vSphere Design Pocketbook

Tweet-Sized Design Considerations for Your Software-Defined Datacenter

Frank Denneman
Duncan Epping
Cormac Hogan
Jason Nash
Vaughn Stewart
And many others...

pernixdata press

vSphere Design Pocketbook

Tweet-Sized Design considerations for your software defined datacenter

vSphere Design Pocketbook

International Standard Book Number (ISBN:)
978-1484920084

Version: 1.0

About the Authors

Frank Denneman is the technical evangelist of PernixData. He has the honor of promoting the awesome technology that is coming out of PernixData.

Duncan Epping is a principle architect at VMware R&D. His main responsibility is exploring new possibilities with existing products and features, researching new business opportunities for VMware through prototyping new solutions or products.

Cormac Hogan is a Senior Technical Marketing Architect at VMware. In this role he focuses on core VMware vSphere storage and all virtual storage features.

Jason Nash is a Director of Datacenter practice at Varrow. This role allows him to enable, evangelize, and develop new products, technologies, and offerings.

Vaughn Stewart is the Director & Cloud Computing Evangelist at NetApp, where he assists in setting the strategic direction for cloud-based solutions – a rewarding role that blends his engineering interests with his enthusiasm for engaging clients, vendors, and leaders in the IT industry.

Foreword

You could have been a fighter pilot. Or a circus clown. But you choose to do virtualization, and here we are!

I was lucky to have started my career creating VMware's first server hypervisor. As I look back, I see we created an industry segment and made history; and if you've been doing this half as long as I have been, I'm sure you get that sense about your own career too. It is not often that technology touches lives in such a significant way. Virtualization has now become a shared context around which we have all created and nurtured ever-lasting human relationships. What we have jointly achieved as a community in just a decade is one for the books, and we should be proud.

This book is a testament to, and a direct consequence of the ties that bind us together as the virtualization community. How else do you explain two soccer fans, two American football fans, and a Gaelic football fan working together? I've had the privilege of knowing these fine gentlemen and most of the contributors featured in this book first-hand...

Frank, Duncan, Cormac, Jason, and Vaughn need no introduction. It is so, because each one of them has given back to the virtualization community in big ways – via blogs,

books, talks, and if you are lucky, over dinner. In this book, they take it to the next level by federating their voices with those of their peers. You'll find a lot of technology gems in the following pages and you'll get to know new people. But don't let this be the end. Find these folks and have follow-up conversations; the book simply scratches the surface of what they know. And use it has a basis to come up with your own big ideas.

Most of all, dear reader, it is the stories we live to tell. Let this book be an inspiration for all of us to do things that would make us a bigger part of the grand narrative that is virtualization.

Satyam Vaghani
CTO, PernixData
San Jose, California

Host Design

Sunny Dua, @sunny_dua
vExpert

Scale-up ESXi Hosts for higher density & consolidation. Scale-out for distributed architecture, which improves the recover time objective.

Josh Odgers, @josh_odgers
VCDX

When choosing hardware for use as an ESXi host, do so with your expected workloads in mind.

Consider NUMA node size, CPU scheduling for larger VMs, a balance between CPU & RAM to avoid underutilization & ensure sufficient I/O & network throughput is available.

Josh Odgers, @josh_odgers
VCDX

To ensure cluster efficiency, the sizing of the ESXi host is essential.

Ensure you have small enough hosts to align with a scale out methodology to maximize DRS (cluster) efficiency while having large enough hosts for your workloads & to get optimal performance.

Yury Magalif, @YuryMagalif
Principal Architect

When sizing and scaling your environment take in to consideration using local SSD drives for host local caching solutions like vFlash Read Cache and PernixData.

Duncan Epping, @DuncanYB
VCDX

When sizing your hosts weigh the cost of many “small hosts” against a few “large hosts”.

Although a few large hosts could be more efficient and lead to reduced TCO, what would be the cost associated when one of those hosts fails?

Sunny Dua, @sunny_dua
vExpert

Size ESXi hosts CPU & RAM keeping following things in mind ~ 20% headroom for peaks, VMkernel overhead, planned downtimes and future growth.

Valentin Bondzio
Escalation Engineer VMware

Make sure to enable "OS Controlled" Power Management in every host's BIOS to grant the ESXi host authority over P-States."

Raymon Epping, @repping
vExpert

When acquiring new hardware investigate integrated management and firmware upgrade capabilities of your hardware platform.

Simplified hardware management will lead to reduced total cost of ownership.

Carel Maritz, @carelmaritz
VCAP5-DCA

Before installing ESXi, download the latest drivers and firmware for your host and all of its components.

Ensure the firmware version is on the VMware HCL and flash your server before the install. Use VUM to install the drivers or streamline them in to your ISO.

Ather Beg, @AtherBeg
vExpert

Before installation, check BIOS for Intel VT-x, AMD-V, EPT, RVI & Hyperthreading etc. Also, disable any unneeded hardware e.g. serial or parallel ports.

The latter will reduce unnecessary interrupts.

Marko Drechsel

Never design without a remote management solution (e.g. iLO).

It could be your live safer when network connection to your ESXi Management VMkernel is lost.

Raymon Epping, @repping
vExpert

Ensure syslog is correctly configured for your virtual infrastructure and log files are offloaded to a safe location outside of your virtual infrastructure.

This will allow the possibility of performing a root cause analysis in case disaster strikes.

Jason Nash, @TheJasonNash
VCDX

Consult with your server vendor on proper memory DIMM size, placement, and type to get optimal performance. While other configurations may work they can greatly impact memory performance.

Andy Daniel, @vNephologist
System Engineer PernixData

When deploying SSD in servers, research read and write performance specifications. Many SSDs are heavily read biased.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

When designing your server platform, take peak network bandwidth requirements into account - not only for the NICs, but also for the PCI Bus.

Eric Sloof, @esloof
VCI

When you are sizing a host, memory and CPU are important but also take a look at network bandwidth and storage bandwidth.

Keep the four host resources well balanced, an imbalance leads to an increase of total cost of ownership.

Matthew Puckett
Virtualization Engineer

For ease of troubleshooting when implementing changes to your environment; back up the host's ESX.conf to allow to do a "diff" of the current vs your backup to easily identify the root cause.

Riccardo Ventura
vExpert

Ensure that you install the ESXi Dump Collector.

This will allow you to configure ESXi hosts to send and save critical diagnostics information after a crash (PSOD), which can help when doing a root cause analysis.

Todd Mace, @mctodd
vExpert

To avoid possible time drifts use NTP for time synchronization on ESXi hosts as NTP synchronizes forward or backwards.

NTP will also compensate for natural clock drift, if time source is unreachable.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

If vCenter is a member of a MS AD, sync the ESXi hosts with NTP to the same time source as the PDC emulator of the AD domain.

This will ensure identical timestamps within the virtual infrastructure.

Josh Coen, @joshcoen
vExpert

Use two single port HBA/CNAs over one dual port HBA/CNA in order to take advantage of multiple PCI buses and increased availability (card failure).

Sachin Bhowan, @sbhowan
VCDX

In Blade environments, consider host placement across connected and stacked enclosures for availability.

Ronny Steiner, @roooooonny
VCAP5-DCD

Ensure that you configure a persistent scratch location for hosts using non-persistent storage (USB/SD-card, Auto Deploy).”

This will help you to upload a coredump to VMware support and finding the root cause.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Ensure that Node Interleaving setting in the BIOS is set to disabled on a NUMA system architecture.”

This allows ESXi to build a System Resource Allocation Table (SRAT) to understand local and remote memory.

Chris Wahl, @ChrisWahl
VCDX

Make sure to size your host with NUMA in mind.

If the workload is not NUMA aware, it may be best to find a CPU with a higher quantity of cores to avoid making a wide VM.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

ESXi Server supports Intel's Hyper-Threading (HT) technology, HT should be enabled as default to offer increase of scheduling options to the ESXi CPU scheduler.

Jason Nash, @TheJasonNash
VCDX

Keep NIC and HBA card placement the same across servers to ensure consistent device naming by ESXi.

Jon Owings, @jon_2vcps
vExpert

Create automated build process to meet security compliance requirements. Manual installations introduce human error.

Ather Beg, @AtherBeg
vExpert

Use consistent naming and make sure all hosts have their FQDNs registered in DNS, with reverse lookups.

This will remove any potential issues due to lookup failures.

Andy Daniel, @vNephologist
System Engineer PernixData

When deploying SSD in servers, proper storage controller configuration is essential to maximize performance.

Cluster Design

Massimiliano Mortillaro, @darkkavenger
vExpert

Always use resource pools to define reservations.

VM reservations are impractical in larger environments, they can be forgotten and they can compete with resource pool reservations if poorly placed.

David Stamen, vExpert
@iamddavee

When deploying business critical highly available applications make sure to create DRS rules to make sure the two highly available virtual machines do not run on the same host.

This configuration avoids a single point of failover

Jonthan Frappier, @jfrappier
vExpert

Create separate compute and storage clusters for management systems to provide resiliency to production infrastructure and reduce resource contention.

Jason Nash, @TheJasonNash
VCDX

When spanning a vSphere cluster across multiple blade chassis and using application clustering consider using DRS affinity rules.

This prevents all nodes in an application cluster from being on the same chassis in case of failure. But use affinity rules sparingly!

Nick Anderson, @speakvirtual
vExpert

Resource Pools are not folders!

Use them for grouping VMs according to the resource priority desired during contention. Any other organizational method can lead to unexpected performance results.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

It's not recommended deploying virtual machines at the same hierarchical level as resource pools.

In this scenario a single virtual machine could receive as many resources as a complete pool of virtual machines in times of contention.

Christian Guerreiro, @tecnoqinteressa

Separate your workloads by using Resource Pools instead of clusters if you have just a few hosts, or if you need more flexibility on maintenance and failure situations.

Mario Mack, @vMario156
vExpert

Ensure to use a DRS Affinity Rule to pin your vCenter Server VM to a small number of ESXi hosts (especially in large clusters).

In case of a vCenter outage you don´t need to search for it in the entire cluster.

Romain Decker, @woueb
VCAP

In a non-converged environment with FC storage, prefer Isolation Response "Leave Powered On" to preserve your workload.

It is unlikely that your network could be unavailable at the same time than FC.

Martijn Baecke, @baecke
VCDX

Enable EVC by default on your vSphere clusters.

This will allow newer processor architecture of the same processor family (Intel or AMD) to be added in the cluster in the future.

Josh Odgers, @josh_odgers
VCDX

When designing for Disaster Recovery, using SRM over vMSC or Stretched Clusters.

These configurations help simplify the solution & ensure you have the ability to do DR testing, including flexibility to define VM startup orders / dependancies & support greater distances than vMSC solutions.

Sunny Dua, @sunny_dua
vExpert

Create Separate ESXi Cluster for IT Applications (Management) and Business Applications to ensure Security, Isolation, and Operational Efficiency.

Kenneth van Ditmarsch, @VirtualKenneth
VCDX

When using vSphere Auto Deploy in environments where multiple vSphere Cluster are needed. Consider using the hardware Asset Tag to group (categorize) the ESXi Hosts to limit the number of rule set patterns and ease administration.

Yury Magalif, @YuryMagalif
Principal Architect

Size VMs per their present workload, not the future workload, because resources are easy to change.

Eelco Kos, @eelcokos
VCP

When working with blades don't put all you hosts of a single vSphere cluster in one enclosure.

You can avoid creating a SPOF and increase redundancy by placing your blades in multiple enclosures.

Sachin Bhowan, @sbhowan
VCDX

Use auto deploy for resource clusters and not management cluster.

Auto deploy depends on vCenter Server and the Auto-Deploy server to be available before the Management Cluster can come-up (chicken and egg scenario!).

Josh Odgers, @josh_odgers
VCDX

Where very high availability and/or non-disruptive VM migration between datacenters is required use vMSC.

vMSC can provide excellent disaster avoidance, reduce the licensing costs (as SRM is not used) & tolerate a catastrophic loss of storage at either site without downtime.

Duncan Epping, @DuncanYB
VCDX

Configuring restart priority of a VM is not a guarantee that VMs will actually be restarted in this order.

Ensure proper operational procedures are in place for restarting services or VMs in the appropriate order in the event of a failure.

Ronny Steiner, @roooooonny
VCAP5-DCD

When deploying a stretched cluster environment, deploy your hosts always with equal numbers and size across both datacenters.

Make sure the overall load does not increase above 50% usage to maintain availability during a full site failover.

Duncan Epping, @DuncanYB
VCDX

For your "Management Network" portgroup ensure to combine different physical NICs connected to different physical switches.

This will increase resiliency and decrease chances of an HA false positive.

Ronny Steiner, @roooooonny
VCAP5-DCD

When deploying a stretched cluster environment, make sure to enable permanent device loss detection on all hosts in the cluster to ensure a VM is killed during a PDL condition.

On top, enable HA advanced setting "das.maskCleanShutdownEnabled" to make sure a killed VM will be restarted on another host after a PDL condition occurred.

Duncan Epping, @DuncanYB
VCDX

In a metro-cluster / geographically dispersed cluster the number of vSphere HA heartbeat datastores is set to four.

Manually select site local datastores, two for each site, to maintain heartbeating even when sites are isolated.

Jason Nash, @TheJasonNash
VCDX

If using DPM and WoL remember that hosts are contacted on their vMotion interfaces so the NICs associated with vMotion must support WoL and must be part of the same layer 2 domain.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

When designing a Cluster, take the vMotion bandwidth requirement into account.

By providing enough bandwidth, the cluster can reach a balanced state more quickly, resulting in better resource allocation (performance) for the VMs.

vCenter Design

Sunny Dua, @sunny_dua
vExpert

Size vCenter database appropriately to avoid performance issues.

The built in SQL Express database should not be used in production. It has a limit to support 5 hosts and 50 virtual machines.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Do not disable DRS or use Host-VM affinity rules to reduce movements of vCenter.

Just document the datastore vCenter is located. When using Storage DRS, set SDRS automation level to manual for vCenter VM.

Paul Gifford, @cloudcanuck
Senior Systems Engineer

With vSphere 5.1 deploy Single Sign-On in a multisite configuration when vCenter Server Linked Mode is a requirement.

Duncan Epping, @DuncanYB
VCDX

Understand the impact of virtualizing vCenter. Ensure it has high priority for restarts and ensure that services which vCenter Server depends on are available: DNS, AD and database.

Josh Odgers, @josh_odgers
VCDX

When SRM is used for DR, running SRM on a separate Windows instance to vCenter ensures that installation/maintenance of or issues with SRM including Storage Replication adapters (SRAs) can be done without impacting the management of the vSphere environment.

Romain Decker, @woueb
VCAP

Try to define configuration models for your Virtual Machines (CPU, memory, etc.) as much as possible.

This will decrease management overhead and improve your capacity planning.

Sunny Dua, @sunny_dua
vExpert

With stateless hosts & management network on DVS, vCenter Server should have the highest HA Restart Policy to ensure it brings up the stateless deployment infrastructures for VMs to be powered on.

Raymon Epping, @repping
vExpert

When designing for availability it is recommended to use a separate management cluster for your vCenter instance and components like SSO, Web Client and Inventory Service.

This to avoid dependency on the platform you are managing.

Martijn Baecke, @baecke
VCDX

Limit the use of snapshots in VMware vSphere. If you want to use them only use them for short-term usage.

Sachin Bhowan, @sbhowan
VCDX

Consider localized database per functional server (vCenter, VRMS, VUM, vCloud, SSO) for availability, management and operational flexibility when upgrading.

Paul Gifford, @cloudcanuck
Senior Systems Engineer

With vSphere 5.1 deploy Single Sign-On in a multisite configuration when vCenter Server Linked Mode is a requirement.

Bonnie Bauder, @bonniebauder
vExpert

For simplicity, use folders in the VMs and Templates view to set permissions rather than at a per VM level for large environments.

Romain Decker, @woueb
VCAP

If your vCenter Server is hosted with other Virtual Machines, set CPU and Memory shares to High.

This will help providing vCenter Server the resources it needs in times of contention.

Josh Odgers, @josh_odgers
VCDX

Where possible in large enterprise environments, keep vCenter as Vanilla as possible by scaling out rather than having all roles such as Inventory, SSO, VUM, SRM on the one Windows instance.

Each role has different availability requirements & scaling out will increase resiliency & simplify maintenance.

Sachin Bhowan, @sbhowan
VCDX

Note vCenter Resource Sizing (vCPU,RAM, connected clients) when adding additional management components (SRM, vCOPS, VUM).

Martijn Baecke, @baecke
VCDX

DNS need to be setup correctly on every vSphere infrastructure component. DNS should be able to answer every request : forward, reverse, short and long.

Eelco Kos, @eelcokos
VCP

Run vCenter as a VM to make use of the advantages that virtualization brings!

Advantages such as: High availability, flexibility and simple VM hardware changes

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Review how fast you can recover your vCenter. Besides determining the RPO and RTO but also the MTTR (Mean Time to Repair).

How will you manage your environment between the point at which the failure is first discovered until the point at which the vCenter returns to operation?

Kenneth van Ditmarsch, @VirtualKenneth
VCDX

Understand the impact of multiple idle sessions within vCenter.

Design management scripts that help you optimize vCenter performance by killing idle sessions.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

When sizing vCenter take into account the number of clusters and virtual machines.

DRS migration calculations for a large number of virtual machines can impact the performance of vCenter.

Duncan Epping, @DuncanYB
VCDX

When the ability to migrate between physical datacenters is required plan accordingly.

Your Datacenter object is the boundary for migrations and dvSwitch connectivity. Multiple clusters under a single Datacenter maybe preferred.

Yury Magalif, @YuryMagalif
Principal Architect

Before you install vCenter SSO, Web Client, or vCenter Server appliance, verify all machines have their clocks synchronized this will avoid unpredictable results during the installation and configuration.

Raymon Epping, @repping
vExpert

When deploying vCenter Server ensure to configure the “Managed IP Address”.

This address is used by many vCenter add-ons, like vSphere Replication, and a missing entry will lead to errors during installation and configuration.

Abdullah Abdullah, @do0dzZZ
vExpert

Install the VSA cluster service outside of the cluster so that it can monitor the health correctly and there is no circular dependency.

Michael Montague, @phatmike128
VCP

Snapshots are not backups!

Ensure that appropriate vCenter alarms are configured to monitor VM snapshots. In a large environment with multiple VM administrators' snapshots can be troublesome if not managed correctly.

Duncan Epping, @DuncanYB
VCDX

When designing for availability vCenter Heartbeat should be considered.

If you consider vCenter Heartbeat make the distinction between uptime requirements for your service, workloads and management tools from an SLA perspective.

Ather Beg, @AtherBeg
vExpert

In addition to a domain account, also have a local account added to the vi-admins local group on the vCenter server.

This allows login if domain services are not available for some reason.

Raymon Epping, @repping
vExpert

Backup vCenter Server on a regular basis and test the recovery procedure.

Loss of vCenter Server could lead to loss of resource pools, distributed switches and difficulty to recovery in a disaster scenario.

Jason Nash, @TheJasonNash
VCDX

vCenter and its related components need to be treated as a Tier 1 application and considered in backup, DR, and BCP planning.

Network & Security Design

Jason Nash, @TheJasonNash
VCDX

Given the recovery features in later versions of the VDS (such as Rollback) there is no need to keep Management or other traffic types on a standard vSwitch.

Your environment can be 100% Distributed Switch.

Prasenjit Sarkar, @stretchcloud
vExpert

Use a NIC that supports TCP checksum offload, Jumbo Frames, and NetQueue to achieve better performance.

Duncan Epping, @DuncanYB
VCDX

When designing limits and shares with Network IO Control realize that limits apply to a NIC pair and shares apply to a NIC Port!

As such configured values should be based on those limitations.

Robert Edwards, @bobbygedwards
vExpert

Separate the networks for vSphere management, virtual machine connectivity, storage and vMotion traffic. Even if it is just logical separation using VLANs.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

If possible select "route based on Physical NIC load" load-balancing policy.

The option "Route based on physical NIC load" takes the virtual machine network I/O load into account and tries to avoid congestion by dynamically reassigning and balancing the virtual switch port to physical NIC mappings.

Cormac Hogan, @vmwarestorage
vExpert

A common issue with jumbo-frame configurations is that the MTU value on the switch isn't set correctly.

In most cases, this must be higher than the hosts and storage that are typically set to 9,000. Switches must be set higher, for example, to 9,198 or 9,216, to account for IP overhead.

Sunny Dua, @sunny_dua
vExpert

Use virtual infrastructure such as vShield App to segregate DMZ traffic from Production Network, if allowed by your security and regulatory policies.

This is more efficient than using separate hosts.

Simon Eady, @simoneady
Bristol VMUG Leader

When using 1Gb Ethernet you should have at least one addition network card to complement the onboard NICs for redundancy and spread teams across them.

Jon Owings, @jon_2vcps
vExpert

When using Distributed Virtual Switching set the NFS/iSCSI and vMotion traffic on separate active uplinks.

Use this with NIOC to not only get bandwidth control in the host but also in upstream switching.

Cedric Megroz, @cmegroz
vExpert

To simplify anti-malware security in your VMs it is recommend using a vShield Endpoint solution.

These are agentless solutions that monitor I/O at the hypervisor and are available from many common anti-malware vendors.

Romain Decker, @woueb
VCAP

In your network configuration it is recommended to set Failback to No for the “Management Network".

This way the vSphere host won't think management is up when a port is transitioning through Spanning-Tree modes or is flapping.

Alexandro Sousa
VCP

Due to large server configurations available today and also Storage vMotion capabilities, is important to consider the use of multiple NICs to handle Maintenance Mode vMotion traffic.

Riccardo Ventura
vExpert

If using multiple dual/quad 1Gb network cards in your sever, separate the uplinks for each vSwitch on different cards.

If a card goes dead, none of your services go down (e.g. iSCSI, VM Network, etc.)

Josh Odgers, @josh_odgers
VCDX

When using IP Storage, configuring multiple das.isolationaddress with the IP of your NFS/iSCSI server & setting Isolation response to Shutdown.

This ensures if an isolation event is detected VMs are Shutdown to allow HA to recover them in a timely manner.

Jason Nash, @TheJasonNash
VCDX

Use Active/Unused uplink configurations on your port-groups if you need to physically separate traffic when using the VDS.

Jason Nash, @TheJasonNash
VCDX

Set the switch ports that connect to your vSphere hosts to enable "portfast" or be in "edge" mode so that they are not temporarily blocked by Spanning-Tree during maintenance or reconnection.

Romain Decker, @woueb
VCAP

Do not use Jumbo Frames (for IP Storage, or vMotion) if there is no direct requirement.

The gains are nominal for the complexity introduced.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

When designing Multi-NIC vMotion configuration for the cluster, all VMkernel interfaces participating in the vMotion must have the IP address from the same IP subnet.

This could have impact on current configurations if not enough IP-addresses are available in current VLAN.

Marko Drechsel

Define accurate your requirements, even more if somebody else is responsible for the network design.

Define Definitions. A "trunk" in Cisco world is something else than a "trunk" in HP world.

Eric Sloof, @esloof
VCI

When a host is equipped with two 10GB network adaptors, use a distributed switch with network IO Control and configure limits and shares for the different types of traffic used.

Gabe Rosas, @networkdojo
Network Virtualization Engineer at VMware vCHS

Consider deploying Cluster VM monitoring for Edge Gateway deployments.

Use vCenter Log Insight to alert on VM resets matched in the FDM log.

Josh Odgers, @josh_odgers
VCDX

For converged environments with IP Storage, configuring Network I/O control so IP Storage has >= highest share value than Network Traffic while vMotion, ESXi Mgmt., FT have <=25% share value of IP Storage will ensure their is minimal performance impact in the event of contention.

Andrea Mauro, @Andrea_Mauro
VCDX

Virtual appliances are nice. But you must thread them as “black-box” (that it’s the concept of an appliance).

So for example you cannot apply specific hardening at appliances OS layer (or application layer) and security must be managed by a right network design.

Theo van Drimmelen, @tdrim
Cloudspecialist

When security is important, use named accounts, complex passwords, only AD accounts, delete accounts for people that left your organisation, and consider using 2 factor-based authentication.

Sachin Bhowan, @sbhowan
VCDX

Consider the back-end switch backplane and packets/per/second specifications when dealing with IP Storage connectivity.

It is possible to overrun the buffers and backplane in many switches and cause drops, which greatly impact storage performance.

Gabe Rosas, @networkdojo
Network Virtualization Engineer at VMware
vCHS

vCNS Edge Gateway HA should be configured with a 15 second Dead Timer and use a dedicated Port-Group for increased HA stability between active and standby edges.

Romain Decker, @woueb
VCAP

Do not use Jumbo Frames (for IP Storage, or vMotion) if there is no direct requirement.

The gains are nominal for the complexity introduced.

Sunny Dua, @sunny_dua
vExpert

Use virtual infrastructure such as vShield App to segregate DMZ traffic from Production Network, if allowed by your security and regulatory policies.

This is more efficient than using separate hosts.

Alessandro Bruno, @alexbindo
System Engineer

vCenter and SQL maintenance may impact VM creation and operations if you are using the VDS. Plan accordingly.

Jason Nash, @TheJasonNash
VCDX

The only type of teaming where you should enable a port-channel on your physical switches is "Router Based on IP Hash".

All other types require no special switch configuration.

Chris Wahl, @ChrisWahl
VCDX

When enabling multi-NIC vMotion, make sure to also use egress traffic shaping or switch level burst control.

This mitigates multi-to-single host vMotion traffic flooding.

Raymon Epping, @repping
vExpert

When designing for availability ensure to have multiple physical NIC ports as part of a single vSwitch.

Preferably of different brands so that in the case of a driver failure network traffic can use the remaining physical NIC.

Ather Beg, @AtherBeg
vExpert

When "Promiscuous Mode" is required for IPS etc., it's best to create a separate vSwitch and relevant policies applied.

It keeps the traffic and policies isolated from the rest.

Duncan Epping, @DuncanYB
VCDX

When using port channels in a virtual environment all NICs are required to be “active” and need to belong to the same port channel.

Active-standby scenarios are not supported and can cause unnecessary downtime.

Jason Nash, @TheJasonNash
VCDX

Set the switch ports that connect to your vSphere hosts to enable "portfast" or be in "edge" mode so that they are not temporarily blocked by Spanning-Tree during maintenance or reconnection.

Josh Odgers, @josh_odgers
VCDX

vMotion & FT traffic is not encypted, as a result it is a security vulnerability.

By using a dedicated non-routable VLAN for vMotion & another for Fault Tolerance ensures the traffic is secure as well as ensuring it is not impacted by other traffic in the broadcast domain.

Todd Mace, @mctodd
vExpert

When using Microsoft NLB, it's a good idea to select "No" as an option for Notifying Switches for VM MAC address changes.

This will prevent flooding the switch with unnecessary traffic and possible vMotion operation failures.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Take VM sizing and Activity into account when sizing your vMotion network.

Having sufficient bandwidth for vMotion will result in faster migration times. This impacts daily operations such as host evacuation or DRS load balancing.

Duncan Epping, @DuncanYB
VCDX

When designing a multi vMotion VMkernel network, which shares NICs with other types of traffic, leverage Network IO Control to avoid a self-inflected Denial of Service attack during a vMotion.

A single vMotion can consume all bandwidth.

Gabe Rosas, @networkdojo
Network Virtualization Engineer at VMware
vCHS

vCNS Edge Gateway HA should be configured with a 15 second Dead Timer and use a dedicated Port-Group for increased HA stability between active and standby edges.

Storage Design

Cormac Hogan, @vmwarestorage
vExpert

A device with an MRU path policy (active/passive array or ALUA array) should be presented on 4 paths on all hosts, 2 paths to the active controller and 2 paths to the passive controller, to avoid path thrashing.

Andreas Lesslhumer, @lessi001
vExpert

Take care that you align the GOS partitions on the virtual hard disks for all your virtual machine.

Doing so avoids performance issues!

Vaughn Stewart, @vStewed
vExpert

If you use VM clones and snapshots, consider storage arrays that support vStorage APIs. Hardware-accelerated clones and snapshots provide greater performance and scalability.

Jon Owings, @jon_2vcps
vExpert

Use your Storage Vendor provided vCenter Plugins for NFS datastore provisioning to easily set vendor best practices settings on Storage and Hosts.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Try to re-format your LUN to VMFS5 rather than upgrade to VMFS5! Upgrading VMFS3 to VMFS5 preserves old block size.

If VMFS3 volume had a blocksize other than 1MB storage vMotion performance between upgraded volume and new VMFS5 volume is severely impacted.

Jon Owings, @jon_2vcps
vExpert

Use SIOC (storage IO control) and Storage DRS to manage the placement and performance of the (monster) VM's.

Todd Mace, @mctodd
vExpert

For effective VMware VAAI functions in your storage design, ensure all datastores have same block size. This will result in faster performance (HW Offload Assistance) and streamlined operations.

Andy Daniel, @vNephologist
System Engineer PernixData

When using an NFS datastore as a persistent scratch or log location, mount it via IP address to avoid a DNS dependency.

Cormac Hogan, @vmwarestorage
vExpert

If increasing the number of NFS volumes mounted on an ESXi host via NFS.MaxVolumes, be sure to increase the TCP/IP Heap Size accordingly. Or else you may run out of TCP/IP heap.

Martijn Baecke, @baecke
VCDX

Isolate IP storage traffic as one does for FC.

This means isolating the NFS & iSCSI traffic on a dedicated VLAN (IEEE 802.1Q) or separate physical switches if VLANs are unavailable.

Josh Odgers, @josh_odgers
VCDX

When designing your VMware solution to use Storage DRS where Auto Tiering is enabled on your disk system OR intelligent caching solutions are used, ensure the I/O metric is DISABLED.

The array will then manage performance and SDRS can manage capacity by initial placement.

Bonnie Bauder, @bonniebauder
vExpert

Use the vSphere client to format your datastores to ensure proper alignment of the VMFS file system.

Duncan Epping, @DuncanYB
VCDX

When designing your storage infrastructure ensure you design for capacity, performance and business continuity.

With business continuity referring to availability and recovery time objective. Large LUNs take a long time to restore!"

Christian Guerreiro, @tecnoqinteressa

Create LUNs/datastores with enough space to fit your bigger VMDK (up to 2TB) so you'll have more flexibility when relocation is needed.

Raymon Epping, @repping
vExpert

When designing for storage performance understand that the distance to storage and RAID type used can have an impact on performance.

Using a virtual storage appliance and local SSDs, or a flash based local caching solution, could lead to increased performance and decreased latency.

Bas Raayman, @basraayman
vExpert

If EtherChannel is not available, run NFS traffic over multiple separate subnets.

This avoids using only one VMkernel port for all NFS traffic, or using the gateway to route NFS traffic.

Vaughn Stewart, @vStewed
vExpert

VMware only supports Microsoft Cluster Services with FC; however, your storage vendor may provide support for iSCSI guest initiators enabling IP based solutions.

Larry Smith Jr., @mrlesmithjr
vExpert

Use storage profiles for VM/VMDK placement. Assign profiles based on storage characteristics.

Allows for easy workload placements based on storage IO/Tiering/etc. needs and always keep them in compliance.

Josh Odgers, @josh_odgers
VCDX

When designing your VMware solution to use Storage DRS where Auto Tiering is enabled on your disk system OR intelligent caching solutions are used, ensure the I/O metric is DISABLED.

The array will then manage performance and SDRS can manage capacity by initial placement.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Enable SIOC and use percentage based congestion threshold.

By using its algorithm, SIOC determines peak throughput and allows you to leverage full potential of datastore without having to calculate optimal queue depth lengths manually.

Duncan Epping, @DuncanYB
VCDX

When implementing Storage DRS, leverage information provided through VASA to create your Datastore Clusters.

This will help ensure that your virtual machines are provisioned on the correct storage tier.

Chris Wahl, @ChrisWahl
VCDX

Rename all of the host local datastores from the default “datastore (” format to “hostname-local” format. This clearly identifies the local storage.

Josh Odgers, @josh_odgers
VCDX

For VMware View solutions, ensure you design to support initial deployment, refresh, recompose and rebalance operations with minimal impact to your ESXi hosts and disk subsystem.

This can be done by choosing NFS and ensuring VAAI and VCAI support.

Andreas Lesslhumer, @lessi001
vExpert

Take care that you align the GOS partitions on the virtual hard disks for all your virtual machine

Doing so avoids performance issues!

Vaughn Stewart, @vStewed
vExpert

ALUA established optimized and non-optimized paths.

Target Port Groups determine the active paths, both optimized and none. Multi-pathing sets how traffic will be delivered over the active optimized paths during normal access and the non-optimized failures during failure conditions.

Cormac Hogan, @vmwarestorage
vExpert

A device with a FIXED path policy (active/active) should be presented on at least 2 paths on all hosts, one to each controller presenting the LUN, for failover purposes.

Jason Nash, @TheJasonNash
VCDX

Confirm the proper PSP configuration any time you perform an upgrade on your storage array. It can change depending on version.

Raymon Epping, @repping
vExpert

When designing for storage performance understand that the distance to storage and RAID type used can have an impact on performance.

Using a virtual storage appliance and local SSDs, or a flash based local caching solution, could lead to increased performance and decreased latency.

Cormac Hogan, @vmwarestorage
vExpert

RDMs continue to rely on a consistent LUN ID. When the RDM is presented to multiple hosts, ensure that the LUN is presented in a consistent manner.

Jason Nash, @TheJasonNash
VCDX

If using array-based auto-tiering storage pools you cannot use Storage DRS to automatically load-balance datastores for latency performance.

Raymon Epping, @repping
vExpert

Storage IO Control is disabled on Datastores by default. When using Enterprise Plus and guaranteeing tenants resources is important ensure to enable it on each and every datastore.

This will prevent a single VM claiming all storage resources.

Ather Beg, @AtherBeg
vExpert

When multipathing FC to storage, use single-initiator zoning.

A single-initiator zone decrease operational complexity and prevents unnecessary RSCNs, allowing uninterrupted storage access.

Prasenjit Sarkar, @stretchcloud
vExpert

For an effective Storage IO Control, apply it to all datastores sharing the same physical spindles as Datastore performance impact happens because of sharing same spindles.

Cormac Hogan, @vmwarestorage
vExpert

NFS datastores must be mounted with root privileges on the ESX host.

Failure to do so will allow you to mount the datastore, but you will not be able to create any files on it."

Vaughn Stewart, @vStewed
vExpert

Ensure the datastores in a datastore cluster reside on different physical disks. This enables additional IO and capacity to be available.

Cormac Hogan, @vmwarestorage
vExpert

iSCSI Port Binding is recommended over NIC teaming with iSCSI since port binding allows the host to multipath & respond to SCSI events for failover. NIC teaming does not.

Cormac Hogan, @vmwarestorage
vExpert

When using RDMs with Microsoft Cluster Services, use the perennially reserved flag to avoid slow boots due to SCSI Reservations.

This will speed up the boot time of your ESXi host.

Frank Denneman, @FrankDenneman
Tech Evangelist PernixData

Use Storage DRS!

Its initial placement feature will find the best location for your virtual machine removing the painful operational task of datastore selection during VM provisioning.

Cormac Hogan, @vmwarestorage
vExpert

When enabling iSCSI Port Binding, you must have the host and target on the same subnet.

iSCSI Port Binding does not allow iSCSI traffic to route.

Vaughn Stewart, @vStewed
vExpert

Using data deduplication along with flash in the array or host could help driving down storage requirements and costs while ensuring performance.

Charlie Gautreaux, @gchuckman
Systems Engineer by day, VMUG leader by night.

When using NFS, remember to calculate and set NFS.HeartbeatDelta, NFS.HeartbeatFrequency, NFS.HeartbeatMaxFailures, NFS.HeartbeatTimeout Also take into account HA failover timeouts as well.

Words of Wisdom

Jonthan Frappier, @jfrappier
vExpert

Create proper documentation for all infrastructure including physical server, network and cabling diagrams. Label all equipment and cables.

Sean Crookston, @seancrookston
vExpert

Expect the unexpected. Ensure you have properly designed a process for backing up your virtual machines at both the virtual machine and guest level.

Doug Baer, @dobaer
VCDX

Always be mindful of your design goals: consider *why* you are making a decision and ensure that you understand both its implications and dependencies.

Vaughn Stewart, @vStewed
vExpert

When there is conflicting information, use your storage vendor provided guidelines.

They are specific to their storage technology & platforms where as VMware info is broad and meant to advise to all arrays.

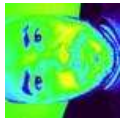
Lieven D'hoore, @ldhoore
vExpert

Use a documented naming convention for datacenters, clusters, hosts, VMs, templates, datastores and switches. Avoid adding physical locations into the naming as these might change (vMotions e.g.).

All this will help later on when automating your environment through scripts.

Bobby Stampfle, @bobbyfantast1c
vExpert

Model your design after Noah's Ark, 2 of everything! Redundancy.



Cedric Megroz, @cmegroz
vExpert

Before every installation and upgrade, it's a good idea to control the compatibility matrix of all the vSphere infrastructure components.

Think about the firmware of your Storage array, Fabric Switch, HBA firmware, BIOS Server, and all the Service Pack of yours software. Only one link to know: www.vmware.com/go/hcl.

Andrea Mauro, @Andrea_Mauro
VCDX

Keep it simple whenever possible, simple architectures are easiest to implemented, managed and makes troubleshooting easier!

Duncan Epping, @DuncanYB
VCDX

A great design start with opening communication with your design partners. Without your security, network, storage, application owners and operational team buy-in your project is doomed to fail.

Jason Nash, @TheJasonNash
VCDX

Just because something is a "Best Practice" doesn't mean it is correct for your design. Sometimes to meet a requirement it is necessary to go against a "Best Practice".

Josh Coen, @joshcoen
vExpert

When the zombies attack, Jason Nash's house will be a safe place to find refuge; go to Jason Nash's house

Duncan Epping, @DuncanYB
VCDX

Document your design decisions and include justification, constraints, risks and requirements.

Although you may know the infrastructure inside out, the person next to you may not.

Neil Koch, @neilkochAUS
VCP

Be sure to have a complete list of your production vSphere virtual machine IP addresses in the event you cannot access vCenter. Also document the boot-up order in case of a full datacenter failure!

Jason Nash, @TheJasonNash
VCDX

A simple way to test correct jumbo frame MTU size through the network is to use the "ping -d -s 8000" command.

This sends an 8K ping and tells devices not to fragment it. Make sure and size the -s option for your particular MTU and transport requirements.

Donny Parrott
Solution Architect

Do not start your design until you know and understand the metrics, scope, security, availability, use case, constraints and opportunities.

Donny Parrott
Solution Architect

As service delivery is the primary deliverable, engage early in service design to inform the infrastructure architecture.

Raymon Epping, @repping
vExpert

We all love technical facts and the physical aspect of a design, but a conceptual design and logical design should always be your foundation.

Abdullah Abdullah, @do0dzZZ
vExpert

Always consider configuring alerts, just not as a side task when implementing the project but emphasis this task as it will save you time when troubleshooting later on.

Kevin Monser, @monser
VCP

Run an exhaustive test on the host memory, using a looping test that will run continuously for at least a couple of days to ensure all hardware functions correctly.

Charlie Gautreaux, @gchuckman
Systems Engineer by day, VMUG leader by night.

Always remember to ensure the vSphere database stats roll-up jobs complete successfully. Otherwise stats will not be collected and you won't notice until you need them.