

WHITE PAPER

ESX Server™ 2

Mainframe-Class Virtual Machines for the Most Demanding Environments



NIC TEAMING IEEE 802.3ad



NIC Teaming IEEE 802.3ad

Summary

This tech note describes the NIC (Network Interface Card) teaming capabilities of VMware ESX Server™ 2 including its benefits, performance impact, and how ESX Server customers can take advantage of this new feature.

What is NIC Teaming?

NIC teaming allows users to group two or more physical NICs into a single logical network device called a bond. Once a logical NIC is configured, the virtual machine is not aware of the underlying physical NICs. Packets sent to the logical NIC are dispatched to one of the physical NICs in the bond and packets arriving at any of the physical NICs are automatically directed to the appropriate logical NIC.

NIC Teaming in VMware ESX Server 2 supports the IEEE 802.3ad static link aggregation standard. Customers need to configure teaming before they use any logical interface.

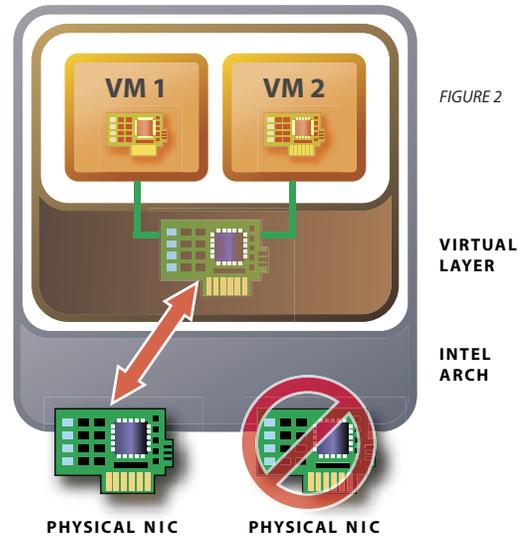
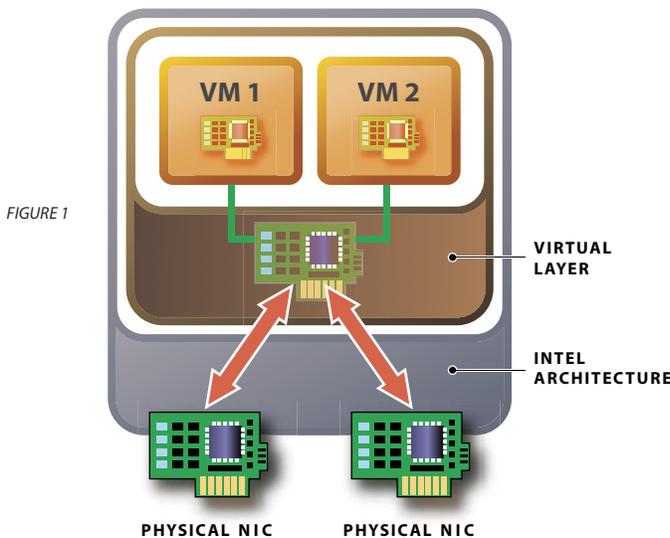
Benefits of NIC Teaming:

- **Load Balancing**

Outgoing traffic is automatically load-balanced based on destination address between the available physical NICs. Incoming traffic is controlled by the switch routing the traffic to the ESX Server, and hence ESX Server has no control over which physical NIC traffic arrives. Load balancing on incoming traffic can be achieved by using and configuring a suitable network switch. (See figure 1)

- **Fault Tolerance**

If one of the underlying physical NICs is broken or its cable has been unplugged, ESX Server 2 will detect the fault condition and automatically move traffic to another NIC in the bond. This capability eliminates a single point of failure for any one physical NIC and makes the overall network connection fault-tolerant. (See figure 2)



- **Transparent NIC Teaming Configurations Inside Virtual Machines**

Traditional implementations of NIC teaming require installation and configuration of a new driver in the OS. However, ESX Server 2 allows users to take full advantage of the standard NIC teaming without installation of a teaming driver. The guest OS still sees one NIC (vance or vmxnet device) and a driver in the vmkernel to handle all the NIC teaming jobs.

Performance Considerations

By installing multiple physical NICs on a server and grouping them into a single virtual interface, users can achieve greater throughput and performance from a “single” network connection. The magnitude of that performance gain, however, depends on many factors, including:

1. CPU-bound Servers

If the server is already CPU-bound, then simply adding more network capacity will not significantly increase overall network throughput as reassembling out-of-order packets, memory copies, and interrupt services are all CPU-intensive operations. A rule of thumb is that for every one bit per second of network data processed, one hertz of CPU cycle is required. For instance, a 2GHz CPU running at full and dedicated utilization would be needed to drive one Gigabit Ethernet card at line rate. Therefore, without enough CPU cycles, packets will be delayed regardless of NIC teaming.

2. Network Interfaces

VMware engineering lab results show that NIC teaming benefits fast ethernet network throughput more than gigabit ethernet network throughput in a typical server configuration. However, gigabit ethernet cards may be used with NIC teaming, and may also be mixed with Fast Ethernet cards as needed.

3. NIC teaming does not always scale linearly

Aside from the constraints we mentioned above, having “n” number of cards to work together does not always translate to “n” times performance gain. For example, in a TCP session each TCP segment has a segment ID, so if the packet is lost or delayed on the network, the receiver could be aware of it and ask the sender to retransmit the packet. If the segments arrive at the destination out of order, the receiver will have to reorder the segment. This TCP segment reordering process can be a time consuming operation and is beyond the scope of this article. Operating systems handle TCP out-of-order segment differently and in some cases, there could be a noticeable performance penalty from one OS to the other.

If two packets of a TCP session go through different physical NICs, there is no guarantee that the first packet will get to the first-hop router or switch earlier than the second packet because different NICs may have different delay times. Different types of cards could certainly contribute to this discrepancy, but even the same type of NIC could require variable amounts of time to process a packet due to different queue lengths.

NIC teaming in ESX Server 2 avoids the outbound packet out-of-order issue by always keeping the packets from the same TCP session on the same physical NIC. If a particular switch supports NIC teaming, it will also ensure that the TCP session segment stays on the same switch port.

Usage Scenarios

The following two scenarios illustrate the benefits of the VMware NIC teaming design in ESX Server 2.

1. A RedHat 7.0 Linux virtual machine runs on ESX Server. The underlying physical NICs are e1000 cards, which do not have Linux drivers supporting NIC teaming for this version of RedHat Linux. ESX Server 2 NIC teaming does not require each guest OS to have a separate NIC teaming driver so this legacy OS user can still realize the benefits of teaming two e1000 cards.
2. A Microsoft® Windows® Server 2003 virtual machine is running a web server on an ESX Server. The underlying physical NICs are 3com 3c90x 10/100 cards. The system administrators is unable to find a NIC teaming driver, and even if they could, they would have to schedule down-time for the web server to add NICs into a team. Now, with ESX Server NIC teaming, neither a guest OS teaming driver nor the down-time is necessary.

Frequently Asked Questions

- Q:** Can the ESX Server NIC Teaming work with a regular switch?
- A:** We recommend switches that are compatible with 802.3ad, however, ESX Server 2 NIC Teaming will work with regular switches and will still support the outbound load balancing and fail-over. Performance is not guaranteed if users deploy enterprise switches without link aggregation features primarily because the inbound traffic may come from unpredictable ports.
- Q:** How is inbound traffic handled after a bond is created? Your document says I can bond 10 physical NICs and get load balancing from ESX Server, but what about inbound?
- A:** The outbound load balancing is based on the layer 3 address for IP traffic and round-robin for non-IP traffic. The inbound load balancing is dependent upon the switches, simply because we have no control over the switch side. You can configure a Cisco EtherChannel switch and let it send packets to one of the teaming ports based on MAC address.
- Q:** Can you mix Gigabit and Fast Ethernet NICs in the same teaming bond?
- A:** Yes. However, we do not prioritize based on type or speed of the NIC card.
- Q:** Can I do NIC teaming on the Console OS?
- A:** No, NIC teaming is not currently supported in the VMware ESX Server 2 Console Operating System.

- Q:** Why am I getting duplicate packets from the NIC teaming interface?
- A:** If hubs are used, then users may see duplicate packets sent to a virtual machine. In general, non-switching hubs are not recommended for use with NIC teaming.
- Q:** Do I need to configure switches that support NIC teaming or 802.3ad link aggregation?
- A:** Typically 802.3ad compatible switches do not have link aggregation settings by default. Users must configure their switches explicitly.
- Q:** Why am I only seeing traffic on one of the vmnics in a bond?
- A:** The NIC teaming implementation determines the vmnic for the TCP/IP traffic based on the IP address. If all packets are destined for the same, or just a few, IP addresses, all the packets may stay on the same vmnic.
- Q:** Does NIC teaming work with both vmxnet and vlnic?
- A:** Yes, NIC teaming works with both vmxnet and vlnic.
- Q:** Can a user still use vmnic0 once it is configured in a bond?
- A:** No. If users already have a vmnic0 in their VM configuration and later provision the vmnic0 to be part of bond0, they will have to change the VM configuration explicitly before booting the VM.

(Footnotes)

1. In ESX Server 2, users can Team between 2 to 10 physical NIC's in one "bond".

V00014-20001205



VMware, Inc. 3145 Porter Drive Palo Alto CA 94304 USA Tel 650-475-5000 Fax 650-475-5001 www.vmware.com
Copyright © 2003 VMware, Inc. All rights reserved. Protected by one or more of U.S. Patent Nos. 6,397,242 and 6,496,847; patents pending. VMware, the VMware "boxes" logo, GSX Server and ESX Server are trademarks of VMware, Inc. Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation. Linux is a registered trademark of Linus Torvalds. All other marks and names mentioned herein may be trademarks of their respective companies.

