



Microsoft Virtualization

Ing. Ondřej Ševeček | GOPAS a.s. |

MCM:Directory | MVP:Enterprise Security | Certified Ethical Hacker |

ondrej@sevecek.com | www.sevecek.com |

GOPAS: info@gopas.cz | www.gopas.cz | www.facebook.com/P.S.GOPAS

Technické informace

- 6 částí po 30 minutách
- Introduction to Hyper-V in Windows 2012 R2
- Basic networking
- "Migrations"
 - snapshot
 - backup and VSS
 - storage migration
 - replication
 - live migration
- Clustering
- Virtual Machine Manager
- Network virtualization



Technické informace

- 3 týdny :-)
- výuka: 9.00—12.00, 13.00—16.00
- přestávky
- nápojové automaty
- kouření
- oběd
- mobilní telefony



Microsoft OS virtualization history

- Software emulation (no hardware support)
 - Virtual PC 2005
 - Virtual Server 2005
 - Virtual PC 2007
 - Windows XP Mode for Windows 7
- Hypervisor platform based Hyper-V (CPU support)
 - Windows 2008 (v6.0)
 - Windows 2008 R2 (v6.1)
 - Windows 2012 (v6.2) and Windows 8
 - Windows 2012 R2 (v6.3) and Windows 8.1



Terminology

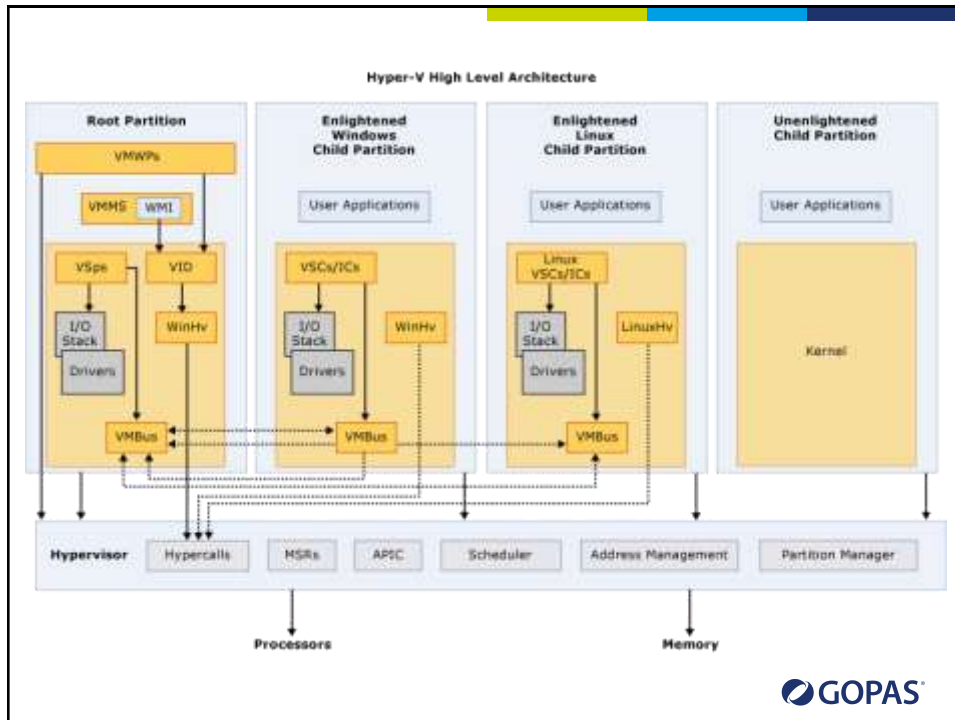
- Root partition, parent partition, parent, host, hardware
- Child partition, child, guest, VM



Hypervisor and hardware virtualization

- Chipset support for virtualized partitions
- Hypervisor
 - the small OS running directly on CPU and handling memory, CPU and interrupts
 - **parent** vs. **child/guest** partitions
- Guest partition
 - virtual view of CPUs, although instructions run directly
 - does not handle interrupts directly
 - second level virtual address translation by CPU
 - virtual view of devices over VMBus (if IC installed)
 - hardware requests are processed by drivers installed in **parent** partition





Hypervisor

- Loaded by BOOTMGR
- BCDEDIT
 - HypervisorLoadType = Auto

VMWP.exe worker processes

- Started by VMMS.exe
 - Hyper-V Virtual Machine Management service (vmms)
- Manage the child partitions
- Performance counters for VMs from parent
- Run under virtual user identity (= VM ID)
 - `get-vm | fl *`
- Perform some emulated device actions
- Works with IC



Integration components (old ones)

- Host/guest on supported versions to work correctly
- Heartbeat
 - used by failover cluster
 - `Get-VMIntegrationService <vmname>`
- Time synchronization
 - disable?



Integration components (Data Exchange since 2012)

- "key-value pair exchange"
 - requires "Hyper-V Data Exchange Service" (vmickvpexchange)
- Data exchange in **guest**
 - intrinsic sent to host: HKLM\Software\Microsoft\Virtual Machine\Auto
 - ♦ key = REG_SZ = value
 - custom sent to host: HKLM\Software\Microsoft\Virtual Machine\Guest
 - receive from host: HKLM\Software\Microsoft\Virtual Machine\External
- Data exchange from **host**
 - .GuestIntrinsicExchangeItems
 - .GuestExchangeItems

```
'client7' | % { Get-WmiObject -Namespace root\virtualization\v2 -Class Msvm_ComputerSystem -Filter "ElementName='$_'" } | % { $_.GetRelated("Msvm_KvpExchangeComponent").GuestIntrinsicExchangeItems | Select @{ n = 'What' ; e = { ([xml] $_).SelectSingleNode("/INSTANCE/PROPERTY[@NAME='Name']/VALUE").'#text' } }, @{ n = 'Value' ; e = { ([xml] $_).SelectSingleNode("/INSTANCE/PROPERTY[@NAME='Data']/VALUE").'#text' } } }
```



Integration components (Guest Services since 2012 R2)

- "out-of-band management"
- disabled by default
 - when enabled, the **guest** service "Hyper-V Guest Service Interface" (vmicguestinterface) starts
- Copy-VMFile
 - yet the only service available
 - copy a file from **host** to **guest**



BIOS

- American Megatrends BIOS
 - VM Generation 1
 - boot from IDE and MBR only
- UEFI BIOS
 - VM Generation 2
 - boot from SCSI and MBR or GPT



Virtual memory in Hyper-V

- Physical RAM
 - [gwmi Win32_PhysicalMemory](#)
- Physical RAM [memory pages](#) mapped directly by CPU
 - sum of all physical memories cannot be higher than real physical RAM
- Windows 2008 R2+ supports [dynamic memory](#)
 - second level virtual address translation
- If code touches a page which is not present in RAM
 - guest OS hard page-faults it back
 - hypervisor hard page-faults it back



Dynamic memory support

- Host must be Windows 2008 R2 SP1+
- Backed by paging file on the host
- Only supported guests Windows 2003+
 - some service packs required
- Performance counters
 - Hyper-V Dynamic Memory
- Reserve memory for the host?
 - HKLM\Software\Microsoft\Windows NT\CurrentVersion\Virtualization
 - MemoryReserve = DWORD = [MB]
 - Windows 2012 with better calculation



Logical CPU support

- Physical hardware
 - socket
 - logical processor core
 - hyper-threading
 - [gwmi Win32_Processor](#)
 - ◆ NumberOfCores
 - ◆ NumberOfLogicalProcessors
- Different guest support



Logical CPU support by guest OS

| | 2000 client | 2000 server | XP client | XP 64bit client | 2003 server | Vista client | 2008 server | 7 client | 2008 R2 server | 8+ client | 2012+ server |
|---------|-------------|-------------|-----------|-----------------|-------------|--------------|-------------|----------|----------------|-----------|--------------|
| 2008 | 1 | 1 | 2 | 2 | 2 | 2 | 4 | 4 | 4 | 4 | 4 |
| 2008 R2 | 1 | 1 | 2 | 2 | 2 | 2 | 4 | 4 | 4 | 4 | 4 |
| 2012 | no support | no support | 2 | 2 | 2 | 2 | 8 | 4 | 64 | 32 | 64 |
| 2012 R2 | no support | no support | 2 | 2 | 2 | 2 | 8 | 4 | 64 | 32 | 64 |

- Cannot assign more virtual CPUs than the number of logical CPUs on host



Memory and CPU support

| | Host | Guest |
|-----------------|-----------------|-------|
| Windows 2008 | 1 TB, 64 cores | 64 GB |
| Windows 2008 R2 | 1 TB, 64 cores | 64 GB |
| Windows 2012 | 4 TB, 320 cores | 1 TB |
| Windows 2012 R2 | 4 TB, 320 cores | 1 TB |



Emulated vs. synthetic vs. SR-IOV devices

- Emulated devices are software emulated by [vmwp.exe](#) (except IDE)
 - "compatible" from below with the built-in OS drivers
 - IDE controllers (ataapi.sys, intelide.sys)
 - Standard VGA adapter
 - Legacy network adapter (Intel 21140 PCI 100tx Ethernet)
 - floppy, mouse, keyboard, COM ports, BIOS (KB2844106)
- Synthetic devices can communicate over VMBus directly
 - require custom MS drivers from IC
 - Microsoft Hyper-V Video (HyperVideo.sys)
 - Microsoft Hyper-V SCSI Controller (storvsc.sys)
 - Microsoft Hyper-V Network Adapter (netvsc60.sys)
- Single-root IO virtualization
 - hardware supported virtualization (extension to PCIeexpress spec)
 - requires "reservation" of the device for a particular child partition
 - drivers from child partition access the device directly
 - [SR-IOV NIC traffic does not go through virtual switch](#)



Monitoring VM performance

- Cannot monitor much from parent partition
 - VMWP.exe for emulated devices only
- Performance counters
 - Hyper-V ...
- Example
 - parent: [Hyper-V Hypervisor Logical Processor - Guest Run Time](#)
 - child: Processor Time



Powershell scripts for performance testing

```
# writing big data into a file for a long time
(1..100000) | % { 'a' * 100000 | out-file c:\test.txt -Append }

# infinite loop
while ($true) { $a = 5 }

# UDP receiving server
$srv = New-Object System.Net.Sockets.UdpClient(33333)
[System.Net.IPEndPoint] $sender = $null
while ($true) { [void] $srv.Receive([ref] $sender) }

# UDP sending client
$cln = New-Object System.Net.Sockets.UdpClient
$bfr = [System.Text.AsciiEncoding]::ASCII.GetBytes(('a' * 3000))
while ($true) { [void] $cln.Send($bfr, $bfr.Length, '10.10.0.102', 33333) }
```



Powershell scripts for performance testing

```
# rewriting a big file indefinitely
$fil = [IO.File]::Open('c:\data\test.txt', 'Open')
[byte[]] $bytes = (1..255)
while ($true) { for ($i = 0; $i -lt $fil.Length - $bytes.Length; $i +=
$bytes.Length) {
    [void] $fil.Seek($i, 'Begin')
    [void] $fil.Write($bytes, 0, $bytes.Length)
}

# Generate about 1 GB RAM requirement
[System.Collections.ArrayList] $list = @()
(1..5000) | % { [void] $list.Add((1..6500)) }
```



Enhanced Session Mode

- RDP through VMBus
- Hyper-V on Windows 2012 R2+
- Client on Windows 8.1+ or Windows 2012 R2+
- RDP Enabled in VM
 - user member of [Remote Desktop Users](#) or [Administrators](#) in the VM

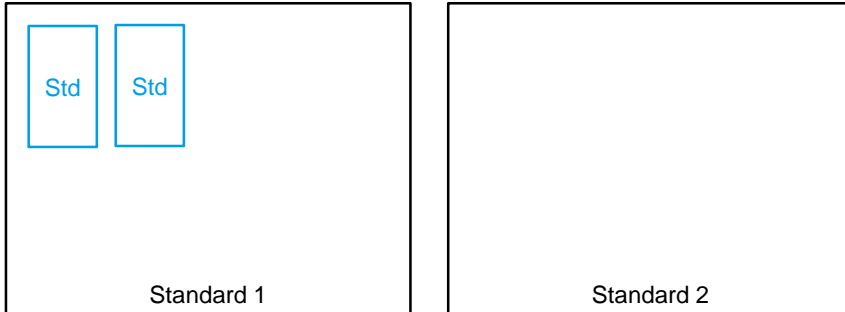


Licensing

- Special licensing
 - no server roles on host except for Hyper-V
- Standard base OS
 - up to 2 CPU sockets
 - 2 virtual OSE
- [Datacenter](#) base OS
 - unlimited virtual OSE
 - 4x more expensive than Standard => [better for 6+ VMs](#)
- OEM editions do not support migrations



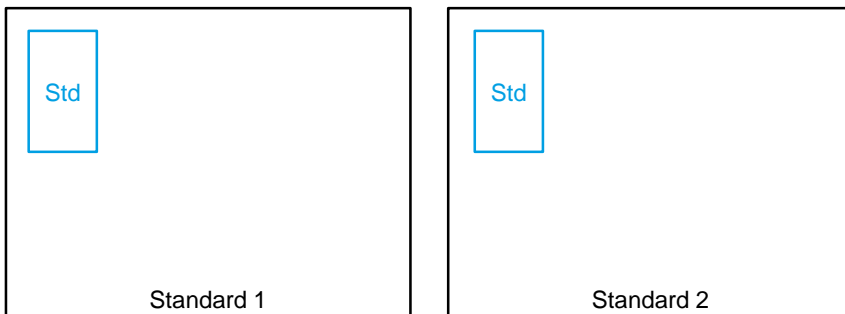
Licensing 2x Standard edition and "migrations"



2x Standard
with migrations



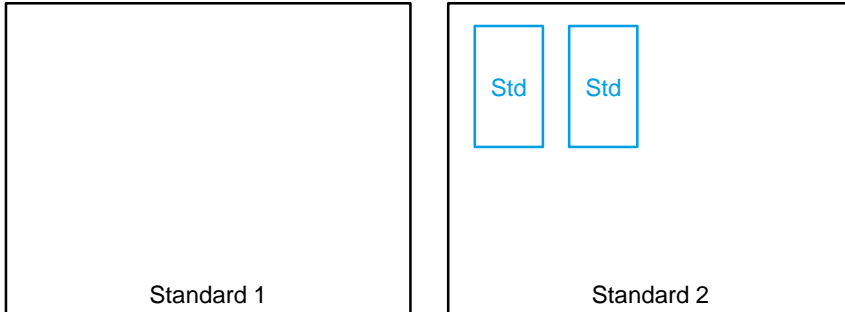
Licensing 2x Standard edition and "migrations"



2x Standard
with migrations



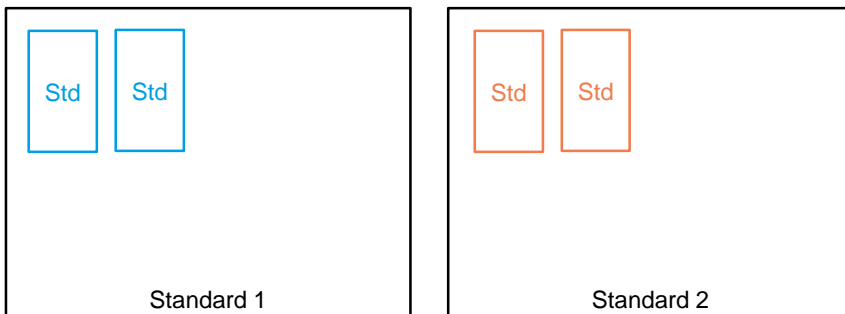
Licensing 2x Standard edition and "migrations"



2x Standard
with migrations



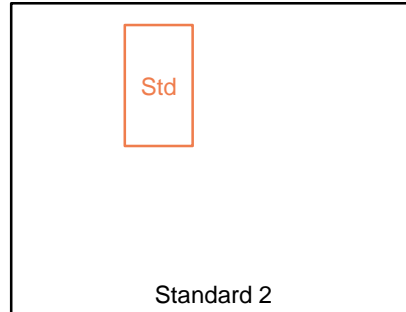
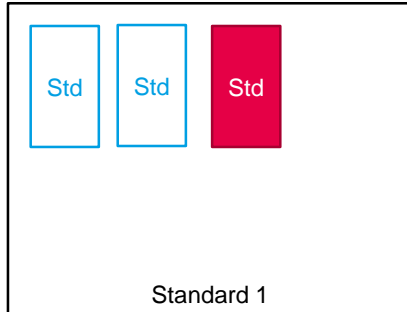
Licensing 2x Standard edition and "migrations"



2x Standard
no migrations



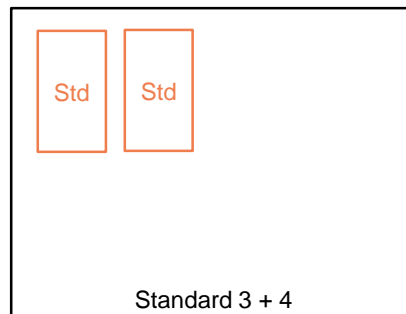
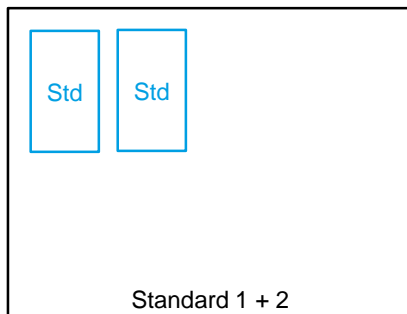
Licensing 2x Standard edition and "migrations"



2x Standard
no migrations



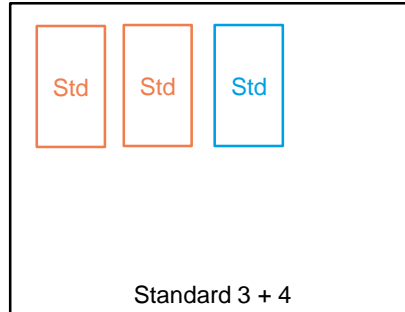
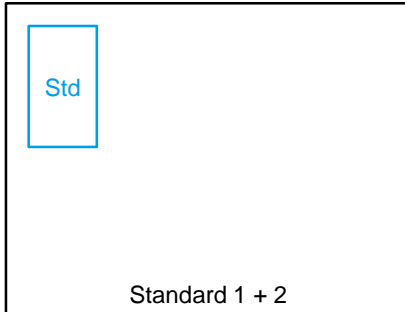
Licensing 4x Standard edition and "migrations"



4x Standard
with migrations



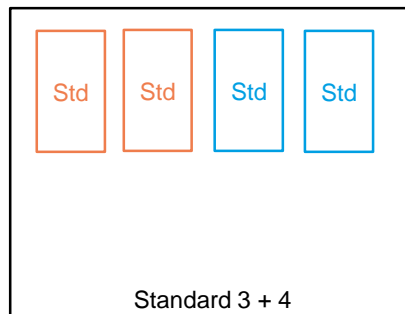
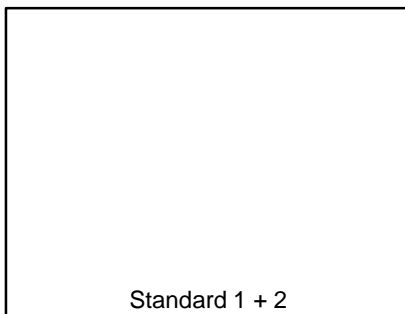
Licensing 4x Standard edition and "migrations"



4x Standard
with migrations



Licensing 4x Standard edition and "migrations"



4x Standard
with migrations



Automatic VM activation (AVMA)

- Host Windows 2012 R2+ Datacenter
 - must be activated
 - any activation method MAK, KMS, OEM
- Guest Windows 2012 R2+
 - slmgr.vbs /ipk AVMAKey
 - <http://technet.microsoft.com/en-us/library/dn303421.aspx>
 - Datacenter = Y4TGP-NPTV9-HTC2H-7MGQ3-DV4TW
 - Standard = DBGBW-NPF86-BJVTX-K3WKJ-MTB6V



Storage scalability per VM

- 4x IDE (emulated)
- 4x SCSI (synthetic)
 - 64x LUNs per SCSI



Virtual hard-disk formats

- Sector based data storage
 - no compression
 - dynamically expanding
 - differencing (copy on write from a base disk)
- Compare with WIM image
 - file based
 - compressed
 - more images in a single file
 - single instance storage per file
- VHD
 - can be mounted with Windows 6.1+
 - supports mounting over SMB
 - does not mount from encrypted/compressed folders
- VFD
 - supported by Hyper-V only, cannot be mounted
- VHDX
 - Windows 2012 and newer



Differencing and dynamic virtual disks

- Base disk
 - identified by UUID
 - last-modified timestamp
 - can be marked read/only
- Differencing (child) disk
 - referenced parent/base UUID and timestamp
 - copy on write
 - cannot remap
 - cannot modify base disk
 - up to 8 children on Windows 2008
 - up to 1024 children on Windows 2008 R2+
- Two levels of disk fragmentation



(Virtual) disk terminology

- Physical sector
 - 512 B or 4 kB
- Logical sector size
 - OS representation of the sector size
 - 512 B native
 - 4 kB native (Windows 6.2+)
 - 512E = 512 B logical on 4 kB physical (Windows 6.0+)
 - `fsutil fsinfo sectorinfo <volume:>`
- Data block size
 - Block Allocation Table (BAT) contains offsets
 - unit of expansion for dynamic and differencing disks
 - ♦ BAT not present in fixed size disks
 - 512 or 2 MB or bigger
 - when you create a VHD with Windows 2008 R2, it uses 2 MB by default
 - every data block contains [sector bitmap](#) + data
 - ♦ 1 = sector with data, 0 = sector with all zeros



Disk allocation compared

| | 2040GB Dynamically Expanding VHD in Windows Server 2008 R2 | 2040GB Dynamically Expanding VHD in Windows Server 2008 |
|------------------------------|--|---|
| VHD Capacity | 2040 GB | 2040 GB |
| Data Blocks within VHD | 1 044 480 | 4 177 920 |
| BAT (Block Allocation Table) | Approximately 4M (default 2M/block) | Approximately 16M (default 512K/block) |
| Sector Bit Maps (512B/Block) | Approximately 510 MB | Approximately 2 GB |
| Footer / Header | 2K | 2K |



Virtual disk format features

- VHD
 - max 2 TB, data block size 2 MB
 - can be mounted and/or booted from on Windows 6.1+
- VHDX
 - max 64 TB
 - metadata update log/transactions
 - ♦ protection against power failures
 - 512 B or 4 kB sectors, block size 256 MB
 - ♦ default 4 kB physical, 512 B logical
 - custom user metadata storage
 - ♦ Win32 API only
 - required by 2012 R2 to store virtual disk on a SMB share
 - shared virtual disks in clustered environments
 - ♦ SCSI persistent reservation
- New-Vhd, Get-Vhd
 - requires [Hyper-V-PowerShell](#) feature



Virtual disk conversions (Edit disk)

- Compact
 - update [sector bitmaps](#) and possibly [BAT](#) for sectors with all zeros (only for disks with BAT)
- Convert
 - from dynamically expanding or differencing to fixed
 - ♦ faster, less overhead without BAT
 - from VHD to VHDX
 - ♦ keeps physical/logical/block sizes
- Expand
 - all types



SMBv3 clustered storage (just some notes)

- Scale-out file server
 - balancing over all cluster nodes
- Transparent failover
 - SMB cluster node fails exceptionally - longer timeout
 - SMB cluster node planned failover - transparent
- SMB multi-path
 - more than a single connection over multiple IP addresses
- SMB direct
 - requires NICs with RDMA on both sides

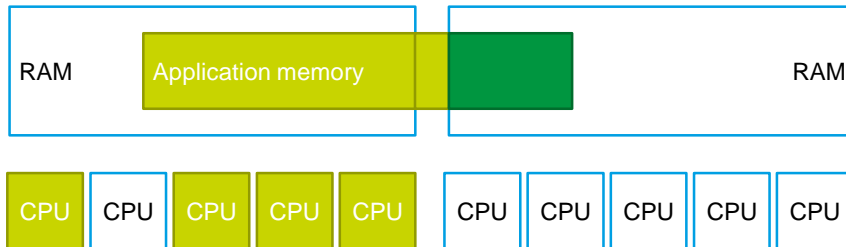


NUMA

- Non-uniform memory architecture
 - CPUs are much faster than memory
 - Only one CPU can access memory at a time
- NUMA nodes
 - pairing of CPUs with RAM ranges (banks?)
 - applications with large RAM demands can span NUMA nodes
- Windows 7 and 2008 R2 support NUMA
 - SQL server since 2005
 - IIS since 2012 (IIS 8.0)
 - Hyper-V parent since Windows 2008 R2
 - Hyper-V guest since Windows 2012

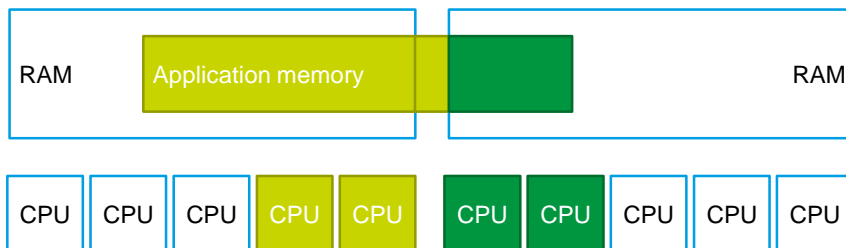


NUMA spanning - remote memory



 GOPAS®

NUMA spanning - only local memory



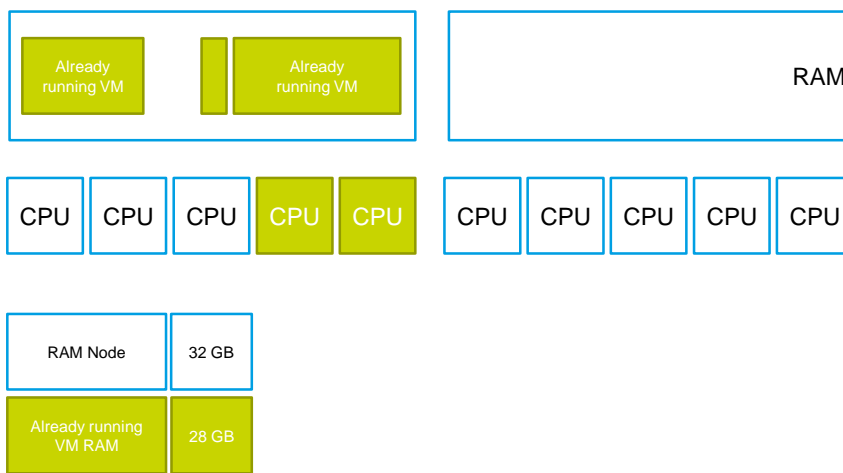
 GOPAS®

NUMA spanning and Hyper-V 2012+

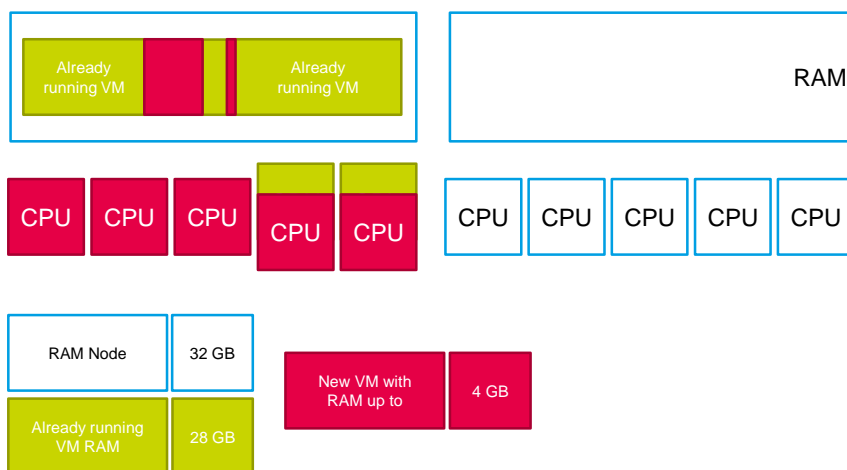
- Physical vs. virtual NUMA nodes
 - virtual NUMA nodes map to physical NUMA nodes
 - [Get-VMHostNumaNode](#)
- Enabled on host
 - will always start/restore/migrate regardless of NUMA nodes
 - suboptimal performance possible
- Disabled on host
 - will not start/restore/migrate if it would require NUMA node spanning
- Performance counters
 - Hyper-V VM Vid Partition - Physical Pages Allocated
 - Hyper-V VM Vid Partition - Remote Physical Pages
- Dynamic memory
 - only a single NUMA node assigned



Hyper-V with NUMA spanning disabled



Hyper-V with NUMA spanning disabled



 GOPAS

Virtual machine generations

- Generation 1
 - Windows 2008 - Windows 2012
 - boot from IDE only
 - American Megatrends BIOS based ([gwmi Win32_BIOS](#))
- Generation 2
 - Windows 2012 R2+
 - VHDX only
 - guest requirements
 - ♦ 64-bit
 - ♦ Windows 6.2+ (Windows 8 and Windows 2012)
 - boot ISO media needs mending on 6.2
 - keyboard is not working in ISO installs :-) on 6.2

 GOPAS

Generation 2 virtual machines

- UEFI BIOS
 - no 2 TB limit for boot partition
 - boot from GPT disks
 - supports VMBus directly
- VMBus drivers at boot phase 0
 - boot from SCSI
 - paging files on SCSI
 - PXE boot with synthetic NIC (IPv6 supported)
- Secure boot
 - UEFI standard
 - OS loader must be signed by trusted CA
- DVD on SCSI
 - hot add/remove
- No floppy, no physical CD/DVD



Converting between generations

- Not supported directly
- Convert VHD to VHDX
 - data disks go nice
 - OS disk needs mending (rather capture WIM image and build new VHDX with GPT and EFI partition)
 - ♦ `imagex /capture or dism /capture-image`
 - ♦ `diskpart (create partition efi)`
 - ♦ `bcdboot /f UEFI`



Networking scalability

- 8 synthetic NICs per VM
- 4 emulated NICs per VM
- unlimited NICs per Hyper-V server
- unlimited switches per Hyper-V server
 - unlimited ports per switch



Virtual switches

- Virtual switch
- External ports
 - create a virtual NIC in host
 - connected to physical adapters
 - can use VLAN ID for parent partition
- Internal ports
 - creates a virtual NIC in host
 - 10 Gbps goes through memory directly
- Private ports
 - VM to VM connection only
 - 10 Gbps goes through memory directly



Switch ports on VMs

- Legacy vs. synthetic
- Connect/disconnect cable
- Static or dynamic MAC address
- VLAN ID
 - 802.1q
 - multiple VLAN IDs with [Set-VMNetworkAdapterVlan - AllowedVlanIdList](#)
- Enable MAC spoofing
- Enable blocking of DHCP server in VM
- Enable blocking of router advertisements in VM
- Enable NIC teaming in guest
 - only for External switch ports
 - Windows 2012 supports NIC teaming (load balancing/failover)
- Port monitor for troubleshooting

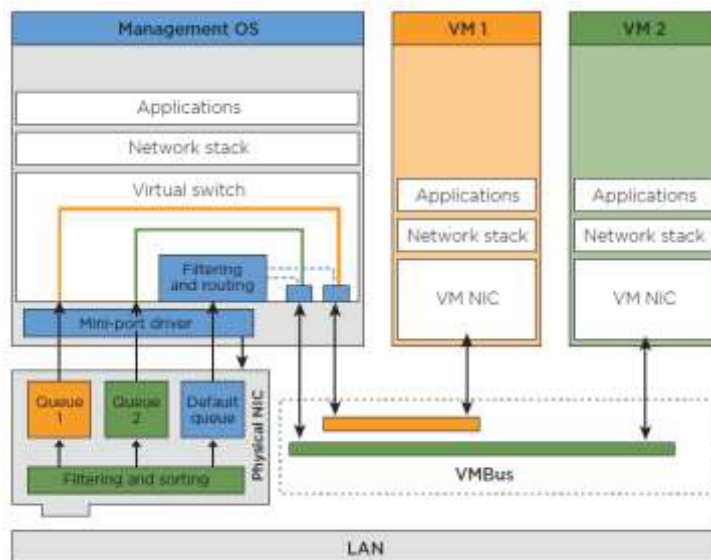


Virtual switch performance

- Counters
 - Hyper-V Virtual Switch
 - Hyper-V Virtual Switch Port
 - CPU and interrupts
- All switching done in parent partition
 - private - [single CPU core](#) regardless number of switches
 - external - distributed per CPU core per virtual port
- Virtual Machine Queue (VMQ)
 - hardware supported MAC/VLAN address queues per VM port on NIC's CPU
 - 2008 R2+, [should disable if not supported by NIC](#)



Virtual Machine Queue (VMQ)



GOPAS

VMQ Enable/Disable

- Physical NIC setting
- Hyper-V setting
 - HKLM\SYSTEM\CurrentControlSet\Services\VMSSMP
 - ♦ `BelowTenGigVmqEnabled` = DWORD = 1/0
 - ♦ `TenGigVmqEnabled` = DWORD = 1/0
- `Set-VMNetworkAdapter` on 2012+
- VMM centrally on a VM
 - "Enable Virtual Network Optimizations"

GOPAS

Multiple external ports on parent partition

- Single physical NIC (teamed)
- Multiple virtual external NICs in parent partition
 - [New-VMNetworkAdapter](#)
- Can be used for simpler clustering
 - management, storage, live migration, etc.



NIC teaming in the host/parent

- Windows 2012+
- Switch independent
 - MAC/IP/TCP/UDP [destination](#) address hash
 - ♦ balances [outbound](#) traffic
 - ♦ [receives on a single](#) NIC only
 - Hyper-V port
 - ♦ balances outbound traffic the same
 - ♦ receives on a single NIC per [VM port](#) (not per VM)
 - Dynamic
 - ♦ combination of both plus rebalancing
- Switch dependent
 - LACP on the switch
 - static link aggregation set on the switch
- Its own VLAN ID technology



Switch independent teaming with Dynamic balancing

- Inbound per VM port
 - the ARP MAC responses go out from one of the NICs with the VM port's source MAC
- Outbound per destination MAC/IP/TCP/UDP hash
 - source MAC address changes to that of the team's physical NIC

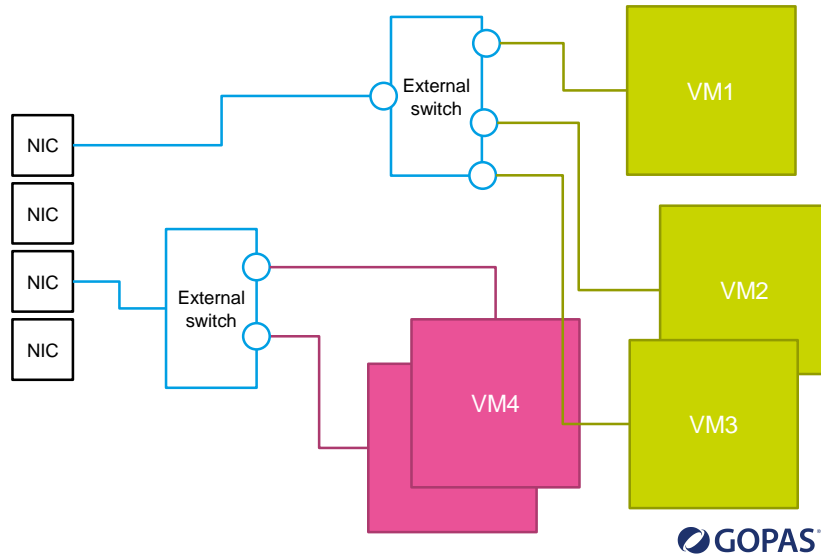


NIC teaming notes with "Dynamic"

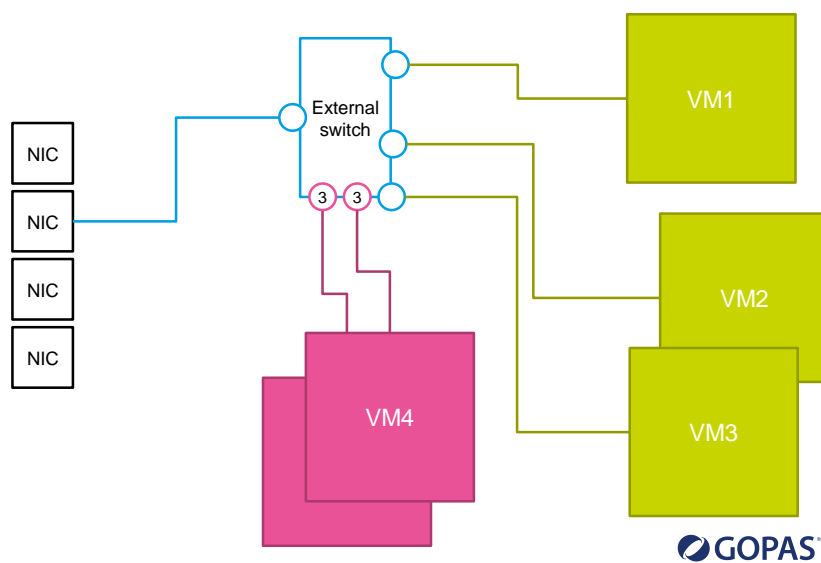
- per VM port not per VM
- single VM NIC cannot exceed 10 Gbps
- different physical NIC speeds not supported
 - balancing does not take speeds in consideration
- team for LM cannot exceed single NIC speed



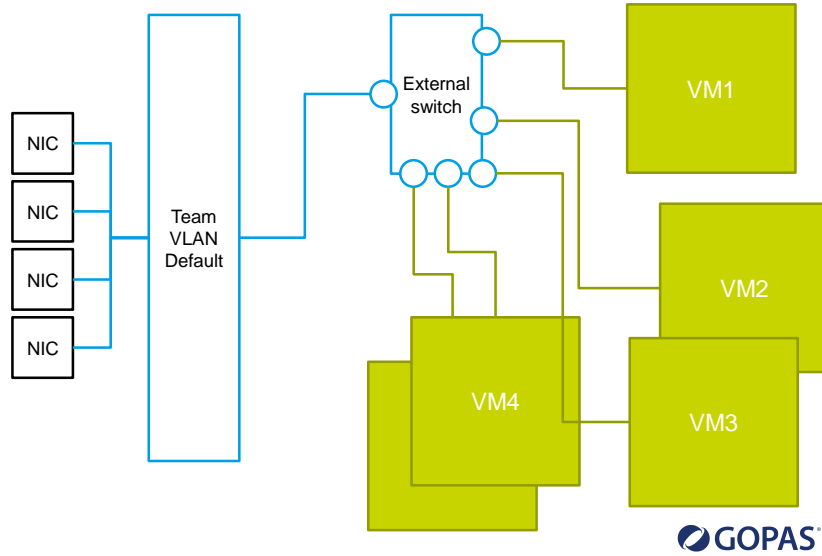
External switches only



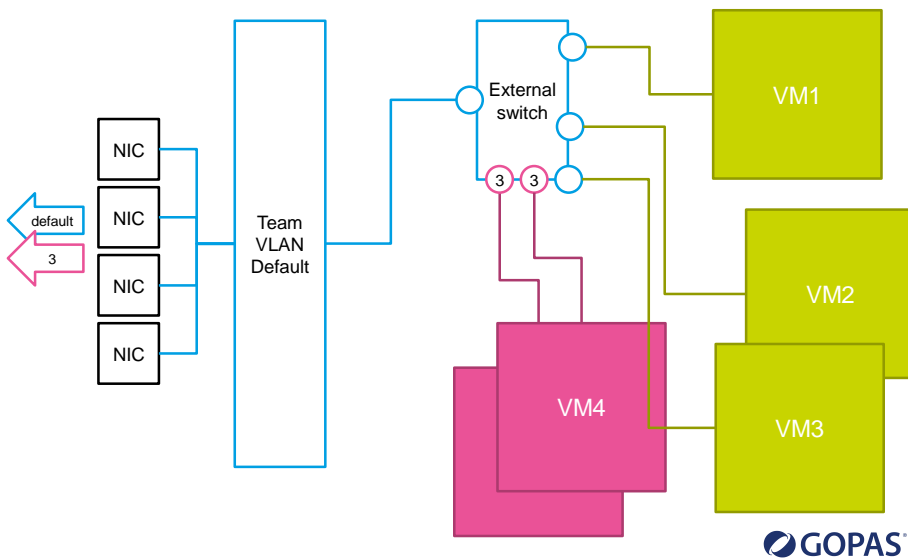
External switches and VLANs



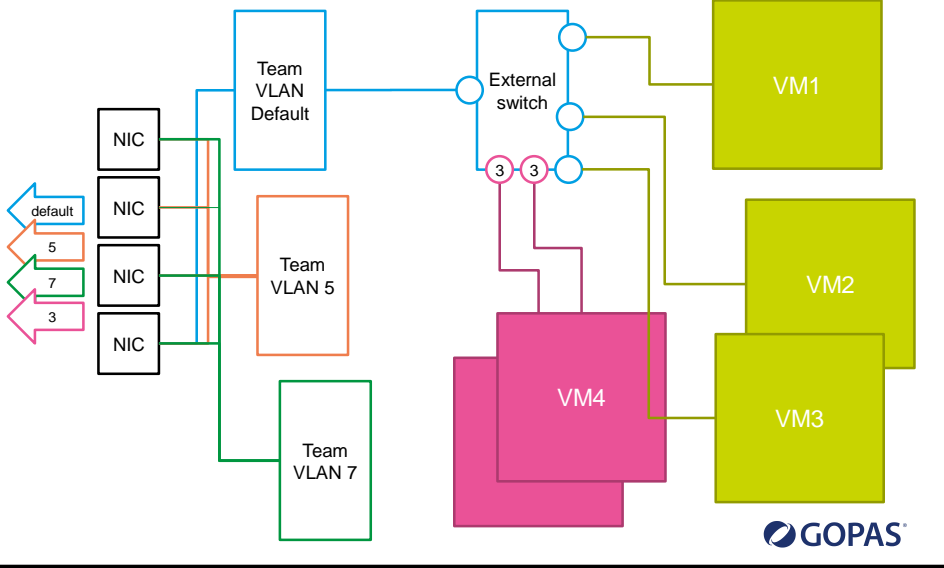
NIC teaming and external switches



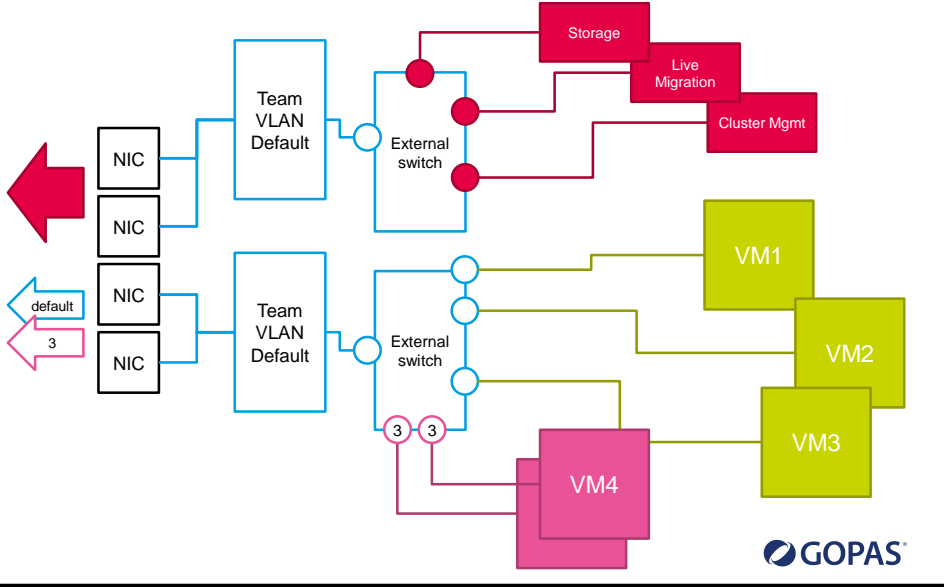
NIC teaming, external switches and VLANs

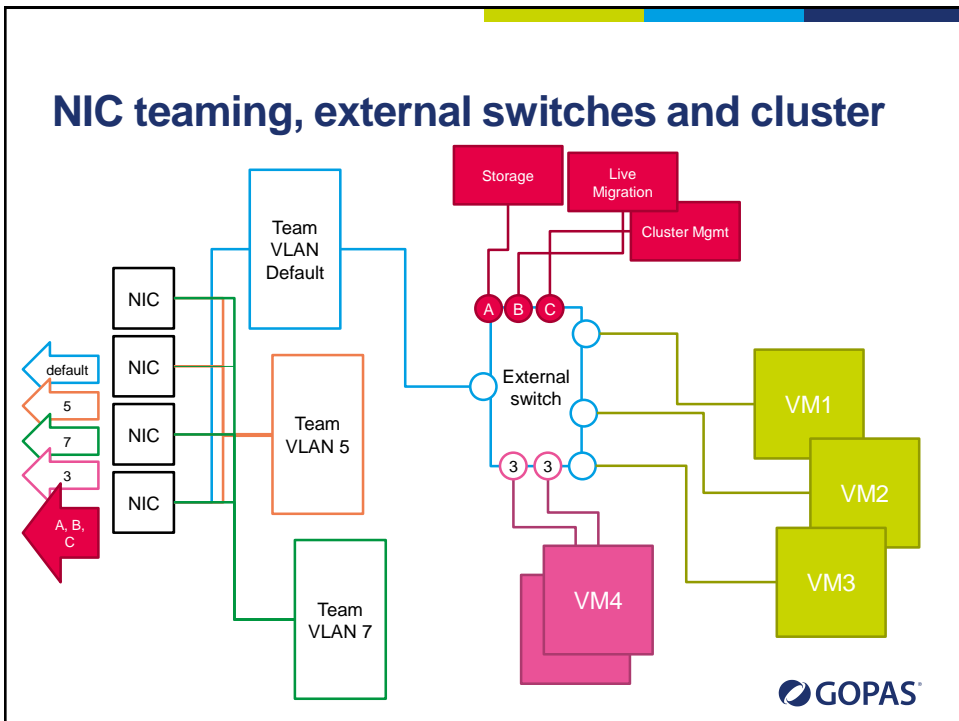
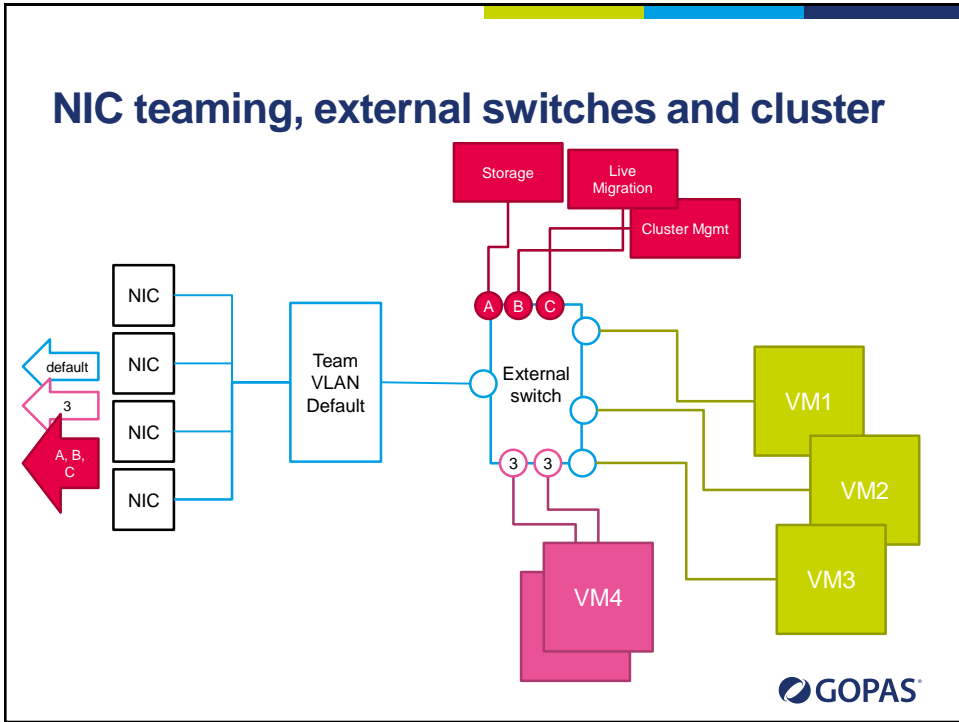


NIC teaming and VLAN teams



NIC teaming, external switches and cluster





NOTES! Received packets discarded!

Team with any NICs **except GT adapter**

Use any type of distribution



NIC bandwidth management

- Maximum
 - upper limit can be enforced
- Minimum
 - cannot guarantee minimum
 - guarantees **proportional** minimum among other machines



Export/Import

- 2012 does not require export in order to import
- Import
 - same IDs
 - different IDs



Snapshots (checkpoints)

- Memory snapshot
 - takes the time to save memory to disk
 - .AVHD(X) is just a latest differencing disk
- Copy (backup) on write block based backup
 - like volume shadow copy (VSS)
- Live vs. offline snapshots
 - live snapshots not supported anywhere
 - offline snapshots **must be supported by applications**



Checkpoints/Snapshots/VSS

| | | | | | | | | | |
|---|---|---|---|---|--|---|---|---|---|
| H | E | L | L | O | | Y | O | U | ! |
|---|---|---|---|---|--|---|---|---|---|



Checkpoints/Snapshots/VSS

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| H | I | | O | N | D | R | E | J | ! |
| | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ | |
| | E | L | L | O | | Y | O | U | |



Checkpoints/Snapshots/VSS

| | | | | | | | | | |
|---|---|---|---|---|--|---|---|---|---|
| H | E | L | L | O | | Y | O | U | ! |
|---|---|---|---|---|--|---|---|---|---|



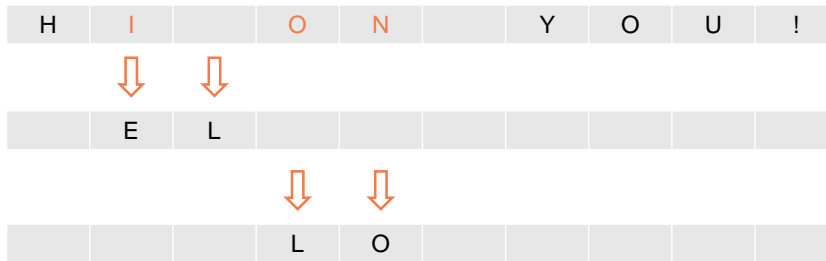
Checkpoints/Snapshots/VSS

| | | | | | | | | | |
|---|---|---|---|---|--|---|---|---|---|
| H | I | | L | O | | Y | O | U | ! |
| | E | L | | | | | | | |

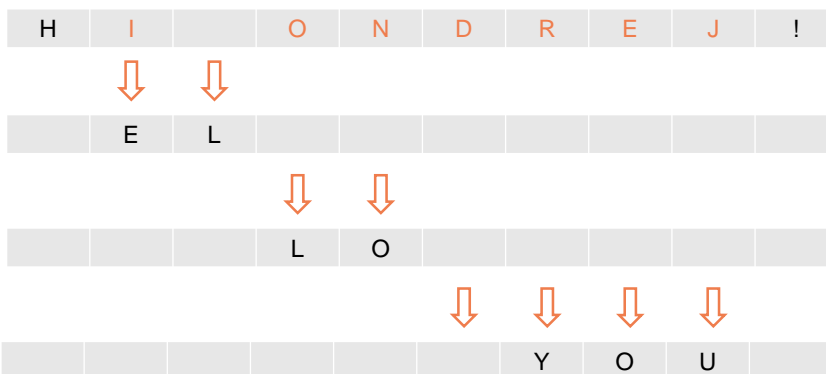
Two orange arrows point downwards from the 'I' and the empty cell following it in the first row to the 'E' and 'L' in the second row.



Checkpoints/Snapshots/VSS



Checkpoints/Snapshots/VSS



Application data inconsistency



Generation ID

- Let the machine know the snapshot has been restored
 - used [just before replication/db-commit](#) etc.
- Device Manager
 - System devices
 - ♦ Microsoft Hyper-V Generation Counter (gencounter.sys)
 - IOCTL to \\.\VmGenerationCounter
- Does not change
 - reboot, resume, host reboot, live migration, fail-over on a cluster
- Changes
 - snapshot, import, copy, clone, backup restored, fail-over after replication



Hyper-V storage migration

- Windows 2012+
 - migrate to SMB share as well
- Copies the disk until mostly done
 - does not use VSS
- Stops writes and redirects them to the new location
- VHD/VHDX and application **consistency guaranteed**
 - outage - pause several tens of milliseconds



Network outages in general

- UDP, ICMP
 - packet losts
 - user must try again
 - DNS, ping
- TCP
 - every packet must be acknowledged
 - retransmissions
 - ♦ SYN - 3, 6, 12 sec = 21 sec.
 - ♦ PSH - 3, 6, 12, 24, 48 sec = 93 sec.



Backup

- Volume Shadow Copy (VSS)
 - snapshot of the parent partition's files
 - calling writers in the child partitions with IC enabled
- System Volume Information
 - on different volume, default on Windows 2012 R2
 - size limit - deletes shadow copy without asking
- DISKPART
 - list shadows all
 - list writers
 - set context persistent, begin backup, add volume, create, end backup
- VSS writer failure in VM fails failure in host



Hyper-V replication

- Windows 2012+
- Copies virtual disk changes in the background
 - once per 15 minutes
- VHD/VHDX consistency guaranteed
- Application data **consistency not guaranteed**
 - does not use VSS primarily (.HRL file)
- Authenticates with Kerberos or TLS/SSL certificate
 - SPN: Hyper-V Replica Service/
 - `netsh http show servicestate`
 - `netsh http show sslcert`



Using Hyper-V replication

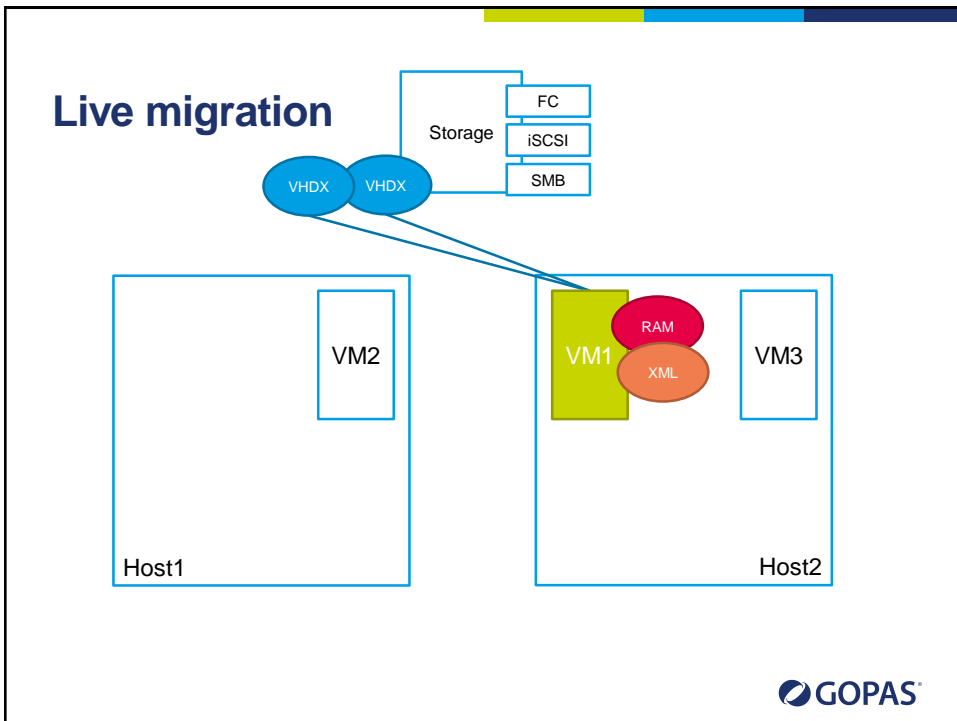
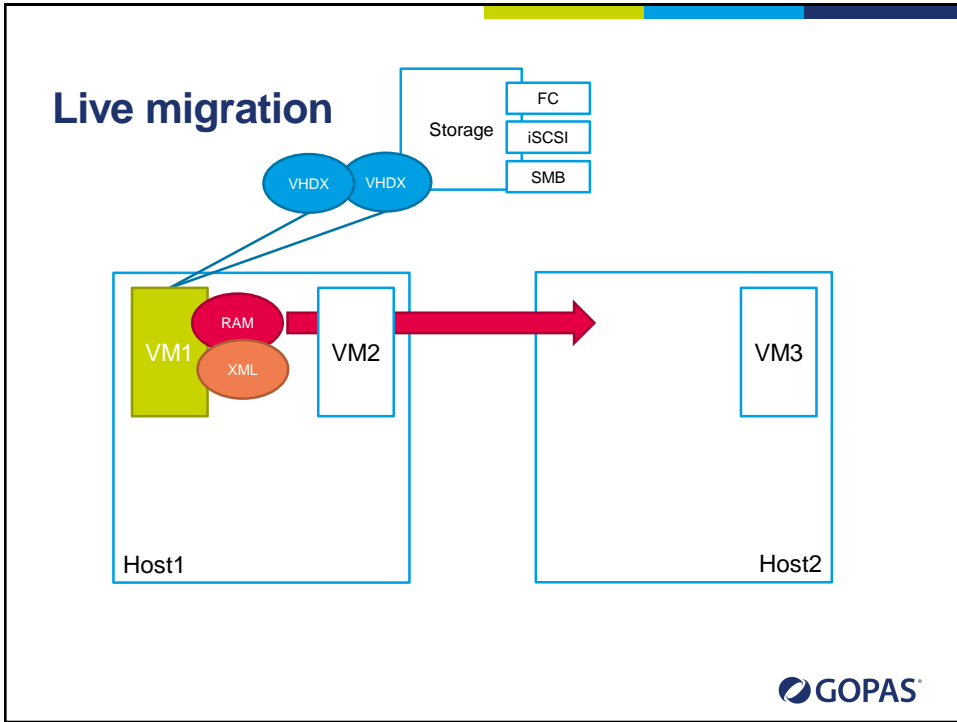
- Cloning
 - make sure you SYSPREP if duplicating
- Migration to a different cluster with minimal interruption
- Planned failover to a backup site
- Recovery **start-easy point** only

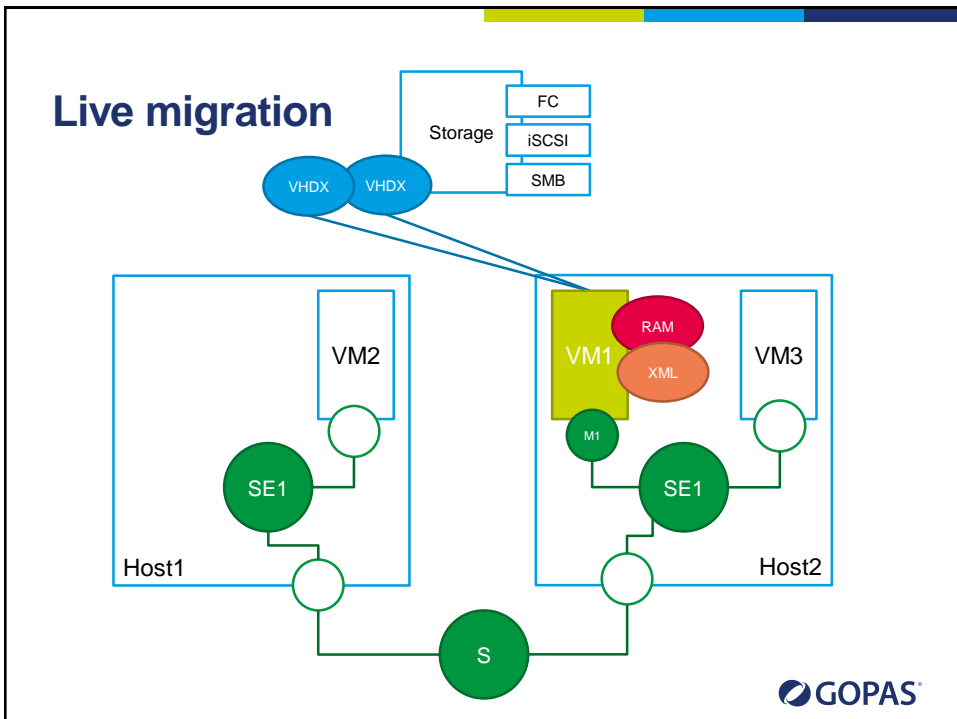
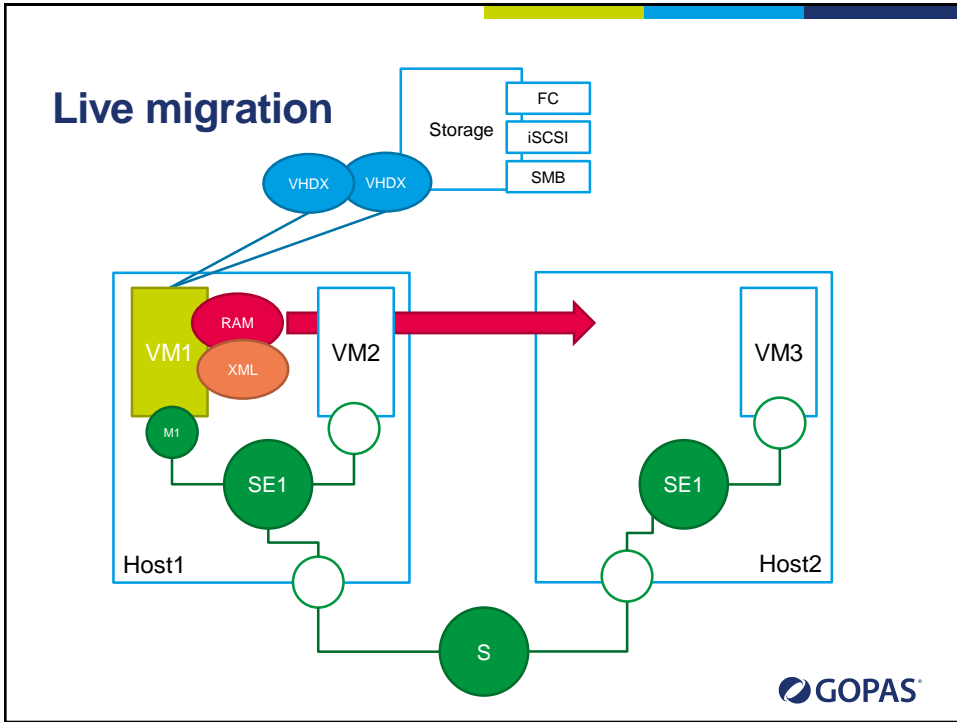


Hyper-V live migration

- Windows 2008 R2
 - only on cluster
- Windows 2012+
 - with any shared storage available
 - cluster or SMB share
- Does not migrate disk data at all
- Copies memory contents until mostly done
 - stops machine for a brief moment to move the rest
 - issues ARP IP conflict detection to update switch ARP tables
- Authenticates with Kerberos constrained delegation (Kerberos only)
 - SPN: Microsoft Virtual System Migration Server/node2
 - SPN: cifs/fsStorage







Operational hints

- Automatic start action
 - delay
- MAC address range
- Cannot export 2008 R2+ and import back to 2008
 - the other way possible
- Generation 2 runs only on 2012 R2
- Script with root\virtualization\v2 since 2012
- Keyboard redirection and mouse release key
- Hot add/remove VHD/VHDX on SCSI controllers since Windows 2008 R2
- Windows 2012 R2+ support extending VHDX when mounted and running in VM
 - must use diskmgmt.msc from the guest as well



Hosting operational hints

- Block DHCP servers
- Block router advertisements
- Enable MAC spoofing for NLB clusters?



SMB file sharing

- SMB 1.0 on Windows 2003-
- SMB 2.0 on Windows Vista and Windows 2008+
 - TCP/network failure resiliency
- SMB 3.0 on Windows 8 and Windows 2012+
 - transparent failover



Failover clustering

- Up to 64 nodes, up to 8000 VMs per cluster
- Manages shared storage
 - SCSI persistent reservation
 - switches the storage among the cluster nodes
- Monitors itself and clustered services
- Determines quorum
 - node majority
 - node majority with witness disk
 - node majority with witness share
 - no majority, only the witness disk



File server clusters

- SMB file server cluster on Windows 2012
 - transparent failover
 - ♦ Resume Key fs filter on the SMB server's persistent storage
 - ♦ SMB witness server and SMB witness client
 - scale-out file server
- Requires SMB 3.0 client
- Performance counters
 - SMB Client Shares
 - SMB Server Shares
 - SMB Server Sessions
- Event logs
 - Application and service logs - Microsoft - SMB Client



Cluster requirements

- No dynamic disks on cluster nodes
 - not even OS disk can be dynamic
 - iSCSI disks can be dynamic
- iSCSI disks in offline mode
- Pre-create cluster's **computer account** in AD
 - Cluster Admins must have **Full control** to the object
- FS witness share with ClusterAccount\$ Full control



Cluster Shared Volumes (CSV)

- Windows 2008 R2 failover cluster
 - only for Hyper-V machines
- Windows 2012+
 - any local NTFS file access
- Use by
 - Hyper-V 2008 R2+
 - SQL server 2014
 - scale-out FS on 2012+



CSV details

- Mounted directly only on a single node
 - Coordinator Node (CN)
 - others are Data Servers (DS)
- C:\ClusteredStorage\Volume
 - can be renamed
- NTFS metadata modifications go through CN
 - create file/folder, delete, change size, etc.
 - hidden SMB share
 - [Get-SmbShare -IncludeHidden](#)
- NTFS data read/write go directly from DS



VM monitoring by failover cluster

- Heartbeat monitoring (Windows 2008 R2+)
 - IC service
 - **corrected with** VM restarted after 1 minute
- Monitoring services inside VMs (Windows 2012+)
 - VM properties in Failover Cluster Manager
 - guest Windows 2012/8+
 - firewall rule enabled in guest (Virtual Machine Monitoring)
 - **first corrective action** restarts the guest OS
 - **second corrective action** also moves the guest to another node
- Protected network (Windows 2012 R2+)
 - VM NIC properties
 - External switches only
 - **corrective action** is to move the machine



Failover cluster woes

- Renaming **VM** does not rename **resource group** nor resources
 - Get-ClusterResource | Get-ClusterParameter VmId
 - Get-Vm
- Deleting VM from Hyper-V leaves the resource **offline** in failover cluster

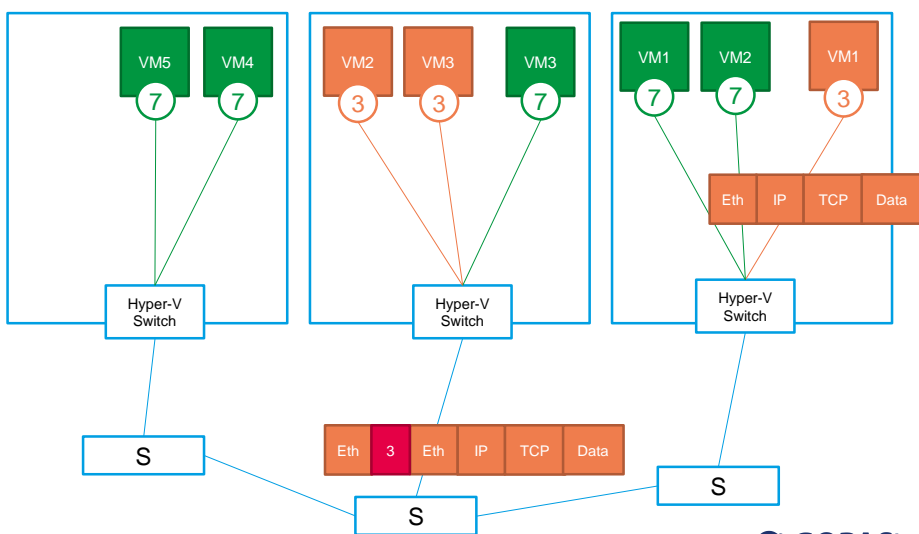


Hyper-V network virtualization

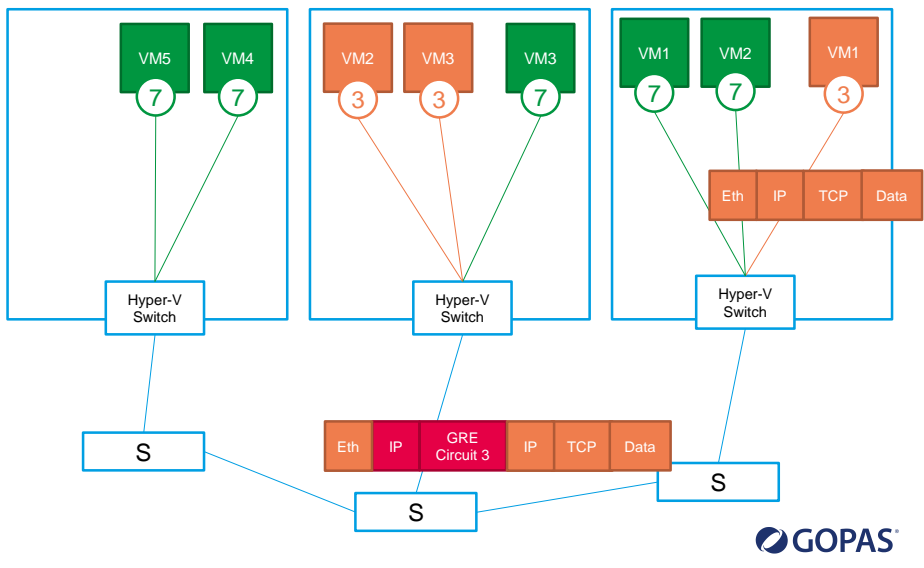
- IP tunnels among Hyper-V hosts
- Tunneled in NVGRE (GRE IP protocol 47)
- Instead of VLANs
 - limited to 4096, usually less
 - require switch support and configuration



NV (Network Virtualization) - 802.1q



NV (Network Virtualization) - NVGRE



GOPAS®

IP addresses in Hyper-V NV

- Virtual subnet (circuit) ID (VSID)
 - 4096 - 16 777 214
- Provider (physical) addresses (PA)
- Customer addresses (CA)

GOPAS®

NV rules

- CA rule (NetVirtualizationLookupRecord)
 - VSID, source CA, source MAC => PA, encap
- Customer route rule (NetVirtualizationCustomerRoute)
 - VSID, RoutingDomainID, IP destination network
 - RoutingDomainID - set of IP subnets that are reachable from each other
- PA rule (NetVirtualizationProviderAddress)
- Provider route rule (NetVirtualizationProviderRoute)
 - PA routing



Verify NV settings

```
Get-NetVirtualizationGlobal
```

```
Get-NetVirtualizationLookupRecord
```

```
Get-NetVirtualizationCustomerRoute
```



Configuring network virtualization

Do this for all physical NICs (switches)

```
Get-NetAdapter

# on Windows 2012 only
# always enabled on Windows 2012 R2+
Enable-NetAdapterBinding <host's NIC name> -ComponentID ms_netwv
```



Configuring network virtualization

Do this for each VM in each network

```
$vmNICPort = Get-VMNetworkAdapter -VMName <the VM> | where
{$_ .MacAddress -eq <VM's MAC> }

# disable with VirtualSubnetId = 0
Set-VMNetworkAdapter -Input $vmNICPort -VirtualSubnetID <some ID>
```

- Stored withing VM XML config
 - Live migration migrates the setting as well



Configuring network virtualization

Do this per each external switch

```
$nic = Get-NetAdapter <Host's physical NIC name>

New-NetVirtualizationProviderAddress
    -InterfaceIndex $nic.InterfaceIndex
    -ProviderAddress <hyper-V host's IP address>
    -PrefixLength 24

# unnecessary if no routing necessary on local subnet
New-NetVirtualizationProviderRoute
    -InterfaceIndex $nic.InterfaceIndex
    -DestinationPrefix 0.0.0.0/0
    -NextHop <hyper-v host's default gateway>
    # or 0.0.0.0 for directly attached
```



Configuring network virtualization

Do this same on both Hyper-V nodes

```
# call this twice for both Hyper-V hosts' IP address
New-NetVirtualizationLookupRecord
    -CustomerAddress <internal virtual IP address from VM>
    -VirtualSubnetId <some subnetID>
    -MACAddress <VM NIC MAC>
    -ProviderAddress <Hyper-V host's IP address>
    -Rule TranslationMethodEncap
    # use -Type L2Only with -CustomerAddress 0.0.0.0

New-NetVirtualizationCustomerRoute
    -RoutingDomainId [GUID]::NewGUID()
    -VirtualSubnetId <some subnetID>
    -DestinationPrefix <customer address subnet 10.10.0.0/16>
    -NextHop 0.0.0.0 # = directly attached
    -Metric 255
```



Multitenant VPN

```
# On HOST:
#
# Get-VmNetworkAdapterIsolation
# Get-VmNetworkAdapterRoutingDomainMapping
#     -IsolationId VirtualSubnetID
#     -RoutingDomainId
#     -RoutingDomainName
# Get-NetCompartment

$red = 'RED Lenin'
$blue = 'Blue'

Add-WindowsFeature -Name Remoteaccess -IncludeAllSubFeature -IncludeManagementTools
Import-Module RemoteAccess ; Install-RemoteAccess -MultiTenancy ; Start-Service
RemoteAccess

Enable-RemoteAccessRoutingDomain -Name BLUE -Type Vpn
Enable-RemoteAccessRoutingDomain -Name RED -Type Vpn

Set-RemoteAccessRoutingDomain -Name BLUE -IPAddressRange 192.168.30.211,192.168.30.2229 -
TenantName BLUEZakaznik
Set-RemoteAccessRoutingDomain -Name RED -IPAddressRange 192.168.30.211,192.168.30.2229 -
TenantName REDZakaznik

Set-VpnAuthType -Type ExternalRadius -RadiusServer DC1RADIUSSErver -SharedSecret
'Pa$$w0rd'
# RADIUS attribute Standard:Class:TenantName = BLUE
```

