vmworld 2005
virtualize now

las vegas • october 18-20, 2005

 Derived from VMware Education Services "Virtual Infrastructure"
See http://www.vmware.com/services/education.html for course schedule
Networking Virtual Machines Under ESX Server: Part 1
Session PAC195-A

Mark Brunstad, Technical Instructor
Vmworld 2005
mbrunstad@vmware.com

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Objectives

- Understand ESX Server’s typical virtual/physical network configuration
- Identify the virtual network hardware available for installation in VMs (virtual NICs)
- Distinguish between the 3 general types of virtual switches used by VMs in an ESX Server
- Demonstrate how virtual switches are attached to physical NICs to form vmnics or bonds
- Examine ESX Server’s load balancing/switch failover options (out-mac, out-ip, Standby)
- Understand the VLAN feature of virtual switches
- Discuss COS and VMkernel NIC sharing

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Virtual/Physical Network Configuration

Virtual NICs

VM

NIC

VM

NIC

VM

NIC

VM

NIC

Virtual Switch (internal-only)

Virtual Switch (NIC team)

Virtual Switch

Service Console

vmkernel

VMotion

Physical NICs

VMotion

Production LANs

Management LAN

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Virtual Network Hardware

- Two choices for virtual NICs installed in a virtual machine:
  - vlan
    - Default virtual NIC
    - Compatible with AMD PCNET-32 controllers
    - Lower performance and higher CPU overhead
    - Compatible with all guest OSs supported in ESX Server
  - vmxnet
    - High-performance, idealized NIC written by VMware
    - Should be used for high-performance applications, especially for Gigabit networking or when using virtual networks
    - Requires VMware-supplied driver in guest OS
    - Driver installed along with VMware Tools

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
What are the VMware Tools?

- Device drivers for
  - virtual video card
  - mouse
  - optimized BusLogic SCSI
  - memory management
  - Heartbeat service
  - vmxnet

- VMware Tools installs into guest OS like an application
Virtual NIC: vmxnet

Driver available for Windows and Linux

How to install:
1. Install VMware Tools, shutdown
2. Configure the virtual machine’s NIC as vmxnet, restart

Guest OS finds the driver in its driver cache
MAC Addresses for Virtual Interfaces

- VMware uses two assigned Vendor ID ranges:
  00:0c:29:***:* (Generated), 00:50:56:***:* (Fixed)
- By default, MAC addresses for virtual NICs are calculated based on the virtual machine's UUID, which is created using
  - Physical host computer’s UUID
  - VM .vmx location)
- Mac Address = OUI + last 3 octets of the UUID
  - Ex: OUI = 00:0c:29:(XX.YY.ZZ)
  - uuid = 56 4d e0 c3 c7 1f 4d 95-c8 1e 8b bc 1c 94 92 80
  - MAC Address = 00:0c:29:94.92.80
Fixed MAC Addresses

- Edit the VMs .vmx file.

  Ethernet<x>.addressType = "generated"
  to
  Ethernet<x>.addressType = "static"

- Replace:

  ethernet<x>.generatedAddress = "00:0C:29:EB:52:80"
  with
  ethernet<x>.address = "00:50:56:{00-3F}:XX:YY"
Configuring Physical NICs

Auto-negotiate is the default; fix speed and duplex for production if necessary

MUI: Options>Network Connections>Physical Adapters
Connecting the vNIC to a vSwitch

- Choose virtual switch type
  - No outbound adapters
  - One outbound adapter
  - Two or more outbound adapters
  - Do not attach VM vNICs directly to a VMotion switch. It’s for the kernel!

- By default, these virtual networks are not directly accessible by the Service Console

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Virtual Switches With No Outbound Adapters (‘vmnets’)  

Each switch is an internal LAN, implemented entirely in software by the VMkernel

- Provides internal networking for the virtual machines of single ESX Server system
- Up to 32 virtual machines can connect to one vmnet
- Traffic shaping is not supported

Derived from VMware Education Services ‘Virtual Infrastructure’  
See http://www.vmware.com/services/education.html for course schedule
Example: One-Box Firewall Environment

- Virtual switch with one outbound adapter is DMZ
- Back-end applications are secured behind internal-only switches

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Example: Web, Application, Database Environment

-derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Virtual Switches With One Outbound Adapter (‘vmnics’)

Each connects virtual machines to one specific physical NIC

- Up to 32 virtual machines can connect to one vmnic
  - Zero collisions on internal traffic
- Each virtual NIC will have its own MAC address
- Bandwidth per virtual machine can be controlled with traffic shaping
Virtual Switches With Two or More Outbound Adapters (‘bonds’)  
Each connects virtual machines to an 802.3ad NIC team

- Up to 32 virtual machines can connect to one bond
- Each Virtual NIC will have its own MAC address
- Improved network performance by network traffic load balancing
- Redundant NIC operation
- Bandwidth per virtual machine can be controlled with traffic shaping
Example:
High Performance/Availability Applications

- Automatic network load balancing
- Redundant network connectivity

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Creating a NIC Team

- Up to 10 physical NICs can be placed in one team
- Up to 8 teams can be created
- A physical NIC that is used in a team can no longer be assigned stand-alone

This server has an unassigned adapter. Add it to the virtual switch.
Load Balancing/Switch Failover Modes

**out-mac**
- Default: each virtual machine’s outbound traffic is mapped to a specific physical NIC based on the virtual machine’s MAC
  - Low overhead
  - Compatible with all switches
  - May not spread traffic out evenly

**out-ip**
- A NIC for each outbound packet is chosen based on its source and destination IP addresses
  - Better distribution of traffic
  - Slightly higher CPU overhead
  - Not compatible with all switches; requires 802.3ad link aggregation support

**Standby**
- The bond will use one “home” NIC exclusively until that NIC fails, then fail over to another
  - Useful for falling back on a backup network path

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
NIC Team Configuration: Out-Mac Mode

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
NIC Team Configuration: Out-IP Mode

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
NIC Team Configuration: Standby Mode (home_link)

[Diagram of network setup with labels: ESX Server, Virtual Machine, switch, NIC, router, Internet, Client, Waiting]

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Assigning NIC Teaming Modes

- To **dynamically** apply a NIC teaming mode without rebooting, issue one of these commands as root:

  - `echo "nicteaming load-balance out-ip" > /proc/vmware/net/bond0/config`
  - `echo "nicteaming load-balance out-mac" > /proc/vmware/net/bond0/config`
  - `echo "nicteaming home_link vmnic5" > /proc/vmware/net/bond0/config`

*Changes will not persist over a server reboot!*

- `home_link` supersedes other modes
Permanent Configuration

- To choose a mode permanently, you must edit the file /etc/vmware/hwconfig
- To install out-ip mode for a given NIC team, bond0 for example:
  - delete any existing line that starts with: nicteam.bond0.load_balance_mode
  - add a new line like this: nicteam.bond0.load_balance_mode = "out-ip"
Permanent Configuration, cont.

- To return to the default, replace out-ip with out-mac.
- To designate one link as the home link and all others as failover links, add a line like so:
  - `nicteam.bond0.home_link = "vmnic5"`
- The presence of a home_link entry overrides any load_balance_mode specification.
- The hwconfig file is only read upon reboot. Therefore you must reboot ESX Server to see these changes take effect.
TANSTAAFL! Networking No Exception

- Remember, networking has a CPU cost
  - Every bps of bandwidth = one Hertz of CPU
  - A 2GHz CPU running at full utilization is needed to support a 1GB Ethernet card at full line rate
- A CPU-bound machine may not be able to take advantage of the increased bandwidth offered to the network by NIC teaming
- High performance applications need extra CPU if they are also heavy network bandwidth users

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Virtual LANs (VLANs)

- Virtual LANs (VLANs) allow the creation of multiple logical LANs within or across physical network segments
- VLANs free network administrators from the limitations of physical network configuration
- VLANs provide several important benefits
  - Improved security: the switch only presents frames to those stations in the right VLANs
  - Improved performance: each VLAN is its own broadcast domain
- ESX Server includes support for IEEE 802.1Q VLAN tagging
VLANs in ESX Server

- Virtual Switch Tagging (VST)
  - Packets leaving a virtual machine are tagged as they pass through the virtual switch
  - Packets are cleared (untagged) as they return to the virtual machine
  - Little impact on performance

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
VST VLANs Are Implemented by Port Groups

- ESX Server implements VST VLANs in terms of “port groups”
  - Each port group has a VLAN ID and a label
  - Each port group is a new subsidiary virtual switch, accessed by its label

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Sharing VMkernel NICs With the Service Console

- When physical hardware has too few NICs, the Service Console can get a virtual NIC connected to a virtual switch
  - Advantages: lower hardware cost; Service Console can use a NIC team
  - Disadvantage: extra dependency on VMkernel; blurred line between management and production

Derived from VMware Education Services ‘Virtual Infrastructure’
See http://www.vmware.com/services/education.html for course schedule
Sharing COS Adapter (Step 1)

- From the ESX Server Console:

1) Review the NIC adapters currently seen by the VMkernel.

```bash
# grep vmnic /etc/vmware/hwconfig
```

```
| devicenames.002.02.1.nic = "vmnic0"
| devicenames.007.01.0.nic = "vmnic1"
```
Sharing COS Adapter (Step 2)

2) Edit the /etc/modules.conf file to prevent loading and configuration of the driver for eth0, thus disabling eth0 on the Service Console.

```bash
# vi /etc/modules.conf

# Comment out the "eth0" entry

| alias parport_lowlevel parport_pc
| # alias eth0 bcm5700
| alias eth1 bcm5700
| alias eth2 bcm5700
```

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Sharing COS Adapter (Step 3)

3) Run `vmkpcidivy -i` and assign the Service Console NIC to the VMkernel.

The NIC assigned to the Service Console may vary between systems.

When this step is executed, it will result in a Service Console without a NIC.

You must have a console connection to the physical server to continue.
Sharing COS Adapter (Step 4)

4) Next, load the vmxnet_console driver. This can be done from the command-line however, the change will not persist across reboots

# insmod vmxnet_console devName="vmnic2"
Sharing COS Adapter (Step 5)

5) To make these changes persist, you must edit /etc/rc.d/rc.local. Add the entry for loading the vmxnet_console driver and an entry for each NIC added to the Service Console. For example:

```
# vi /etc/rc.d/rc.local

-----------------------
| insmod vmxnet_console devName="vmnic2"
| ifup eth0
```
Sharing COS Adapter (Step 6-7)

- 6) Reboot the server so as to test the new configuration.

- 7) Run the `ifconfig -a` command to verify that the eth0 interface is configured:
Summary

- Choose the right virtual hardware
  - vlance for portability
  - vmxnet for performance

- Choose the right virtual network connectivity
  - An internal-only virtual switch (vmnet)
  - A virtual switch patched to a physical NIC (vmnic)
  - A virtual switch w/two or more physical NICs (bond)
  - VLAN (port group)

- Use virtual switches to build sophisticated network architectures
  - Firewalls, NAT, intrusion detection systems, test environments

Derived from VMware Education Services ‘Virtual Infrastructure’
See [http://www.vmware.com/services/education.html](http://www.vmware.com/services/education.html) for course schedule
Questions?
This presentation covers the current versions of our products. Details about future releases of our products are available in select sessions at VMworld, including:

**PAC879:** The Next Phase of Virtual Infrastructure: Introducing ESX Server 3.0 and VirtualCenter 2.0

**PAC177:** Distributed Availability Services Architecture

**PAC484:** Consolidated Backup with ESX Server: In-Depth Review

**PAC485:** Managing Data Center Resources Using the VirtualCenter Distributed Resource Scheduler

**PAC532:** iSCSI and NAS in ESX Server 3