resource tab.	location Configuration traffic types are propagated Switch. te management, go to Softw	Virtual Machines		
dvPortGroup2	Network Resource Pool	Host Limit - Mbit/s	Physical Ada., Shares			
	FT Traffic	Unlimited	Normal 50			
	ISCSI Traffic	Unlimited	High 100			
	vMotion Traffic	Unlimited	Normal 50			
	Management Traffic	Unlimited	Normal 50			
	NFS Traffic	Unlimited	Normal 50			
	Virtual Machine Traffic	Unlimited	Low 25			
۱ <u>۲</u>						

What happens if we turn off IOMETER?

• VM traffic consumes the link up to capacity of ~890Mbps (no competing flows)

Now, flip the shares around

- Set iSCSI to 25 Shares \rightarrow 20% of bandwidth (~180Mbps)
- Set VM Traffic to 100 shares \rightarrow 80 % of bandwidth (~710 Mbps)

Network I/O Control – Demonstration

Now, Introduce a "Limit" on VM traffic of 500 Mbps



Confidential

Now, disable Network I/O Control

• iSCSI and VM traffic compete for link and become erratic.

<u>Most applicab</u>le to 2x 10GigE link deployment—converged traffic types (vMotion, FT, VM, iSCSI, etc)

But, ... can be used with 1GigE links (per demonstration)



NIC Teaming Enhancements—Load Based Teaming (LBT)

DHCP-VM02 Settings	Distributed Virtual Port Group configuration panel	
General Policies Security Traffic Shaping VLAN <u>Teaming and Failover</u> Miscellaneous Advanced	Policies Teaming and Failover Load Balancing: Network Failover Detection: Notify Switches: Failback: Failback: Failover Order Select active and standby dvUplinks. During a failover, standby dvUplinks activate in the order specified below. Name Active dvUplinks dvUplink1 dvUplink3 dvUplink4 Standby dvUplinks	Note: adjacent physical switch configuration is same as other teaming types (except IP-hash). i.e. same L2 domain

Load Based Teaming (LBT) – only available with vDS

- Existing vSphere 4.0 teaming modes use <u>static</u> virtual port to pNIC allocation—teaming algorithms do not account for load
- LBT Goal: avoid congestion by balancing load on the team
- LBT invoked if <u>saturation</u> detected on Tx or Rx (>75% mean utilization over 30s period)
- 30 sec period—long period avoids mac address flapping issues with adjacent physical switches
- · Note: Current implementation does not take NetIORM shares into account

IPv6—Progress towards full NIST "Host" Profile Compliance

VI 3 (ESX 3.5)

• IPv6 supported in guests

vSphere 4.0

- IPv6 support for
 - ESX 4
 - vSphere Client
 - vCenter Server
 - vMotion
 - IP Storage (iSCSI, NFS) EXPERIMENTAL
- Not supported for vSphere vCLI, HA, FT, Auto Deploy

vSphere 4.1

- NIST compliance with "Host" Profile (<u>http://www.antd.nist.gov/usgv6/usgv6v1.pdf</u>)
- Including IPSEC, IKEv2, etc.
- Not supported for vSphere vCLI, HA, FT

Cisco Nexus 1000V—Planned Enhancements

Easier software upgrade

- In Service Software Upgrade (ISSU) for VSM and VEM
- Binary compatibility
- Weighted Fair Queuing (s/w scheduler)
- **Increased Scalability**
- **SPAN to and from Port Profile**
- **VLAN** pinning to PNIC
- Installer app for VSM HA and L3 VEM/VSM communication
- Start of EAL4 Common Criteria certification
- 4094 active VLANs
- Scale Port Profiles > 512

Storage



Confidential

Storage I/O Control

vStorage API for Array Integration (VAAI)

New Performance Monitoring Statistics



The I/O Sharing Problem

- Low priority VM can limit I/O bandwidth for high priority VMs
- Storage I/O allocation should be in line with VM priorities



What you want to see







Confidential

Setting I/O Controls



Enabling Storage I/O Control



vmware[®]

🚱 YSITAUTO1 - vSphere Client								_ 🗆 🗙
Eile Edit View Inventory Administration Plug-ins Help								
🖸 💽 🔥 Home 🕨 🛃	Inventory 🕨 📔 Datas	stores				🛃 🗸 Search Inven	tory	Q
□ ✓ YSITAUTO1 □ IO-DRS □ datastore1 (1) □ LocalDisk □ pa-group-stautomatio □ pvt_lun250gb ③ SharedStorage □ test	pvt_lun250gb Getting Started Setting Started Image: Setting Started <t< th=""><th>Summary Virtual Machir State Powered On Powered On</th><th>es Hosts Perfo Status Normal Normal</th><th>vrmance Configuration Tasks & Even Na Host vcuiqa010.eng.vmware.com vcuiqa012.eng.vmware.com</th><th>nts Alarms Permis ame, State, Host or Gue Shares Value 1500 3500</th><th>sions Storage Vie st OS contains: • [Limit - IOPs Unlimited Unlimited</th><th>ws Datastore % Sha 30 70</th><th>Clear ares</th></t<>	Summary Virtual Machir State Powered On Powered On	es Hosts Perfo Status Normal Normal	vrmance Configuration Tasks & Even Na Host vcuiqa010.eng.vmware.com vcuiqa012.eng.vmware.com	nts Alarms Permis ame, State, Host or Gue Shares Value 1500 3500	sions Storage Vie st OS contains: • [Limit - IOPs Unlimited Unlimited	ws Datastore % Sha 30 70	Clear ares
Recent Tasks Name, Target or Status contains:								
Name	Target	Status	Details	Requested St	art Ti ▽ 🛛 Start Time	e Cor	mpleted Time	▲
Reconfigure virtual machine	io-rm_vm1	📀 Completed		8/25/2009 6:	52:20 PM 8/25/2009	9 6:52:20 PM 8/2	5/2009 6:52:22 PM	
Reconfigure virtual machine	io-rm_vm2	Completed		8/25/2009 6:4	46:25 PM 8/25/2009	96:46:25 PM 8/2	5/2009 6:46:27 PM	
Reconfigure virtual machine	io-rm_vm2	Completed		8/25/2009 6:4	46:15 PM 8/25/2009	9 6:46:15 PM 8/2	5/2009 6:46:18 PM	•
🚰 Tasks 🞯 Alarms							Admin	istrator //

Allocate I/O Resources

Shares translate into ESX I/O queue slots

- VMs with more shares are allowed to send more I/O's at a time
- Slot assignment is dynamic, based on VM shares and current load
- Total # of slots available is dynamic, based on level of congestion





Storage I/O Control in Action: Example #1

Two VMs running IOmeter* on two hosts

- 16 KB random reads
- VM1 : 1000 shares
- VM2 : 2000 shares

Result: VMs get IOPS in ratio 1:2

	Without Stora	age I/O Control	With Storage I/O Control		
	IOPS	IOmeter Latency (ms)	IOPS	IOmeter Latency (ms)	
VM1	1500	20	1080	31	
VM2	1500	21	1900	16	

Storage I/O Control in Action: Example #2

Two Windows VMs running SQL Server on two hosts

- 250 GB data disk, 50 GB log disk
- VM1: 500 shares
- VM2: 2000 shares

Result: VM2 with higher shares gets more orders/min & lower latency!

	Without Storage I/O Control		With Storage I/O Control		
	Orders/ Minute	Processing Time (ms)	Orders/ Minute	Processing Time (ms)	
VM1 (500 Shares)	8800	213	7000	275	
VM2 (2000 Shares)	8500	220	12400	150	
Aggregate	17300		19400		

Step 1: Detect Congestion



Congestion signal: ESX-array response time > threshold

- Default threshold: 35ms
- We will likely recommend different defaults for SSD and SATA
- Automated tiered storage arrays are NOT SUPPORTED with 4.1

Changing default threshold (not usually recommended)

- Low latency goal: set lower if latency is critical for some VMs
- High throughput goal: set close to IOPS maximization point

What are Automated Tiered Storage

LUNs/blocks migrated between tiers automatically & transparently

- Tiers: SSD, FC, SAS, SATA
- Examples: Compellent, EMC FAST, Dell Equallogic

Static latency threshold may not work well for such systems

- Threshold too low unnecessary throttling
- Threshold too high congestion not detected

Certification program to assess compatibility

- <u>Targeted at automated tiered systems only</u> tell us who you are!
- Systems supporting only manual migration between tiers do not need certification
- Cert program to be launched on or after vSphere 4.1 GA
 - May initially be limited to a subset of automated tiered storage partners

Achieve compatibility by dynamically changing threshold via vSphere API

vSphere 4.1 Documentation also describes use of this features in the Resource Management Guide Chapter 4 (pages 37-39)

Also check out the VMWorld 2009 presentation video:

http://vmworld.com/docs/DOC-3855

Additional Collateral planned for release at GA

- Technical White paper on Concepts and Deployment Considerations for SIOC
- Frequently Asked Questions
- Datasheet or webpage content



vStorage APIs for Array Integration (VAAI)

Improves performance by leveraging efficient array-based operations as an alternative to VMware host-based solutions Three Primitives include :

- 1. Full Copy Xcopy like function to offload work to the array
- 2. Write Same -Speeds up zeroing out of blocks or writing repeated content
- 3. Atomic Test and Set Alternate means to locking the entire LUN

Helping function such as:

- Storage vMotion
- Provisioning virtual machines from Template
- Improves thin provisioning disk performance
- VMFS share storage pool scalability
- •Requires firmware from Storage Vendors (6 participating, see later slides)
- •Supports block based storage only in the 4.1 release
Benefits

- Reduce installation time
- Standardize to ensure efficient management, protection & control

Challenges

- Requires a full data copy
- 100 GB template (10 GB to copy): 5-20 minutes
- FT requires additional zeroing of blocks

Improved Solution

- Use array's native copy/clone & zeroing functions
- Up to 10-20x speedup in provisioning time

<u>⊂</u> lone	
Conver <u>t</u> to Virtual Machine.	
Deploy Virtual Machine from	n this Template
Add Permission	Ctrl+P
Rena <u>m</u> e	
Remove from Inventory	
Delete from Disk	

Storage vMotion with Array Full Copy Function

Benefits

- Zero-downtime migration
- Eases array maintenance, tiering, load balancing, upgrades, space mgmt

Challenges

- Performance impact on host, array, network
- Long migration time (0.5 2.5 hrs for 100GB VM)
- Best practice: use infrequently

Improved solution

Use array's native copy/clone functionality



VAAI Speeds Up Storage vMotion - Example



Copying Data – Optimized Cloning with VAAI



VMFS directs storage to move data directly

- Much less time!
 - Up to 95% reduction
- Dramatic reduction in load on:
 - Servers
 - Network
 - Storage



- A number of VMFS operations cause the LUN to temporarily become locked for exclusive write use by one of the ESX nodes, including:
 - Moving a VM with vMotion
 - Creating a new VM or deploying a VM from a template
 - Powering a VM on or off
 - Creating a template
 - Creating or deleting a file, including snapshots
- A new VAAI feature, atomic_test_and_set allows the ESX Server to offload the management of the required locks to the storage and avoids locking the entire VMFS file system.

Makes VMFS more scalable overall, by offloading block locking mechanism

Using Atomic Test and Set (ATS) capability provides an alternate option to use of SCSI reservations to protect the VMFS metadata from being written to by two separate ESX Servers at one time.





For more details on VAAI

vSphere 4.1 Documentation also describes use of this features in the ESX Configuration Guide Chapter 9 (pages 124 - 125)

Listed in TOC as "Storage Hardware Acceleration"

Three setting under advanced settings:

- DataMover.HardwareAcceratedMove Full Copy
- DataMover.HardwareAcceratedInit Write Same
- VMFS3.HarwareAccerated Locking Atomic Test Set

Additional Collateral planned for release at GA

- Frequently Asked Questions
- Datasheet or webpage content

Partners include : Dell/EQL, EMC, HDS, HP, IBM and NetApp

* Will only support block based storage in 4.1

New Performance Monitoring Statistics

Ele Edit View Terminal Tabs Help 1:51:51am up 1:34, 157 worlds; CP SCSI stats: 0 conflicts, 0.00 con Cache hit: 0.00%, Cache miss: 0.00% nBuffers: 2048, maxFlushQuantum: 10	Terminal	 Compreherror performance monitoring Heteroge
DEVICE PATH/WORLD/PARTITI	DN DQLEN WQLEN ACTV QUED %USD LOAD CMDS/s READS/s WRITES/s MBREAD/s MBWRTN/s DAVG/cmd KAVG/cmd GAVG/cmd QAVG/cmd	environme
mpx.vmbba0:C0:T - mpx.vmbba1:C0:T - t10.ATA ST3 - {NFS}nfsMount0 - {NFS}nfsMount1 - {NFS}nfsMount2 - {NFS}nfsMount3 - {NFS}nfsMount3 - {NFS}nfsMount4 - {NFS}nfsMoun4	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	• Tools sup • GU comp • Cor drill-c
C C Arrest Street Contraction	entory b 🛐 Hosts and Custers .	
Image: State of the state o	10.20.104.57 YMware ESXi, 4.1.0, 00000 Evaluation (60 days remaining)	
10.20.104.57	Battorg Staticel Summary Versual Machined Resource Allocation Performance Configuration Totals 5 Events Allorm Permissions Statice Detastore/Real-time, 7/17/2009 3:10:20 PM - 7/17/2009 4:10:20 PM Chart Options Switch for: Datastore Image: Configuration Switch for: Datastore Image: Configuration Switch for: Datastore Image: Configuration Image: Configuration </td <td>Features: •Throughpu • Da^a • Sto • Da^a • VM •Realtime a •Esxtop (fo</td>	Features: •Throughpu • Da ^a • Sto • Da ^a • VM •Realtime a •Esxtop (fo

• Comprehensive host & VM storage performance statistics enable proactive monitoring to simplify troubleshooting

- Heterogeneous customer storage environments supported (FC, iSCSI, NFS)
- Tools support varied usage scenarios
 - GUI for trending and user-friendly comparative analysis
 - Command-line for scripting/ drill-down at host

- •Throughput and latency statistics for:
 - Datastore per host
 - Storage adapter & path per Host
 - Datastore per VM
 - VMDK per VM
- Realtime and historical trending (vCenter)
 Esxtop (for ESX) and resxtop (for ESXi)

Statistics targeted in solution: vCenter & esxtop

Inventory object	Per Component	Statistic	FC/NFS/ iSCSI	vCenter?	Esxtop?
Host	datastore	Throughput, latency	All	~	♦</td
	Storage adapter	Throughput, latency	FC*	\checkmark	Available today
	Storage path	Throughput, latency	FC*	~	Available today
	LUN	Throughput, latency	FC, iSCSI**	Available today	Available today
VM	Datastore	Throughput, latency	AII	~	<
	VMDK	Throughput, latency	All	~	~

Not available in this timeframe: Aggregation at cluster level in vCenter (possible through APIs)

*Network-based storage (NFS, iSCSI) I/O breakdown still being researched

** Not applicable to NFS; datastore is the equivalent

ESXTOP publishes throughput and latency for LUN, if datastore has only one LUN then LUN will be equal datastore

						Terminal								
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>T</u> erminal Ta <u>b</u> s <u>H</u> elp														
1:51:51am up 1:34, 157 worlds; CPU SCSI stats: 0 conflicts, 0.00 confl Cache hit: 0.00%, Cache miss: 0.00%, nBuffers: 2048, maxFlushQuantum: 1024	load avera icts/sec, fssRead: , flush in	age: 0. 0.00 i 0.00, nterval	.04, usec/ fssW l: 30	0.03, cmd /rite: 0000 m	0.02 0.00 s	9								
DEVICE PATH/WORLD/PARTITION	DQLEN WQ	LEN ACT	rv qu	JED %U	SD LO	DAD CMDS/s	READS/s	WRITES/s N	1BREAD/s I	MBWRTN/s D	AVG/cmd K/	AVG/cmd G	GAVG/cmd Q	AVG/cmd
mpx.vmhba0:C0:T -	1	-	0	Θ	Θ Θ.	.00 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mpx.vmhba1:C0:T -	4	-	0	Θ	Θ Θ.	.00 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
t10.ATAST3 -	1	-	0	0	00.	.00 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
{NFS}nfsMount0 -	-	•	1	-	-	- 90.14	40.06	50.08	0.16	0.20	-	-	0.01	-
{NFS}nfsMount1 -	-	-	0	-	-	- 0.00	0.00	0.00	0.00	0.00	-	-	0.00	-
{NFS}nfsMount2 -	-	-	Θ	-	-	- 0.00	0.00	0.00	0.00	0.00	-	-	0.00	-
{NFS}nfsMount3 -	-	-	0	-	-	- 0.00	0.00	0.00	0.00	0.00	-	-	0.00	
{NFS}nfsMount4 -	-	-	0	-	-	- 0.00	0.00	0.00	0.00	0.00	-	-	0.00	
{NFS}nfsMount5 -	-	-	0	-	-	- 0.00	0.00	0.00	0.00	0.00	-	-	0.00	

vmware[®]

Datastore Activity Per Host



Other Host Stats

Customize Performance Chart					
Saved Chart Settings: Default Chart Options Chart Options Chart Type	gs at	startup			
CPU	_) Stacked Grapi	n (Per VM)		
					All None
		Units Number	Internal Name numberRead	Collection Level	
€ System	n	Millisecond KBps KBps Number Millisecond	totalReadLatency read write numberWrite totalWriteLatency	2 2 2 2 2	
Counter Description Counter Description Counter Description Rollup: Average Rate of reading data from the datastore To: 7/17/2009 4:12 PM					Statistics Type: Rate
					Manage Chart Settings Save Chart Settings
Help					OK Cancel Apply

vmware[®]

Datastore Activity per VM



Confidential

vmware[®]

Virtual Disk Activity per VM



ESXi



Confidential

Boot From SAN

- Fully supported in ESXi 4.1
- Was only experimentally supported in ESXi 4.0
- Boot from SAN supported for FC, iSCSI, and FCoE

Scripted Installation

- Numerous choices for installation
 - Installer booted from
 - CD-ROM (default)
 - Preboot Execution Environment (PXE)
 - ESXi Installation image on
 - CD-ROM (default), HTTP / HTTPS, FTP, NFS
 - Script can be stored and accessed
 - Within the ESXi Installer ramdisk (default)
 - On the installation CD-ROM
 - HTTP / HTTPS, FTP, NFS
 - Configuration script ("ks.cfg") can include other scripts
 - Preinstall
 - Postinstall
 - First boot

New Feature: Full Support of Tech Support Mode

Troubleshooting Mode Option	NS	Modify Tech Support timeout	
Enable Local Tech Support Disable Remote Tech Support Modify Tech Support timeout Restart Management Agents	t (SSH) t	Modify the amount of time before Tech Support Mode logins are automatically disabled. Maps Hardware Status	
ecurity Profile			
Services		Refresh	Properties
I/O Redirector (Active Directory Network Login Server (Active Dir Ibtd	Can	enable in vCenter or DCUI	
I/O Redirector (Active Directory Network Login Server (Active Dir Ibtd Local Tech Support Local Security Authentication Ser NTP Daemon VMware vCenter Agent Remote Tech Support (SSH) Direct Console UI	Can	enable in vCenter or DCUI	

Lockdown Mode:

Disabled

New Feature: Full Support of Tech Support Mode



Tech Support Mode is not supported unless used in consultation with VMware Tech Support.

esx23b.vmworld.com login: _

Big Scary Warning will be removed...

...but admin will be notified when active

] 🛃 tm-vcenter05.vmworld.com	esx23b.vmworld.com VMv Summary Virtual Machine	vare E5Xi, 4.1.0, 235786 s Performance Configuration Tasks (& Events Alarms Permissions Maps	Hardware Status				
esx23a.vmworld.cor esx23b.vmworld.cor esx24a.vmworld.cor	esx23a.vmworld.cor esx23b.vmworld.cor esx24a.vmworld.cor The Local Tech Support Mode for the host esx23b.vmworld.com has been enabled							
esx24b.vmworld.cor	General	General Resources						
	Manufacturer: Model:	IBM IBM eServer BladeCenter H	CPU usage: 161 MHz	Capacity 8 x 2.5 GHz				





- Timeout automatically disables Tech Support Mode (local and remote)
- Running sessions are not terminated
- All commands issued in Tech Support Mode are sent to syslog

Recommended uses

- Support, troubleshooting, and break-fix
- Scripted deployment preinstall, postinstall, and first boot scripts

Discouraged uses

- Any other scripts
- Running commands/scripts periodically (cron jobs)
- Leaving open for routine access or permanent SSH connection



Provides authentication for all local services

- vSphere Client
- Other access based on vSphere API
- DCUI
- Tech Support Mode (local and remote)

Has nominal Active Directory groups functionality

- Members of "ESX Admins" AD group have Administrative privilege
- Administrative privilege includes:
 - Full Administrative role in vSphere Client and vSphere API clients
 - DCUI access
 - Tech Support Mode access (local and remote)

Storage

- esxcli swiscsi session: Manage iSCSI sessions
- esxcli swiscsi nic: Manage iSCSI NICs
- esxcli swiscsi vmknic: List VMkernel NICs available for binding to particular iSCSI adapter
- esxcli swiscsi vmnic: List available uplink adapters for use with a specified iSCSI adapter
- esxcli vaai device: Display information about devices claimed by the VMware VAAI (vStorage APIs for Array Integration) Filter Plugin.
- esxcli corestorage device: List devices or plugins. Used in conjunction with hardware acceleration.

Network

• esxcli network: List active connections or list active ARP table entries.

Storage

• NFS statistics available in resxtop

Virtual Machine

- esxcli vms: Forcibly stop virtual machines that do not respond to normal stop operations, by using kill commands.
- NOTE: designed to kill VMs in a reliable way (not dependent upon wellbehaving system)

Additional commands for troubleshooting

- vscsiStats
- nc (netcat)
- tcpdump-uw



Ability to totally control local access via vCenter Server

- DCUI
- Lockdown Mode (disallows all access except root on DCUI)
- Tech Support Mode (local and remote)
- If all configured, then **no local activity possible** (except pull the plugs)

Security Profile		
Services	Refresh	Properties
I/O Redirector (Active Directory		
Network Login Server (Active Dir		
lbtd		
Local Tech Support		
Local Security Authentication Ser		
NTP Daemon		
VMware vCenter Agent		
Remote Tech Support (SSH)		
Direct Console UI		
Lockdown Mode		Edit

Lockdown Mode:

Disabled

Deployment Options

- Boot from SAN
- Scripted Installation (a la "Kickstart")

Centralized updating of 3rd party code

• VMware Update Manager can deploy drivers, CIM providers, other modules

Improved Local Authentication

- Built-in Active Directory Service
- DCUI and Tech Support Mode access by any authorized user (not just root)

Easier CLI options for troubleshooting

- Full support of Tech Support Mode both local and remote (via SSH)
- Additional commands in Tech Support Mode: vscsiStats, nc, tcpdump-uw, etc.
- Additional management options in vCLI: SCSI, VAAI, Network, VM

Better control over local activity

- DCUI and Tech Support Mode is configurable in vCenter Server
- Total host lockdown possible
- Activity in Tech Support Mode is sent to syslog

Deployment Options

- Boot from SAN
- Scripted Installation (a la "Kickstart")

Centralized updating of 3rd party code

• VMware Update Manager can deploy drivers, CIM providers, other modules

Improved Local Authentication

- Built-in Active Directory Service
- DCUI and Tech Support Mode access by any authorized user (not just root)

Easier CLI options for troubleshooting

- Full support of Tech Support Mode both local and remote (via SSH)
- Additional commands in Tech Support Mode: vscsiStats, nc, tcpdump-uw, etc.
- Additional management options in vCLI: SCSI, VAAI, Network, VM

Better control over local activity

- DCUI and Tech Support Mode is configurable in vCenter Server
- Total host lockdown possible
- Activity in Tech Support Mode is sent to syslog

Availability – HA & FT DRS & DPM



Confidential

Availability Feature Summary

HA and DRS Cluster Limitations

High Availability (HA) Diagnostic and Reliability Improvements

Fault Tolerance (FT) Enhancements

vMotion Enhancements

- Performance
- Usability
- Enhanced Feature Compatibility

DRS Host Affinity

DPM Enhancements

Data Recovery Enhancements

Increased cluster limitations

- Cluster limits are now unified for HA and DRS clusters
- Increased limits for VMs/host and VMs/cluster
- Cluster limits for HA and DRS:
 - 32 hosts/cluster
 - 320 VMs/host (regardless of # of hosts/cluster)
 - 3000 VMs/cluster
- Note that these limits also apply to post-failover scenarios. Please be sure that these limits will not be violated even after the maximum configured number of host failovers.

HA and DRS Cluster Limitations



5-host cluster, tolerate 1 host failure

- vSphere 4.1 supports 320 VMs/host
- Supports 320x5 VMs/cluster? NO
- Cluster can only support 320x4 VMs

5-host cluster, tolerate 2 host failures

- Supports 320x5 VMs/cluster? NO
- Cluster can only support 320x3 VMs





HA Healthcheck Status

 HA provides an ongoing healthcheck facility to ensure that the required cluster configuration is met at all times. Deviations result in an event or alarm on the cluster.

HA Operational Status

 A new Cluster Operational Status window displays more information about the current HA operational status, including the specific status and errors for each host in the HA cluster.

VMware HA	
Admission Control: Current Failover Capacity:	Enabled 2 hosts
Configured Failover Capacity:	1 host
Host Monitoring: VM Monitoring: Application Monitoring:	Enabled Disabled Disabled
Advanced Runtime Info Cluster Operational Status	

Improved HA-DRS interoperability during HA failover

 DRS will perform vMotion to free up contiguous resources (i.e. on one host) so that HA can place a VM that needs to be restarted

HA app awareness – expose APIs for 3rd party app developers

Fault Tolerance (FT) Enhancements



FT fully integrated with DRS

 DRS load balances FT Primary and Secondary VMs. EVC required.

Versioning control lifts requirement on ESX build consistency

 Primary VM can run on host with a different build # as Secondary VM.

Events for Primary VM vs. Secondary VM differentiated

• Events logged/stored differently.



vMotion Enhancements

- Significantly decreased the overall migration time (time will vary depending on workload)
- Increased number of concurrent vMotions:
 - ESX host: 4 on a 1 Gbps network and 8 on a 10 Gbps network
 - Datastore: 128 (both VMFS and NFS)

otatus	Details
In Progress	s
10% 🔳	Migrating the active state of Virtual Machine
12% 🔳	Migrating the active state of Virtual Machine
36% 💻	Migrating the active state of Virtual Machine
35% 🛄	Migrating the active state of Virtual Machine
35% 💻	Migrating the active state of Virtual Machine
37% 💻	Migrating the active state of Virtual Machine
61% 💶	Migrating the active state of Virtual Machine
61%	Migrating the active state of Virtual Machine

Maintenance mode evacuation time is greatly decreased due to above improvements

C	Change EVC Mode
	◯ Disable EVC
1	vMware EVC Mode: AMD Opteron™ Gen. 3 w/o "3DNow!"
	Description
	Applies the baseline feature set of AMD Opteron™ Generation 3 ("Greyhound")
	processors, with "3DNow!" support removed, to all hosts in the cluster.
	Hosts with the following processor types will be permitted to enter the cluster:
	future AMD processors
	Compared to the AMD Opteron™ Generation 2 EVC mode, this EVC mode exposes
	additional CPU features including SSE4A, MisAlignSSE, POPCOUNT and ABM (LZCNT).
	For more information, see Knowledge Base article 1003212.

Usability Improvements

- Preparation for AMD Next Generation w/o 3DNow!: Future AMD CPUs may not support 3DNow!. To prevent vMotion incompatibilities, a new EVC mode is introduced.
- Better handling of powered-on VMs: vCenter Server now uses a live VM's CPU feature set (instead of host's CPU features) to determine migration into an EVC cluster. This will provide better granularity in error detection.
Usability Improvements

• VM's EVC capability: The Virtual Machines tab for hosts and clusters now displays the EVC mode corresponding to the features used by VMs.

	Getting Started Summary Virtual Machines Performance Configuration Tasks & Events Alarms Permissions Maps Storage Views Hardware Status								
l	Name, State or Guest OS contains: -							Clear	
l	Name	State	Status	Provisioned Space	Used Space	Host CPU - MHz	Host Mem - MB	Guest Mem - %	EVC Mode
	ig4-SLES11-64	Powered On	📀 Normal	8.50 GB	8.50 GB	120	332	75	Intel® Xeon® Core™2

• VM Summary: The Summary tab for a VM lists the EVC mode corresponding to the features used by the VM.

General		
Guest OS:	Suse Linux Enterprise 11 (64-bit)	
VM Version:	7	
CPU:	4 vCPU	
Memory:	512 MB	
Memory Overhead:	237.37 MB	
VMware Tools:	Out of date	
IP Addresses:	10.115.161.77	View al
DNS Name:	prme-evc-dhcp77	
EVC Mode:	Intel® Xeon® Core™2	
State:	Powered On	
Host:	intel-merom-1.eng.vmware.com	
Active Tasks:		

• Earlier Add-Host Error detection: Host-specific incompatibilities are now displayed prior to the Add-Host work-flow when adding a host into an EVC cluster.

Rule	×				
Rule	DRS Groups Manager				
Give ti Then,	he new rule a name and choose its type from the menu below. select the entities to which this rule will apply.				
Name					
Tures	'				
Virtual Machines to Hosts					
	Groups				
Clust	er Vm Group:				
Must	run on hosts in group				
Cluster Host Group:					
1	Select Cluster Vm Group				

Required rules

- Must run on hosts in group
- Should run on hosts in group
 Must Not run on hosts in group
 Should Not run on hosts in group

Preferential rules

Rule enforcement: 2 options

- Required: DRS/HA will never violate the rule; event generated if violated manually. Only advised for enforcing host-based licensing of ISV apps.
- Preferential: DRS/HA will violate the rule if necessary for failover or for maintaining availability



VM-VM anti-affinity rule enhancement

Rules

VM-VM anti-affinity rules can now incorporate more than 2 VMs

Confidential

DPM Enhancements

Scheduling Distributed Power Management

- Turning on/off DPM is now a scheduled task
- DPM can be turned off prior to business hours in anticipation for higher resource demands



Disabling DPM brings hosts out of standby

 Eliminates risk of ESX hosts being stuck in standby mode while DPM is disabled. Ensures that when DPM is disabled, all hosts are powered on and ready to accommodate load increases.

VMware Data Recovery: New Capabilities



Backup and Recovery Appliance

- Support for up to 10 appliances per vCenter Server instance to allow protection of up to 1000 virtual machines
- File Level Restore client for Linux VMs

VMware vSphere 4.1

- Improved VSS support for Windows 2008 and Windows 7: application level quiescing

Destination Storage

- Expanded support for DAS, NFS, iSCSI or Fibre Channel storage plus CIFS shares as destination
- Improved deduplication performance

vSphere Client Plug-In

- Ability for seamless switch between multiple backup appliances
- Improved usability and user experience

Management



Confidential

vCenter Server

- 32-bit to 64-bit data migration
- Enhanced Scalability

Update Manager

Host Profile Enhancements

Orchestrator

Active Directory Support (Host and vMA)

Management Assistant (vMA)

Scale and readiness

Converter

• Hyper-V Import

Virtual Serial Port Concentrator (VSPC)

vCenter Server – Migration to 64-bit

vCenter Server MUST be hosted on 64-bit Windows OS

• 32-bit OS NOT supported as a host OS with vCenter vSphere 4.1

Why the change?

- Scalability is restricted by the x86 32 bit virtual address space and moving to 64 bit will eliminate this problem
- Reduces dev and QA cycles and resources (faster time to market)

Two Options

- 1. vCenter Server in a virtual machine running 64-bit Windows OS
- 2. vCenter Server install on a 64-bit Windows OS

Best Practice – Use Option 1

	vSphere 4	vSphere 4.1	Ratio
VMs per host	320	320	1x
Hosts per cluster	32	32	1x
VMs per cluster	1280	3000	3x
Hosts per VC	300	1000	3 x
Registered VMs per VC	4500	15000	3x+
Powered-On VMs per VC	3000	10000	3 x
Concurrent VI Clients	30	120	4x
Hosts per DC	100	500	5x
VMs per DC	2500	5000	2x

Update Manager

- vSphere Client						
tion Plug-ins Help						
entory 🕨 🎁 Hosts and Clusters				67 - Se	arch Inventory	Q
ue-vc-server.eng.vmware.com ¥Mware vCe	nter Server, 4.0.0,	151679				• 🔺
Getting Started Datacenters Virtual Machine	es 🛛 Hosts 🔪 Tasks &	Events Alarms Permissions	Maps Update Mana	iger		- 1
	News seeks			Care Albert Units	A darie Mierry	
		Name concar	ins:	Ciear .	ocari Accacri Help	Admini view
Attached Baseline Groups	<u> </u>	Attached Baselines	~	Туре	Host Compliance	# Hosts
All Groups and Independent Baselines					All Applicable Hosts	11
		Critical Host Patches		Patch	Mon-Compliant	5
	1	Upgrade to 4.0 build 152112		Upgrade	2 Linknown	6
					 Compliant 	o I
					0%	
					Compliant	
					. (
1						
All Groups and Independent Baselines ->	All -> 🗏 All Ap	plicable Hosts				
Host Name	Patch Status	Upgrade Status	Attached Baselines		Last Patch Sca	n
ueteam-esx-05.eng.vmware.com	② Unknown	((1) Critical Host Patche	es	,	~
10.20.134.35	😵 Non-complian	t ((1) Critical Host Patche	es	3/11/2009 3:53	23 PM
ueteam-esx-02.eng.vmware.com	② Unknown	((2) Critical Host Patche	es, Non-Crit	ical H	
ueteam-esx-03.eng.vmware.com	Non-complian	t ((2) Critical Host Patche	es, Non-Crit	ical H 3/11/2009 3:53	23 PM
ender-esx.eng.vmware.com	Non-complian	t 😢 Unknown ((2) Critical Host Patche (1) Critical Host Datche	es, Non-Crit	cal H 3/11/2009 3:5: 3/14/2000 3:5:	223 PM
Contwernissceng, winware.com	Norecompliant	C Compilant ((1) Chucai Host Patche	75	3/11/2009 3.3.	
	(2) Linknown	(Critical Host Patche 	ac.		
uebeam-esx-07.eng.vmware.com	 Unknown Unknown 	(Critical Host Patch. Non-Critical Host P. 	es atches		
ueteam-esx-07.eng.vmware.com	② Unknown③ Unknown	(Critical Host Patch Non-Critical Host P 	atches		
ueteam-esx-07.eng.vmware.com	 Unknown Unknown 	(Critical Host Patch (1) Non-Critical Host P. 	es atches	🚺 Stage 🔌	Remediate
ueteam-esx-07.eng.vmware.com Hide Hosts	 O Unknown O Unknown 	(Critical Host Patchi (1) Non-Critical Host P. 	es atches	🚺 Stage 🔌	Remediate
Useban-esc-07.eng.rmmete.com Hide Hosts Sahus	O Unknown Unknown Details	(((1) Critical Host Patch (1) Non-Critical Host P	es atches	Start Time	Remediate
Hide Hosts	O Unknown Unknown Details	Initiated by vCenter Serve	(1) Critical Host Patchi (1) Non-Critical Host P r Requested	es atches Start Ti 11:27:41	Stage Start Time 3/27/2009 11:27:41	Remediate
Hide Hosts Status Status Completed Completed Completed	O Unknown Unknown Details	Initiated by vCenter Serve System	(1) Critical Host Patchi (1) Non-Critical Host P # Requested #ver.e 3/27/2009	es atches Start Ti 11:27:41	Start Time	Remediate Completed T . 3/27/2009 1
Hide Hosts Status Status Completed Completed Completed Completed	 ② Unknown ③ Unknown Oetails 	(Initiated by vCenter Serve System 🕜 ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P rr Requested srver.e 3/27/2009	es atches I Start Ti 11:27:41	Start Time 3/27/2009 11:27:41	Remediate Completed T . 3/27/2009 1
Hide Hosts Status rtin Win XP Completed	 ② Unknown ③ Unknown ② Details 	(Initiated by VCenter Serve System 🕜 ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P r r Requested srver.e 3/27/2009	es atches Start Ti 11:27:41	Stage ▼ Start Time . 3/27/2009 11:27:41	Remediate Completed T . 3/27/2009 1
Hide Hosts Trin Win XP Completed Com	 ② Unknown ③ Unknown ② Details 	Initiated by VCenter Serve System 🕜 ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P r Requested srver.e 3/27/2009	es atches I Start Ti 11:27:41	Stage Start Time 3/27/2009 11:27:41	Remediate Completed 1 . 3/27/2009 1
Hide Hosts Status rdn Wn XP To Completed Pre-remediation Check Summary	 Ø Unknown Ø Unknown 	Initiated by VCenter Serve System 🕜 ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P r r Requested stver.e	es atches Start Ti 11:27:41	Stage (Remediate Completed T . 3/27/2009 1
totames.cd?.eng.vmware.com Hide Hosts Status rtin Wn XP Completed Pre-remediation Check Summary	Ø Unknown Ø Unknown Details	() () Initiated by VCenter Serve System 🔐 ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P rr Requested srver.e 3/27/2009	es atches Start Ti 11:27:41	Stage Start Time . 3/27/2009 11:27:41	Remediate Completed T . 3/27/2009 1
	Ø Unknown Ø Unknown Details	Initiated by VCenter Serve System P us-vc-se Other or different sur	(1) Critical Host Patch (1) Non-Critical Host P r r Requested stver.e	95 atches Start Ti 11:27:41	Stage Stage Start Time 3/27/2009 11:27:41	Remediate Completed T . 3/27/2009 1
teleamershift of any immediate in a status Tenemediate in a status Clusters to remediate: 10	Ø Unknown Ø Unknown Details	Initiated by VCenter Serve System @ ue-vc-se	(1) Critical Host Patch (1) Non-Critical Host P rr Requested rver.e	95 atches Start Ti 11:27:41		Remediate X Completed T . 3/27/2009 1
Status Status Completed Pre-remediation Check Summary Clusters to remediate: 2 Hosts to remediate: 10	Ø Unknown Ø Unknown Details	Initiated by VCenter Serve System Provide Universe Other or different sur	r Cricial Host Patch r Requested r 3/27/2009 mmary info?	es atches	Stage Start Time . 3/27/2009 11:27:41	Remediate X Completed T . 3/27/2009 1

Cluster	Host	VM	Problem	Resolution
VIM cluster	10.20.135.35	erider-www	A removable device is attached.	Disconnect removable device.
VIM cluster	ueteam-esx	Test VM	HA admission control policy will be violated.	Disable HA admission control.
My Cluster			DPM is enabled.	Disable DPM on the cluster.
My Cluster			EVC is disabled.	Enable EVC on the cluster.

- Central automated, actionable VI patch compliance management solution
- Define, track, and enforce software update compliance for ESX hosts/clusters, 3rd party ESX extensions, Virtual Appliances, VMTools/VM Hardware, online*/offline VMs, templates
- Patch notification and recall
- Cluster level pre-remediation check analysis and report
- Framework to support 3rd party IHV/ISV updates, customizations: mass install, /update of EMC's PowerPath module
- Enhanced compatibility with DPM for cluster level patch operations
- Performance and scalability enhancements to match vCenter

Host Profiles

- Cisco support
- PCI device ordering (support for selecting NICs)
- iSCSI support
- Admin password (setting root password)
- PSA configuration

Orchestrator

- provides a client and server for 64-bit installations, with an optional 32-bit client.
- performance enhancements due to 64-bit installation

Management - Other New Features (Continued)

Active Directory Support (Host and vMA)

- No need to manage user accounts on ESX or ESXi "stateless"
- Match Hyper-V capability (it can do this today)

Management Assistant

- Improved authentication capability Active Directory support
- Transition from RHEL to CentOS

Converter

Allows import of VMs from Hyper-V host

Virtual Serial Port Concentrator

Traditional low-bandwidth, secure remote console approach for managing servers



Thank You



© 2009 VMware Inc. All rights reserved