Disclaimer

• This presentation may contain product features that are currently under development.
• This overview of new technology represents no commitment from VMware to deliver these features in any generally available product.
• Features are subject to change, and must not be included in contracts, purchase orders, or sales agreements of any kind.
• Technical feasibility and market demand will affect final delivery.
• Pricing and packaging for any new technologies or features discussed or presented have not been determined.
High Availability is Part of IT Business Continuity
High Availability is Part of IT Business Continuity

- High Availability
- Disaster Recovery
- Data Protection

vSphere HA
High Availability is Part of IT Business Continuity

- vSphere HA
- vSphere FT

High Availability

Disaster Recovery

Data Protection
Agenda

1. What's new
2. Failure Events
3. Best Practices
   - Networking and Storage
   - HA and VSAN
   - Host Isolation Response
   - Admission Control
4. Tech Previews
vSphere HA – What’s New in 5.5

• Protection for VSAN VMs
• AppHA Integration
• VM-VM Anti-affinity rule
vSphere HA Recap

- vSphere HA minimizes unplanned downtime
- Provides automatic VM recovery in minutes
- Protects against 3 types of failures

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Connectivity</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host failures</td>
<td>Host network isolated</td>
<td>GuestOS hangs/crashes</td>
</tr>
<tr>
<td>VM crashes</td>
<td>Datastore incurs PDL</td>
<td>Application hangs/crashes</td>
</tr>
</tbody>
</table>

- Does not require complex configuration changes
- OS and application-independent
HA Cluster

• Cluster of ESXi hosts
  – One of the hosts is elected as master

• Heartbeats via network and storage to communicate availability

• HA Network i.e. network used by HA agents
  – Management network (or)
  – VSAN network (if VSAN is enabled)
Failure Events
Host Failure
Host Failure

Master declares slave host dead
Host Failure

New master elected and resumes master duties
Network Partition
Host Isolation
Best Practices

Networking and Storage

HA and VSAN
Host Isolation Response
Admission Control
Networking Recommendations

- Redundant HA Network
- Fewest hops possible
- Consistent portgroup names, network labels
- Route based on originating port ID
- Failback policy = No
- Enable PortFast, Edge, etc.
- MTU size the same
Networking Recommendations

- Disable Host Monitoring if network maintenance
  - Reconfigure HA on cluster after network maintenance
- vmknics for vSphere HA on separate subnets
- Specify additional network isolation address
  - Use HA advanced options
- Each host can communicate with all other hosts

Keep things simple
Storage Recommendations

• Storage Heartbeats
  – All hosts in cluster should see the same datastores

• Choose a heartbeat datastore that is
  – Fault isolated from HA network
  – Resilient to failures

• Override auto-selected datastores if necessary
Best Practices

Networking and Storage
HA and VSAN
Host Isolation Response
Admission Control
### HA and VSAN

<table>
<thead>
<tr>
<th>vSphere HA Network</th>
<th>VSAN Disabled</th>
<th>VSAN Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management network</td>
<td>VSAN network</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heartbeat Datastores</th>
<th>VSAN Disabled</th>
<th>VSAN Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any datastore mounted by &gt;1 host</td>
<td>Only traditional datastores (no VSAN)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Host Isolation</th>
<th>VSAN Disabled</th>
<th>VSAN Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can’t ping isolation addresses, management network inaccessible</td>
<td>Can’t ping isolation addresses, VSAN network inaccessible</td>
<td></td>
</tr>
</tbody>
</table>
HA and VSAN
Heartbeat Datastore Recommendations

• Heartbeat datastores are not necessary in a VSAN cluster
  - However, they are useful in some scenarios

• Add a non-VSAN datastore to cluster hosts if VM MAC address collisions on the VM network are a significant concern
  - In the absence of heartbeat datastores, FDM master will likely restart isolated VMs resulting in two copies of each VM

• Choose a datastore that is fault isolated from VSAN network
HA and VSAN

Host Isolation Address Recommendations

- Isolation Address
  - For example, use the default gateways of the VSAN networks
  - Isolation addresses are set using the HA advanced option `das.isolationAddressX`

- Configure HA to not use the default management network gateway
  - This is done using the HA advanced option `das.useDefaultIsolationAddress=false`

- If isolations and partitions are possible
  - Ensure one set of isolation addresses will be accessible during a partition
HA and VSAN
Host Isolation Address Recommendations (Continued)

• If the VSAN network is non-routable
  – provide pingable isolation addresses on the VSAN subnet
  – use (subset of VSAN network) IP addresses of cluster hosts as isolation addresses

• Each VSAN network should be on unique subnet
  – Using the same subnet for two VMkernel networks can cause unexpected results
  – For example, vSphere HA may fail to detect VSAN network isolation events

• More details here:
Best Practices

Networking and Storage
HA and VSAN
Host Isolation Response
Admission Control
Determining HA Host Isolation

Connected Slave

Lost connection to master

Election

Cannot talk to other HA agents and ping isolation address

Declares itself as Isolated

Datastore heartbeats

Applies isolation response to VMs

Notifies Master
Host Isolation Response

• To delay response in HA 5.1+, use das.config.fdm.isloationPolicyDelaySec

• Isolation responses
  – Leave Powered On (default with 5.x)
  – Shutdown (default with 4.x)
  – Power Off

Q: Which one should you use? It depends…
## Isolation Response Setting: Primary Decision Inputs

<table>
<thead>
<tr>
<th>Applies to which VMs</th>
<th>Host will likely retain access to a VM storage?</th>
<th>VMs will likely retain access to VM network?</th>
</tr>
</thead>
</table>

- In a cluster with VSAN and Non-VSAN VMs
  - Evaluate for each type of VM
    - VSAN VMs will lose access to their storage on a host isolation
  - Use per-VM overrides along with cluster defaults if needed
Isolation Response Setting: Case 1 – VMs are Fine

<table>
<thead>
<tr>
<th>Applies to which VMs</th>
<th>Host will likely retain access to a VM storage?</th>
<th>VMs will likely retain access to VM network?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Recommendation

- Use **Leave powered on**
- VM is running fine. Why power it off?
Isolation Response Setting: Case 2 – Network Is Important

<table>
<thead>
<tr>
<th>Applies to which VMs</th>
<th>Host will likely retain access to a VM storage?</th>
<th>VMs will likely retain access to VM network?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Recommendation

- **Shutdown** if VM network access is important. Otherwise, **Leave Power On**
- Shutdown allows HA master agent to restart the VM

vmworld’2014
Isolation Response Setting: Case 3 – Total Isolation

<table>
<thead>
<tr>
<th>Applies to which VMs</th>
<th>Host will likely retain access to a VM storage?</th>
<th>VMs will likely retain access to VM network?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

If an isolation event will likely affect all hosts when it occurs

• Recommendation: **Leave Power On**
• HA master agent will not restart any VMs in this situation
Isolation Response Setting: Cases 4 and 5 – Loss of VM Storage

<table>
<thead>
<tr>
<th>Applies to which VMs</th>
<th>Host will likely retain access to a VM storage?</th>
<th>VMs will likely retain access to VM network?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Non-VSAN</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VSAN and non-VSAN</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Must consider other factors when decision on the setting to use
Isolation Response Setting: Cases 4 and 5 – Loss of VM Storage

1. Is HB Datastore accessible?
   - NO
   - YES

2. Is Memory State Important?
   - NO
   - YES

3. Power Off (or) Leave Powered On
   - YES
   - NO

4. Leave Powered On
5. VMs retain access to VM network?
   - NO
   - YES

6. Power Off

First VM instance could cause network conflicts

Master cannot start VM unless powered off
Best Practices

Networking and Storage
HA and VSAN
Host Isolation Response
Admission Control
vSphere HA Admission Control

• You can reserve resources in case of host failures
• Ensures resources are available to restart VMs
  – Satisfy reservations and memory overhead
• No guarantee that the VMs perform well after a failure
• Work in progress to close this gap
  – Fling for capacity planning and impact assessment 😊
    • see Tech Preview/Demo
  – Group Discussion BCO3430-GD
How to do Admission Control

• Select the appropriate admission control policy
• Enable DRS to maximize likelihood that VM resource demands are met
• Simulate failures to test and assess performance
  – Use maintenance mode
  – Use the impact assessment fling?
• Make adjustments if
  – VMs are not restarted
  – Desired performance is not realized
Reducing Performance Impediments

• Maximize utility of the remaining (healthy) hosts
  – Enable DRS in automatic mode

• Maximize the hosts a given VM can run on
  – For example, limit the use of VM to Host “required” affinity rules

• Ensure sufficient resources to meet VM demand
  – Move some VMs to another cluster or add hosts

• Ensure critical VMs get the resources they need
  – Adjust VM shares, reservations, and HA restart priorities
vSphere HA Admission Policies

1. Percentage of Cluster Resources
2. Number of Hosts
3. Dedicated Failover Hosts

Next
- Policy details
- When useful
- Recommendations
Admission Control Policy

1. Percentage of cluster resources

- Define failover capacity by reserving a percentage of the cluster resources.
- Reserved failover CPU capacity: 25 % CPU
- Reserved failover Memory capacity: 25 % Memory

- vSphere HA
  - Protected
  - CPU Failover Threshold: 25 %
  - Memory Failover Threshold: 25 %
Admission Control Policy Recommendations

1. Percentage of cluster resources

• Often the best choice
• Maximizes use of cluster resources prior to a failure
• Use when reservations vary considerably and/or there are VMs with large reservations
• Recalculate when hosts are added to cluster
  – N+1: 6 hosts \(\rightarrow\) 1/6 (17%); 10 hosts \(\rightarrow\) 1/10 (10%)
Admission Control Policy

2. Number of Hosts

<table>
<thead>
<tr>
<th>Advanced Runtime Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slot size</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total slots in cluster</td>
</tr>
<tr>
<td>Used slots</td>
</tr>
<tr>
<td>Available slots</td>
</tr>
<tr>
<td>Failover slots</td>
</tr>
<tr>
<td>Total powered-on virtual</td>
</tr>
<tr>
<td>machines in cluster</td>
</tr>
<tr>
<td>Total hosts in cluster</td>
</tr>
<tr>
<td>Total good hosts in cluster</td>
</tr>
</tbody>
</table>
Admission Control Policy Recommendations

2. Number of Hosts

• Maximizes chance of restarting VMs with reservations
• Avoids fragmentation
• Often the most conservative policy
• If a concern, Failover Host policy may be better
• If you use this policy
  – Let HA calculate settings
  – Use reservations sparingly
Admission Control Policy Recommendations

3. Dedicated failover host(s)

• Best if have VMs with large reservations
• Impact to VMs on other hosts minimized
• However..
  – Restart time can be longer
  – Failover host(s) are idle prior to failure
• If you use this policy…
  – select largest host(s) as failover hosts to ensure all VMs can restart
Tech Previews

FT and HA
Tech Preview 1 – FT

vSphere Availability Portfolio

Coverage

Application

Guest OS

VM

Hardware

App Monitoring APIs

Guest Monitoring

Fault Tolerance

Infrastructure HA

Downtime

none

minutes
A Clean Slate

FT LOGGING

10 GigE

FT protocol
Tech Preview 2 – HA

VM Component Protection – Storage

• Problem:
  – Host has storage-connectivity loss
    • APD: All Paths Down
    • PDL: Permanent Device Loss
  – Difficult to manage VMs running on APD/PDL affected hosts

• Approach:
  – Affected VMs are proactively terminated and restarted on healthy hosts
Tech Preview 2 – HA
VM Component Protection – Storage

VMware ESX
SAN
NFS
VMware ESX

VM
VM
VM
VM
vSphere HA is Turned ON

Routine information for vSphere HA is reported under vSphere HA Monitoring.

Host Monitoring

ESXi hosts in this cluster exchange network heartbeats. Disable this feature when performing network maintenance that may cause isolation responses.

Host Monitoring is Turned ON

Host Hardware Monitoring - VM Component Protection

ESXi hosts have the capability to detect various failures that do not necessarily cause virtual machines to go down, but may deem them unusable (e.g., losing network/disk communication).

Protection against Storage Connectivity Loss is Turned ON

Virtual Machine Monitoring

VM Monitoring restarts individual VMs if their VMware Tools heartbeats are not received within a set timeout. Application Monitoring restarts individual VMs if their in-guest application heartbeats are not received within a set timeout.

VM Monitoring is Turned OFF

- Virtual Machine Response
  - Expand for details
- Admission Control
  - Expand for details
- Datasources for Heartbeats
  - Expand for details
- Advanced Options
  - None
Tech Preview 3 – HA
Admission Control Fling - vRAS

• vRAS – vSphere Resource and Availability Service
• vRAS assess the impact of host failures and VM migration on resources using DRS dump files (which contain cluster snapshot)
• Sample what-if scenarios
  – Host failures in a HA cluster
  – Put hosts into maintenance mode
  – Demand for resources increase – are all VMs still happy?
• Works with vSphere 5.0, 5.1, and 5.5
vRAS - Demo
Admission Control Fling - vRAS

Upload file

Estimate performance Impact

VM Group Details
If a host goes down...
what will happen to your VMs?

Simulate Now

Simulate the failure of hosts in a cluster
See which VMs will be restarted by VMware HA
Understand impact on resource availability
More vSphere HA and FT at VMworld

- VMware BCDR demo booth on show floor
- High Availability Group Discussion – BCO3430-GD
- “Ask the Experts” – all week
- This session repeated tomorrow/Wednesday at 2 PM
Thank You
GS Khalsa – @gurusimran - gkhalsa@vmware.com
Manoj Krishnan – @manojkkkrish - krishnanm@vmware.com
Thank You
Fill out a survey

Every completed survey is entered into a drawing for a $25 VMware company store gift certificate
vSphere HA Best Practices and FT Tech Preview

GS Khalsa, @gurusimran, VMware, Inc
Manoj Krishnan, @manojkkkrish, VMware, Inc
Additional Slides

• Stretch clusters
• HA and VSAN recommendations
• Admission control – Number of hosts explained
• HA and FT tech previews
Stretched-Cluster HA Recommendations

• Use DRS Affinity Group for site awareness
  – Use “Should” Rules (not “Must” rules)
• Avoid VMotion, Storage VMotion across sites
• Set Storage DRS to Manual
• Keep multi-VM applications in the same site
• Set HA Admission Control to 50% CPU, Memory
• Use four HA heartbeat datastores (two per site)
Stretched-Cluster HA Recommendations

• Two isolation addresses – one per site
• Use HA Restart priorities
• No guarantee HA will restart all VMs
• Don’t host vCenter, witness server in cluster
• Test the various failover scenarios
• Credit where credit is due:

http://www.vmware.com/resources/techresources/10299
HA with VSAN – Keep in Mind…

- Minimum cluster size is three hosts
- HA and VSAN use same network
- HA uses port 8182 on VSAN network
- Tag network port groups first
- Isolation address unchanged if vSAN is enabled
- HA does not use VSAN for DS heartbeats
- Keep VM files (.vmx and .vmdk) together
Admission Control

Number of Hosts

• Uses concept of slot sizes
Admission Control

- Number of Hosts: Slot sizes explained
  - No (explicit) CPU and memory reservations
    - 32 MHz, 0 MB memory + memory overhead are used
  - Example slot size = 32 MHz, 49 MB memory
Admission Control

Number of Hosts: Slot sizes explained

Reservation: 2 GHz
1024 MB

Reservation: 1 GHz
2048 MB

Advanced Runtime Info

- Slot size: 2000 MHz
- Total slots in cluster: 12
- Used slots: 2
- Available slots: 7
- Failover slots: 3
- Total powered-on virtual machines in cluster: 2
- Total hosts in cluster: 4
- Total good hosts in cluster: 4
Admission Control

Number of Hosts: Slot sizes explained

**Reservation:**
- 2 GHz
- 1024 MB

**Reservation:**
- 1 GHz
- 2048 MB
Admission Control

Number of Hosts: Slot sizes explained

**Reservation:**
- 2 GHz
- 1024 MB

**Reservation:**
- 1 GHz
- 2048 MB (plus overhead)
Admission Control

Number of Hosts: Slot sizes explained
Admission Control

Number of Hosts: Slot sizes explained

<table>
<thead>
<tr>
<th>Slot size</th>
<th>2000 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total slots in cluster</td>
<td>12</td>
</tr>
<tr>
<td>Used slots</td>
<td>2</td>
</tr>
<tr>
<td>Available slots</td>
<td>7</td>
</tr>
<tr>
<td>Failover slots</td>
<td>3</td>
</tr>
<tr>
<td>Total powered-on virtual</td>
<td>2</td>
</tr>
<tr>
<td>machines in cluster</td>
<td></td>
</tr>
<tr>
<td>Total hosts in cluster</td>
<td>4</td>
</tr>
<tr>
<td>Total good hosts in cluster</td>
<td>4</td>
</tr>
</tbody>
</table>

---

**vmworld 2014**
Admission Control

Number of Hosts: Slot sizes explained

### Advanced Runtime Info

- **Slot size**: 2000 MHz, 2095 MB
- **Total slots in cluster**: 12
- **Used slots**: 2
- **Available slots**: 7
- **Failover slots**: 3
- **Total powered-on virtual machines in cluster**: 2
- **Total hosts in cluster**: 4
- **Total good hosts in cluster**: 4
Admission Control

Number of Hosts: Slot sizes explained

![Diagram showing virtual machines and slot sizes explained in a table.](image)
Admission Control

Number of Hosts: Slot sizes explained

### Advanced Runtime Info

<table>
<thead>
<tr>
<th>Slot size</th>
<th>2000 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2095 MB</td>
</tr>
</tbody>
</table>

- Total slots in cluster: 12
- Used slots: 2
- Available slots: 7
- Failover slots: 3
- Total powered-on virtual machines in cluster: 2
- Total hosts in cluster: 4
- Total good hosts in cluster: 4
Admission Control

Number of Hosts

- Define failover capacity by static number of hosts.
  - Reserved failover capacity: 1 Hosts

Slot size policy:

- Cover all powered-on virtual machines
  - Calculate slot size based on the maximum CPU/Memory reservation and overhead of all powered-on virtual machines.
- Fixed slot size
  - Specify the slot size explicitly.
  - CPU slot size: 200 MHz
  - Memory slot size: 256 MB

VMs requiring multiple slots: 0/13

View Calculate
Admission Control

Number of Hosts

![Fixed slot size](image)

- **CPU slot size:** 3000 MHz
- **Memory slot size:** 4096 MB

![Advanced Runtime Info](image)

- **Slot size:** 3000 MHz, 4096 MB
- **Total slots in cluster:** 8
- **Used slots:** 2
- **Available slots:** 4
- **Failover slots:** 2
- **Total powered-on virtual machines in cluster:** 2
- **Total hosts in cluster:** 4
- **Total good hosts in cluster:** 4
vSphere HA/FT Tech Previews

• Virtual Machine Component Protection (VMCP)
  – Fine-grained controls for VM restart policy
  – Queries destination host(s) for storage health

• SMP Fault Tolerance (FT)
  – Protect VMs that have more than one vCPU
  – Session BCO5065 (Multiprocessor FT Tech Preview)

• Demos at VMware BCDR booth on show floor
Additional Slides

• Extra slides
  – Don’t want to delete them – they might be usable elsewhere…
vSphere App HA

• Planned and unplanned application downtime
Recommendations: Networking
Recommendations: Storage

- Storage Heartbeats
  - HA selects two datastores by default
vSphere 5.1 HA Enhancements

• Auto Deploy integration
• Admission control slot size configurable
• Permanent Device Loss (PDL) and All Paths Down (APD) handling (vSphere 5.0 U1, 5.1, 5.5)
• Application monitoring SDK change
PDL and APD Handling

• `disk.terminateVMOnPDLDefault`
  - Ensures VM is killed when PDL occurs
  - VM killed when it issues I/O

• `das.maskCleanShutdownEnabled`
  - 5.0 U1 default “False” – Recommendation: “True”
  - HA can restart VM killed by PDL
  - VM powered off from APD also restarted
vSphere HA Best Practices and FT Tech Preview

Gurusimran Khalsa, VMware, Inc
Manoj Krishnan, VMware, Inc