Design tips for VMware vSphere 4

Gabrie van Zanten
GabesVirtualWorld.com

Created with Google Wave

- I would like to thank for their input:

Scott Herrold	@vmguru
Rob Markovic	@vRobM
Dave Graham	@DaveGraham
Kendrick Coleman	@KendrickColeman
Gregg Robertson	@GreggRobertson5
Simon Long	@SimonLong_
Hany Michael	@hany_michael

Business requirements

- Mix business units on one datacenter, cluster, host
  - Is networking infra mixed across business units? Why not your virtualization layer?
  - VMs are fully isolated
    - VLANs in virtual switches (layer2)
    - vShield zones to protect VMs
    - NPIV for direct storage mapping
    - See links 1 & 2
- SLA differences between business units
  - Mixing Service Levels often best way to go
    - Meet SLA's by guaranteeing cpu and memory shares
    - Most optimize use of resources
  - Separating Service Levels
    - Leads to unbalanced infrastructure

Business requirements

- Protection against failures
  - Backups
    - File level, Application and VM backups
    - Long downtime
  - High Availability
    - Protects host and VM failures with downtime
  - Fault Tolerance
    - Protects host and VM failures without downtime
  - Microsoft Clustering Services
    - Protects host, VM and application failure with very little downtime
  - Datacenter failover
    - Protects against complete datacenter outage
    - Downtime varies on business continuity needs

Virtualization Strategy

- Virtualizing applications
  - Wait for vendor support or not?
  - Get “The Power from Above”
  - Not the “low hanging fruits”
  - Aim BIG

Security requirements

- Mixing networks
  - DMZ and internal LAN on one host
  - VMs and host crossing networks
  - Human errors
- Virtualizing DNS, DHCP and ADir
  - Not a technical issue
  - Rely on in case of Disaster Recovery

Capacity planning

- Capacity inventory
  - Monitor physical servers for at least a month
  - Include business peaks
- Special hardware requirements
  - Dongles
  - USB over network
- Predict growth of capacity.
  - VM sprawl needs attention
  - Storage-needs grow fast
- Datacenter capacity
  - Enough capacity during migration

Design decisions

- What size of ESX hosts will you buy?
  - Use capacity planning results
  - Build several scenarios with different hardware
  - Few very big hosts or many smaller hosts
  - Remember HA limits in a cluster
    - Max 32 hosts per cluster
    - 1280VMs per cluster
    - 100VMs per host if <= 8 hosts in cluster
    - 40 VMs per host if > 8 hosts in cluster
    - See link 3
  - Primary HA nodes in a blade center

Design decisions

- How to run vCenter Server as a VM?
  - Fully supported to run virtual.
  - vCenter Server VM best practices:
    - Disable DRS
    - Run VM only on first and second host
    - Leave one COS out DvSwitch
    - Set HA to high priority
    - vCenter Server startup
      - Active Directory
      - DNS
      - SQL
      - vCenter
  - See links 4 & 5

Design decisions - Storage

- Storage
  - Do read: “VMFS best practices and counter FUD”
    - Not true -> VMFS doesn’t scale because of locking
    - Not true -> 2TB is a show stopper
    - Not true -> Spanning VMFS across multiple extents is bad
    - Not true -> You can only have 32 VMs per VMFS
    - (link 6)

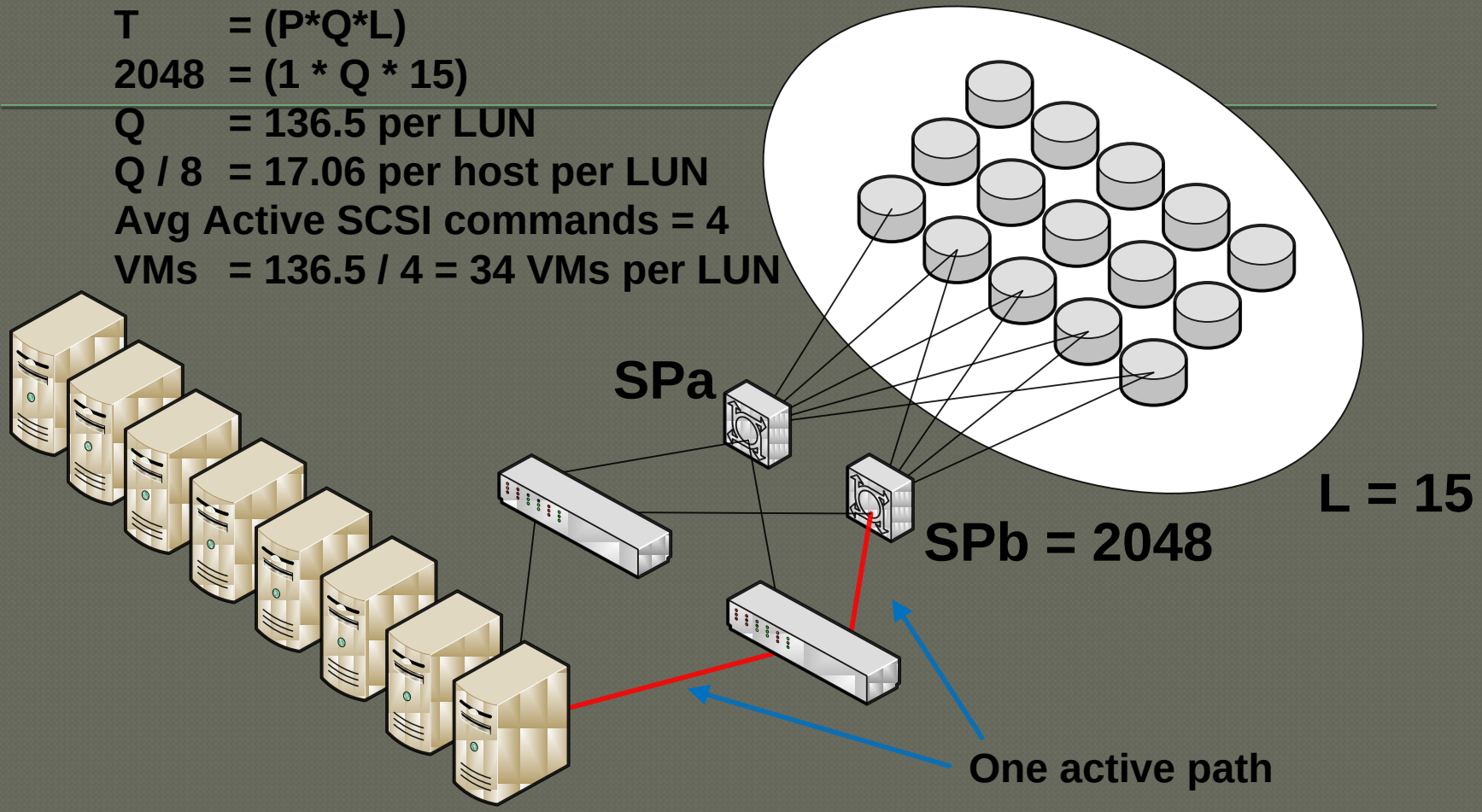
Table 2. Calculating Load on a VMFS volume for Sample Configurations

Maximum Outstanding I/O Recommended for Array, per LUN (*n*)	Average Active SCSI Commands per Virtual Machine to the Shared VMFS (*a*)	LUN Queue depth on each ESX Host (*d*)	Maximum Number of Virtual Machines per ESX Host on the Shared VMFS *m=(d/a)*	Maximum Number of Virtual Machines on the Shared VMFS (*n/a*)
256	4	32	8	64
256	4	64	16	64
1024	4	64	16	256
256	32	32	1	8
1024	32	32	1	32
1024	1	32	32	1024

Design decisions - Storage

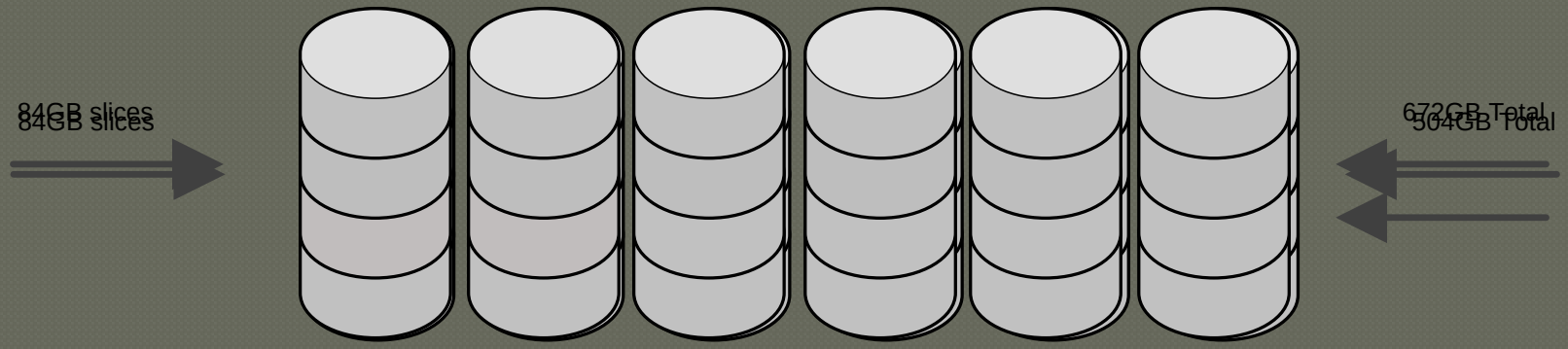
- Designing your LUNs
  - Number of VMs per LUN
    - Understand queue depths (link 7 & 8)
    - Target Port Q => Paths * Queue depth * LUNs
    - Max nr VMs per LUN = Queue depth / Avg Active SCSI commands

T = (P*Q*L)
2048 = (1 * Q * 15)
Q = 136.5 per LUN
Q / 8 = 17.06 per host per LUN
Avg Active SCSI commands = 4
VMs = 136.5 / 4 = 34 VMs per LUN



Design decisions - Storage

- LUN size
  - LUN size = (Max nr VMs * AvgSize) + 20%)
  - Round up to next 25GB
  - 20% for swap and snapshots
- Spindle trashing



Design decisions - Storage

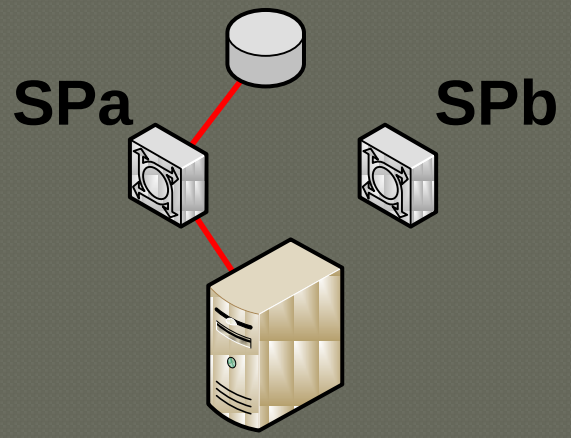
- Choose the correct VMFS block sizes

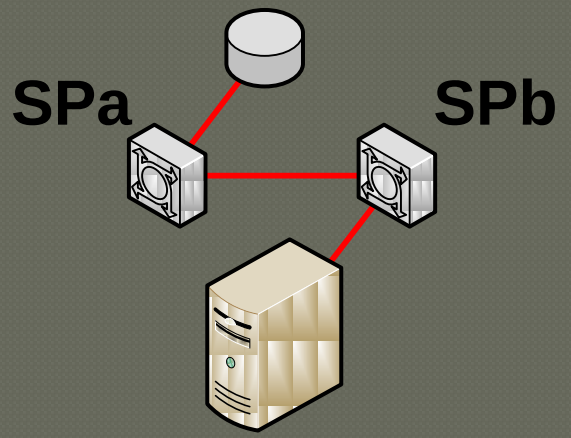
Block Size	Max File Size
1MB	256GB - 512B
2MB	512GB - 512B
4MB	1024GB - 512B
8MB	2048GB - 512B

- Snapshots always grow 16MB blocks
- Thin-provisioned disks grow VMFS block size blocks
- Small files always 64KB sub-blocks
- With 8MB blocks less SCSI locks

Design decisions - Storage

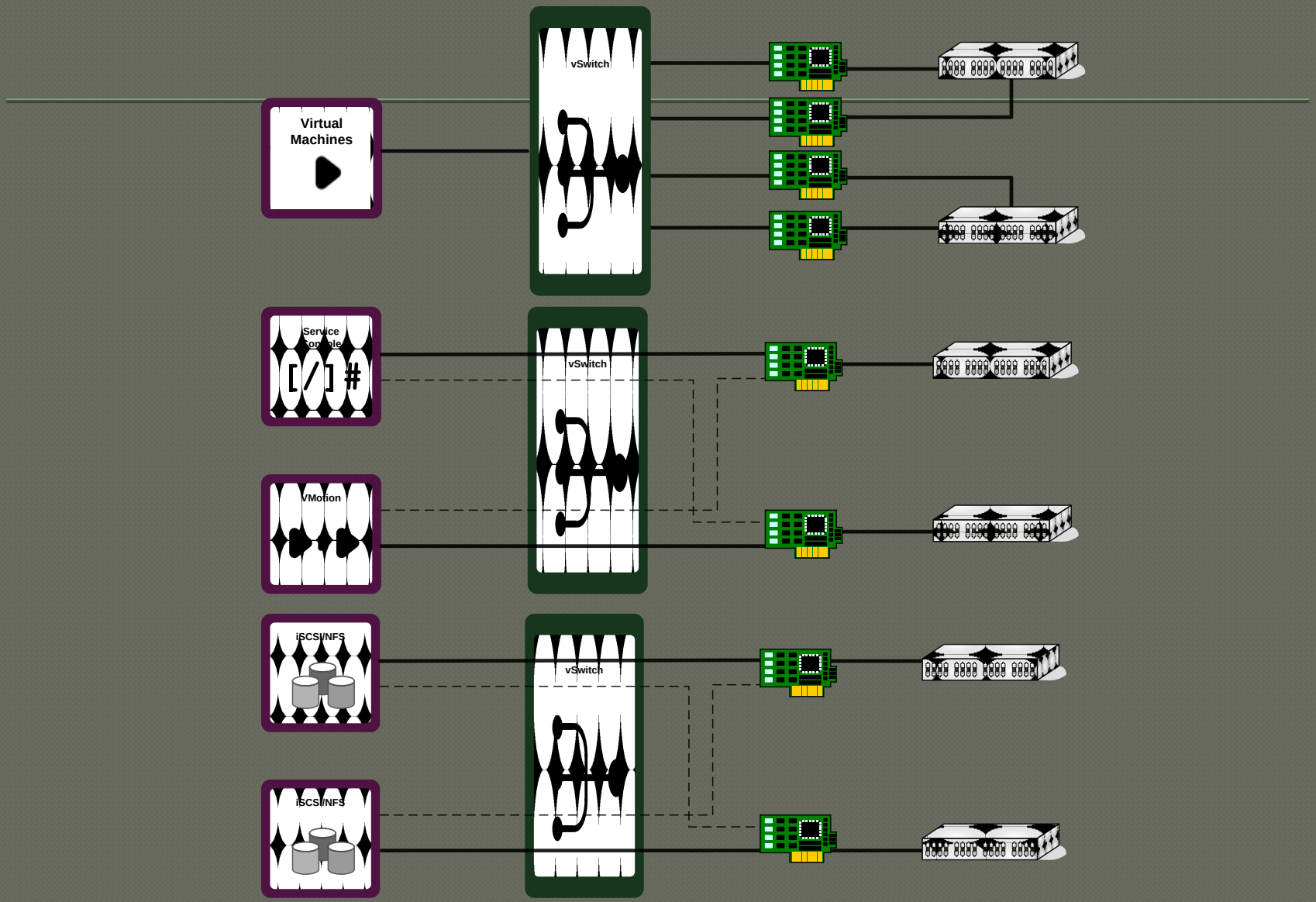
- Storage paths design
  - IBM prefered path, EMC prefered owner / current owner
  - Asymmetric Logical Unit Access (new in vSphere 4)
  - MTU & RR know optimal path
  - Round Robin preferred unless SAN guide
    - (link 9)



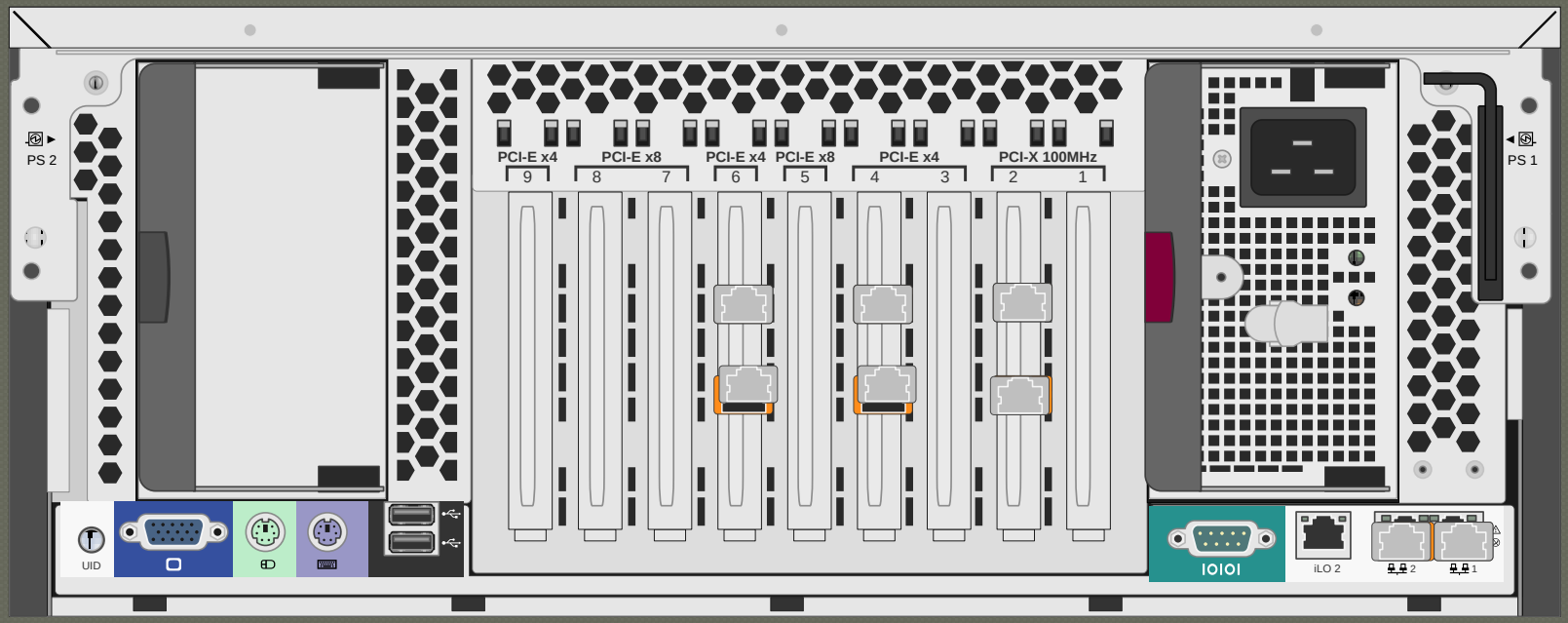


Design decisions – Networking

- Avoid using Link Aggregation (LACP)
  - Very limited use case, rarely 1Gbps bottleneck
  - Difficult to configure, many dependencies
  - Configuration errors cause big problems
  - One conversation cannot exceed nic limit
- Redundant network design
  - Use active and standby adapters over more portgroups
  - Spread over physical switches
- Single, dual and quad port NICs

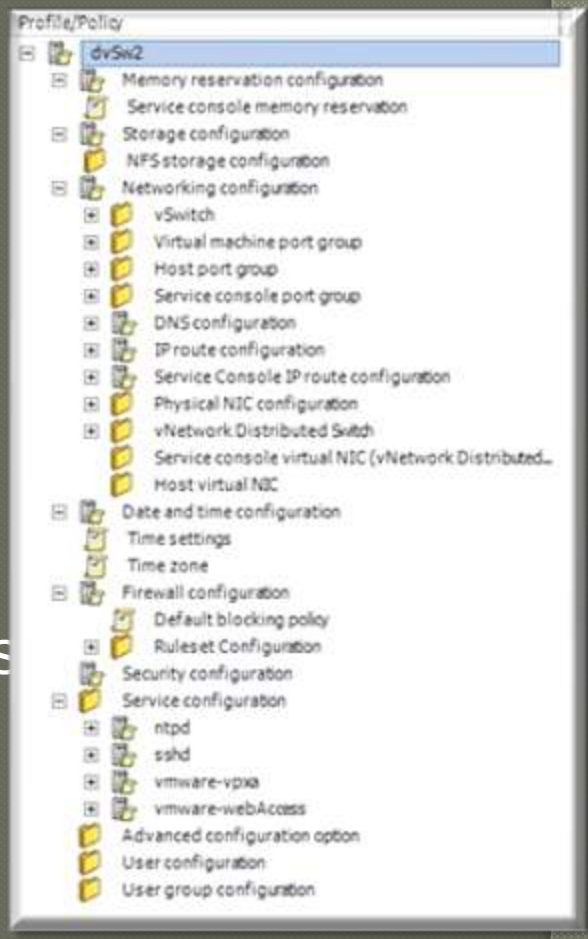

Virtual Machines
vSwitch
Service Console
vSwitch
VMotion
iSCSI/NFS
vSwitch
iSCSI/NFS

Design decisions – Networking



ESX Automated install

- Automated
  - Each host should be equal
  - No need for local settings backup
  - Re-install quick and easy
  - Powershell very powerful
- Host profiles or UDA
  - Host profiles needs (manual) install firs
  - Host profiles only with Enterprise Plus
  - UDA works for ESX 3.x & 4.x
  - UDA is free ☺ UDA is free ☹



ESX Automated install

- Example UDA script:
  - “Scripted install of ESX 3.0” (link 10)
  - “Using Linux shell scripts to automatically create VMware ESX 3.5 hosts” (link 11)
  - Settings made:
    - Service Console / VMotion
    - All portgroups & VLANs
    - Active Directory authentication (only user)
    - DNS, NTP, firewall
    - Unloads VMFS-2 driver

Questions

- Gabrie van Zanten
- http://www.GabesVirtualWorld.com
- TheGabeMan@gmail.com
- @GabVirtualWorld

Links

- Important information sources:
- **http://www.Yellow-Bricks.com**
  - Duncan Epping (VMware PSO)
- **http://virtualgeek.typepad.com**
  - Chad Sakac (Storage / EMC)
- **http://vlp.vsphere-land.com**
  - Eric Siebert Top 20 / 100 blogs

Links

1. http://www.vmware.com/pdf/TechTrendNotes.pdf
2. http://www.vmware.com/pdf/vi3_security_architecture_wp.pdf
3. http://www.yellow-bricks.com/2009/05/25/max-amount-of-vms-per-host
4. http://www.boche.net/blog/index.php/2009/10/09/virtualizing-vcenter-with-vds-catch-22/
5. http://www.yellow-bricks.com/2009/10/09/best-practices-running-vcenter-virtual-vsphere/
6. http://virtualgeek.typepad.com/virtual_geek/2009/03/vmfs-best-practices-and-counter-fud.html
7. http://frankdenneman.wordpress.com/2009/03/04/increasing-the-queue-depth/
8. http://www.yellow-bricks.com/2009/07/07/max-amount-of-vms-per-vmfs-volume/
9. http://www.yellow-bricks.com/2009/09/29/whats-that-alua-exactly/
10. http://www.gabesvirtualworld.com/?p=21
11. http://www.gabesvirtualworld.com/?p=424