



# An Introduction to NIC Teaming with Lenovo Networking Switches

---

**Shows how Network Interface Card (NIC) teaming can be configured**

---

**Includes examples for Linux, Windows, and VMware**

---

**Describes how Ethernet switches must be configured for different teaming modes**

---

**Shows use cases for NIC teaming**

**Scott Lorditch**



# Abstract

This paper describes the network interface card (NIC) teaming modes on common operating systems and how Lenovo Networking switches must be configured to work with those modes. This paper is intended for those users who install and configure servers and networks.

At Lenovo Press, we bring together experts to produce technical publications around topics of importance to you, providing information and best practices for using Lenovo products and solutions to solve IT challenges.

For more information about our most recent publications, see this website:

<http://lenovopress.com>

# Contents

Introduction .....	3
NIC teaming modes .....	3
Operating system support for teaming .....	3
Considerations for dual-homed servers .....	12
Summary .....	13
Author .....	13
Notices .....	14
Trademarks .....	15

# Introduction

There are several network interface card (NIC) teaming modes that can be used with Lenovo servers. Rack-mounted servers are often “dual-homed” that connect to two different switches. Flex System™ and BladeCenter® servers are dual-homed by default and connect to two different switches across the midplane of the chassis. For rack and chassis servers, two parallel NIC ports can go to the same switch.

The various NIC teaming options and how to configure them are described in this paper.

## NIC teaming modes

The following teaming modes are available:

- ▶ Active/standby

One NIC is active at a specific time and carries all traffic to and from the server; one or more NICs are standing by. If there is a failure on the primary NIC, one of the standby NICs takes on the entire load.

Some Operating Systems unconditionally fail back to the primary NIC if it becomes available while others can be configured to fail back or not.

The two or more ports from the server can be connected to the same switch but are more commonly connected to different switches, which makes the server dual-homed.

- ▶ Active/active

Two or more NICs are active at the same time. Various techniques for allocating traffic among the active NICs are available from different operating systems. The most commonly used technique uses MAC or IP addresses as parameters for a hashing algorithm. This technique ensures that packets for the same flow (source/destination pair) use the same physical port. These techniques typically require that the ports are attached to the same switch or to a set of switches that share a data plane through a technique, such as stacking or Flex Interconnect Fabric.

- ▶ Active/active for hypervisors

This technique assigns one physical NIC to a guest operating system for sending and receiving traffic. This assignment persists except if the NIC fails, then the guests are reassigned to other active NICs. This teaming mode does not require any special switch configuration; it is available on VMware and on Microsoft Hyper-V.

## Operating system support for teaming

The various commonly used operating systems have different ways of configuring NIC teaming or bonding, which are described next.

### Teaming in Linux

Linux distributions included the bonding driver for several years. The driver supports modes active/standby and active/active and several other modes (for example, send all packets on all physical NICs), which often are not seen in use and can cause issues with upstream switches and routers.

Bonding mode=1 is active/standby bonding, in which one NIC is designated as the primary in the configuration of the driver. This parameter specification is stored in different places for different Linux distributions.

Linux also supports round-robin transmission of packets (mode=0). This mode does not ensure that packets that are part of the same flow use the same physical port and can cause packets to arrive at the destination out of sequence. Bonding mode=2 uses a hashing algorithm to allocate flows to physical ports and is preferable.

LACP support also is available with the bonding driver (mode=4), which is an active/active mode and requires LACP configuration on the switch ports.

## Configuration of Linux bonding

Current Linux distributions include the “bonding” driver, which must be loaded and its options specified to make it active. This activation can be done in two ways: there are commands, such as **ifenslave**, and options on the IP command that can dynamically create a bond, which does not survive a reboot unless a start script reissues the same commands.

More common is the use of driver configuration files, which are found in `/etc/modprobe.d/*.conf` in text format (or a slightly different directory in some distributions). For more information about the supported options, see this website:

<https://www.kernel.org/doc/Documentation/networking/bonding.txt>

The following key options are available:

- ▶ **mode**

The mode=0 is round-robin; mode=1 is active/standby; mode=2 is active/active with XOR hashing; and mode=4 is similar to mode=2 but uses LACP rather than static aggregation. These options are the options that are most commonly used.

- ▶ **miimon**

This option specifies link monitoring frequency, which is measured in milliseconds. This option determines how quickly the system becomes aware of a port that is down.

- ▶ **downdelay**

This option is how long to wait after link monitoring shows a port to be down before failover is triggered.

- ▶ **primary**

This option identifies which port is active in an active/standby configuration.

- ▶ **xmit\_hash\_policy**

This option identifies how to perform hashing in active/active mode. It is analogous to the hashing options on Lenovo® switches. The following policy options are available:

- layer2 (MAC addresses)
- layer2+3 (MAC and IP)
- layer3+4 (IP and port)

An `ifconfig-xxx` file also must be configured or generated for each bonding team. These files are referred to as “network scripts” and their directory location depends on the distribution that is used.

The bonded devices often are of the form `bond<x>`; therefore, a file might be `ifconfig-bond0`. These files are similar to files for physical interfaces, such as `ifconfig-eth0` with some additions (as shown in bold in Example 1 on page 5). The bonding options are not required here, but the options do take precedence over those options in the `modprobe` file.

A sample file is shown in Example 1, but there are differences in different distributions.

*Example 1 Sample ifconfig-xxx file*

---

```
BOOTPROTO="static"
BROADCAST="<mask for broadcasts>"
IPADDR=<address>
NETMASK="<mask>"
NETWORK="<address of network origin>"
STARTMODE="onboot"
BONDING_MASTER="yes"
BONDING_MODULE_OPTS="mode=active-backup miimon=100"
BONDING_SLAVE0="eth0"
```

---

Also, the ifconfig-eth<x> files for the members of a bond often have lines of this form added, as shown in Example 2.

*Example 2 Additions to the ifconfig-eth<x> file*

---

```
SLAVE=yes
MASTER=bond<x>
```

---

## Teaming in Windows Server

Versions of Windows Server before 2012 use utilities that are provided by the NIC vendors such as Emulex, QLogic, Intel to implement NIC teaming. These utilities support active/active and active/standby teaming modes. These utilities often do not function unless at least one NIC that is managed by them is from the same vendor; sometimes, they do not team NICs unless at least one member of the team is from that vendor.

### Emulex OneCommand

The Emulex OneCommand utility supports active/active and active/standby teaming and Link Aggregation Control Protocol (LACP).

The Team Type option is the counterpart of the Linux “mode=” parameter and provides for active/standby (failover), static aggregation (802.3ad), and LACP, as shown in Figure 1.

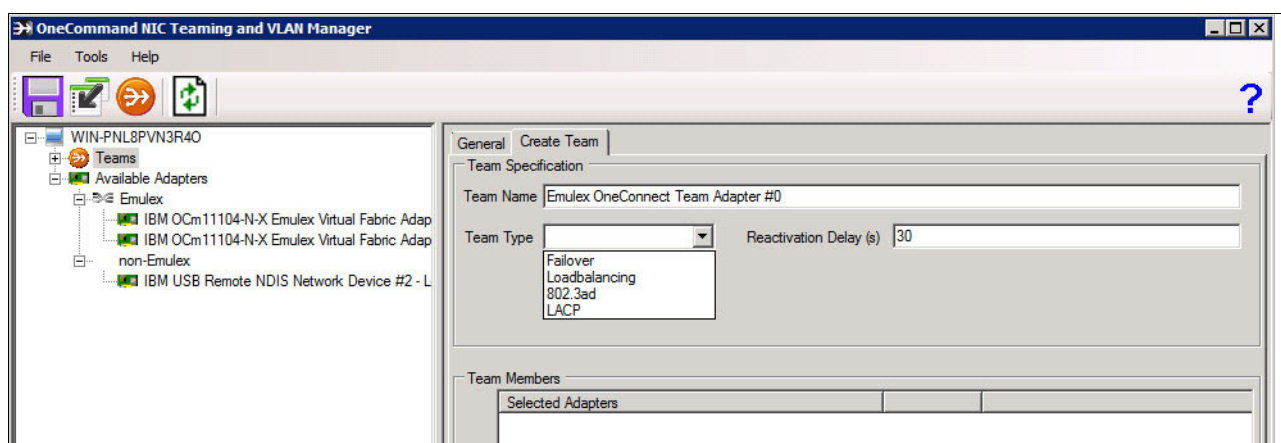


Figure 1 Team Type parameter in Emulex OneCommand Manager

VLANs can be added by using this utility, as shown in Figure 2. When a VLAN interface is created, the “native” VLAN option is automatically cleared. If you want an interface for the native VLAN and other VLANs, ensure that the **Enable Native VLAN** option is selected before you select **Apply**.

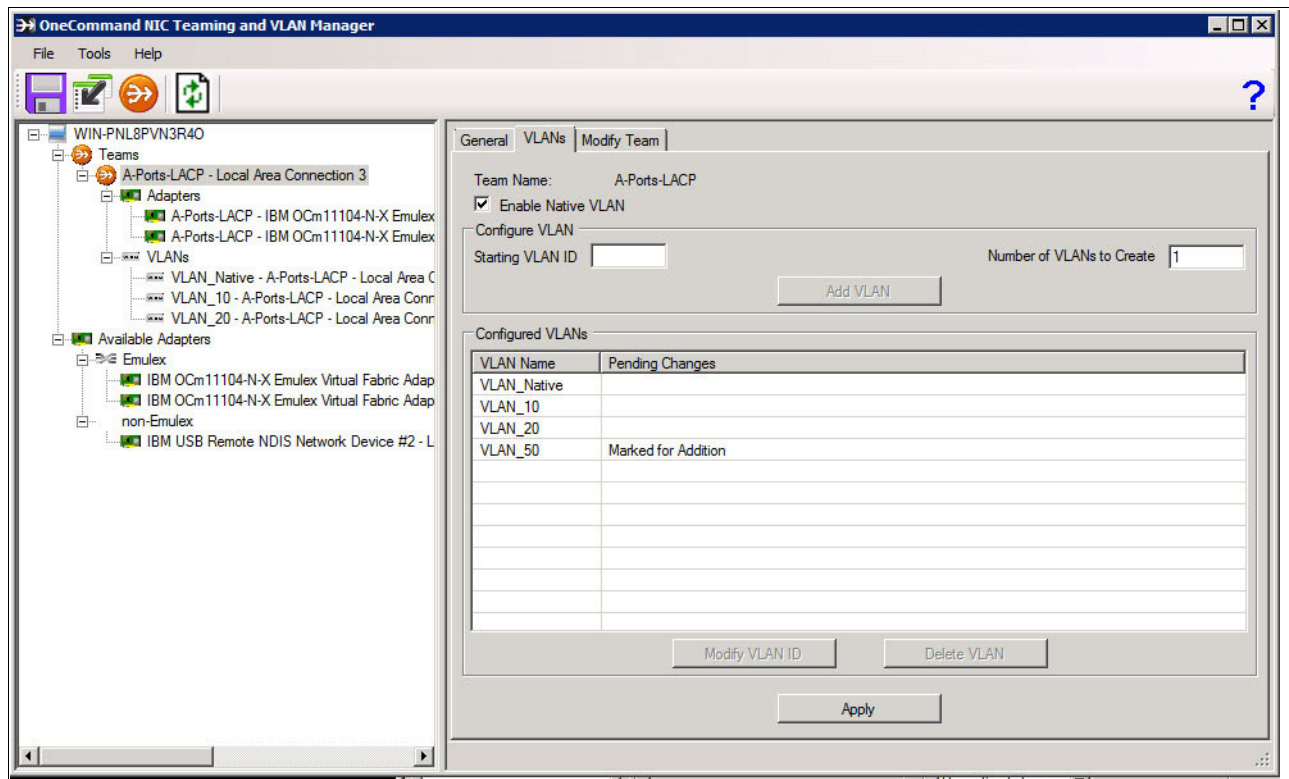


Figure 2 Adding VLANs

## QLogic Control Suite

The QLogic Control Suite utility has very similar functionality to Emulex OneCommand. The team type parameter as shown in Figure 3, with very similar options. Some of the options are unavailable when the adapter has storage functionality (FCoE, iSCSI) enabled.

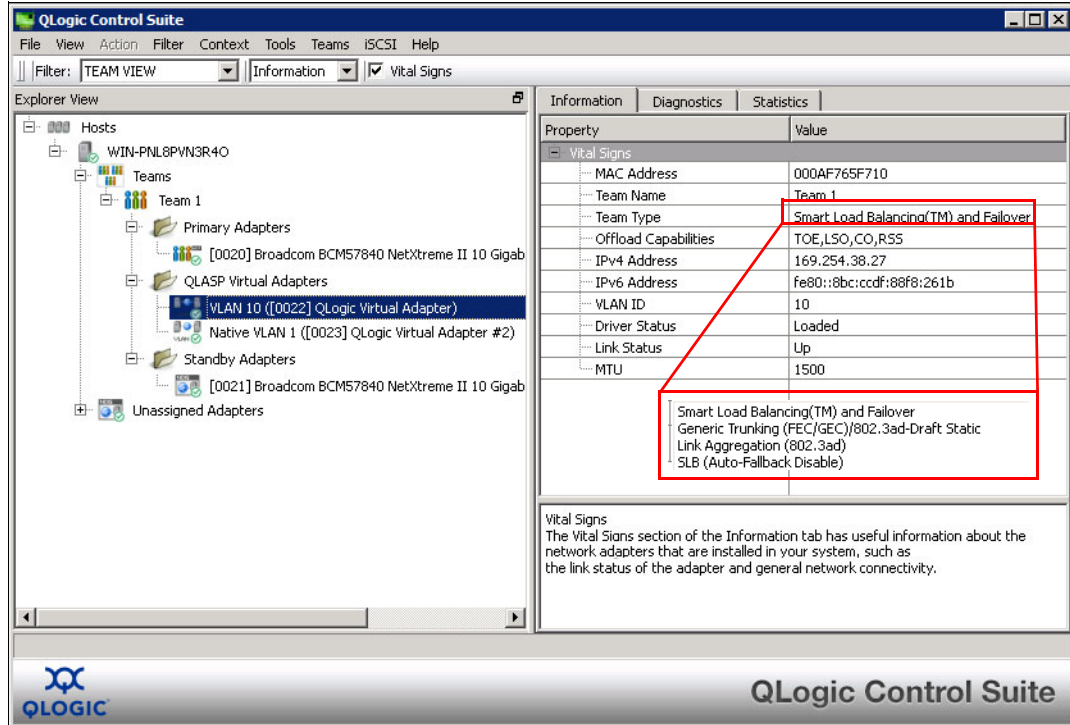


Figure 3 QLogic teaming setting

VLANs are managed by the QLogic utility as shown in Figure 4. Both tagged and untagged VLANs are supported. If it is desired to have a native (untagged) VLAN as well as additional tagged VLANs, then the native VLAN must also be explicitly added using the control suite.

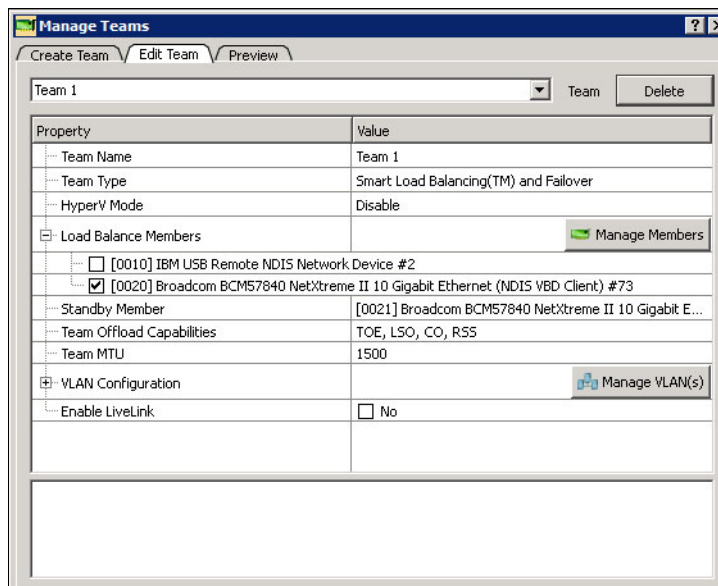


Figure 4 QLogic VLANs

## Windows Server 2012

Windows Server 2012 features its own built-in utility, which can be started from a power shell window or from the desktop. This utility supports active/active and active/standby teaming and per-guest teaming when Hyper-V is in use.

As is the case with the QLogic and Emulex utilities, after the teams and VLANs are created, they are added to the Network Connections folder with the physical NIC ports. If IP addresses or other properties are to be assigned to a team or VLAN, the appropriate items in that folder must have their properties configured.

The Windows Server 2012 teaming properties window is shown in Figure 5.

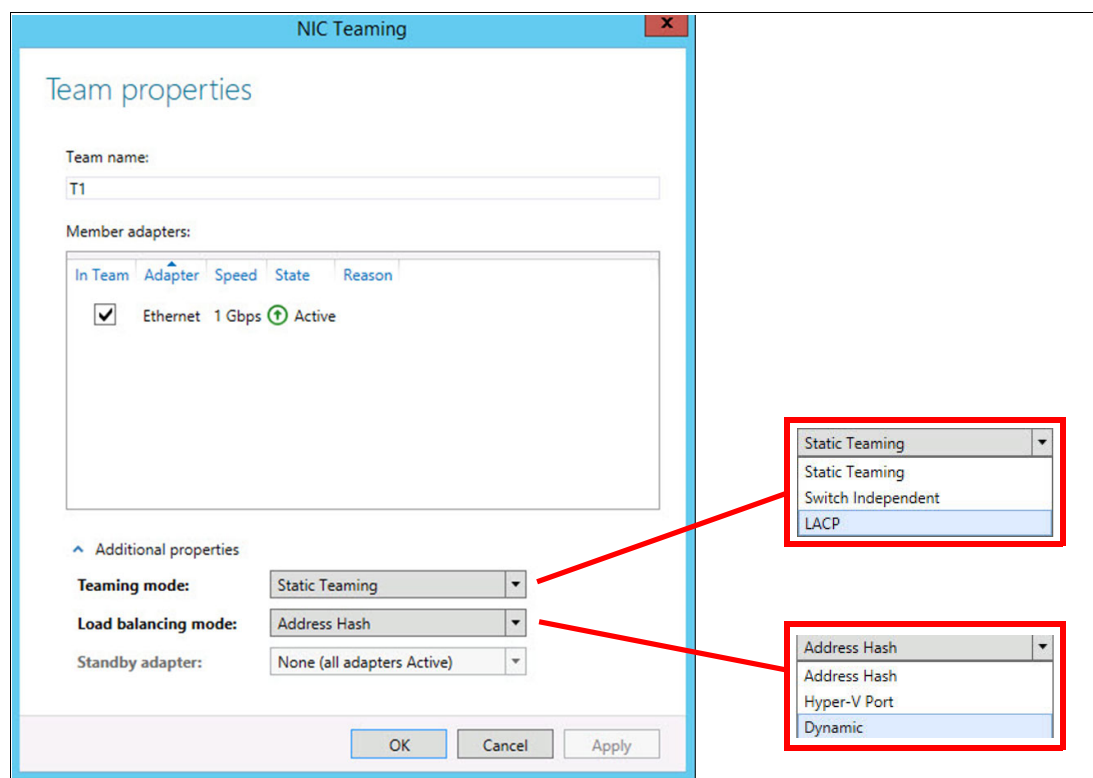


Figure 5 Windows Server 2012 native teaming

As shown in Figure 5, the following options are available:

- Static teaming versus LACP
- Load balancing mode
- Specification of standby adapters

Consider the following points:

- Switch Independent mode can be used to support active/standby teaming and Hyper-V mode.
- In most cases, Address Hash is used here, unless Hyper-V is deployed. Hyper-V mode assigns guest machines to one member of the team.



Multiple VLANs that use 802.1q tagging can be added to a team, as shown in Figure 6. An entire team (“primary interface”) also can be assigned to a specific VLAN; the default is for the native (untagged) VLAN to be used.

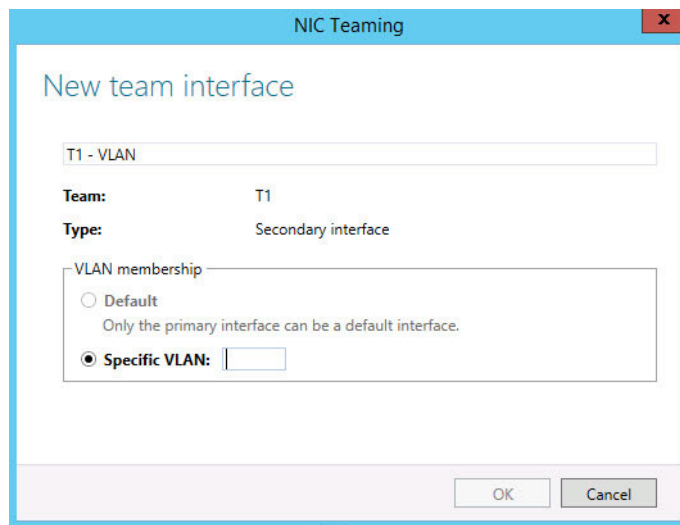


Figure 6 VLAN tagging

## Teaming in VMware

VMware supports teaming modes with more versatility than any other operating system.

NICs are assigned as active or standby and multiple standby NICs are allowed. Teaming can use hashing (as described in “NIC teaming modes” on page 3), and can still be configured for extra NICs in a specific port group to be set as standby if all of the active NICs fail.

LACP is supported only with VMware’s Distributed Virtual Switch (DVS). Third-party virtual switch software offerings might also have LACP support.

The starting point for all network configuration is from VCenter (or the new web configurator) under the Configuration tab. The display resembles the display that is shown in Figure 7 on page 10.

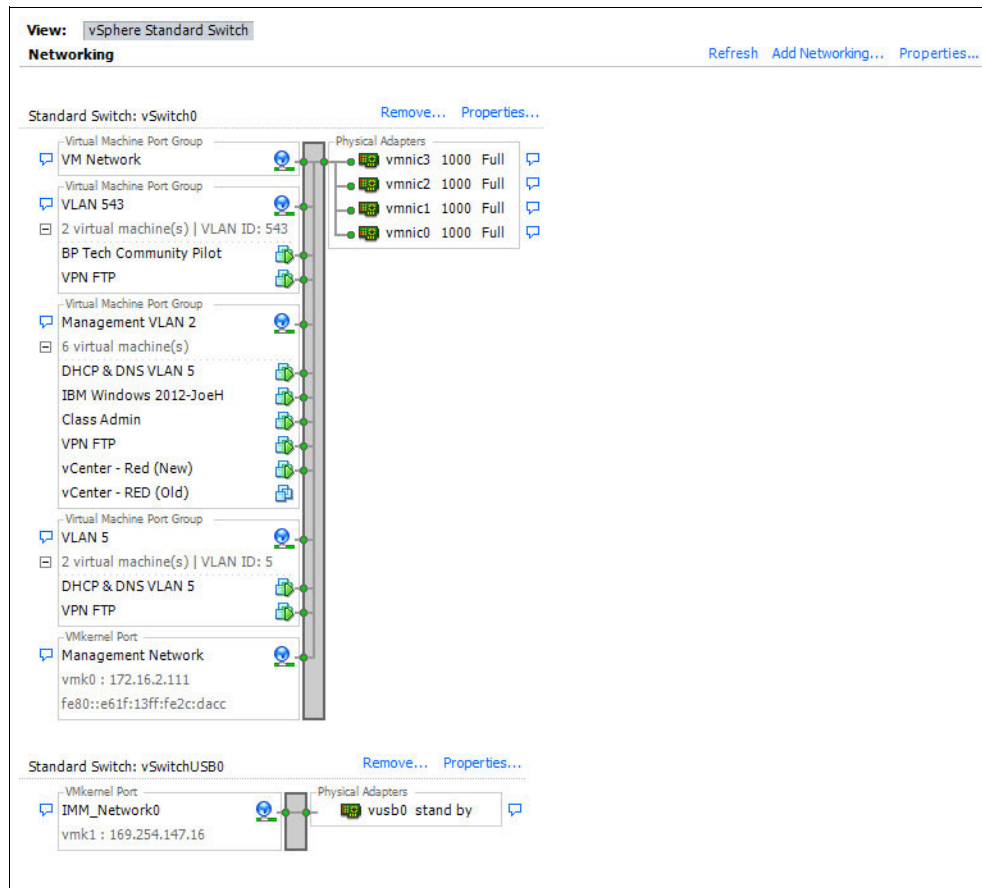


Figure 7 VMware Network Display: Standard Switch

NIC Teaming in VMware is an option that is assigned at the level of a port group, which can be a VMkernel group (management, VMotion, and so on) or a virtual machine group, which services guest virtual machines, as shown in Figure 8.

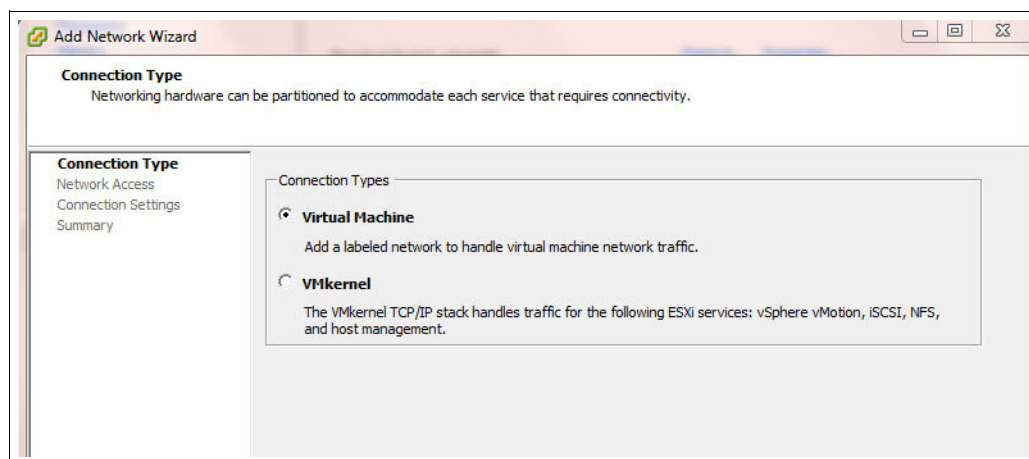


Figure 8 VMware: Add Network Wizard

Physical NICs are assigned to a port group from an existing or newly created virtual switch, as shown in Figure 9. An NIC can be owned by only one virtual switch at a time; including standard and distributed switches.

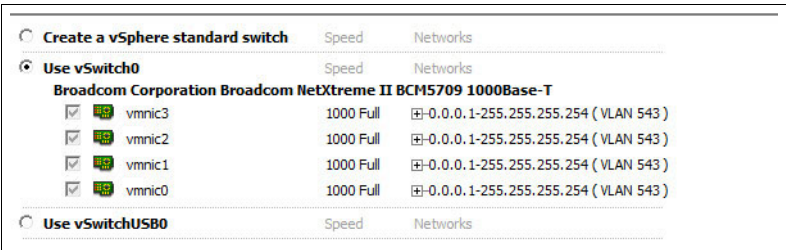


Figure 9 Assigning NICs to a virtual switch

After the physical NICs are assigned to a port group, the guest VMs that use those NICs also must be attached to the port group, as shown in Figure 10.

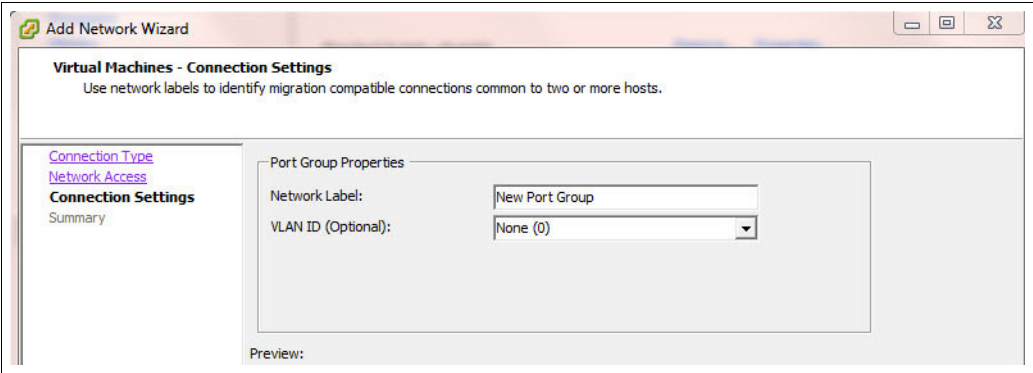


Figure 10 Assigning a VM to a port group

A port group cannot have a VLAN ID and use the native VLAN from an external physical switch. It can be assigned to a single VLAN by using 802.1q tagging, or it can participate in all available VLANs. In the last instance, guest machines that are attached to the port group are responsible for handling tagging.

NIC Teaming (or aggregation) is handled at the level of the port group, as shown in Figure 11.

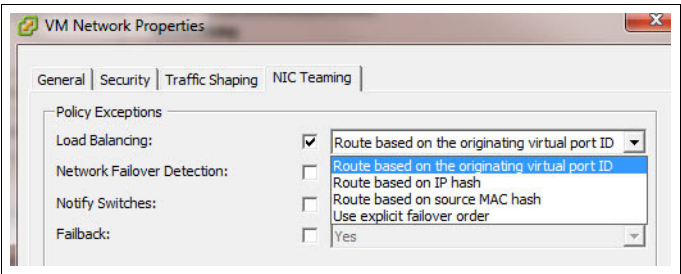


Figure 11 NIC Teaming mode selection

As shown in Figure 11, the default option (Route based on the originating virtual port ID) has the effect of assigning guests to individual NICs. Because the MAC address of each guest flows only in and out of a single port, no special configuration is required on external switches and multiple NICs can be active at the same time. This option is suitable for environments in which the guest machines use roughly the same amount of bandwidth.

IP hash and MAC hash are useful options to provide better use balancing across the available physical NIC ports when some guests use more bandwidth than other users. These options are analogous to the hashing options that can be part of configuration of Lenovo switches.

Whichever option is used, adapters can be selected as active or standby, as shown in Figure 12. Standby adapters are used only when no active adapters are operational. An active/standby team similar to those teams in other operating systems can be implemented by using this feature.

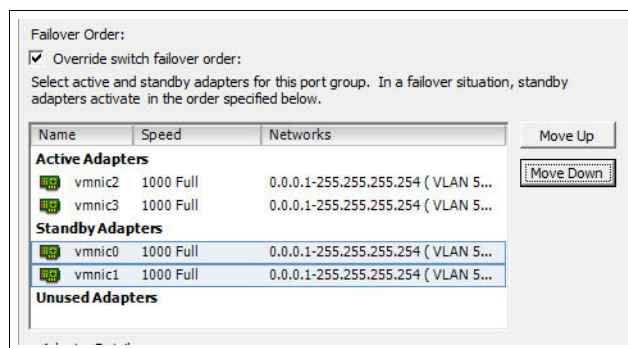


Figure 12 NIC Teaming failover mode selection

## Considerations for dual-homed servers

Servers in Flex System and BladeCenter chassis are intended to have access to at least two embedded switches; therefore, they are dual-homed (or more) unless a customer chooses not to purchase a redundant switch.

**Note:** Servers which have four or more NICs can have multiple NICs to the same switch, and those NICs can be treated in the same way as any NICs which connect to the same switch.

Because the first two NICs in a chassis server connect to different switches, there are some constraints on how the switches must be configured to enable some of the available teaming modes. These same constraints apply to rack-mounted servers, when they are connected to more than one switch.

The following constraints are featured:

- Active/standby teaming often does not require any special switch configuration.
- Active/active teaming as done by hypervisors where guest virtual machines are assigned to a particular physical NIC also do not require special switch configuration. The use of these teaming modes with the switch configuration options that are required for the usual active/active teaming modes can cause problems.

For example, if the default teaming mode is chosen in VMware, the ports on the switches that attach to the physical NICs must not be aggregated together. Doing so can cause some guest VMs to be unreachable.

- The usual active/active teaming modes require link aggregation of the ports on the switches to mirror the aggregation that is used on the server. This configuration can be done as a static port channel or with LACP.

NIC ports on Flex System servers can include multiple ports to the same switch module and ports that are connected to two distinct switch modules. When ports to more than one switch module are teamed together, the switch modules must share a data plane so that they function as a single logical switch.

This configuration can be done by using one of the following methods:

- ▶ The switches can be stacked together with traditional stacking or SI4093 modules can be part of a Flex Interconnect Fabric. If this issue is the case, the ports are configured as a link aggregation as part of the single configuration of the stack or fabric. This configuration can be done as a static or LACP aggregation by using the usual commands.
- ▶ The server-facing ports can be configured as a vLAG instance. This configuration requires all of the usual vLAG configuration, including the use of a cross-link (ISL) between the two switches. Static or LACP aggregations can be configured. vLAG is limited to two switches, while the stacking and fabric options can support all four switches in a chassis (although this configuration is rarely done).

## Summary

Despite all of the different utilities, there are few different ways to use NIC teaming. Active/active teaming is increasingly the norm and can be accommodated by Lenovo switches by using the features that are described in this paper.

## Author

**Scott Lorditch** is a Consulting System Engineer for Lenovo. He performs network architecture assessments and develops designs and proposals for solutions that involve Lenovo Networking products. He also developed several training and lab sessions for technical and sales personnel. Scott joined IBM as part of the acquisition of Blade Network Technologies and joined Lenovo as part of the System x acquisition from IBM. Scott spent almost 20 years working on networking in various industries, as a senior network architect, a product manager for managed hosting services, and manager of electronic securities transfer projects. Scott holds a BS degree in Operations Research with a specialization in computer science from Cornell University.

Thanks to David Watts of Lenovo Press for his contributions to this project.

# Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.  
1009 Think Place - Building One  
Morrisville, NC 27560  
U.S.A.  
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document REDP-5245-00 was created or updated on August 31, 2015.

Send us your comments in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:  
[ibm.com/redbooks](http://ibm.com/redbooks)
- Send your comments in an email to:  
[redbooks@us.ibm.com](mailto:redbooks@us.ibm.com)

## Trademarks

Lenovo, the Lenovo logo, and For Those Who Do are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. These and other Lenovo trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by Lenovo at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of Lenovo trademarks is available on the Web at <http://www.lenovo.com/legal/copytrade.html>.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo(logo)®	BladeCenter®	Lenovo®
Blade Network Technologies®	Flex System™	

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.