



Server Infrastructure  
Days 2013

# Configure Networking in Virtual Environment

Andrea Mauro

@Andrea\_Mauro

MCSE Private Cloud

MCSE Server Infrastructure



#wssid201

3



# Agenda

// WS05

- Introduzione
- Design del networking per Hyper-V
- Cenni su ottimizzazioni e disponibilità
- Implementazione ed esempi

# Elementi di una rete

- NIC (Network Interface Card)
  - pNIC (physical NIC)
  - vNIC (virtual NIC)
  - tNIC (team NIC)
- Switch
  - Physical switch(es)
  - Virtual switch(es)

# Hyper-V Virtual NIC

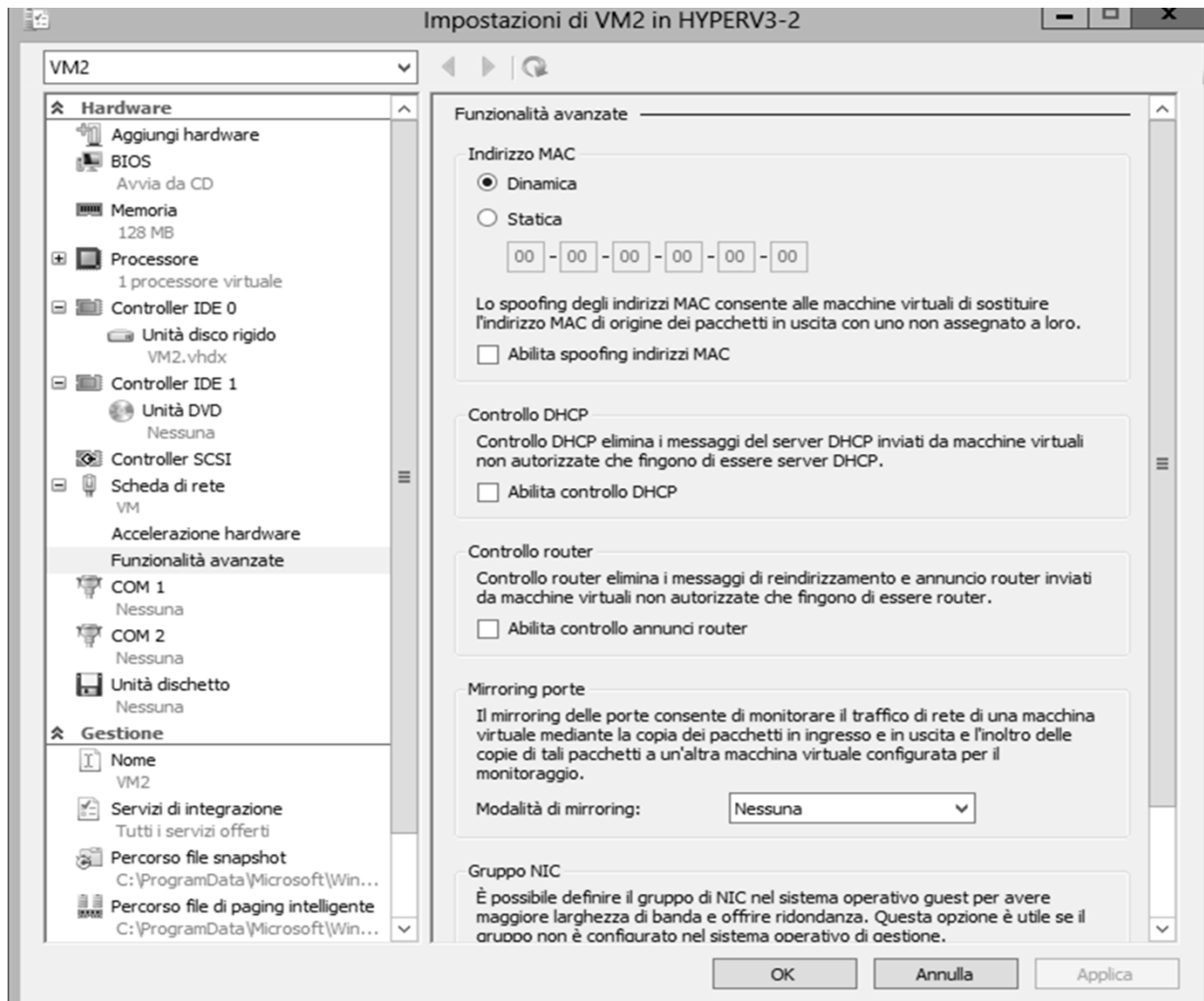
- Legacy interface
  - Utilizza un driver emulato
  - Supporta pre-boot execution (PXE)
  - È compatibile con molti sistemi operativi
  - Ma non supporta sistemi operativi a 64 bit
- Synthetic Interface
  - Richiede Hyper-V integration components
  - Utilizza il synthetic driver stack
  - VMBus che comunica usando shared memory
  - Prestazioni migliori e ottimizzazioni
  - Supporta VLAN tagging e altro

# Hyper-V Offloads

The screenshot shows the 'Impostazioni di VM2 in HYPERV3-2' window. The left sidebar lists hardware components: Aggiungi hardware, BIOS (Avvia da CD), Memoria (128 MB), Processore (1 processore virtuale), Controller IDE 0 (Unità disco rigido VM2.vhdx), Controller IDE 1 (Unità DVD Nessuna), Controller SCSI, Scheda di rete VM, Accelerazione hardware (selected), Funzionalità avanzate (COM 1, COM 2, Unità dischetto), and Gestione (Nome VM2, Servizi di integrazione). The main pane shows 'Accelerazione hardware' settings:

- Accelerazione hardware**  
Specificare le attività di rete di cui è possibile eseguire l'offload a una scheda di rete fisica.
- Coda macchine virtuali**  
Per Coda macchine virtuali è necessaria una scheda di rete fisica che supporti tale funzionalità.  
 Abilita coda macchine virtuali
- Offload attività IPsec**  
Per l'offload di attività IPsec è necessario il supporto di una scheda di rete fisica e del sistema operativo guest.  
Se non sono disponibili risorse hardware sufficienti, le associazioni di sicurezza non verranno scaricate e verranno gestite dal sistema operativo guest nel software.  
 Abilita offload attività IPsec  
Selezionare un valore compreso tra 1 e 4096 per il numero massimo di associazioni di sicurezza scaricate.  
Numero massimo:  Associazioni di sicurezza scaricate
- Single-Root I/O Virtualization**  
La funzionalità Single-Root I/O Virtualization (SR-IOV) richiede hardware specifico. Potrebbe inoltre essere necessario installare altri driver nel sistema operativo guest.  
Se non sono disponibili risorse hardware sufficienti, la connettività di rete verrà fornita attraverso il commutatore virtuale.  
 Abilita SR-IOV

# Hyper-V Network Security





# Hyper-V Bandwidth Management

The screenshot displays the 'Impostazioni di VM2 in HYPERV3-2' window. On the left, the 'Hardware' section is expanded to show the 'Scheda di rete' (Network Adapter) settings, which are currently set to 'VM'. The main pane shows the 'Scheda di rete' configuration options:

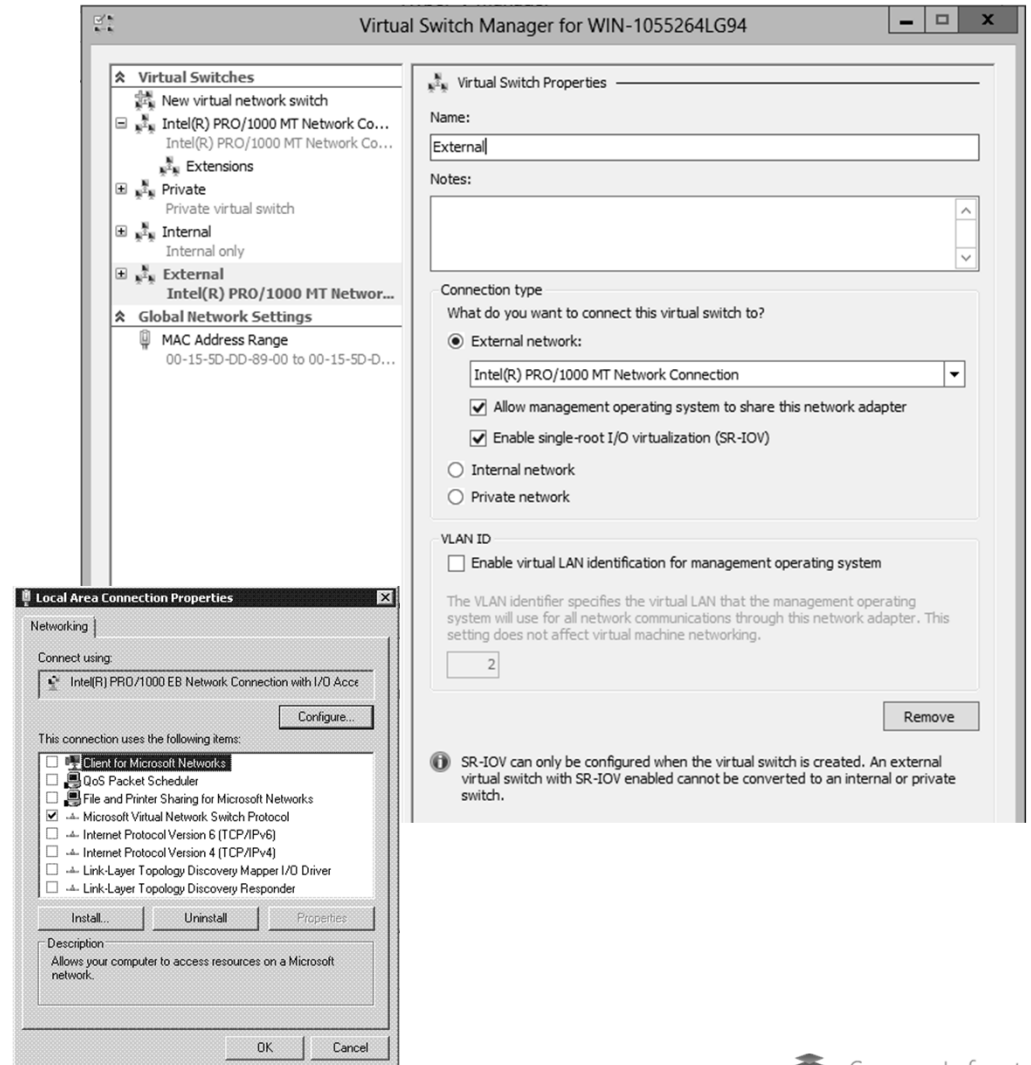
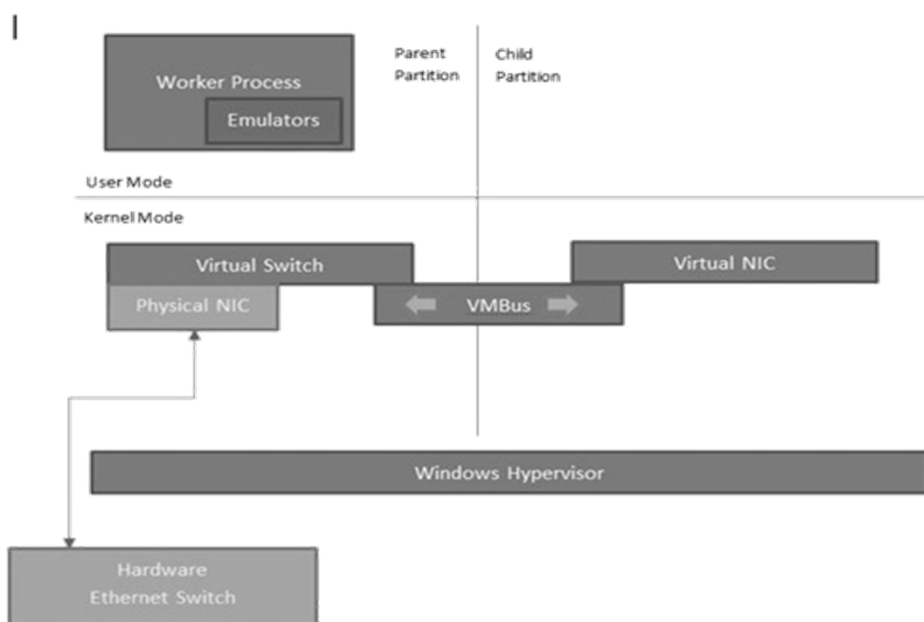
- Scheda di rete:** Specificare la configurazione della scheda di rete o rimuovere la scheda di rete.
- Commutatore virtuale:** VM
- ID VLAN:**  Abilita identificazione LAN virtuale. L'identificatore VLAN specifica quale LAN virtuale verrà utilizzata da questa macchina virtuale per tutte le comunicazioni di rete attraverso questa scheda di rete. Value: 2
- Gestione larghezza di banda:**  Abilita gestione larghezza di banda. Specificare come la scheda di rete utilizza la larghezza di banda della rete. La larghezza di banda minima e la larghezza di banda massima sono misurate in megabit al secondo.
  - Larghezza di banda minima: 0 Mbps
  - Larghezza di banda massima: 0 Mbps

Per lasciare senza restrizioni il valore minimo o massimo, specificare 0.

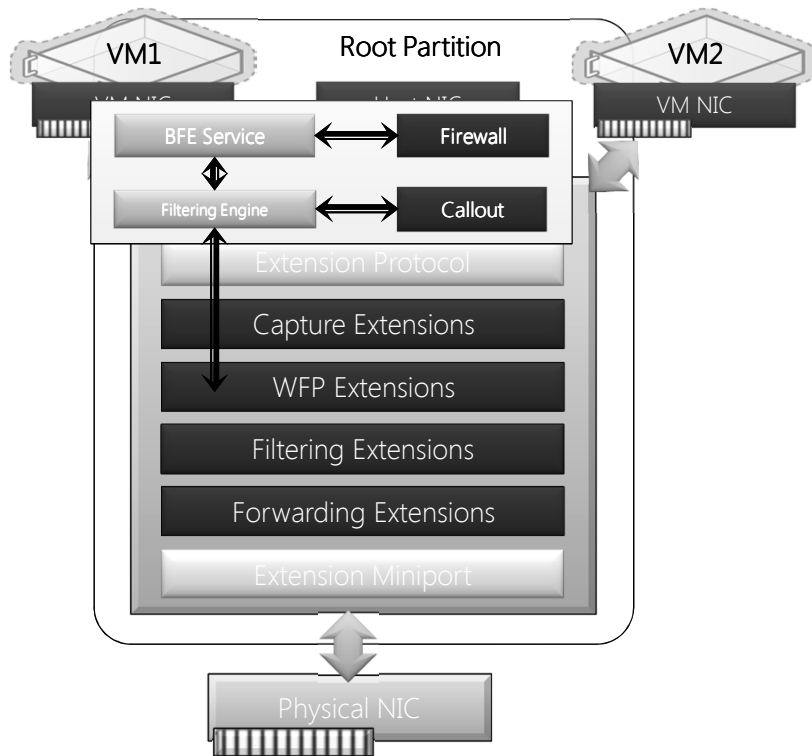
# Virtual Switch

- Configurazione a livello di Hyper-V

- External
- Private
- Internal



# Hyper-V Extensible Switch



- Forwarding extensions can define the next hop (destination) of each packet using WFP APIs
- Filtering extensions can capture and filter traffic using WFP APIs
- Capture extensions do not modify existing extensible switch traffic
- Examples:
  - Example: Cisco Nexus 1000V and UCS
  - Example: Virtual Firewall by 5NINE Software
  - Example: NEC ProgrammableFlow's vPFS OpenFlow

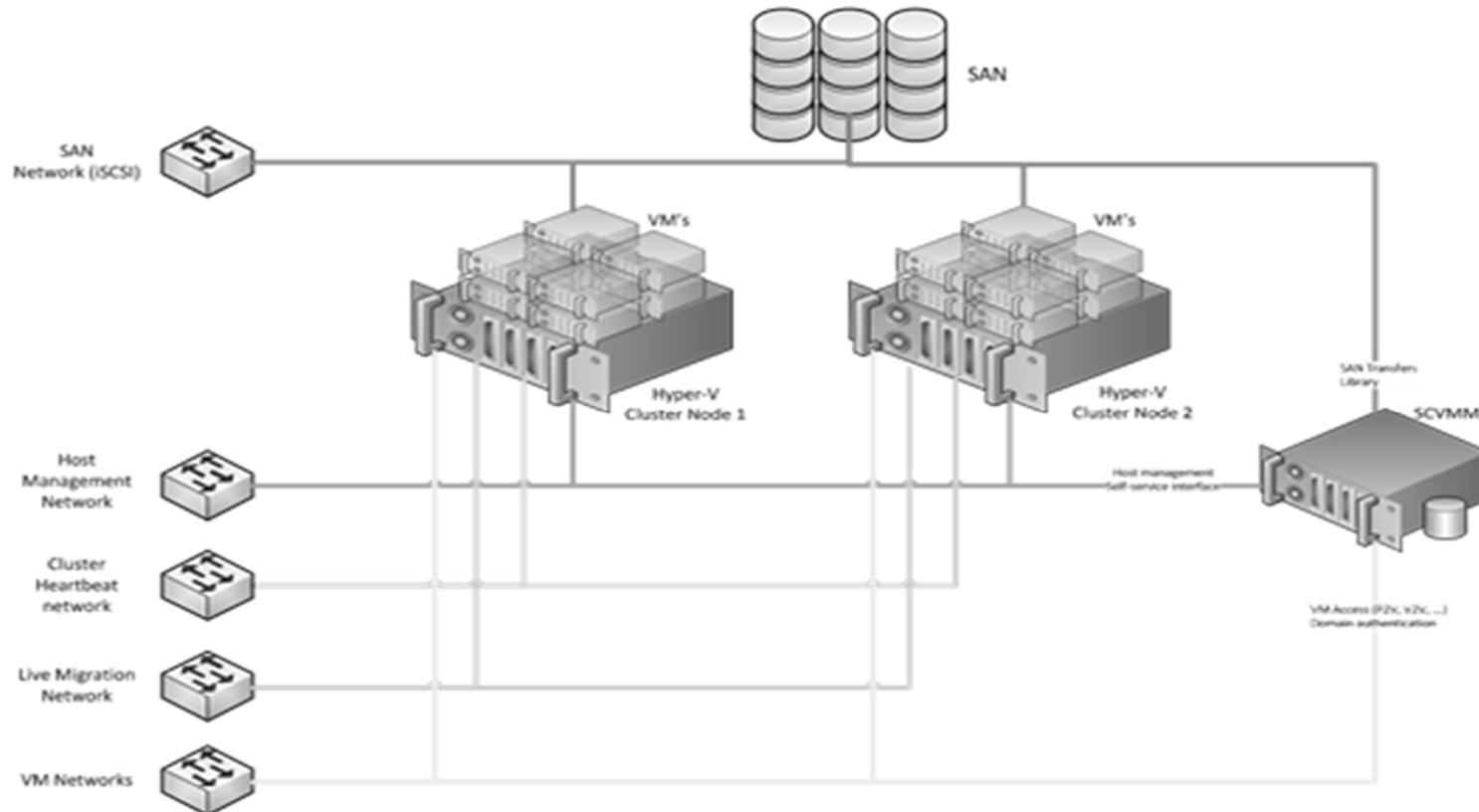


# Cisco Nexus 1000v for Hyper-V

- <http://www.thomasmaurer.ch/2013/06/cisco-nexus-1000v-for-hyper-v-now-available-for-download/>
  - Advanced virtual machine networking based on Cisco NX-OS operating system and IEEE 802.1Q switching technology
  - Cisco vPath technology for efficient and optimized integration of virtual network services
  - Tight integration with System Center Virtual Machine Manager 2012 SP1
  - Layer 2 Switching with Transmit side Rate Limiting
  - Security Policy Mobility, inbuilt support for Private VLANs with local PVLAN Enforcement
  - Provisioning Port Profiles with deep Integration with SCVMM
  - Traffic Visibility, including VM Migration Tracking, NetFlow v.9 with NDE, Cisco Discovery Protocol v.2
  - And more...

# Reti per un cluster Hyper-V

- Per le VM solo reti "public"
- Varie nuove reti necessarie alle funzioni del cluster



# Tipi di reti

- Rete di gestione (Management)
  - Tutto quello che riguarda la parent partition
    - inclusi i backup e le repliche
- Almeno una rete pubblica per le VM
- Rete privata per il Failover Cluster
  - Rete di Heartbeat
- Opzionale (ma consigliate)
  - Rete privata per Live Migration (LM)
  - Rete privata per il CSV
- Se richiesto: rete per IP storage
  - iSCSI
  - SMB3

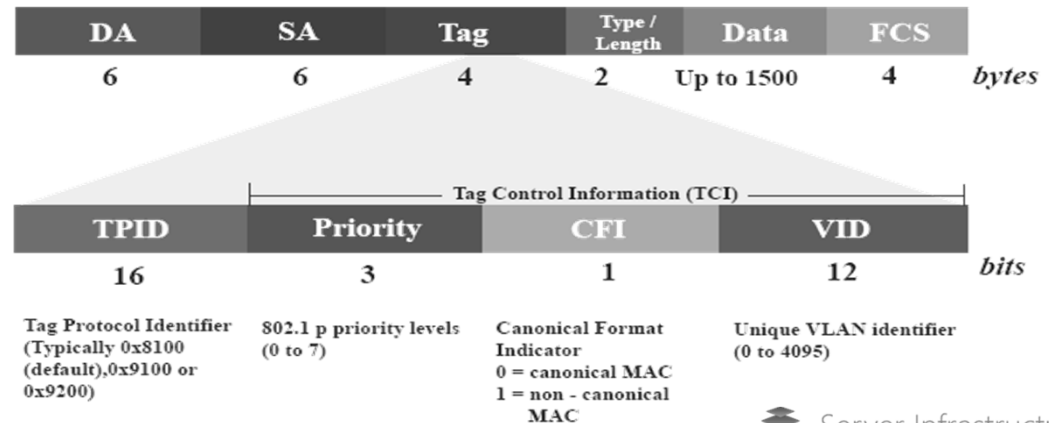
# Perché reti diverse?

- Aspetti di prestazioni e sicurezza
  - Separare i differenti tipi di pacchetti
  - Separare differenti tenant
- Come isolare e come garantire la banda
  - NIC separate
  - NIC condivise ma con
    - VLAN, Port ACLs, QoS, DCB & VM QoS
  - E per il virtual switch?

# Segmentare le reti con le VLAN

- Virtual LAN (802.1Q)
  - Più reti logiche (isolate) su una stessa fisica
  - VLAN ID 1-4094
  - Concetto di pNIC virtuale
  - Richiedono configurazione lato switch
    - Problema di prestazioni?
    - Compatibilità con Jumbo Frames?
- Private VLAN

```
vlan database
vlan 100,200
exit
interface ethernet 1/g1
switchport mode trunk
switchport trunk allowed vlan add 100,102
exit
interface ethernet 1/g2
switchport access vlan 100
exit
```



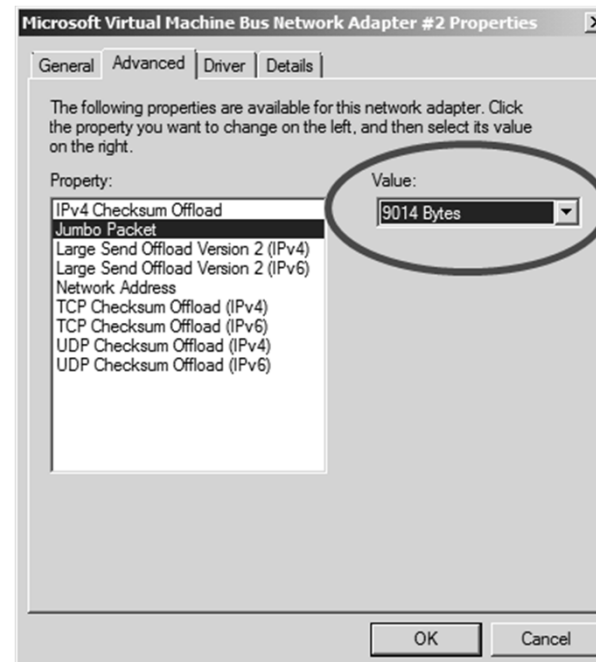


# Jumbo Frames

- Frame layer 2 di MTU maggiore
  - Ethernet frames da 1,500 bytes a ~9k
    - Dimensione esatta?
  - Aumenta throughput
  - Riduce l'utilizzo della CPU per grossi trasferimenti
- Dove usarlo?
  - iSCSI
  - LM?
- Considerazioni

```
interface ethernet 1/g1
  mtu 9216
exit

C:\>ping -l 9000 -f IP
```



# Tipi di NIC

- 1 Gb Ethernet
  - Normalmente adeguate per molti workload
- 10 Gb Ethernet
  - Ottime prestazioni
  - RDMA opzionale (per SMB 3.0 file access)
  - Diversi tipi di offload
  - Quality of Service (DCB) & Flexible bandwidth allocation
  - ...Ma soluzione ancora relativamente costosa
- InfiniBand (32 Gb e 56 Gb)
  - Prestazioni elevate, latenza minima
  - RDMA incluso (for SMB 3.0 file access)
  - ...Ma la gestione è diversa dall'Ethernet
- Convergenti



1GbE



10 GbE



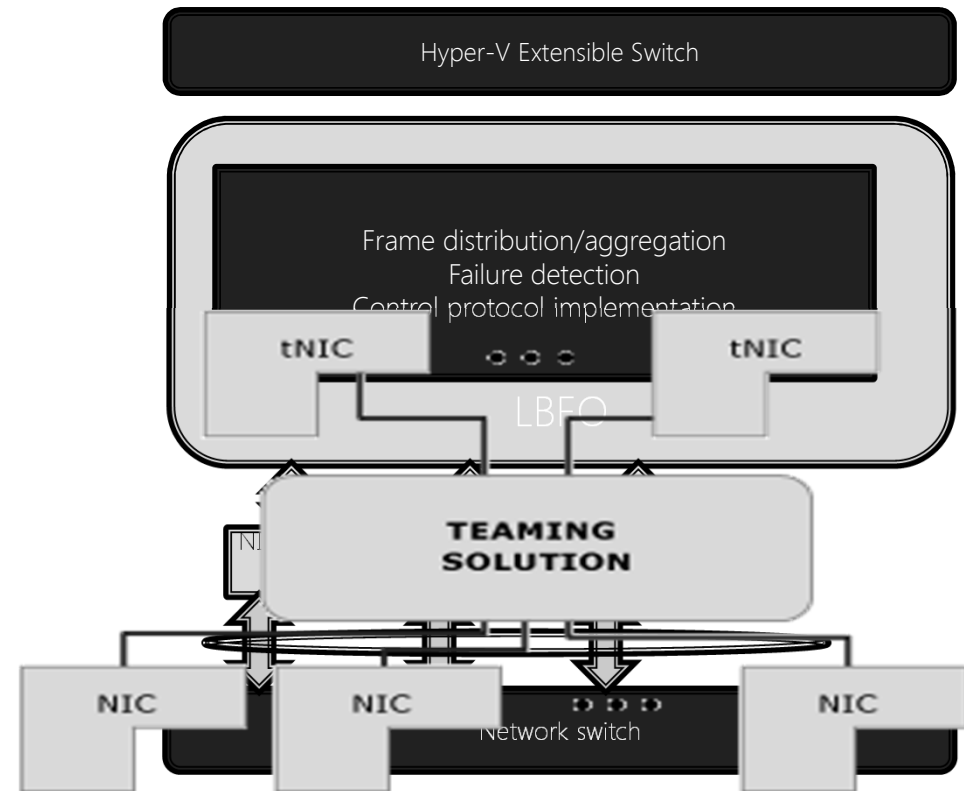
InfiniBand

# Ridondanza delle NIC

- Alcune reti richiedono ridondanza
  - Altre no (come quella di heartbeat)
- Ridondanza -> almeno 2 NIC
- Prestazioni -> molte NIC
- Teaming o non teaming?
  - Integrato in Windows Server 2012

# NIC Teaming

- Virtualizzazione delle NIC fisiche sotto forma di una scheda logica (team)
- Altri nomi
  - NIC bonding
  - Network adapter teaming
  - Load balancing and failover
  - ...



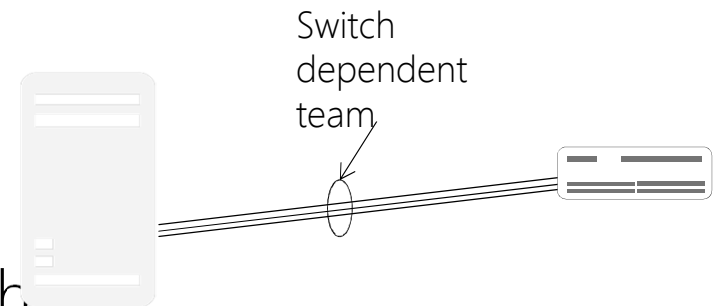
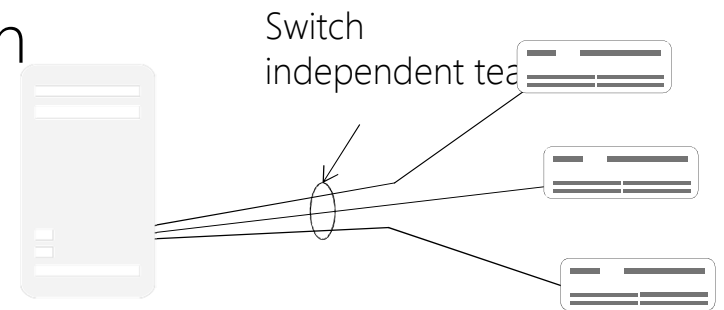
# NIC Teaming - Terminologia

- Funzionalità di Windows Server 2012
  - Load Balancing/Failover (LBFO)
- Team
  - L'intero «costrutto»
- Team member
  - NIC, pNIC, Network Adapters
- Team interface
  - tNIC, Team NIC

```
New-NetLbfoTeam MyTeam NIC1,NIC2  
Add-NetLbfoTeamNIC -Team NetTeam01 -VlanID 10
```

# Switch independent vs. dependent

- Switch independent mode
  - Doesn't require any configuration of a switch
  - Protects against adjacent switch failures
- Switch dependent modes
  - Generic or static teaming
    - IEEE 802.3ad draft v1
  - Dynamic teaming
    - IEEE 802.1ax
    - Also known as LACP or 802.3ad
- Requires configuration of the adjacent switch



# LAG, Channel, LACP, Etherchannel, ...

- Alcuni standard
- Diverse implementazioni
  - Cisco's proprietary Port Aggregation Protocol (PAgP)
  - IEEE standard Link Aggregation Protocol (LACP)
- Collegamento switch to switch
  - Switch fisici o logici («stack»)

```
interface ethernet 1/g23
channel-group 1 mode on
exit
interface ethernet 1/g24
channel-group 1 mode on
exit
interface port-channel 1
switchport mode trunk
exit
```

# Load balancing

- In Windows Server 2012 sono supportate le seguenti tecniche di distribuzione del traffico
  - Hyper-V switch port
    - VM have independent media access control (MAC) addresses, the MAC address of the virtual machine can provide the basis for dividing traffic
    - Windows Server 2012 uses the Hyper-V switch port as the identifier rather than the source MAC address
  - Hashing
    - The components that can be used as inputs to the hashing function include the following:
      - Source and destination MAC addresses
      - Source and destination IP addresses, with or without considering the MAC addresses (2-tuple hash)
      - Source and destination TCP ports, usually used along with the IP addresses (4-tuple hash)



# Tipi di switch e di load balancing

	<b>Address Hash</b>	<b>Hyper-V port</b>
Switch Independent	Sends on all active members, receives on one member (primary member)	Sends on all active members, receives on all active members, traffic from same port always on same NIC
Switch Dependent	Sends on all active members, receives on all active members, inbound traffic may use different NIC than outbound traffic for a given stream (inbound traffic is distributed by the switch)	All outbound traffic from a port will go on a single NIC. Inbound traffic may be distributed differently depending on what the switch does to distribute traffic

# Teaming vs. NIC optimizations

Feature	Comments
RSS	Programmed directly by TCP/UDP when bound to TCP/UDP.
VMQ	Programmed directly by the Hyper-V switch when bound to Hyper-V switch
IPsecTO, LSO, Jumbo frames, all checksum offloads (transmit)	Yes – advertised if all NICs in the team support it
RSC, all checksum offloads (receive)	Yes – advertised if any NICs in the team support it
DCB	Yes – works independently of NIC Teaming
RDMA, TCP Chimney offload	No support through teaming
SR-IOV	Teaming in the guest allows teaming of VFs
Network virtualization	Yes

# Teaming in Windows Server 2012

**ADAPTERS AND INTERFACES** TASKS ▾

Network Adapters | Team Interfaces |

Adapter	Speed	State	Reason
▲ Available to be added to a team (4)			
NIC1	1 Gbps		
NIC2	1 Gbps		
NIC3	1 Gbps		
NIC4	1 Gbps		

- Add to New Team
- Add to Selected Team
- Disable
- Properties

^ Additional properties

**Teaming mode:** Switch Independent ▾

**Load balancing mode:** Address Hash ▾

**Standby adapter:** None (all adapters Active) ▾

**Primary team interface:** MyTeam: Default VLAN

**NIC Teaming** ⌵

### New team

Team name: MyTeam

Member adapters:

In Team	Adapter	Speed	State	Reason
<input type="checkbox"/>	NIC1	1 Gbps		
<input checked="" type="checkbox"/>	NIC2	1 Gbps		
<input checked="" type="checkbox"/>	NIC3	1 Gbps		
<input type="checkbox"/>	NIC4	1 Gbps		

⌵ Additional properties

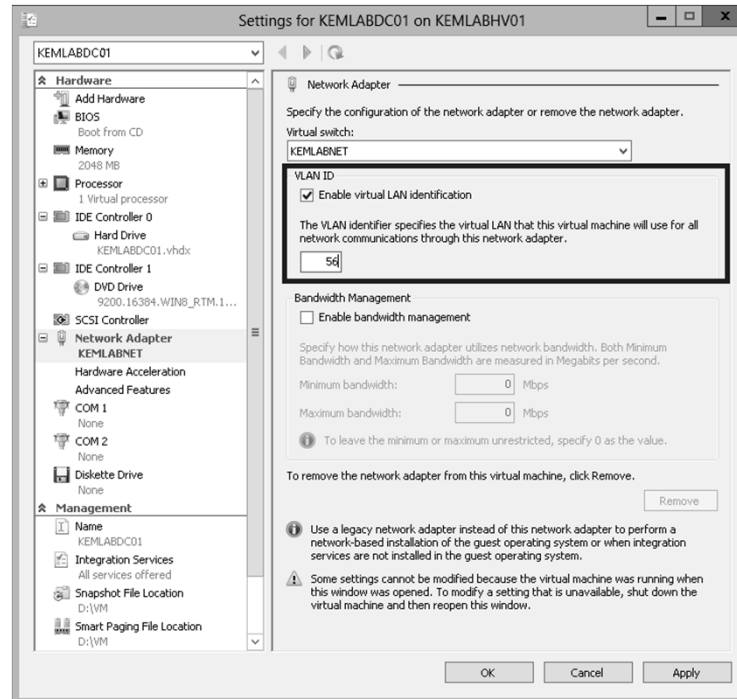
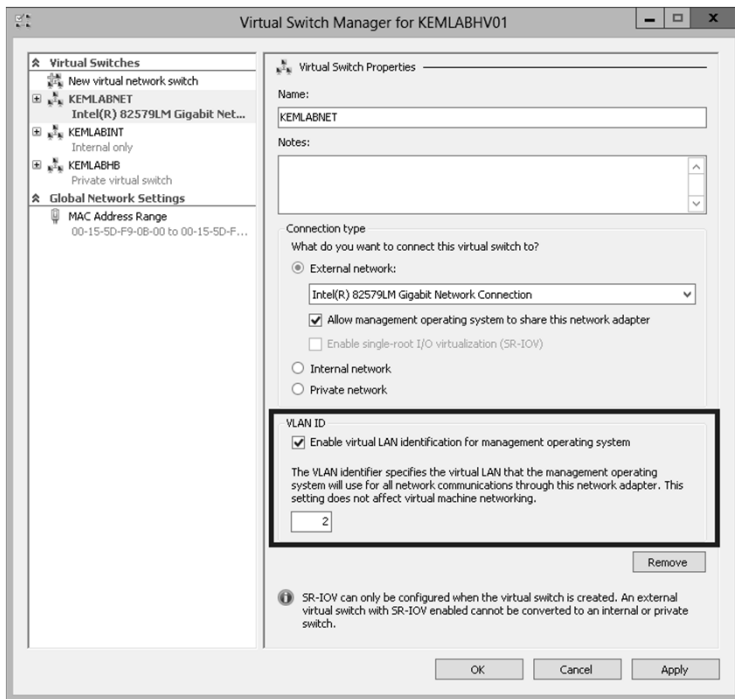
OK Cancel

# Teaming di NIC... ma non sempre

- Alcuni casi nel quale non va usato
  - Rete «privata» del Fail-Over Cluster
  - Storage SAN basato su iSCSI
    - Si utilizza il multi-path
  - Storage NAS basato su SMB3
    - Si utilizza SMB Multichannel

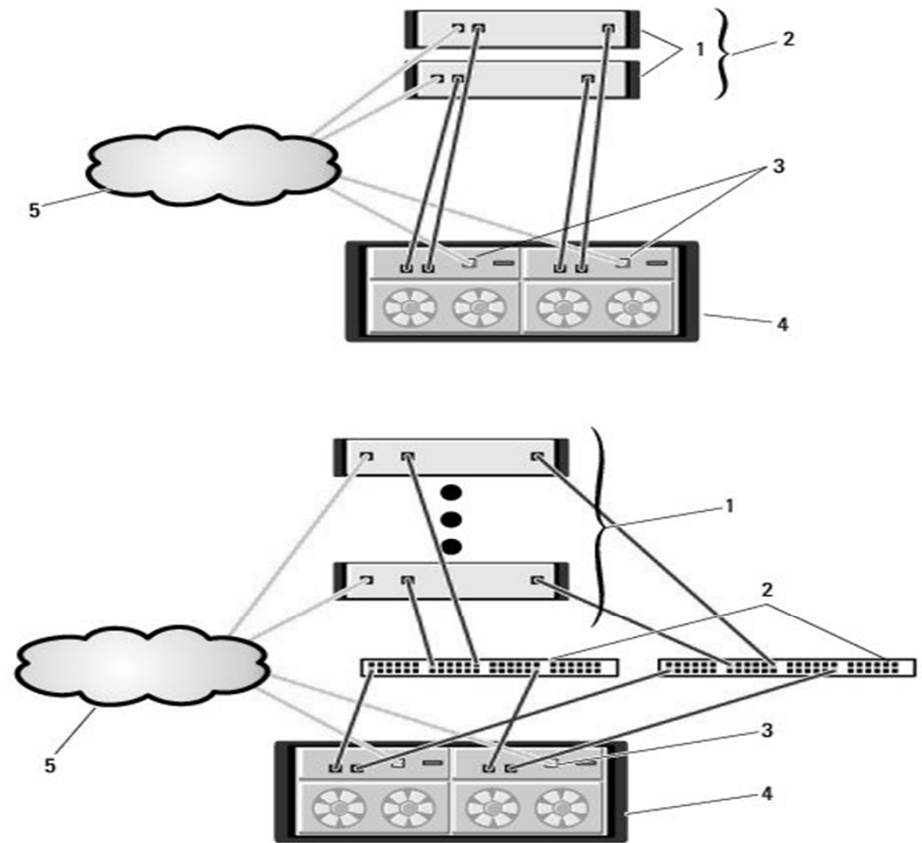
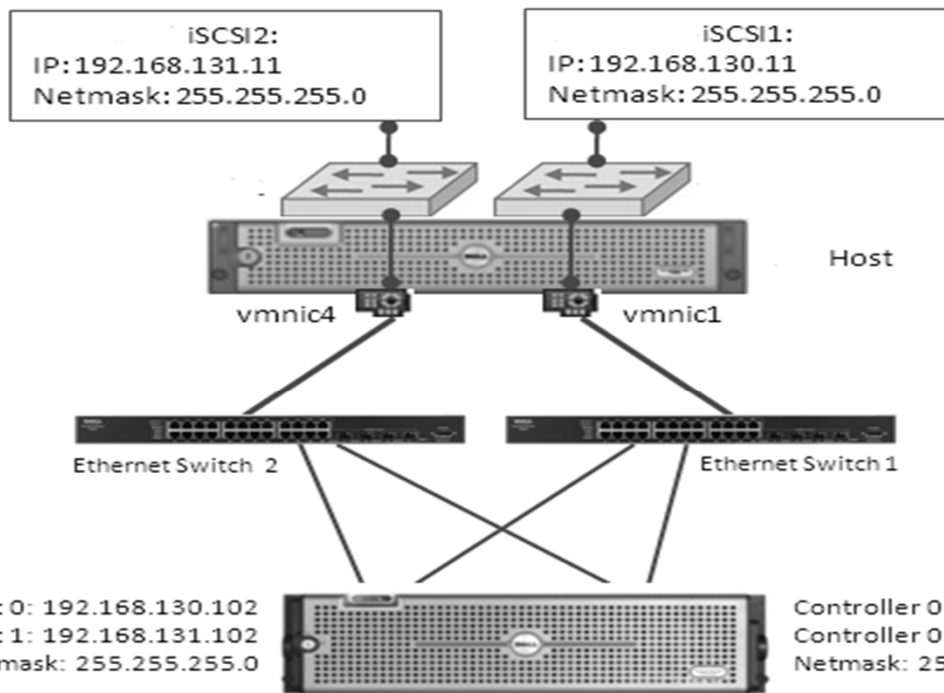
# VLAN dilemma

- Quando non usarle?
- VLAN nel team o nel virtual switch?



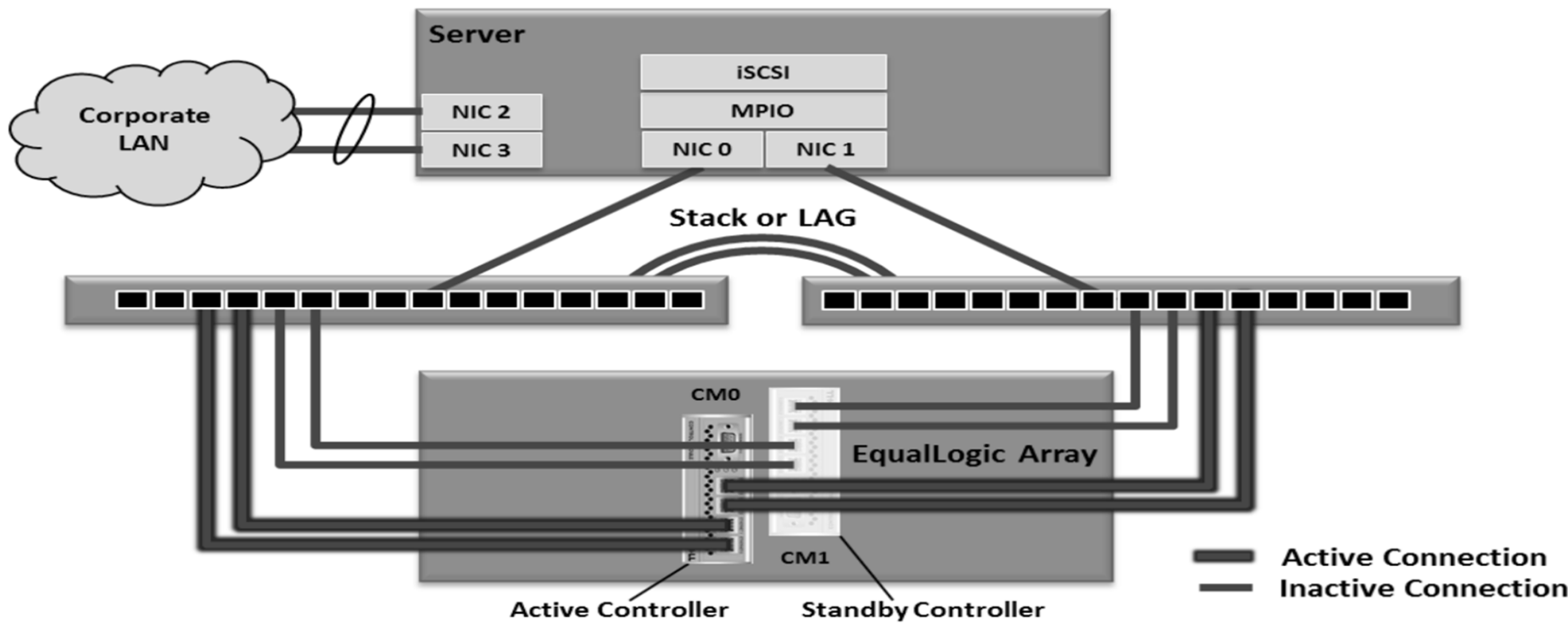
# Multi-path

- Topologia simil-FC
- Reti separate



# Multi-path

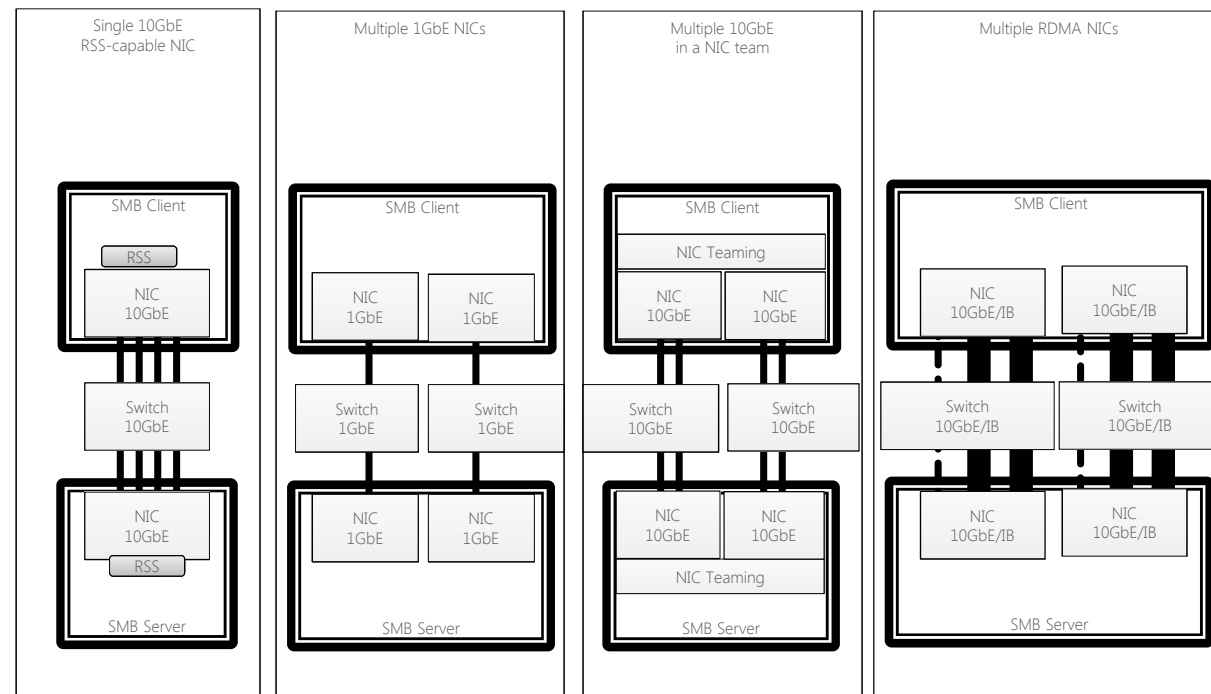
- Soluzione specifica per alcuni storage iSCSI
- Rete flat



# SMB Multichannel

- Soluzione di “multi-path” per SMB3
- Automatic Failover
  - SMB Multichannel implements end-to-end failure detection
  - Leverages NIC teaming (LBFO) if present, but does not require it
- Full Throughput
  - Bandwidth aggregation with multiple network interface cards (NIC)
  - Multiple CPUs cores engaged when using Receive Side Scaling (RSS)
- Automatic Configuration
  - SMB detects and uses multiple network paths

Sample Configurations



Vertical lines are logical channels, not cables



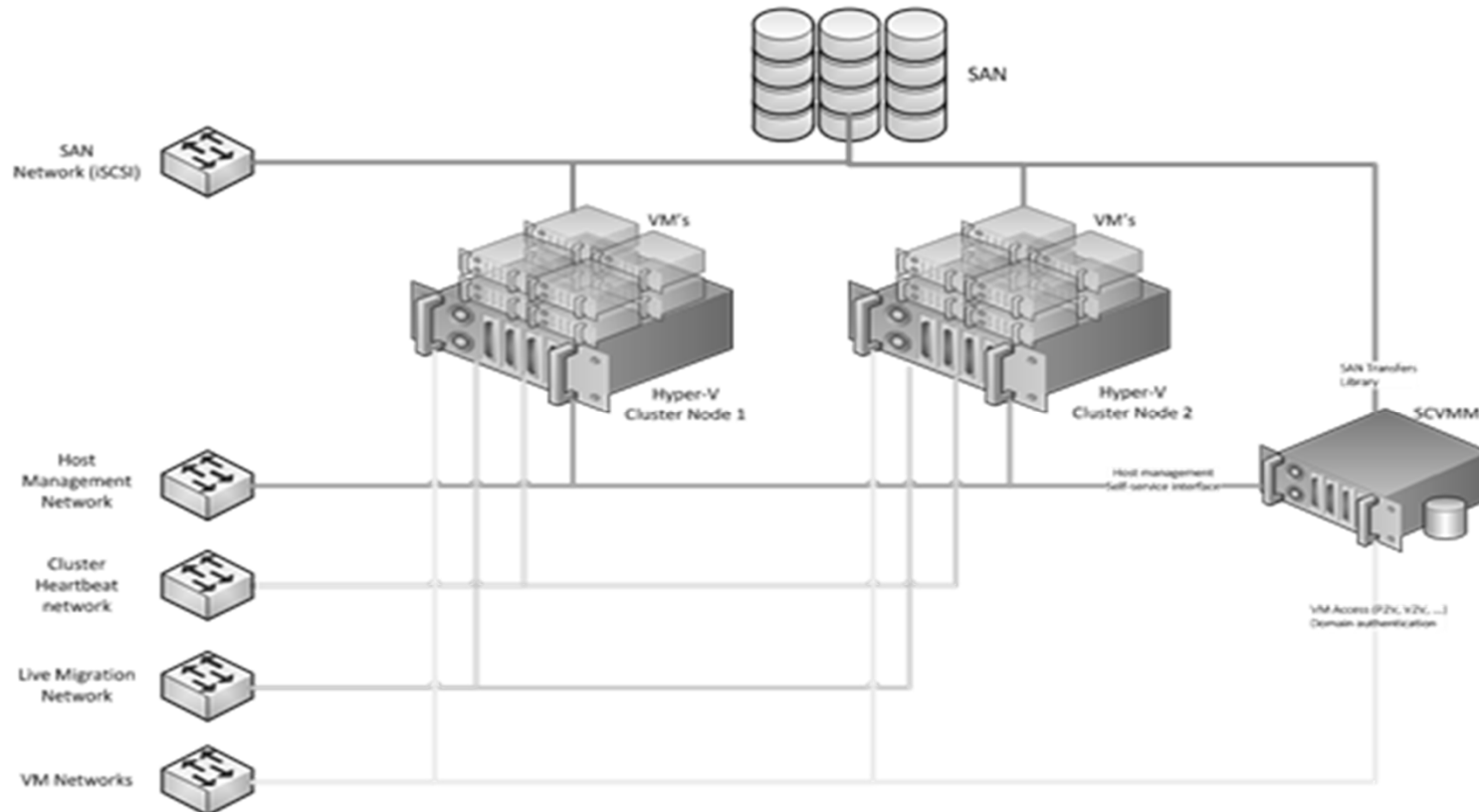
# Spanning Tree Protocol

- Lo spanning tree è un algoritmo utilizzato per realizzare reti complesse (a livello fisico) con percorsi ridondanti
- Tempi di convergenza?
- Necessario?
- Configurazione delle porte degli host

```
interface ethernet 1/g1
  spanning-tree portfast
exit
```

# Reti per un cluster Hyper-V

- Per le VM solo reti "public"
- Varie nuove reti necessarie alle funzioni del cluster



# Tipi di reti necessarie

	Accesso	Sicurezza	Latenza	Banda usata
Gestione	Management	Da gestire	Media	Bassa
VM	Pubblico*	Dipende*	Dipende	Dipende
Privata (cluster)	Isolata	Minima*	Media	Bassa
Live Migration	Isolata	Da gestire	Bassa	Alto
CSV	Isolata	Minima*	Bassa	Basso, a volte alto
iSCSI	Isolata	Minima*	Bassa	Alta
SMB	Isolata	Alta	Bassa	Alta
Backup	Management	Da gestire	Media	Variabile*
Replica	Pubblico*	Alta	Media	Variabile



# Schede di rete necessarie

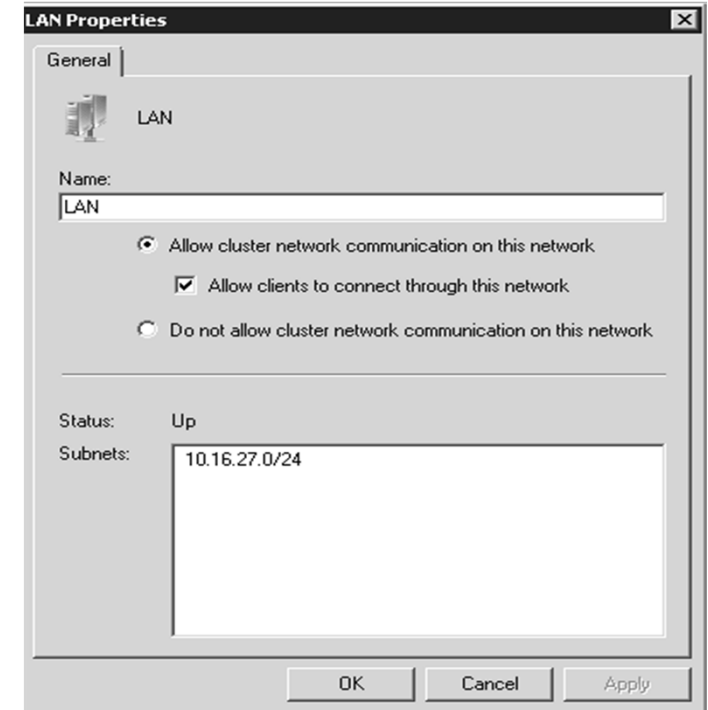
	NIC (min)	NIC (best)	Note
Gestione	1	1?	Necessaria
VM	1	2 o più	Può essere condivisa con la gestione
Privata (cluster)	1	1	Necessaria per un cluster
Live Migration	(1)	1 o più?	Può essere condivisa con la privata
CSV	(1)	1	Può essere condivisa con la privata
iSCSI	(1)	2 o più	Dipende dallo storage, usare il multi-path
SMB	(1)	2 o più	Usare SMB Multichannel
Backup	(1)	2 o più?	Dipende dal software di storage
Replica	(1)	1?	Dipende



# Rete per cluster heartbeat

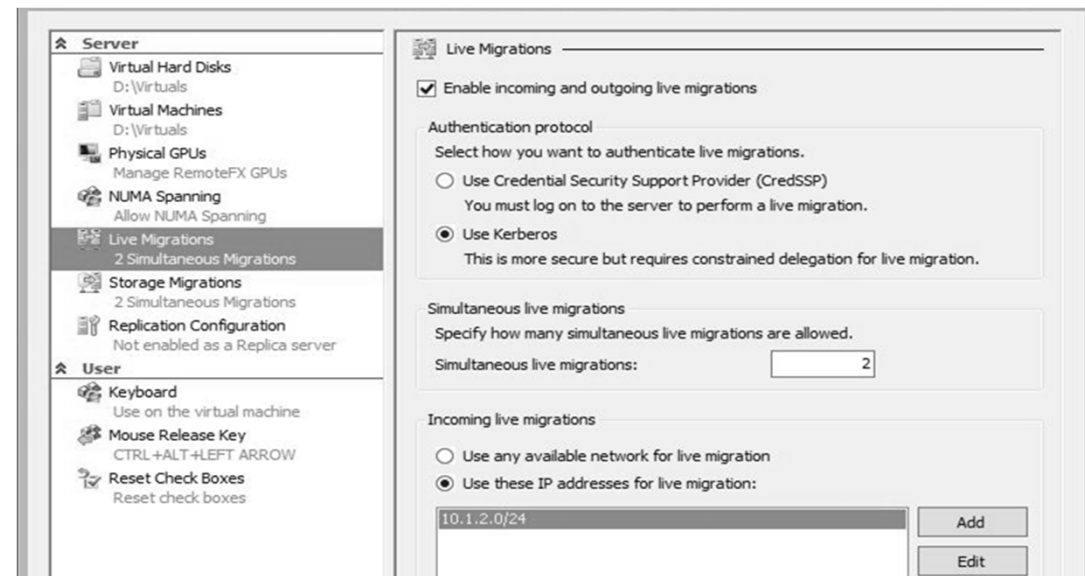
- Configurabile nelle proprietà delle reti del Failover Cluster
- Impostare
  - Private
  - Management come backup
- PowerShell
  - Network Roles
    - 0=Not used, 1=Cluster only, 3=Cluster+Clients

```
(Get-ClusterNetwork Internal).Role = 1
```



# Rete per Live Migration

- Nessun limite teorico
  - Né per LV né per Storage LV
  - Impostare un limite a livello di Hyper-V
- Modalità «shared nothing»
- Ricordarsi dei requisiti di compatibilità di processore
- PowerShell



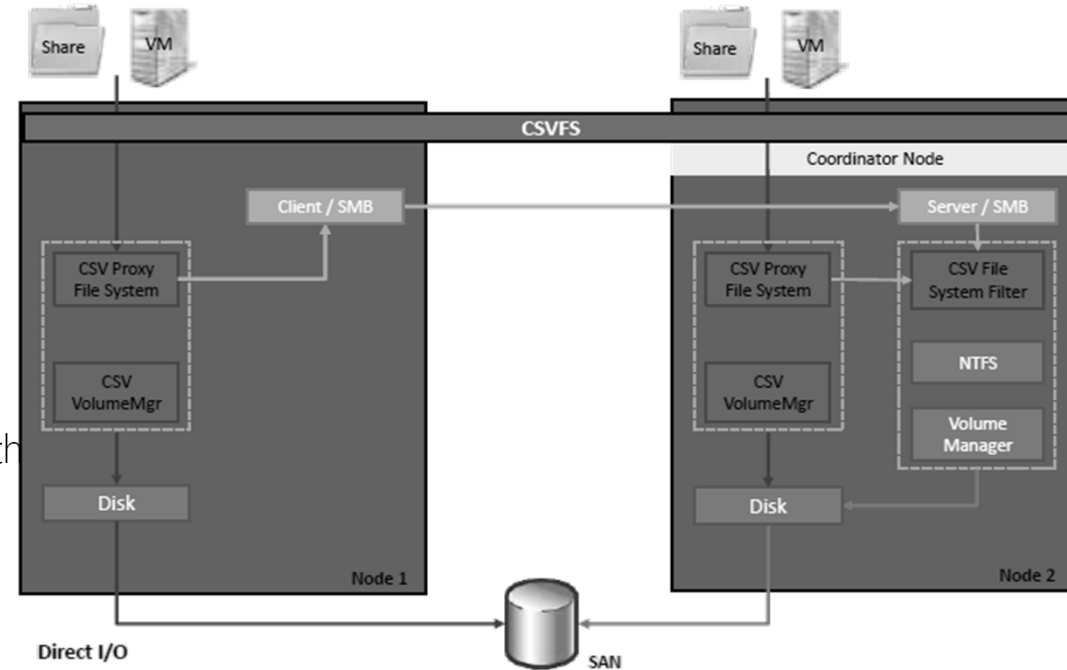
```
Enable-VMMigration
Set-VMMigrationNetwork 192.168.10.1
Set-VMHost -VirtualMachineMigrationAuthenticationType
Kerberos
Get-VMMigrationNetwork
```

# Rete per Storage Live Migration

- Teoricamente non necessaria
  - Quando entrambi gli storage sono visibili dall'host
- Ma per lo Storage vMotion from host to host?
  - Uno storage (o share) condiviso rimane raccomandato
- Notare che può beneficiare delle funzioni di Offload Data Transfer (ODX)

# Rete per CSV

- Cluster Shared Volumes communication
- Necessaria per redirected I/O mode
- Best practice
  - Drive letter of system disk
    - On all nodes, the drive letter for the system disk must be the same
  - Authentication protocol
    - NTLM protocol
    - Kerberos
  - Client for Microsoft Networks and File and Printer Sharing for Microsoft Networks must be enabled
- Scelta in automatico
- Per verificare la rete preferenziale



```
import-module failoverclusters
Get-ClusterNetwork | ft Name, Metric, AutoMetric, Role
```





Questions & Answers

# Grazie!

Non dimenticate di dare  
Il vostro feedback

Commenta la sessione  
tramite l'hashtag #wssid2013

