

HP ProLiant Network Adapter Teaming

White Paper



Legal notices

© Copyright 2004, 2008 Hewlett-Packard Development Company, L.P.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Microsoft®, Windows®, Windows Server 2003 and Windows Server 2008 Enterprise are trademarks of Microsoft Corporation in the U.S. and other countries.

Cisco® and Cisco® EtherChannel® are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

Printed in the US.

Part number: 488077-00A

Contents

Revision history

Revision tables.....	6
----------------------	---

1 Introduction

1-1 Abstract.....	8
1-2 How to read this document	8
1-2-1 Section layout.....	8

2 An executive overview of teaming

2-1 What is HP ProLiant Network Adapter Teaming?	9
2-2 The goals of HP ProLiant Network Adapter Teaming.....	9
2-2-1 Fault tolerance: self-healing network connectivity	9
2-2-2 Load balancing: RAID 5 for server network adapters	10
2-3 Why HP ProLiant Network Adapter Teaming?.....	10

3 Teaming fundamentals for the first-time user

3-1 A technical overview of HP ProLiant Network Adapter Teaming	11
3-1-1 Fault tolerance: Dual Homing or Network Adapter Teaming?.....	11
3-1-2 Load balancing: server-based routing protocol or NIC teaming?	12
3-2 Overview of team types	13
3-2-1 Network Fault Tolerance Only (NFT).....	13
3-2-2 Network Fault Tolerance Only with Preference Order	13
3-2-3 Transmit Load Balancing with Fault Tolerance (TLB).....	13
3-2-4 Transmit Load Balancing with Fault Tolerance and Preference Order.....	14
3-2-5 Switch-assisted Load Balancing with Fault Tolerance (SLB)	14
3-2-6 802.3ad Dynamic with Fault Tolerance.....	14
3-2-7 Automatic	14
3-3 How to choose the best team type.....	15
3-4 Basic deployment steps for HP ProLiant Network Adapter Teaming.....	15
3-4-1 Download	15
3-4-2 Install.....	16
3-4-3 Launch.....	16
3-5 Overview of the HP Network Configuration Utility	17
3-5-1 NCU main window	17
3-5-2 Team Properties window	17
3-5-3 Adapter Properties window	18
3-6 Getting started.....	19
3-6-1 Example deployment of basic teaming	19

4 The mechanics of teaming for the advanced user

4-1 Section objectives and prerequisites	20
4-2 Architecture of HP ProLiant Networking Adapter Teaming	20
4-2-1 The “networking layers” of an operating system	20
4-2-2 Teaming software components	20
4-2-3 HP teaming and Layer 2 versus Layer 3 addresses.....	21
4-3 Types of HP ProLiant network adapter teams.....	22
4-3-1 Network Fault Tolerance Only (NFT).....	22
4-3-2 Network Fault Tolerance Only with Preference Order	29
4-3-3 Transmit Load Balancing with Fault Tolerance (TLB).....	29
4-3-4 Transmit Load Balancing with Fault Tolerance and Preference Order.....	37
4-3-5 Switch-assisted Load Balancing with Fault Tolerance (SLB)	37
4-3-6 802.3ad Dynamic with Fault Tolerance.....	41
4-4 Team types and redundancy mechanisms interoperation	41
4-4-1 Team type and redundancy mechanisms compatibility chart	41
4-5 Team status and icons	42

4-5-1	Network adapter teamed status	42
4-5-2	Team state	42
4-5-3	Team icons.....	43
4-6	HP ProLiant Network Adapter Teaming and advanced networking features	43
4-7	HP ProLiant Network Adapter Teaming configuration and deployment with ProLiant Essentials Rapid Deployment Pack (RDP).....	44

5 Teaming feature matrix

Appendix A: Overview of network addressing and communication

Layer 2 versus Layer 3	46
Address – unicast versus broadcast versus multicast.....	46
Example scenarios of network addressing and communication.....	46
5-1-1 Scenario 1-A: one device PINGs another on the same Layer 2 network.....	47
5-1-2 Scenario 2-A: one device PINGs another on a different Layer 2 network.....	48

Appendix B: Frequently asked questions

Appendix C: Overview of utilities included with HP ProLiant Network Adapter Teaming

Appendix D: ProLiant networking information resources

Product information	55
Whitepapers	55
Network adapter and network adapter teaming drivers	55

Appendix E: Technical Support

Appendix F: Glossary

Figures

Figure 3-1 HP ProLiant Network Adapter Teaming.....	11
Figure 3-2 Using Multihoming for server network redundancy	12
Figure 3-3 Using NIC teaming for server network redundancy	12
Figure 3-4 Teaming types and functionality	13
Figure 3-5 How to choose the best team type to use	15
Figure 3-6 NCU main window.....	17
Figure 4-1 Overview of NFT communication.....	22
Figure 4-2 Scenario 1-B: A device PINGs an NFT team on the same Layer 2 network	23
Figure 4-3 Team Properties window—Settings tab containing the Path Validation settings.....	25
Figure 4-4 Upstream link failures cause server isolation.....	27
Figure 4-12 Overview of TLB communication	30
Figure 4-13 Scenario 1-C: A device PINGs a TLB team on the same Layer 2 network.....	31
Figure 4-15 Scenario 2-C: TLB team using IP address for load balancing algorithm	34
Figure 4-16 Scenario 1-D: TLB team using MAC address for load-balancing algorithm.....	36
Figure 4-17 Overview of SLB communication.....	38
Figure 4-25 Teamed port Information tab	42
Figure 4-26 Team state icons.....	43
Figure 4-27 VLAN tagging used with HP ProLiant Network Adapter Teaming	44
Figure A-1 Scenario 1-A: One device PINGs another on the same Layer 2 network	47
Figure A-2 Scenario 2-A: one device PINGs another on a different Layer 2 network	48

Tables

Table 0-1 Teaming Whitepaper Revision History	6
Table 0-2 Teaming Driver Revision History	6
Table 4-1 Heartbeat frame format	26
Table 4-2 802.1Q tagged heartbeat frame format	26
Table 4-8 Transmit Load Balancing method comparison	32
Table 4-9 Load Balancing based on Destination IP Address.....	33
Table 4-10 Load balancing based on Destination MAC Address.....	35
Table 4-11 Example of load-balancing algorithms.....	39

Revision history

Revision tables

Table 0-1 Teaming Whitepaper Revision History

Date	Author	Major Revision	Minor Revision	Reference Driver Version
March 1, 2008	Daisy M Longmire	9.20		NCDE 9.20
September 9, 2005	Sean McGee	8.10	007	NCDE 8.10
July 21, 2005	Sean McGee	3	158	NCDE 7.86

Table 0-2 Teaming Driver Revision History

Date	Driver Version	Overview of Changes
March 5, 2008	9.20.0.0	Added support for the following: <ul style="list-style-type: none">Windows Server 2008 Removed support for the following: <ul style="list-style-type: none">HP NC31xx Fast Ethernet AdaptersHP ProLiant Essentials Intelligent Networking Pack (Advanced Features)ProLiant Essentials Licensing support for the ProLiant Essentials Intelligent Networking Pack.Windows NT and Windows 2000
August 16, 2005	8.10.0.0	Added support for the following: <ul style="list-style-type: none">NC320i, NC324i, NC325i and NC326i LOMsRead-only web interface accessible through the System Management Homepage.The ability to create a detailed log of your system configuration for submission to third-level technical support personnel.Status of the active path and fast path states, along with member ID information and readability enhancements.Dynamic dual-channel load balancing, so user intervention is minimal and the NIC grouping happens automatically.User-specified preference order for transmit load balancing, so you can select which team member will be primary.
April 29, 2005	8.0.0.0	Added support for the following: <ul style="list-style-type: none">NC370x NICs.Detects IPMI enabled adapters and disables LSO to prevent members from not joining the team.
March 25, 2005	7.86.0.0	Resolved the following issue: <ul style="list-style-type: none">Communication failures when upgrading or creating Switch-assisted Load Balancing with Fault Tolerance (SLB) team types.
December 14, 2004	7.80.0.0	Added support for the following: <ul style="list-style-type: none">NIC Teaming Utilization Tab. This tab displays current and peak speed, throughput, and utilization information for the selected team and its members.Ability to replicate network configuration settings from one ProLiant to a group of target ProLiant servers using the System Insight Manager (HPSIM) group configuration mechanism. Removed support for the following: <ul style="list-style-type: none">HP NC3121 Fast Ethernet Adapter

Table 0-2 Teaming Driver Revision History

Date	Driver Version	Overview of Changes
September 11, 2004	7.71.0.0	Added support for the following: <ul style="list-style-type: none">• NC320T PCI Express Gigabit Server Adapter. Resolved the following issues: <ul style="list-style-type: none">• Blue screen that resulted when the Dual Channel feature was enabled and each channel did not have at least one NIC with link during system initialization.• NCU no longer fails to create teams on a system that has never had this software installed before.
August 18, 2004	7.70.0.0	Added support for the following: <ul style="list-style-type: none">• large-send offload (LSO) and checksum offload advanced teaming features, which can reduce TCP/IP processing overhead.• NC 320T PCI Express Gigabit Server Adapter, NC 150T PCI 4-port Gigabit Combo Switch Adapter, NC 310F PCI-X Gigabit Server Adapter.• 802.3ad dynamic teaming which allows the user to create channels on the switch without configuring the switch.• ProLiant Essentials Licensing support for the ProLiant Essentials Intelligent Networking Pack.• Active Path Failover guarantees the team to use a port that has a path to a selected destination.• Fast Path Failover allows the team to use the most expedient route to a destination.• Dual channel teaming allows receive load balancing on a team connected to two switches.• Advanced teaming allows teams with dissimilar ports to use large-send offload and checksum offload advanced teaming features, which can reduce TCP/IP processing overhead.

1 Introduction

1-1 Abstract

This whitepaper provides a high- and low-level discussion of the technology behind HP ProLiant Network Adapter Teaming for HP ProLiant servers running Microsoft® Windows®. HP ProLiant Network Adapter Teaming is software-based technology used by server administrators and network administrators to increase a server's network availability and performance. HP ProLiant Network Adapter Teaming provides network adapter, network port, network cable, switch, and communication path fault recovery technology, in addition to, transmit and receive load balancing technology.

1-2 How to read this document

This document is divided into three main sections. These sections have been designed to target three specific audiences: the executive, the first time user, and the advanced user. While a reader may skip to the section that is specific to their needs, HP recommends that any prior sections be read first because some information is not repeated in every section.

1-2-1 Section layout

1-2-1-1 Section 2 : An executive overview of teaming

Target audience: Executive

The purpose of this section is to provide a high-level, executive overview and introduction to HP ProLiant Network Adapter Teaming. The main purpose of this section is to provide information for an executive seeking business justification to deploy HP ProLiant Network Adapter Teaming.

1-2-1-2 Section 3 : Teaming fundamentals for the first-time user

Target audience: Beginner to intermediate user

In this section, an intermediate introduction to HP ProLiant Network Adapter Teaming is provided for the first time user or implementer. This section also can be used as a quick reference section for the occasional user. A general overview of the capabilities and options for HP ProLiant Network Adapter Teaming is provided.

1-2-1-3 Section 4 : The mechanics of teaming for the advanced user

Target audience: Intermediate to advanced user

This section provides detailed information about the design, implementation, and configuration of HP ProLiant Network adapter teaming. The purpose of this section is to assist networking specialists, systems engineers, and IM professionals in the design and troubleshooting of environments incorporating this technology in HP ProLiant servers. Topics for this section include advanced technical discussions of all features, failure recovery methods, load balancing logic, network scenario considerations, and so on. The reader should be familiar with the basics of IP communication and addressing, the OSI model, the use of network drivers, and the fundamentals of network switching. Additionally, the reader should be familiar with the terms found in the glossary of this white paper.

This white paper specifically discusses HP ProLiant Network Adapter Teaming for Microsoft Windows Server 2003 and Windows Server 2008. For technical information on network adapter teaming for other operating systems, go to www.hp.com.

Information in this document is specific to the set of drivers contained in release version NCDE 9.20. The NCDE driver release mechanism is used by HP to release all HP ProLiant networking drivers together. Because this white paper is about technology, most of the information is applicable to future releases; however, feature availability, behavior, and defaults may differ slightly between revision levels.

2 An executive overview of teaming

2-1 What is HP ProLiant Network Adapter Teaming?

HP ProLiant Network Adapter Teaming is software-based technology used by server administrators and network administrators to increase a server's network availability and performance. HP ProLiant Network Adapter Teaming provides network adapter, network port, network cable, switch, and communication path fault recovery technology, in addition to transmit and receive load balancing technology.

HP ProLiant Network Adapter Teaming is wholly developed by HP Engineering specifically for HP ProLiant customers.

2-2 The goals of HP ProLiant Network Adapter Teaming

The objective of HP ProLiant Network Adapter Teaming is to provide network fault tolerance and load balancing for HP ProLiant servers. These two objectives are accomplished by "teaming" together two or more server network adapter ports. The term "team" refers to the concept of multiple server network adapters (teamed ports), from the same server, working together as a single server network adapter (commonly referred to as a virtual network adapter).

2-2-1 Fault tolerance: self-healing network connectivity

In today's server environments, fault tolerance is provided at many levels—power supplies, fans, hard drives, processors, memory, and so on. One of the most often overlooked fault-tolerant devices is the server network adapter. Many server administrators (SA) spend thousands of dollars eliminating single points of failure in their servers because they want to make sure their servers are "highly available" to their users. Even when an SA realizes the need for network adapter hardware fault recovery, they don't always consider network cable fault recovery or even switch fault recovery. These aspects of network fault tolerance are often seen as the responsibility of the network engineer.

It is not uncommon for an SA to receive a 2 a.m. pager notification because the server lost network connectivity. For the SA that didn't plan for it, this can be a frustrating experience for the SA and for the server's clients. Regardless of all the redundancy built into the server's other hardware components, a server with single point of failure in its network connectivity is a server outage waiting to happen. The very expensive server can be rendered useless because the server can no longer perform its function on the network.

There are several considerations to keep in mind when planning for server network fault tolerance:

- **Network adapter failures**

Network adapter failures are typically hardware related—the network adapter hardware stopped working. However, these kinds of failures can often be caused by software-related problems (for example, driver failure). Other network adapter failures can be caused by an accident (for example, SA disables wrong adapter during server reconfiguration).

- **Server expansion slot failures**

This is defined as hardware failure of the server's expansion slot (for example, PCI, PCI-X, and so on) in which the network adapter is installed. While extremely uncommon, it can be a single point of failure. The deployment of a network adapter with two ports in a single PCI slot will provide network "port" fault tolerance but will not provide expansion slot fault tolerance.

- **Network cable disconnects**

A network cable disconnect is one type of failure that is often out of the SA's control. This type of failure can be caused by someone tripping over the network cable, by another SA choosing the wrong network cable to unplug from the server, or by a network engineer unplugging the wrong network cables from the switch.

- **Switch failures and misconfigurations**

Switch failures, or misconfigurations, are probably the most over-looked aspects of server network fault tolerance. SAs often expect the network administrator to provide a switch port to the server that has 100% uptime. While network administrators can achieve high availability in the network, they cannot promise 100% availability for the switch and switch port to which the server's network adapter port is connected. Switch failures, like server failures, do occur. Switch misconfigurations, like server misconfigurations, also occur. A prudent SA realizes and accounts for these potential failure scenarios even though they occur "outside" of the server chassis.

- **Upstream network failures**

If failures and misconfigurations can occur on the switch “directly” connected to the server, then failures and misconfigurations can occur on an upstream switch “indirectly” connected to the server. Unfortunately for the server, an upstream failure or misconfiguration has the potential to negatively affect the server as much as a directly connected switch. A server network fault-recovery solution that considers these potential upstream failures and misconfigurations is “enterprise aware,” resulting in a server with increased availability.

2-2-2 Load balancing: RAID 5 for server network adapters

Load balancing for server network adapters refers to the simultaneous use of multiple network ports (network adapters may contain one or more ports each) to provide increased performance for transmitted and received network traffic on an individual server. The term “load balancing” is used to see many different computer and network technologies. A server administrator can load balance data across hard drives using RAID technology or load balance server tasks across a cluster of servers using Microsoft Clustering. Similarly, a server administrator can load balance a single server’s network traffic across two or more network ports within the same server.

There are several considerations to keep in mind when planning for server network load balancing:

- Utilization of designated fault-tolerant resources

Providing network fault tolerance requires additional network adapter/port resources. These designated fault-tolerant resources sit idle until a failure occurs. Load balancing can be implemented to fully utilize the idle network resources that would otherwise remain in a standby state until a failure occurred. The rationale for deploying network load balancing can be compared to the decision to deploy RAID 1 instead of RAID 5. While RAID 5 provides for fault tolerance just like RAID 1, RAID 5 also allows for better performance because all hard drive hardware are in use. Similarly, an SA has the option of using HP ProLiant Network Adapter Teaming to increase server network performance by load balancing server traffic across fault tolerant network adapter resources. Deploying load balancing with fault tolerance versus fault tolerance alone provides better resource utilization.

- Investment protection in older technology

Load balancing is very important in situations where investment protection in older technology (for example, Fast Ethernet – 100 Mb) is desired; yet, increased throughput capabilities are required. Instead of discarding Fast Ethernet network adapters to install Gigabit network adapters because a server needs more throughput, an SA can utilize several Fast Ethernet network adapters to increase server network throughput up to a theoretical maximum of 800 Mbps. Later when the SA has the budget resources to purchase newer technology (for example, Gigabit Ethernet – 1000 Mb), he can replace the team of Fast Ethernet adapters with a team of Gigabit Ethernet adapters. As a result, network adapter load balancing can be used to bridge the gap between cheaper (older) technology and more expensive (newer) technology.

HP ProLiant Network Adapter Teaming provides 1) solutions for both fault tolerance and load-balancing requirements, 2) solutions that allow the server to “heal” itself and restore network connectivity without human intervention, and 3) solutions that allow the server to fully utilize its available network adapter resources.

By utilizing an HP ProLiant Network adapter team of two or more network adapter ports, an SA provides the server with the ability (intelligence) and capability (components) to restore lost network connectivity while simultaneously increasing server network performance. Most of today’s servers ship standard with at least two network adapter ports, many built directly onto the motherboard (in other words, LAN On Motherboard or LOM). The SA need only install HP ProLiant Network Adapter Teaming software, select at least two network adapter ports, and create a team. Everything else works the same, both inside the server (server’s protocol stack) and outside of the server (the server’s clients).

With a small investment of time and a small investment in network adapter hardware, SAs can achieve maximum server network availability and server network performance. This small investment (teaming) assures that the larger investment (the server) is utilized to its fullest.

2-3 Why HP ProLiant Network Adapter Teaming?

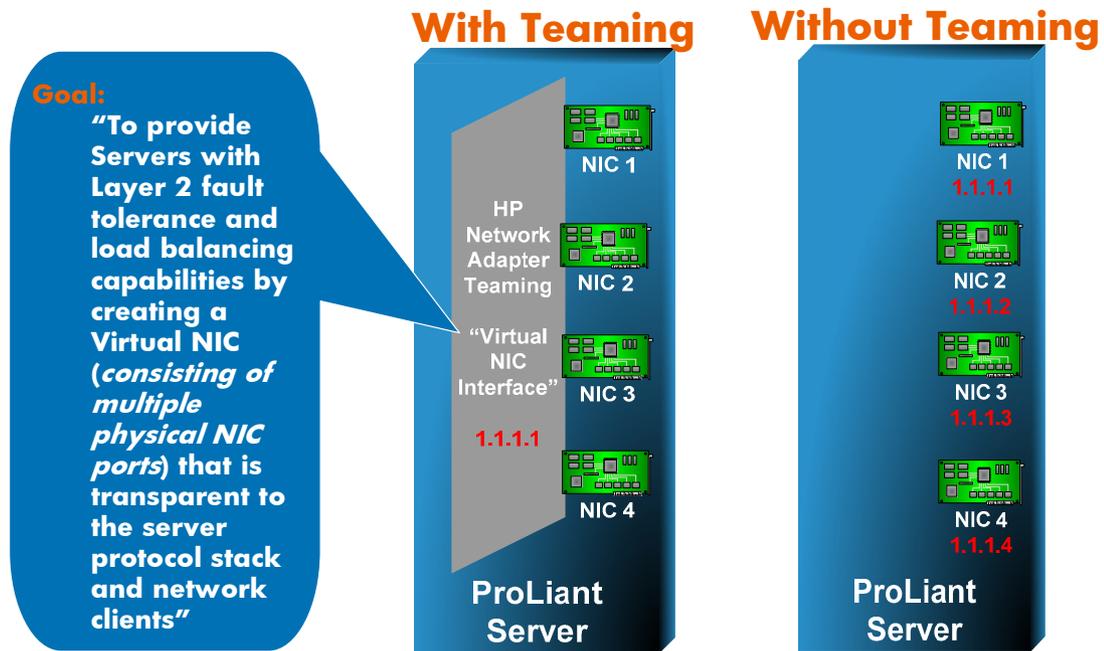
As a company that innovates in server and networking technology, HP is in a unique position to provide fully featured “server networking technology” solutions for HP ProLiant customers. HP has made a significant investment in its Network Adapter Teaming technology through patents and publications and extensive hardware and software development activities. This purposeful investment is a result of HP’s desire to provide HP ProLiant server customers with a consistent and complete suite of advanced server networking technology solutions. In addition, the HP ProLiant Network Adapter Teaming product provides solutions to many server networking problems that are not solved by any other product on the market.

3 Teaming fundamentals for the first-time user

3-1 A technical overview of HP ProLiant Network Adapter Teaming

HP ProLiant Network Adapter Teaming provides fault tolerance and load balancing across a team of two or more network adapter ports. The term “team” refers to the concept of multiple network adapters (teamed ports) working together as a single network adapter, commonly referred to as a virtual network adapter or virtual NIC interface.

Figure 3-1 HP ProLiant Network Adapter Teaming

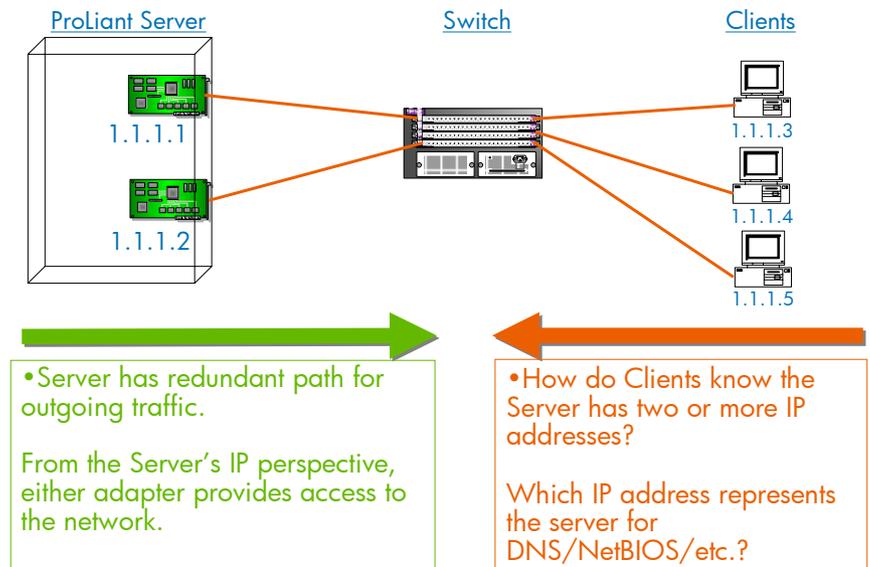


3-1-1 Fault tolerance: Dual Homing or Network Adapter Teaming?

When considering network redundancy for a server, there are two commonly used solutions: dual homing or Network Adapter Teaming (sometimes referred to as bonding).

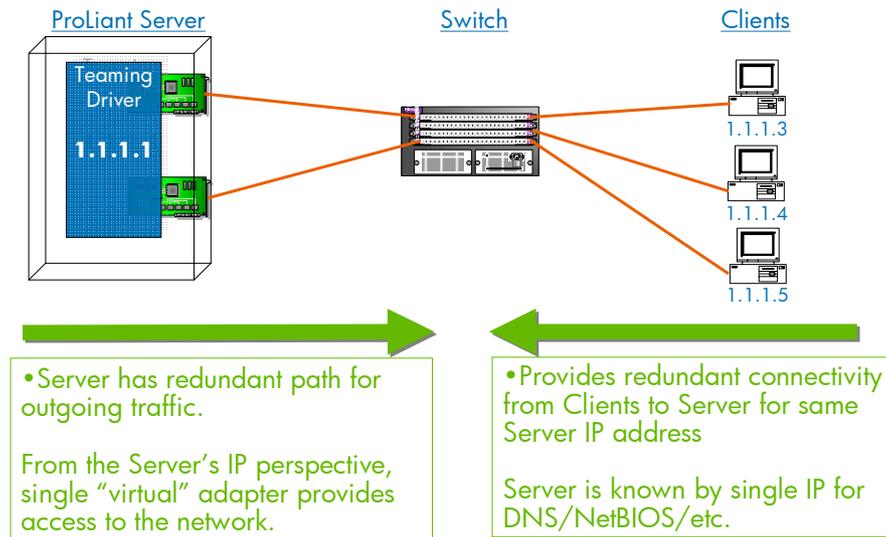
Dual Homing a server involves installing two or more network adapters (NICs) in the server, assigning an individual IP address to each adapter and connecting the adapters to the network (see Figure 3-2). Dual Homing provides path redundancy for the server to the network. However, Dual Homing does not completely solve the network redundancy problem for a server because Dual Homing does not provide IP address redundancy. For example, if the Multihomed server in Figure 3-1 loses network connectivity on NIC 1, any clients that were connected to the server using IP address 1.1.1.1 will lose connectivity. For the clients to reconnect to the server, they would have to discover the alternate IP address somehow and then reinitiate connectivity. While Dual Homing provides semi-transparent redundancy for the server’s outbound connectivity, it does not provide transparent redundancy for clients connecting to the server.

Figure 3-2 Using Multihoming for server network redundancy



In contrast, Network Adapter Teaming provides transparent redundancy for both the server's outbound connectivity and all inbound client connectivity (see Figure 3-3). Full transparent redundancy for the IP address is provided by teaming. A network port failure (or cable failure, or upstream switch failure, and so on) in the server with teaming as identified in Figure 3-1, will be completely transparent to the server's protocol stack and to the server's clients. Any client connections to the server using the server's IP address (bound to the virtual network adapter interface rather than the network adapter hardware) remain active and unaware that an individual network port failed.

Figure 3-3 Using NIC teaming for server network redundancy



3-1-2 Load balancing: server-based routing protocol or NIC teaming?

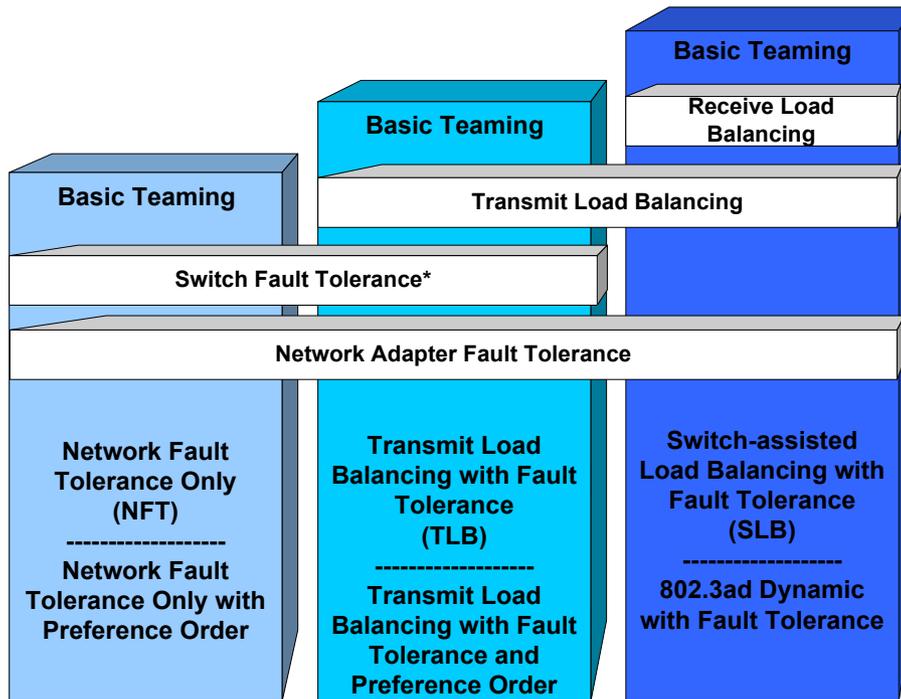
There are two widely used solutions for achieving load balancing for a server's network traffic. An SA can deploy a server-based routing protocol that supports load balancing (for example, OSPF, EIGRP, IS-IS, and so on) or the SA can deploy teaming. Server-based routing protocols are much more difficult to deploy and manage. In addition, they do not allow for load balancing between end devices that do not support the routing protocol (for example, clients on the same local area network (LAN) with the server would have to 'speak' the routing protocol).

Teaming is simpler to understand and easier to deploy than server-based routing protocols. In addition, teaming supports load balancing without requiring any kind of special configuration or feature support on the network devices (for example, clients) with which the team is communicating.

3-2 Overview of team types

HP ProLiant Network Adapter Teaming is a collection of fault-tolerant and load-balancing features that work together in different combinations called team types. These team types provide HP ProLiant Network Adapter Teaming with the flexibility to support almost every network environment. Figure 3-4 provides a graphical overview of which team types support which redundancy and load-balancing features.

Figure 3-4 Teaming types and functionality



* Switch Fault Tolerance means Team supports members connected to more than one switch.

3-2-1 Network Fault Tolerance Only (NFT)

Network Fault Tolerance (NFT) is the foundation of HP ProLiant Network Adapter Teaming. In NFT mode, from two to eight teamed ports are teamed together to operate as a single virtual network adapter. However, only one teamed port—the primary teamed port—is used for both transmit and receive communication with the server. The remaining adapters are considered to be stand-by (or secondary adapters) and are referred to as non-primary teamed ports. Non-primary teamed ports remain idle unless the primary teamed port fails. All teamed ports may transmit and receive heartbeats, including non-primary adapters.

The fault-tolerance feature that NFT represents for HP ProLiant Network Adapter Teaming is the only feature found in every other team type. The foundation of every team type supports NFT.

3-2-2 Network Fault Tolerance Only with Preference Order

Network Fault Tolerance Only with Preference Order is identical in almost every way to NFT with the only difference being that this team type allows the SA to prioritize the order in which teamed ports should be the primary teamed port. This ability is important in environments where one or more teamed ports are more preferred than other ports in the same team. The need for ranking certain teamed ports better than others can be a result of unequal speeds, better adapter capabilities (for example, higher receive/transmit descriptors or buffers, interrupt coalescence, and so on), or preference for the team's primary port to be located on a specific switch.

3-2-3 Transmit Load Balancing with Fault Tolerance (TLB)

Transmit Load Balancing with Fault Tolerance (TLB) is a team type that allows the server to load balance its transmit traffic. TLB is switch independent and supports switch fault tolerance by allowing the teamed ports to be

connected to more than one switch in the same LAN. With TLB, traffic received by the server is not load balanced. The primary teamed port is responsible for receiving all traffic destined for the server. In case of a failure of the primary teamed port, the NFT mechanism ensures connectivity to the server is preserved by selecting another teamed port to assume the role.

3-2-4 Transmit Load Balancing with Fault Tolerance and Preference Order

Transmit Load Balancing with Fault Tolerance and Preference Order is identical in almost every way to TLB with the only difference being that this team type allows the SA to prioritize the order in which teamed ports should be the primary teamed port. This ability is important in environments where one or more teamed ports are more preferred than other ports in the same team. The need for ranking certain teamed ports higher than others can be a result of unequal speeds, better adapter capabilities (for example, higher receive/transmit descriptors or buffers, interrupt coalescence, and so on), or preference for the team's primary port to be located on a specific switch.

3-2-5 Switch-assisted Load Balancing with Fault Tolerance (SLB)

Switch-assisted Load Balancing with Fault Tolerance (SLB) is a team type that allows full transmit and receive load balancing. SLB requires the use of a switch that supports some form of Port Trunking (for example, EtherChannel, MultiLink Trunking, and so on). SLB does not support switch redundancy because all ports in a team must be connected to the same switch. SLB is similar to the 802.3ad Dynamic team type that is discussed later in this white paper.

3-2-6 802.3ad Dynamic with Fault Tolerance

802.3ad Dynamic with Fault Tolerance is identical to SLB except that the switch must support the IEEE 802.3ad dynamic configuration protocol called Link Aggregation Control Protocol (LACP). In addition, the switch port, to which the teamed ports are connected, must have LACP enabled. The main benefit of 802.3ad Dynamic is that an SA will not have to manually configure the switch.

802.3ad Dynamic is a standard feature of HP ProLiant Network Adapter Teaming.

3-2-7 Automatic

The Automatic team type is not really an individual team type. Automatic teams decide whether to operate as an NFT, or a TLB team, or as an 802.3ad Dynamic team. If all teamed ports are connected to a switch that supports the IEEE 802.3ad Link Aggregation Protocol (LACP) and all teamed ports are able to negotiate 802.3ad operation with the switch, then the team will choose to operate as an 802.3ad Dynamic team. However, if the switch does not support LACP or if any ports in the team do not have successful LACP negotiation with the switch, the team will choose to operate as a TLB team. As network and server configurations change, the Automatic team type ensures that HP ProLiant servers intelligently choose between TLB and 802.3ad Dynamic to minimize server reconfiguration.

Select and download the driver labeled "HP ProLiant Networking Software." The downloadable package is called a "Component Pack" and has a file name in the form CPxxxx.exe.

3-4-2 Install

The HP ProLiant Network Adapter Teaming driver is a self-installing executable. Launch the CPxxxx.exe file and follow directions to complete the installation. If error messages are received because of outdated network card drivers, go back to the URL listed in step 3-5-1 and choose the latest drivers for the network card.

3-4-3 Launch

The NCU can be launched from the tray icon (unless disabled) or from the Control Panel applet labeled HP Network.

3-5 Overview of the HP Network Configuration Utility

HP ProLiant Network Adapter Teaming is configured and monitored through a graphical user interface entitled the HP Network Configuration Utility (NCU). NCU is also used for configuring and monitoring individual network adapters.

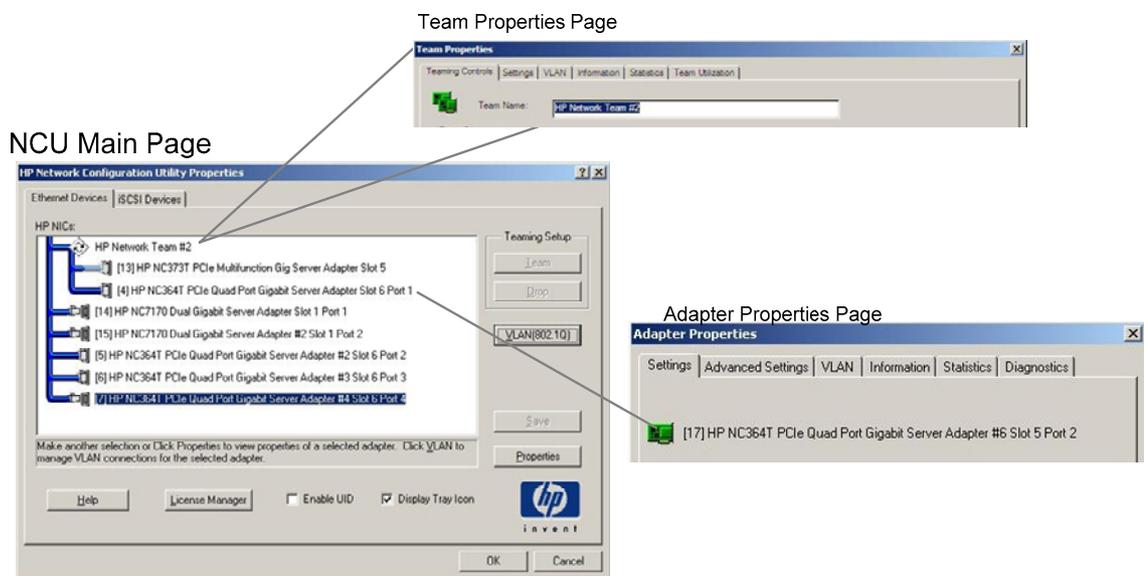
3-5-1 NCU main window

The NCU main window, labeled HP Network Configuration Utility Properties, is used to create and dissolve teams, to monitor the overall health of teams and individual network adapters, and to export team configuration settings as an XML file. The main window is also used to access the Properties windows for team settings and NIC settings.

3-5-1-1 Primary actions from the NCU main window

- Creating a team—Select two or more individual (non-teamed) network adapter ports and then select **Team**.
- Adding a team member—Select an existing team and one or more (non-teamed) network adapter ports and selects **Add**.
- Deleting a team—Select an existing team and then select **Dissolve**.
- Dropping a team member—Select a member of an existing team and then select the **Drop**.
- Team settings—Select an existing team and then select **Properties**.
- Adapter Properties—Select any teamed or non-teamed port and then select **Properties**.
- Viewing Licenses—To view currently installed licenses, select **License Manager**.

Figure 3-6 NCU main window



3-5-2 Team Properties window

The Team Properties window (see Figure 3-6) is used to manage and monitor all team-specific settings and consists of the following individual tabs.

3-5-2-1 Team Properties window tabs

- Teaming Controls—Changes the team name, select team type, select the Transmit Load Balancing algorithm, and change port preference order for NFT with Preference teams.

- Settings—Sets the team’s Locally Administered Address (LAA) MAC address, manage heartbeat timers, and manage the team-level advanced property settings (for example, max frame size, checksum offloading, Large Send Offload, and so on).
- VLAN—Assigns a team to one or more VLANs, and is also used to manually choose which VLANs to use for the default or native VLAN, and VLAN for receive path validation (heartbeats).
- Information—Monitors team status information. Examples of information that can be monitored are current team mode, overall team operating speed and duplex, team state, driver version, primary teamed port, and individual teamed port status.
- Statistics—Monitors real-time team statistics such as number of frames received/transmitted (broadcast versus multicast versus unicast), bytes transmitted/received, number of failovers, heartbeats transmitted/received, and so on.
- Team Utilization—Monitors the effectiveness of load balancing by providing a real-time numeric and graphical view of team-level utilization by percentage and bandwidth, as well as individual teamed port-level usage information.

3-5-3 Adapter Properties window

The Adapter Properties window (see Figure 3-6) is used to manage and monitor all adapter-specific settings and consists of the following individual tabs.

3-5-3-1 Adapter Properties window tabs

- Settings—Sets the individual adapter’s speed and duplex and sets a non-teamed port’s locally administered MAC address (LAA).
- Advanced Settings—Manages per-adapter advanced settings such as max frame size, jumbo packet (Windows Server 2008 only), Receive-Side Scaling (RSS), TCP/IP Offload Engine (TOE), checksum offloading, Large Send Offload, coalesce buffers, transmit/receive descriptors, Wake on LAN, etc. For teamed ports (adapters), many of these properties are only configurable at the team level by using the Settings tab on the Team Properties window described in “3-5-2 Team Properties window.”
- VLAN—Manages VLAN membership for non-teamed ports. Teamed ports do not have this page because VLAN membership is managed at the team level. See the VLAN tab described in “3-5-2 Team Properties window.”
- Statistics—Monitors real-time adapter statistics such as number of frames received/transmitted, bytes transmitted/received, heartbeats transmitted/received for teamed ports, and so on. The Advanced button can be activated to monitor advanced statistics such as CRC errors, collisions, and underruns on a per-adapter basis.
- Information—Monitors adapter status information. Examples of information that can be monitored are current teamed status, current speed/duplex, current MAC address, driver version, part number, bus/iobay/port (for HP BladeSystem servers), and bus/slot/port information, and so on.
- Diagnostics—Runs manual diagnostics on a per-adapter level. The diagnostic tests are manually selected and executed by selecting **Run Tests**. Running Diagnostic tests on an adapter causes its normal network operation to be halted. Diagnostic tests are not recommended for adapters in production.

3-6 Getting started

3-6-1 Example deployment of basic teaming

1. Perform the initial installation.
Complete the “Basic deployment steps for HP ProLiant Network Adapter Teaming” steps in section 3-4 above.
2. Select the network adapter ports for a team.
From the NCU main window, select (by clicking an adapter to highlight it) two or more network adapter ports then click **Team**.
3. Set the team type to **Automatic**.
 - a. Click **Properties** on the NCU main window for the newly formed team (white team icon).
 - b. Under Team Type Selection on the Teaming Controls tab, notice that Automatic (Recommended) is already chosen. This is the default setting for new teams. Note the name assigned to the team in the Team Name window (for example, HP Network Team #1).
 - c. (Optional) Rename the team by changing the text in the Team Name window.
 - d. Click **OK** on the Team Properties window when finished.
4. Apply all changes.
 - a. Click **OK** on the NCU main window.
 - b. Click **Yes** when asked if all configuration changes should be applied.
 - c. Wait until the All Configuration Changes Were Made Successfully dialog box appears.
5. Assign an IP address to the team.
 - a. Open the Network Connections (Network and Dial-up Connections) window from the Control Panel.
 - b. Under Device Name (in the Details view), open the Properties window for the device with the same name as noted in step 3.
 - c. Assign the appropriate IP address information based on the connected network requirements.

4 The mechanics of teaming for the advanced user

4-1 Section objectives and prerequisites

This section is intended to provide an in-depth technical discussion of the mechanics of HP ProLiant Network Adapter Teaming for the advanced user. This section assumes the reader has read the previous sections in this white paper and has hands-on experience in the deployment and usage of HP ProLiant Network Adapter Teaming.

4-2 Architecture of HP ProLiant Networking Adapter Teaming

4-2-1 The “networking layers” of an operating system

Within an operating system (OS), a hierarchy of layers work together to enable one OS to communicate with another OS. Each of these layers performs a separate function and passes information between the layers above and below it. Within Windows Server 2003 and Windows Server 2008, there are four layers that are important to understand when discussing HP ProLiant Network Adapter Teaming: the miniport layer, intermediate layer, NDIS layer, and protocol layer.

- **Miniport layer**
The network adapter driver resides at the miniport layer. This driver is responsible for directly controlling the hardware. It is necessary for basic network adapter functionality and is used even when HP ProLiant Network Adapter Teaming is not deployed. Typically, this driver is written by the OEM vendor of the network adapter hardware. HP network adapter drivers (for example, NXP2NIC.SYS and Q57ND60X.sys) are referred to as miniport drivers.
- **Intermediate layer**
The intermediate layer driver provides a network function, but is not considered a miniport because it does not directly control a piece of hardware. The intermediate layer driver performs a function that is in between the miniport layer and NDIS (OS protocol stack). The networking function that is performed by the intermediate layer is beyond the ability of a miniport layer driver. In this case, HP ProLiant Network Adapter Teaming is considered an intermediate driver (CPQTEAM.SYS) and functions to make several miniport drivers seamlessly work as a single virtual network adapter that interfaces with NDIS. VLANs are also implemented as this layer. Another example of an intermediate driver is the NLB (Network Load Balancing) feature in Microsoft Windows.
- **NDIS**
Microsoft’s implementation of the Network Driver Interface Specification (NDIS), handles communication between the all layers, either miniport drivers or intermediate drivers, and the protocol layer (for example, IP, IPX, and so on). NDIS is more than just a single layer because it operates “around” all the layers to allow them to work together.
- **Protocol layer**
The protocol layer is where IP, IPX, AppleTalk, and the like interface with NDIS. This layer is responsible for Protocol addresses (for example, IP or IPX addresses), and also for translating Layer 3 addresses (in other words, IP addresses) to Layer 2 addresses (in other words, MAC addresses).

In the absence of an intermediate driver, a protocol address is usually assigned to each individual miniport driver. However, when utilizing HP ProLiant Network Adapter Teaming, the protocol address is assigned to a single HP ProLiant Network Adapter Teaming instance that represents the underlying miniports. If more than one HP network adapter team exists in a single server, there will be more than one instance of the HP network adapter team and an individual protocol address will be assigned to each instance.

4-2-2 Teaming software components

HP ProLiant Network Adapter Teaming consists of three components: the miniport driver, intermediate driver, and configuration GUI.

4-2-2-1 Network adapter miniport driver

For Microsoft Windows 2008, the miniport driver used with the HP network adapter is N1G60i32.sys, Q57ND60X.sys, N1E6032.sys, or BXND60X.sys depending on the adapter in use.

For Microsoft Windows 2003, the miniport driver used with the HP network adapter is Q57XP32.SYS, N1E5132.sys, BXND51X.sys, or NXP2NIC.sys (does not support teaming in this release) depending on the adapter in use.

4-2-2-2 Teaming intermediate driver

For Microsoft Windows 2003 and 2008, the intermediate driver is CPQTEAM.SYS, and is used for all teaming functions involving HP NC series adapters.

4-2-2-3 HP Network Configuration Utility

The configuration GUI is called the HP Network Configuration Utility (NCU) and the file name is CPQTEAM.EXE. The configuration GUI is accessible from the Control Panel or from the Tray icon (unless disabled).

These three components are designed to work as a single unit. When one is upgraded all components must be upgraded to the current version. For driver updates to HP network adapters and HP ProLiant Network Adapter Teaming, see <http://h18004.www1.hp.com/support/files/networking/us/index.html>.

4-2-3 HP teaming and Layer 2 versus Layer 3 addresses

When implementing HP ProLiant Network Adapter Teaming, one of the most important concepts to understand is Layer 2 and Layer 3 addresses and the way they are handled. When network adapters are teamed together, they function as a single virtual network adapter. Other network devices (for example, PCs, other servers, routers, and so on) communicating with an HP network adapter team cannot distinguish that they are communicating with more than one network adapter. In addition, HP ProLiant Network Adapter Teaming must maintain strict IEEE standards compliance in its use of Layer 2 and Layer 3 addresses.

For an HP network adapter team to appear as a single virtual network adapter, it is necessary for all networking devices to see the team by a single Layer 2 address and a single Layer 3 address. In other words, when a device is communicating with a team, regardless of the number of network adapters that make up the team, the network device only “sees” one MAC address and one protocol address (for example, IP or IPX). When communicating using IP, this means that a networking device will have only one entry in its ARP (Address Resolution Protocol) cache for an HP network adapter team regardless of the number of network adapter ports that make up the team.

When an HP network adapter team initializes, the teaming driver for each team “reads” the burned in MAC address (BIA) for each network adapter assigned to that particular team. Essentially, the MAC addresses are decoupled from the network adapters and pooled together for use by the teaming driver. The teaming driver picks one MAC address as the team’s MAC address and assigns it to the primary adapter, unless the user has manually set the MAC address (Locally Administered Address) via the NCU. For all team types all ARP replies from the server for this particular HP network adapter team provide this same MAC address as the team’s MAC address. This address does not change unless the team is reconfigured. The teaming driver assigns the remaining MAC addresses to the non-primary adapters.

When a failover event occurs, the MAC addresses of the current primary adapter and one of the non-primary adapters are swapped. The former non-primary adapter becomes the new primary adapter and the former primary adapter becomes a non-primary adapter. By swapping the MAC addresses in this manner, the HP network adapter team is always known by one MAC address and one protocol address. It is unnecessary for protocol addresses to swap during a failover event because the protocol address is directly assigned to the intermediate (teaming) driver and not to the miniport driver.

When transmitting frames, the current primary adapter always transmits using the team’s MAC address as the Layer 2 address and the team’s protocol address as the Layer 3 address. Non-primary adapters always transmit using the MAC address assigned to them by the teaming driver and using the team’s protocol address as the Layer 3 address. For NFT and TLB, the MAC address used by non-primary adapters when transmitting is always different from the primary adapter’s MAC address and is always unique from that of any other non-primary adapter, to comply with IEEE standards. For SLB, the additional switch intelligence allows all teamed ports to transmit using the same MAC address, the team’s MAC address.

A network device communicating with an HP network adapter team may receive frames from more than one network adapter in the same team. When this happens, the network device does not know that more than one MAC address is being used. The important issue is that all frames originating from the same HP network adapter team use the same protocol address. The network device does not know that multiple MAC addresses are coming from the team because MAC headers are stripped off before the frames are processed up the stack by the operating system of the network device. When the operating system receives the frames, they all appear as though they came from the same network adapter. In addition, ARP cache entries are not created by examining the MAC addresses from received frames. ARP cache entries are ONLY created from ARP requests and ARP replies or from manually creating static entries.

4-3 Types of HP ProLiant network adapter teams

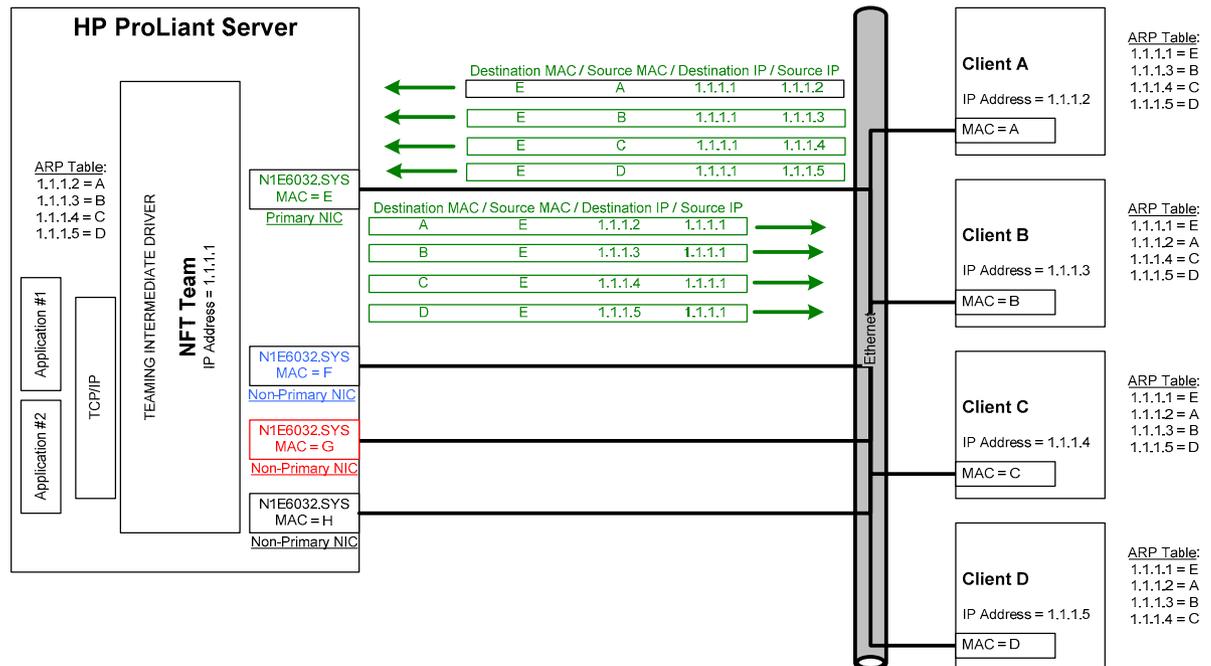
There are seven team types for HP network adapters: Network Fault Tolerance Only (NFT), Network Fault Tolerance Only with Preference, Transmit Load Balancing with Fault Tolerance (TLB), Transmit Load Balancing with Fault Tolerance and Preference Order, Switch-assisted Load Balancing with Fault Tolerance (SLB), 802.3ad Dynamic with Fault Tolerance, and Automatic. While there are seven different team types, there are three base team types from which all other team types are derived. The three base team types are NFT, TLB, and SLB. Respectively, each mode gains in features and incorporates most features from the previous teaming mode. In other words, NFT is the simplest teaming mode, supporting only network adapter fault tolerance. TLB supports network adapter fault tolerance plus load balancing of any traffic being transmitted from the server. SLB supports network adapter fault tolerance, load balancing of any traffic transmitted from the server, plus load balancing of any traffic received by the server.

For a graphical overview of team types and functionality, see Figure 3-4 on page 13.

4-3-1 Network Fault Tolerance Only (NFT)

Network Fault Tolerance (NFT) is the foundation of HP ProLiant Network Adapter Teaming. In NFT mode, two to eight ports are teamed together to operate as a single virtual network port. However, only one network port, the primary port, is used for both transmit and receive communication with the server. The remaining ports are considered to be in stand-by (or secondary ports) and referred to as non-primary ports, and remain idle unless the primary port fails. All ports may transmit and receive heartbeats, including non-primary ports.

Figure 4-1 Overview of NFT communication



4-3-1-1 Network addressing and communication with NFT

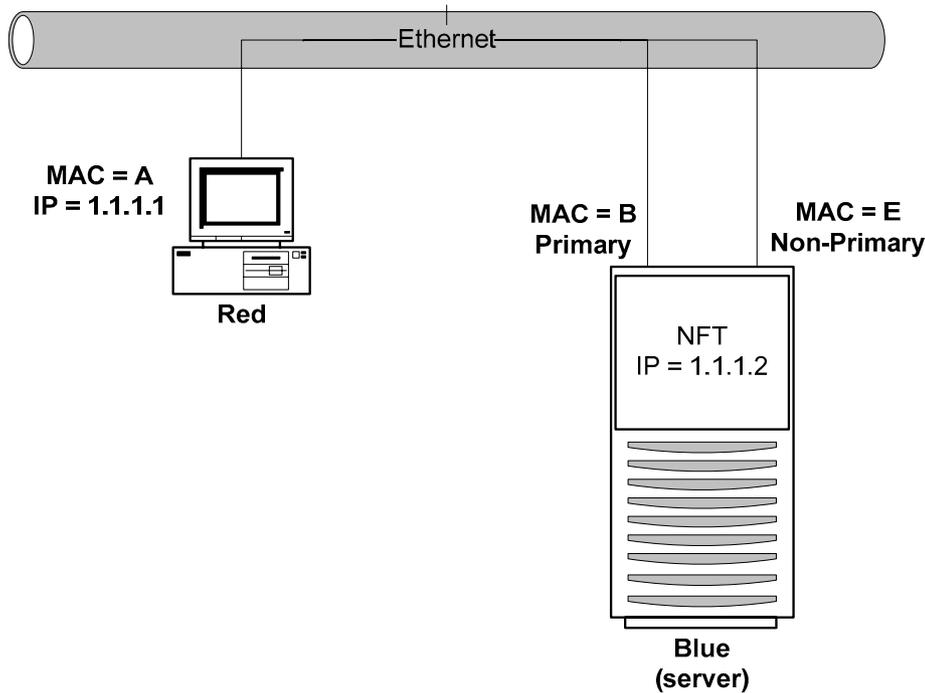
Before learning the specifics of NFT and how it communicates on the network, HP recommends that “4-2-3 HP teaming and Layer 2 versus Layer 3 addresses” be thoroughly reviewed and understood as well as “Appendix A: Overview of network addressing and communication.”

4-3-1-1-1 Scenario 1-B: A device PINGs an NFT team on the same Layer 2 network.

This section builds on the concepts reviewed in “Appendix A: Overview of network addressing and communication,” and describes how NFT functions from the network addressing and communication perspective.

Utilizing a network diagram similar to Figure A-1 on page 23, Blue has been modified to be a server utilizing an HP network adapter team in NFT mode with two network ports in a team (see Figure 4-2). The two network ports have MAC addresses of B and E, and are known by a single Layer 3 address of 1.1.1.2. Network port B has been designated as the primary port in this NFT team.

Figure 4-2 Scenario 1-B: A device PINGs an NFT team on the same Layer 2 network



4-3-1-1-1a) Red transmits a broadcast ARP request asking for Blue’s MAC address.

A user on Red issues the command `ping 1.1.1.2` to initiate a PING to Blue. First, Red determines whether or not Blue is on the same Layer 2 network. After Red has determined that Blue is on the same Layer 2 network, Red must find out what Blue’s MAC address is. First, Red checks its own ARP cache for a MAC address entry matching the IP address of 1.1.1.2. If Red does not have a static entry or an entry cached from a previous conversation with Blue, then it must broadcast an ARP request frame on the network asking Blue to respond and provide its MAC address. Red must broadcast this ARP request because without knowing Blue’s unique MAC address, it has no way of sending a frame directly (unicast) to Blue.

4-3-1-1-1-2b) Blue transmits a unicast ARP reply to Red, providing its MAC address.

Blue sees the ARP request (the frame is received on both the primary and non-primary teamed ports) because the frame is broadcast on the network. However, the team discards all non-heartbeat frames incoming on non-primary ports, and responds with a unicast ARP reply to Red. The ARP reply is transmitted by the primary port (B). In Blue’s ARP reply, Blue provides the MAC address of its teaming driver, which is the same as the current primary port’s MAC address (B) (see “4-2-3 HP teaming and Layer 2 versus Layer 3 addresses”). Blue also notes Red’s MAC address (A) and IP address (1.1.1.1) and enters them into its ARP cache. Red receives the reply and enters the MAC address (B) and the IP address of Blue (1.1.1.2) into its own ARP cache.

4-3-1-1-1-3c) Red transmits a unicast PING request to Blue using Blue’s destination MAC address.

Red can now create a PING request frame using Blue’s MAC address (B). Red sends the PING request to Blue. Blue receives the frame on its primary port (B) and notices that a station with an IP address of 1.1.1.1 is asking for it to respond.

4-3-1-1-1-4d) Blue transmits a broadcast ARP request asking for Red’s MAC address.

Note: The following step may not occur if Blue’s ARP table still contains an entry for Red as a result of steps (a) and (b).

Blue checks its ARP cache for a MAC address entry that matches 1.1.1.1. If Blue does not find one, then Blue broadcasts an ARP request asking for Red’s MAC address.

4-3-1-1-1-5e) Red transmits a unicast ARP reply to Blue providing its MAC address.

Note: The following step will not occur if step (d) does not take place.

Red sees the ARP request and transmits a unicast ARP reply directly to Blue providing its MAC address (A). Blue receives the ARP reply and puts Red’s MAC address (A) and IP address (1.1.1.1) in its ARP cache.

4-3-1-1-1-6f) Blue transmits a unicast PING reply to Red using Red's destination MAC address.

Blue then transmits a unicast PING reply to Red using Red's MAC address (A) and the user sees the PING reply message printed on the screen. This completes the entire conversation.

4-3-1-2 NFT redundancy mechanisms

4-3-1-2-1 Basic redundancy mechanisms

4-3-1-2-1-1 Link loss

Link loss is the most fundamental of all redundancy mechanisms. Link loss literally monitors a teamed port for physical link. If physical link is lost, the teamed port is disabled for use by the team.

4-3-1-2-1-2 Heartbeats

Heartbeats are special multicast frames that HP ProLiant Network Adapter Teaming uses for validating team member network connectivity and for notifying other network equipment of MAC address changes as a result of failover events (see "Path validation functions" for a complete description of the different heartbeat types).

Heartbeat frames contain only Layer 2 addresses (see "Heartbeat frame format" for the complete frame format of a heartbeat) and do not contain any Layer 3 addresses. This means that heartbeat frames are not routable. In other words, heartbeat frames will not be routed (by a router) between team members if the team members are on two different Layer 2 networks joined by a Layer 3 device (router). If the heartbeat frames are not delivered between team members, then erroneous failovers may occur.

Path validation functions

There are three main functions of the path validation frame used for HP ProLiant Network Adapter Teaming: receive validation, transmit validation, and Switch MAC table updates. The same path validation frame format is used for all heartbeat functions and they are indistinguishable from the network's perspective.

Transmit path validation

Transmit path validation is used for checking the transmit path for all teamed ports (primary and non-primary). If a particular teamed port hasn't transmitted any frames within a certain amount of time, the transmit path validation attempts to transmit a heartbeat frame to test the port's ability to transmit. If the port is able to successfully place a frame on the wire without receiving a physical layer error, the test succeeds. This feature can be disabled, but fault tolerance will be reduced.

Receive path validation

Receive path validation is used for checking the receive path for all teamed ports (primary and non-primary). If a particular teamed port hasn't received any frames within a certain amount of time, the receive path validation attempts to validate the ports ability to receive by having another teamed port transmit a heartbeat frame. If the teamed port under test is able to receive the heartbeat frame, the test succeeds. This feature can be disabled, but fault tolerance will be reduced.

Update Switch MAC table with team MAC address

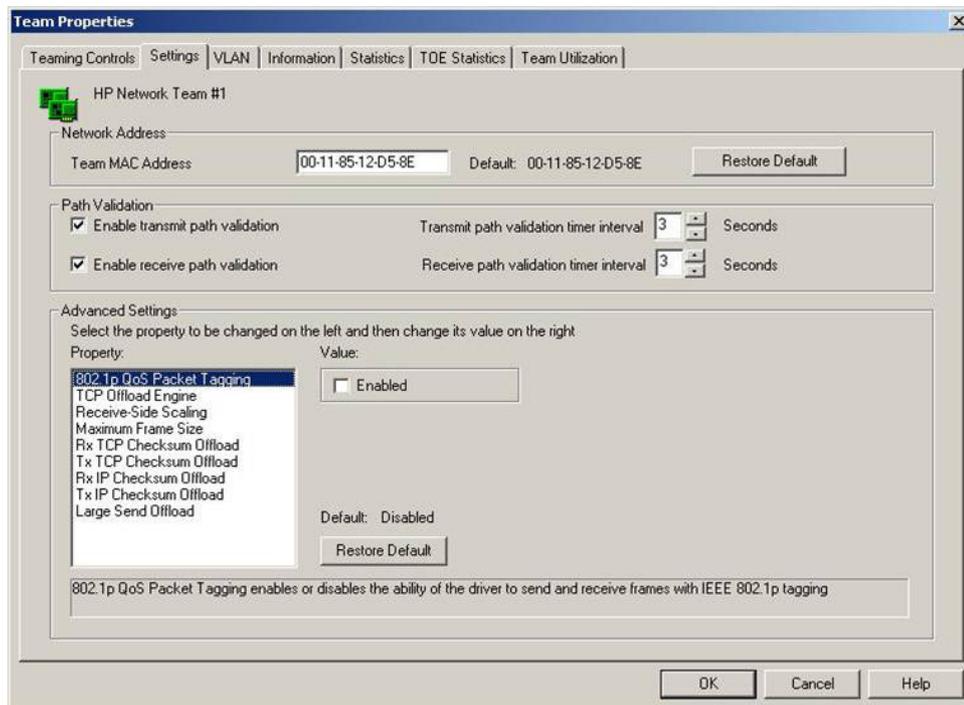
The Primary Port must ensure that the switch has the team's MAC address on the correct port, and that the switch keeps the team's MAC address in its MAC table [or Content Addressable Memory (CAM) table]. The Switch MAC table update heartbeat is used with most team types when a failover occurs, and a non-primary port takes over the primary role. This allows the new primary port to update the Switch MAC table immediately so traffic destined for the team will make a smooth transition and not experience timeouts or failures. This feature cannot be disabled and is always active on certain team types regardless whether receive or transmit path validation heartbeats are enabled or disabled.

Path validation timers

HP ProLiant Network Adapter Teaming utilizes a timer mechanism for checking network connectivity with heartbeats. Two timers are available for custom tuning on the Settings Tab of the Team Properties window in the NCU.

Note: Transmit path validation and receive path validation are both enabled by default. Each has a default value of 3 seconds. If manual configuration is desired, the valid range for both timers is 3 to 60 seconds.

Figure 4-3 Team Properties window—Settings tab containing the Path Validation settings



Transmit path validation

At the transmit path validation timer interval (default = 3 seconds), the team checks to determine if each port has successfully transmitted any frame since the last timer interval. If not, then the port's internal status is degraded by incrementing the port's internal status by 1 (starting at 0). After the port's internal status has degraded to 2 (2 transitions) without transmitting any kind of frame, that port will transmit a heartbeat frame. Therefore, a transmit path validation heartbeat frame may be transmitted once every 6 seconds (if the transmit path validation interval timer is set to the default of 3 seconds) on a port that has not transmitted anything. If the port does successfully transmit a heartbeat or any other kind of frame, its internal status is reset to 0. However, if transmission is not successful, the team places the port in a state called "Failed (TX Heartbeat)" and the team will fail over to a functional non-primary port.

As long as a port remains in the "Failed (TX Heartbeat)" state, it will attempt to transmit a heartbeat frame at every timer interval.

In summary, when using the default transmit path validation timer of 3 seconds, each port may transmit heartbeat frames as often as every 6 seconds and the longest it will take to detect transmit failures will be 9 seconds.

Receive path validation

At the receive path validation timer interval (default = 3 seconds), the team checks to determine if each port has successfully received any frame since the last timer interval. If not, then the port's internal status is degraded by incrementing the port's internal status by 1 (starting at 0). After a port's internal status has degraded to 2 (2 transitions) without receiving any kind of frame, all other ports in the same team will transmit a heartbeat frame with the intention that the port attempting to validate receives will receive at least one of the transmitted heartbeat frames. If the port does receive one of these heartbeats, its internal status is reset to 0. However, if the port still does not receive a frame, then the port is placed in a state called "Failed (RX Heartbeat)" and the team will fail over to a functional non-primary port.

As long as a port remains in the "Failed (RX Heartbeat)" state, all other ports attempt to transmit a heartbeat frame to the failed port at every timer interval.

In summary, when using the default receive path validation timer of 3 seconds, each port may transmit heartbeat frames as often as every 6 seconds and the longest it will take to detect receive failures will be 9 seconds.

Update Switch MAC table with team MAC address

There is no timer associated with this type of heartbeat since it only happens during team initialization or team failover.

Heartbeat frame format

Heartbeat frame size is 84 bytes including the FCS. Heartbeat frames are LLC test frames with 63 bytes of data. The destination address of all heartbeat frames is a Layer 2 multicast address of 03-00-C7-00-00-EE. This is a HP registered multicast address and used only for heartbeat frames. Below is an example of a heartbeat frame.

Table 4-1 Heartbeat frame format

802.2 frame format	Value	84 bytes total
Destination MAC address	03-00-C7-00-00-EE	6 bytes
Source MAC address	"MAC address of Teamed port"	6 bytes
Length	66	2 bytes
DSAP	0xAA	1 byte
SSAP	0xAA	1 byte
SNAP type	Unnumbered, TEST	1 byte
Data	"Teaming Proprietary data"	63 bytes
FCS	"Varies"	4 bytes

Heartbeat frames on a tagged VLAN include an extra 4-byte field used for VLAN identification. The frame size for VLAN tagged heartbeats is 88 bytes. If multiple VLANs are configured for the team, the SA may choose which VLAN interface to use for receive path validation heartbeats via the VLAN Configuration tab. Below is an example of what the heartbeat frame with a VLAN tag would look like.

Table 4-2 802.1Q tagged heartbeat frame format

802.2 frame format	Value	88 bytes total
Destination MAC address	03-00-C7-00-00-EE	6 bytes
Source MAC address	"MAC address of Teamed port"	6 bytes
802.1Q/p tag	"Value of VLAN ID and Priority"	4 bytes
Length	66	2 bytes
DSAP	0xAA	1 byte
SSAP	0xAA	1 byte
SNAP type	Unnumbered, TEST	1 byte
Data	"Teaming Proprietary data"	63 bytes
FCS	"Varies"	4 bytes

As discussed earlier, the basic redundancy mechanisms used by HP ProLiant Network Adapter Teaming are as follows:

- Link loss detection—"Do I have physical link?"
- Transmit path validation heartbeats—"Can I successfully transmit a frame onto the network?"
- Receive path validation heartbeats—"Can I successfully receive any kind of frame from the network?"

These basic mechanisms do not provide for redundancy beyond the network adapter port in the team. In other words, these basic redundancy mechanisms have no ability to detect network conditions that would prevent a particular team member port from being used by the team. If one teamed port is isolated from the rest of the team on another network, the isolated teamed port would most likely pass all three of the basic redundancy mechanisms tests: a) it would have link, b) it could successfully transmit heartbeat frame (it does not care if another team member receives it or not), and c) it could receive any kind of broadcast frame that may be transmitted on the isolated segment (from another network device). As a result, the basic redundancy mechanisms don't provide for intelligently choosing which teamed ports to use or not to use.

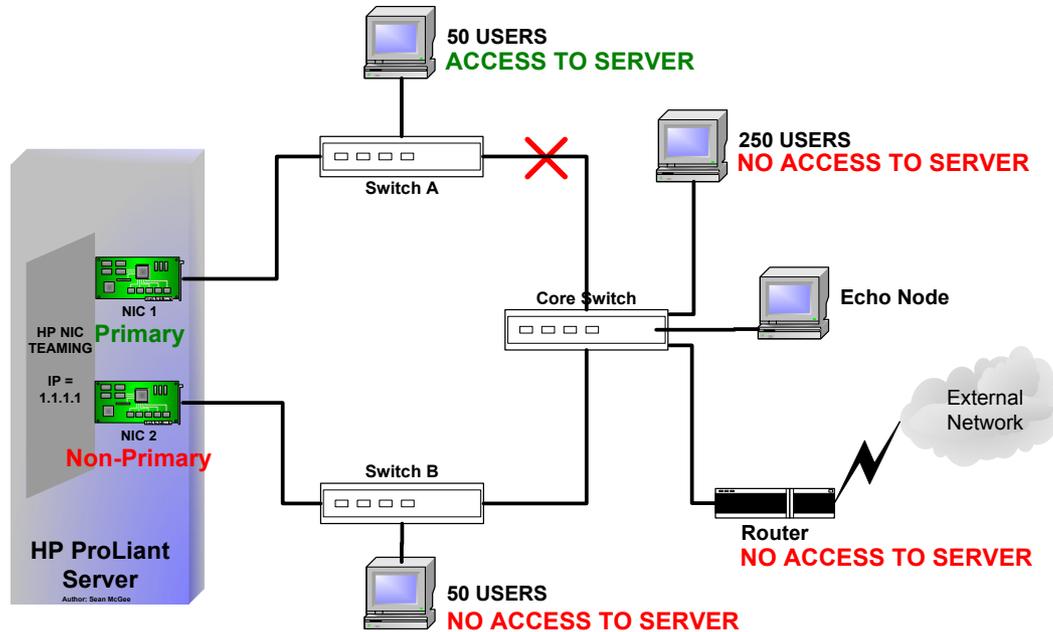
For example, see the diagram of an NFT team in Figure 4-4. The HP ProLiant server's primary team member port is connected to Switch A. Switch A's uplink to the core switch has been unplugged/disconnected. This results in the HP ProLiant server being effectively disconnected from the entire network except for Switch A. Only the 50 users on switch A can access the server, whereas the 50 users on Switch B, the 250 users connected via the core switch, and the router connection to any external network can no longer communicate with the server. Even though a non-primary port is connected to this isolated segment, non-Primary ports in an NFT team are not used for any transmit or receive communication. If this was a TLB team, instead of an NFT team, the server would still

only be able to communicate with the network represented by switch A since only primary ports can receive traffic from the network.

When this type of situation occurs on the network, it would be better for the server to proactively failover and choose a new primary from one of the team ports that still has connectivity with the core segment. In addition, the team should discontinue the use of any teamed ports that do not have connectivity with the core segment to prevent transmitting on a teamed port that cannot reach the intended target (for example, client, router, and so on).

The use of heartbeat frames between teamed ports will not provide enough information for the teaming driver to make the best failover decision. In other words, if the primary port cannot receive heartbeat frames from the non-primary and vice versa, the only thing the teaming driver knows is that the teamed ports are no longer attached to the same network segment. The teaming driver still does not know which network segment is the best segment to choose.

Figure 4-4 Upstream link failures cause server isolation



4-3-1-2-2 Redundancy mechanism comparison chart

	Link loss	TX path validation	RX path validation
Detects Layer 1 (physical layer) problems	✓	✓	✓
Detects Layer 2 (data link layer) problems			✓
Detects Layer 3 (network layer) problems			
Detects failure to receive a frame			✓
Detects failure to transmit a frame		✓	
Detects ports split across different LANs/VLANs			
Validates end-to-end Layer 2 connectivity per teamed port with a designated external device (echo node)			
Chooses teamed port with highest bandwidth path to core switch (STP root)			
Chooses teamed port with lowest hop count to core switch (STP root)			
Passive mechanism	✓		
Active mechanism		✓	✓
Basic redundancy (for free)	✓	✓	✓

	Link loss	TX path validation	RX path validation
Detects bad ports versus good ports	✓	✓	✓
Detects good ports versus best ports			
Mechanism can be manually disabled		✓	✓

4-3-1-2-3 Failover events

4-3-1-2-3-1 Link loss

When a network port is a member of a team and loses physical link (wire fault, lost link light), the teaming driver prevents the port from being used by the team. If this port is in use by the team, the team recovers from the failure based on the ports role in the team (primary or non-primary) and the team's mode (NFT, TLB, or SLB). If this was the last available team member in the team, then the team would fail.

4-3-1-2-3-2 Heartbeats (Receive and Transmit Validation)

The use of heartbeat frames for teamed port failovers was designed to detect network port communication failure even in cases when it maintained physical link. Special heartbeat frames are transmitted and received by teamed ports to validate the transmit and receive paths of each individual port. When a heartbeat frame is transmitted by one teamed port but not received by another teamed port, the teaming driver assumes that one of the ports is having a communications problem. If the primary port in an NFT or TLB team experiences a failover event, a failover will occur to a functional non-primary port.

Heartbeat frames were not designed to monitor networking failures such as loss of connectivity between switches, or loss of client connectivity.

4-3-1-3 NFT network adapter failure recovery

For Network Fault Tolerance Only (NFT) teams, there are two operating modes: Normal Mode and Manual Mode. Normal mode is the default mode that NFT operates in. If the user desires to manually force a failover, Manual Mode on the Team Controls tab on the Team Properties can be used.

- **Manual Mode**
This mode for NFT is used for user-initiated failovers (manual failovers). When used, Manual Mode allows an NFT team to automatically failover during events that normally cause a failover (for example, a cable is unplugged on the primary port of an NFT team). However, Manual Mode also allows the team to manually failover with the click of a button. Manual mode is normally used for troubleshooting purposes (for example, using an analyzer to take an inline network trace). Manual Mode is a temporary mode.
- **Normal Mode**
The second mode available for NFT is Normal Mode. In this mode, an NFT team initiates a failover from the primary port to an operational non-primary port whenever a failover event occurs (see "Failover events") on the primary port. When the failover occurs, the two ports swap MAC addresses so the team remains known to the network by the same MAC address. If the old primary port is restored, it becomes a non-primary port for the team but no MAC address changes are made unless there is another failover event on the primary port.

Note: Failures of non-primary ports do not trigger any type of recovery because non-primary ports are already in standby mode. The only consequence of a failed non-primary port is the possibility of the primary port failing and the team becoming unavailable to the network because both teamed ports are in a failed state. If there are three or more network ports in a team and two ports fail, the team is still available via the third port.

4-3-1-4 NFT applications

NFT is deployed in environments that only require fault tolerance and do not require transmit or receive throughput greater than the capacity of the primary port (for example, a server that requires fault tolerance in case of a network adapter malfunction, but does not have a demand for receiving or transmitting more than the capacity of the primary port.)

4-3-1-5 Recommended configurations for NFT environments

HP recommends the following:

- Path validations should be enabled (default).

- MAC addresses should not be manually set to a locally administered address (LAA) via the Microsoft Adapter Properties user interface. An administrator should not implement LAAs on individual ports that are members of a team; otherwise teaming may not function correctly. Setting an LAA for the team is permitted via the NCU.
- Spanning Tree's blocking, listening, and learning stages should be disabled, or bypassed, on all switch ports to which an HP ProLiant Network adapter team port is attached. These stages are not needed when a non-switch networking device (for example, server) is attached to the switch port. HP ProCurve switches have a feature called STP Fast Mode that is used to disable these Spanning Tree stages on a port-by-port basis. Cisco switches have an equivalent feature called PortFast.
- Team members can be split across more than one switch in order to achieve switch redundancy. However, all switch ports that are attached to members of the same team must comprise a single broadcast domain (in other words, same VLAN). Additionally, if problems exist after deploying a team across more than one switch, all team members should be reattached to the same switch. If the problems disappear, then the cause of the problem resides in the configuration of the switches and not in the configuration of the team. If switch redundancy is required (team members are attached to two different switches), then HP recommends that the switches be deployed with redundant links between them and Spanning Tree be enabled (or other Layer 2 redundancy mechanisms) on the ports that connect the switches. This helps prevent switch uplink failure scenarios that leave team members in separate broadcast domains.

4-3-2 Network Fault Tolerance Only with Preference Order

Network Fault Tolerance Only with Preference Order is exactly like Network Fault Tolerance Only except that it allows the SA to rank teamed ports with a user ranking. This user ranking (or preference) is used by the teaming driver as a criterion in deciding which teamed port should be the primary port. Teamed ports are ranked higher or lower than each on the Team Controls tab on the Team Properties window.

For all other aspects of this team type operation, see 4-3-1 "Network Fault Tolerance Only (NFT)" on page 22.

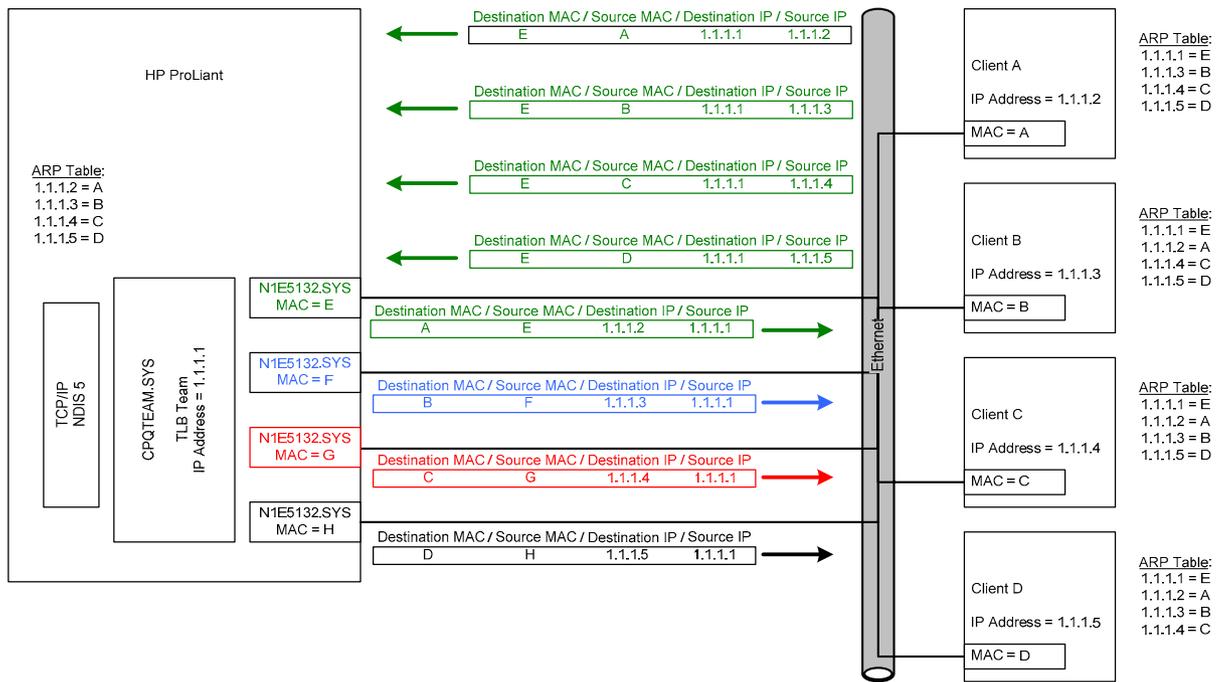
4-3-2-1 NFT with Preference Order applications

Network Fault Tolerance Only with Preference Order is mainly used in teaming configurations where one or more teamed ports are better than others. For example, an NFT team of a Gigabit network adapter and a Fast Ethernet adapter could utilize Preference order to rank the Gigabit adapter higher than the Fast Ethernet adapter. As long as the Gigabit adapter is in a good state, it will rank higher than the Fast Ethernet adapter and will be the team's primary port.

4-3-3 Transmit Load Balancing with Fault Tolerance (TLB)

Transmit Load Balancing mode, previously known as Adaptive Load Balancing (ALB), incorporates all the features of NFT, plus Transmit Load Balancing. In this mode, two to eight ports may be teamed together to function as a single virtual network port. The load-balancing algorithm used in TLB allows the server to load balance traffic transmitted from the server. However, traffic received by the server is not load balanced, meaning the primary teamed port is responsible for receiving all traffic destined for the server (see Figure 4-5). As with NFT, there are two types of team members, primary and non-primary ports. The primary port transmits and receives frames and the non-primary ports only transmit frames.

Figure 4-5 Overview of TLB communication



4-3-3-1 Network addressing and communication with TLB

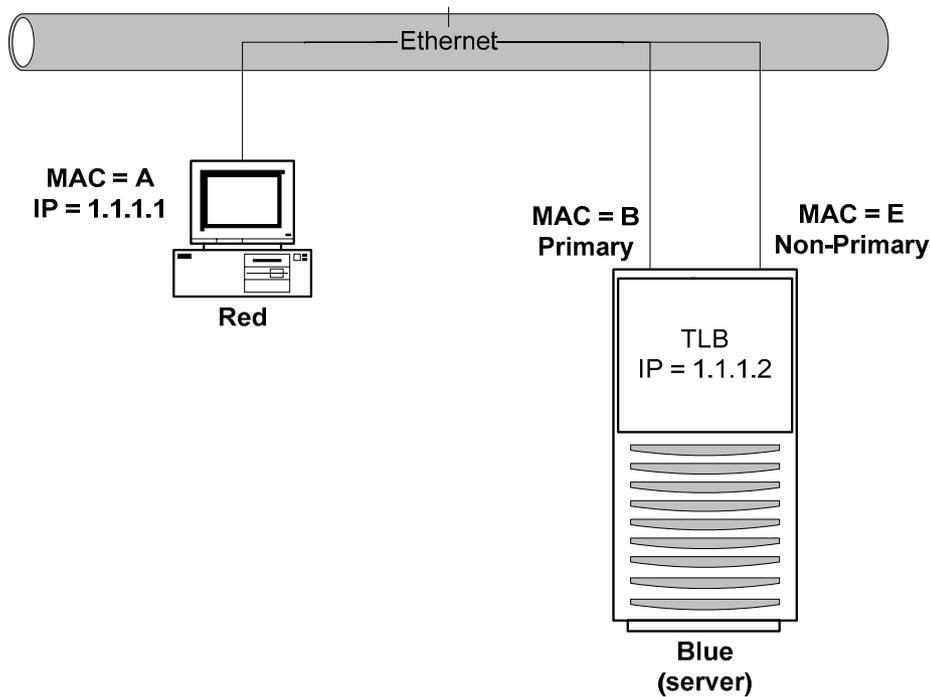
Before learning the specifics of TLB and how it communicates on the network, HP recommends that 4-2-3 “HP teaming and Layer 2 versus Layer 3 addresses” be thoroughly reviewed and understood, as well as “Appendix A: Overview of network addressing and communication.”

4-3-3-1-1 Scenario 1-C: a device PINGs a TLB team on the same Layer 2 network.

This section builds on the concepts reviewed in Scenario 1-A of “Example scenarios of network addressing and communication” in Appendix A, and describes how TLB functions from the network addressing and communication perspective.

Utilizing a network diagram similar to Figure A-1 on page 47, Blue has been modified to be a server utilizing an HP ProLiant Network adapter team in TLB mode with two network ports in a team (see Figure 4-6). The two network ports have Layer 2 addresses of MAC B and MAC E, respectively, and are known by a single Layer 3 address of 1.1.1.2. Network port B has been designated as the primary port in this NFT team.

Figure 4-6 Scenario 1-C: A device PINGS a TLB team on the same Layer 2 network



4-3-3-1-1-a) Red transmits a broadcast ARP request asking for Blue’s MAC address.

A user on Red issues the command `ping 1.1.1.2` to initiate a PING to Blue. First, Red determines whether or not Blue is on the same Layer 2 network.

After Red has determined that Blue is on the same Layer 2 network, Red must find out what Blue’s MAC address is. First, Red checks its own ARP cache for a MAC address entry matching the IP address of 1.1.1.2. If Red does not have a static entry or an entry cached from a previous conversation with Blue, then it must broadcast an ARP request frame on the network asking Blue to respond and provide its MAC address. Red must broadcast this ARP request because, without knowing Blue’s unique MAC address, it has no way of sending a frame directly (unicast) to Blue.

4-3-3-1-1-2b) Blue transmits a unicast ARP reply to Red, providing its MAC address.

Blue sees the ARP request (the frame is received on both the primary and non-primary teamed ports because the frame is broadcasted onto the network. However, the team discards all non-heartbeat frames incoming on non-primary ports), and responds with a unicast ARP reply to Red.

In Blue’s ARP reply, Blue provides the MAC address of its teaming driver, which is the same as the current primary port’s MAC address (B) (see “HP teaming and Layer 2 versus Layer 3 addresses”). Blue also takes note of Red’s MAC address (A) and IP address (1.1.1.1) and enters them into its ARP cache. Red receives the reply and enters the MAC address (B) and the IP address of Blue (1.1.1.2) into its own ARP cache.

4-3-3-1-1-3c) Red transmits a unicast PING request to Blue using Blue’s destination MAC address.

Red can now create a PING request frame using Blue’s MAC address (B). Red sends the PING request to Blue. Blue receives the frame on its primary port (B) and notices that a station with an IP address of 1.1.1.1 is asking for it to respond.

4-3-3-1-1-4d) Blue transmits a broadcast ARP request asking for Red’s MAC address.

Note: The following step may not occur if Blue’s ARP table still contains an entry for Red as a result of steps (a) and (b).

Blue checks its ARP cache for a MAC address entry that matches 1.1.1.1. If Blue does not find one, then Blue broadcasts an ARP request asking for Red’s MAC address.

4-3-3-1-1-5e) Red transmits a unicast ARP reply to Blue providing its MAC address.

Note: The following step will not occur if step (d) does not take place.

Red sees the ARP request and transmits a unicast ARP reply directly to Blue providing its MAC address (A). Blue receives the ARP reply and puts Red's MAC address (A) and IP address (1.1.1.1) in its ARP cache.

4-3-3-1-1-6f) Blue transmits a unicast PING reply to Red using Red's destination MAC address.

The final step in the conversation is for Blue to transmit a PING reply to Red. However, because Blue's team is running TLB, it must make a load-balancing decision before transmitting the PING reply. The load-balancing decision is made by using Red's MAC address or Red's IP address. After Blue decides which network port to use, it transmits a unicast PING reply to Red using Red's MAC address (A).

If Blue chooses to transmit from the primary port, Red will receive a PING Reply from Blue with a source MAC address of B, destination MAC address of A, a source IP address of 1.1.1.2 and a destination IP address of 1.1.1.1. However, if Blue chooses to transmit from the non-primary port, Red will receive a PING reply from Blue with a source MAC address of E, destination MAC address of A, a source IP address of "1.1.1.2" and a destination IP address of 1.1.1.1. Either way, Red only distinguishes that it received a PING reply from the Layer 3 address, 1.1.1.2 (see "Transmit Load Balancing Methods (algorithms)" for a complete discussion). The user sees the PING reply message printed on the screen. This completes the entire conversation.

4-3-3-2 Transmit Load Balancing Methods (algorithms)

All load-balancing team types (TLB, SLB, and 802.3ad Dynamic) load balance transmitted frames. There is a fundamental decision that must be made when determining load balancing mechanisms: whether or not to preserve frame order.

Frame order preservation is important for several reasons – to prevent frame retransmission because frames arrive out of order and to prevent performance-decreasing frame reordering within OS protocol stacks. In order to avoid frames from being transmitted out of order when communicating with a target network device, the team's load-balancing algorithm assigns "outbound conversations" to a particular teamed port. In other words, if frame order preservation is desired, outbound load balancing by the team should be performed on a conversation-by-conversation basis rather than on a frame-by-frame basis.

To accomplish this, the load-balancing device (either a team or a switch) needs information to identify conversations. Destination MAC address, Destination IP address, and TCP Connection are used to identify conversations.

It is very important to understand the differences between the load-balancing methods when deploying HP ProLiant Network Adapter Teaming in an environment that requires load balancing of routed Layer 3 traffic. Because the methods use conversations to load balance, the resulting traffic may not be distributed equally across all ports in the team. The benefits of maintaining frame order outweigh the lack of perfect traffic distribution across teamed ports' members.

Implementers of HP ProLiant Network Adapter Teaming can choose the appropriate load balancing method via the NCU.

Table 4-3 Transmit Load Balancing method comparison

	Preserves frame transmission order	Load balances through router	Guarantees equal load balancing	HP recommendation
Automatic	✓	✓		✓
TCP Connection	✓	✓		
Destination IP address	✓	✓		
Destination MAC address	✓			
Round Robin			✓	

4-3-3-2-1 TLB Automatic method

Automatic is a load-balancing method that is designed to preserve frame ordering.

This method will load balance outbound traffic based on the highest layer of information in the frame. For instance, if a frame has a TCP header with TCP port values, the frame will be load balancing by TCP connection (see "TLB TCP Connection method" below). If the frame has an IP header with an IP address but no TCP header, then the frame is load balanced by destination IP address (see "TLB Destination IP Address method" below). If the frame does not have an IP header, the frame is load balanced by destination MAC address (see "TLB Destination MAC Address method" below).

Automatic is the HP-recommended setting for outbound load balancing. Although in the current product Automatic mode is identical to TCP Connection mode, future releases may augment the Automatic mode. By deploying this method now, future upgrades will automatically take advantage of the new intelligence.

4-3-3-2-2 TLB TCP Connection method

TCP Connection is also a load-balancing method that is designed to preserve frame ordering.

This method will load balance outbound traffic based on the TCP port information in the frame's TCP header. This load-balancing method combines the TCP source and destination ports to identify the TCP conversation. Combining these values, the algorithm can identify individual TCP conversations (even multiple conversations between the team and one other network device). The algorithm used to choose which teamed port to use per TCP conversation is similar to the algorithms used in the "TLB Destination IP Address method" and "TLB Destination MAC Address method" sections below.

If this method is chosen, and the frame has an IP header with an IP address but not a TCP header, then the frame is load balanced by destination IP address (see "TLB Destination IP Address method" below). If the frame does not have an IP header, the frame is load balanced by destination MAC address (see "TLB Destination MAC Address method" below).

4-3-3-2-3 TLB Destination IP Address method

Destination IP Address is a load-balancing method that will attempt to preserve frame ordering.

This method makes load-balancing decisions based on the destination IP address of the frame being transmitted by the teaming driver. The frame's destination IP address belongs to the network device that will ultimately receive the frame. The team utilizes the last three bits of the destination IP address to assign the frame to a port for transmission.

Because IP addresses are in decimal format, it is necessary to convert them to binary format. For example, an IP address of 1.2.3.4 (dotted decimal) would be 0000 0001 . 0000 0010 . 0000 0011 . 0000 0100 in binary format. The teaming driver only uses the last three bits (100) of the least significant byte (0000 0100 = 4) of the IP address. Utilizing these three bits, the teaming driver consecutively assigns destination IP addresses to each functional network port in its team starting with 000 being assigned to network port 1, 001 being assigned to network port 2, and so on. Of course, how the IP addresses are assigned depends on the number of network ports in the TLB team and how many of those ports are in a functional state (see Table 4-4).

Table 4-4 Load Balancing based on Destination IP Address

Two-port team		Three-port team	
Destination IP	Transmitting port	Destination IP	Transmitting port
000	network port 1	000	network port 1
001	network port 2	001	network port 2
010	network port 1	010	network port 3
011	network port 2	011	network port 1
100	network port 1	100	network port 2
101	network port 2	101	network port 3
110	network port 1	110	network port 1
111	network port 2	111	network port 2
Four-port team		Five-port team	
Destination IP	Transmitting port	Destination IP	Transmitting port
000	network port 1	000	network port 1
001	network port 2	001	network port 2
010	network port 3	010	network port 3
011	network port 4	011	network port 4
100	network port 1	100	network port 5
101	network port 2	101	network port 1
110	network port 3	110	network port 2

Table 4-4 Load Balancing based on Destination IP Address

Two-port team		Three-port team	
111	network port 4	111	network port 3

If the Destination IP Address algorithm is chosen, and the frame does not have an IP header, the frame is load balanced by destination MAC address (see “TLB Destination MAC Address method” below).

4-3-3-2-3-1 Scenario 2-C: a TLB team using IP address-based load balancing

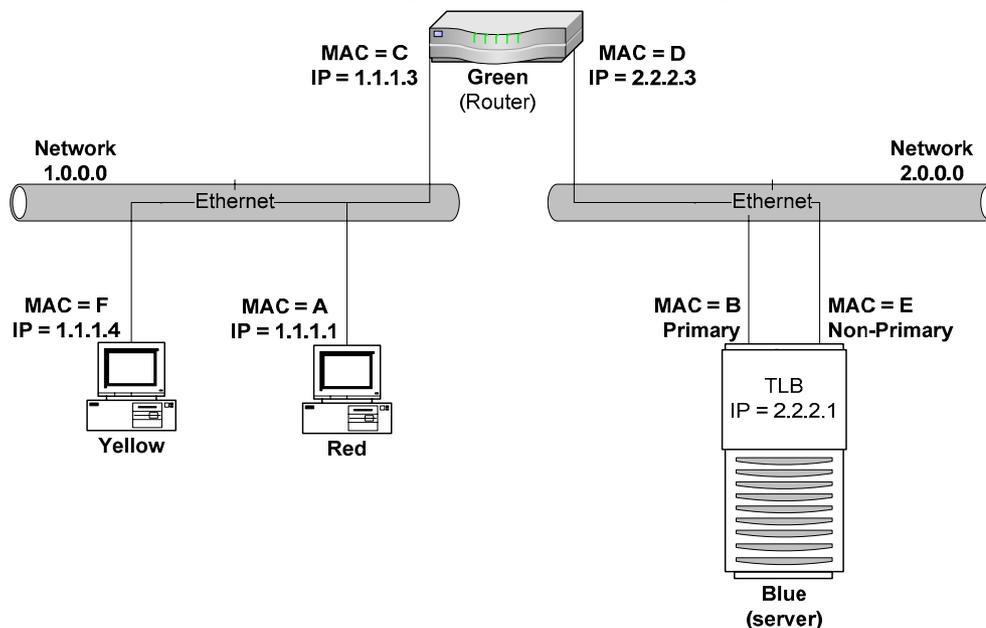
Using the concepts reviewed in Scenario 2-A of “Example scenarios of network addressing and communication” in Appendix A and Figure 4-7, this section describes how TLB IP address-based load-balancing functions.

Beginning at the point in Scenario 2-A where Blue/2.2.2.1 transmits the PING reply to Red/1.1.1.1, Blue must decide whether to use network port B or E. Blue’s teaming driver calculates using the IP address of Red (1.1.1.1) because Red is the frame’s destination. Because a dotted decimal 1.1.1.1 is equal to 0000 0001 . 0000 0001 . 0000 0001 . 0000 0001 in binary, and the last three bits (001) are used to determine the transmitting network port (see Table 4-4 – Two-port team), 001 is assigned to network port 2 (or the non-primary port). Therefore, when communicating with Red, Blue will always use the non-primary port to transmit frames.

If Blue transmits a frame to Yellow, the same calculation must be made. Yellow’s IP address in dotted decimal is 1.1.1.4 and equal to 0000 0001 . 0000 0001 . 0000 0001 . 0000 0100 in binary. Blue’s teaming driver will again use the last three bits to determine which network port will transmit the frame. Referring to Table 4-4 – Two-port team, 100 is assigned to network port 1 (or the primary port). Therefore, when communicating with Yellow, Blue will always use the primary port to transmit frames.

It is important to note that if an implementer uses the MAC address load balancing algorithm for the network in Figure 4-7, load balancing will not function as expected, and traffic will not be load balanced using all teamed ports. Because Blue transmits all frames destined for Red and Yellow via Green (Blue’s gateway), Blue uses Green’s Layer 2 address (MAC) as the frame’s Destination MAC Address but uses Red’s and Yellow’s Layer 3 addresses (IP) as the frame’s Destination IP Address. Blue never transmits frames directly to Red’s or Yellow’s MAC address because Blue is on a different Layer 2 network. Because Blue always transmits to Red and Yellow using Green’s MAC address, the teaming driver will assign all conversations with clients on Network 1.0.0.0 to the same network port. When an HP ProLiant Network adapter team needs to load balance traffic that traverses a Layer 3 device (router), at a minimum IP address-based load balancing should be used. Ideally, Automatic should be used.

Figure 4-7 Scenario 2-C: TLB team using IP address for load balancing algorithm



4-3-3-2-4 TLB Destination MAC Address method

Destination MAC Address is another load-balancing method that will attempt to preserve frame ordering.

This algorithm makes load-balancing decisions based on the destination MAC address of the frame being transmitted by the teaming driver. The destination MAC address of the frame is the MAC address that belongs to

the next network device that will receive the frame. This next network device could be the ultimate destination for the frame or it could be an intermediate router used to get to the ultimate destination. The teaming driver utilizes the last three bits of the destination MAC address and assigns the frame to a port for transmission.

Because MAC addresses are in hexadecimal format, it is necessary to convert them to binary format. For example (see Table 4-5), a MAC address of 01-02-03-04-05-06 (hexadecimal) would be 0000 0001 – 0000 0010 – 0000 0011 – 0000 0100 – 0000 0101 – 0000 0110 in binary format. The teaming driver load balances based upon the last three bits (110) of the least significant byte (0000 0110 = 06) of the MAC address. Utilizing these three bits, the teaming driver will consecutively assign destination MAC addresses to each functional network port in its team starting with 000 being assigned to network port 1, 001 being assigned to network port 2, and so on. Of course, how the MAC addresses are assigned depends on the number of network ports in the TLB team and how many of those ports are in a functional state.

Table 4-5 Load balancing based on Destination MAC Address

Two-port team		Three-port team	
Destination MAC	Transmitting port	Destination MAC	Transmitting port
000	network port 1	000	network port 1
001	network port 2	001	network port 2
010	network port 1	010	network port 3
011	network port 2	011	network port 1
100	network port 1	100	network port 2
101	network port 2	101	network port 3
110	network port 1	110	network port 1
111	network port 2	111	network port 2

Four-port team		Five-port team	
Destination MAC	Transmitting port	Destination MAC	Transmitting port
000	network port 1	000	network port 1
001	network port 2	001	network port 2
010	network port 3	010	network port 3
011	network port 4	011	network port 4
100	network port 1	100	network port 5
101	network port 2	101	network port 1
110	network port 3	110	network port 2
111	network port 4	111	network port 3

4-3-3-2-4-1

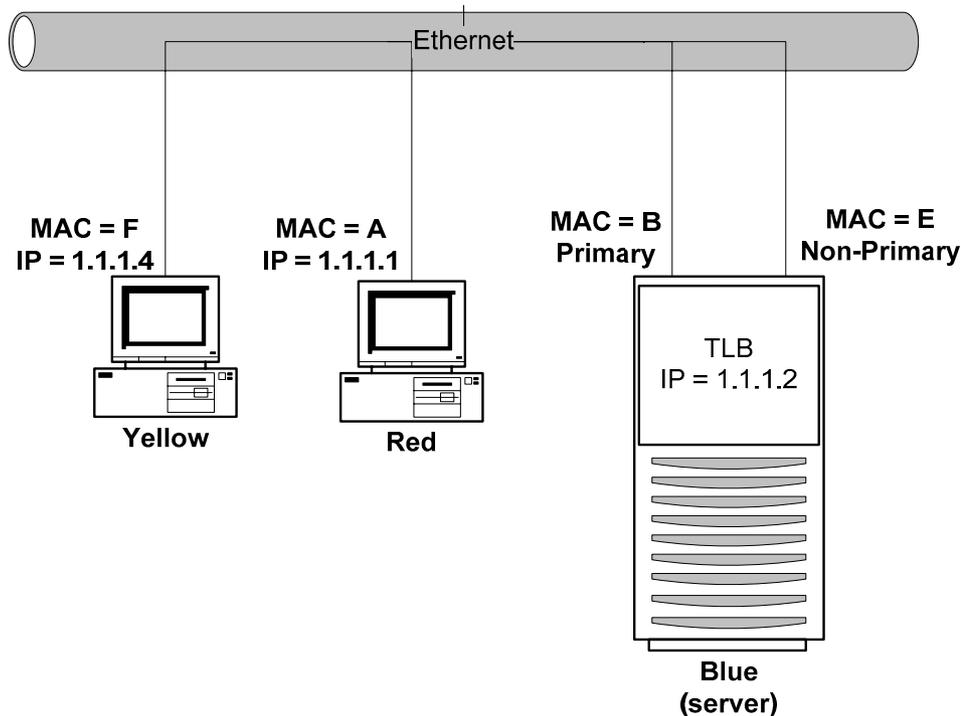
4-3-3-2-4-2 Scenario 1-D: A TLB team using Destination MAC Address load-balancing method

Taking the concepts reviewed in Scenario 1-A of “Example scenarios of network addressing and communication” in Appendix A, this section describes how TLB MAC addressed based load-balancing functions.

Beginning at the point in Scenario 1-A where Blue/1.1.1.2 transmits the PING reply to Red/1.1.1.1, Blue must decide whether to use network port B or E. Blue’s teaming driver calculates using the MAC address of Red (A) because Red is the frame’s destination. Because a hexadecimal A is equal to 1010 in binary, and the last three bits (010) are used to determine the transmitting network port (see Table 4-5 – Two-port team), 010 is assigned to network port 1 (or the primary port). Therefore, when communicating with Red, Blue will always use the primary port to transmit frames.

If Blue transmits a frame to Yellow, the same calculation must be made. Yellow’s MAC address is hexadecimal F, which is equal to 1111 in binary. Blue’s teaming driver will again use the last three bits to determine which network port will transmit the frame. Referring to Table 4-5, for a team with two network ports, 111 is assigned to network port 2 (or the non-primary port). Therefore, when communicating with Yellow, Blue will always use the non-primary port to transmit frames.

Figure 4-8 Scenario 1-D: TLB team using MAC address for load-balancing algorithm



In summary, special consideration should be given when choosing the MAC address-based load-balancing algorithm in an environment where the server and clients are separated by a Layer 3 device such as a router. In such an environment, the server must communicate with the clients via the router (set as the server's default gateway). When communicating with the clients, the server sends all traffic to the router, which then sends the traffic to the clients. If MAC address-based load balancing is selected, all traffic destined for clients is transmitted using the same teamed port in the load balancing team and is not load balanced. This occurs because the server must use the router's MAC address as the Layer 2 address in every frame while it uses the client's IP address as the Layer 3 address in the same frame. Because the router's MAC address is used in every frame, the MAC address-based load-balancing algorithm chooses the same teamed port for all transmitted traffic.

A better option is to choose the Automatic load-balancing algorithm described in "4-3-3-2-1 TLB Automatic method". The team's transmit load balancing will be based on either the IP address of the clients or the TCP ports for the connection rather than on the router's MAC address (which is the same for all connections through the router). In hybrid environments in which the server is communicating with clients on its own network (Layer 2), as well as clients on the other side of a router (Layer 3), HP recommends using Automatic method (see "4-3-3-2-1 TLB Automatic method"), TCP Connection method (see "4-3-3-2-2 TLB TCP Connection method"), or Destination IP Connection method (see "4-3-3-2-3 TLB Destination IP Address method").

4-3-3-2-5 TLB Round Robin method for outbound load balancing

Round Robin is a load-balancing method that will NOT preserve frame ordering.

This method is the simplest of all methods. It load balances every outbound frame out every operational teamed port on a frame-by-frame basis. Absolutely no frame ordering is maintained. All teamed ports are equally used.

HP recommends that the implications of this method of load balancing be carefully considered before deployment.

4-3-3-3 TLB redundancy mechanisms

The basic and advanced redundancy mechanisms used by TLB are identical to NFT. See "4-3-1-2 NFT redundancy mechanisms" for more information.

4-3-3-4 TLB network adapter failure recovery

With TLB, the recovery mechanism provided is very similar to the NFT failover methods discussed previously. In a two-port TLB team, the primary port receives all data frames, while the non-primary port receives and transmits only heartbeat frames. With TLB, all teamed ports are capable of transmitting data frames. In the event of a failover, the non-primary port becomes the primary port and assumes the MAC address of the team. In effect, the two ports swap MAC addresses. The new primary port now receives and transmits all data frames. If the old

primary port is restored, it becomes a non-primary port for the team. It will now only receive heartbeat frames and will be capable of transmitting data frames. If a non-primary teamed port fails in a two-port team, the data frames being transmit load balanced by the failed teamed port are transmitted by the primary port. If a non-primary teamed port is restored, it remains a non-primary teamed port, and the team will resume load balancing data frames on that port. No MAC address changes are made when a non-primary port fails or is restored.

4-3-3-5 TLB applications

TLB is deployed in environments that require fault tolerance and additional *transmit* throughput greater than the capacity of the primary port. However, TLB environments do not require *receive* throughput greater than the capacity of the primary port. For example, in a database server whose primary role is that of transmitting data to clients, receive throughput requirements may be much smaller than the transmit requirements because database requests require less bandwidth than transmitting database content.

4-3-3-6 Recommended configurations for TLB environments

HP recommends the following:

- Rx and TX Path Validation should be enabled (default)
- MAC addresses should not be manually set to a locally administered address (LAA) via the Microsoft Adapter Properties User Interface. A user should not implement LAAs on individual network ports that are members of a team; otherwise teaming may not function correctly. Setting an LAA for the team is permitted via the NCU.
- Spanning Tree's blocking, listening, and learning stages should be disabled, or bypassed, on all switch ports to which a teamed port is attached. These stages are not needed when a non-switch networking device (for example, server) is attached to the switch port. HP ProCurve switches have a feature called STP Fast Mode that is used to disable these Spanning Tree stages on a port-by-port basis. Cisco switches have an equivalent feature called PortFast.
- Team members can be split across more than one switch in order to achieve switch redundancy. However, all switch ports that are attached to members of the same team must comprise a single broadcast domain (in other words, same VLAN). Additionally, if problems exist after deploying a team across more than one switch, all team members should be reattached to the same switch. If the problems disappear, then the cause of the problem resides in the configuration of the switches and not in the configuration of the team. If switch redundancy is required (team members are attached to two different switches), then HP recommends that the switches be deployed with redundant links between them and Spanning Tree be enabled (or other Layer 2 redundancy mechanisms) on the ports that connect the switches. This helps prevent switch uplink failure scenarios that leave team members in separate broadcast domains.
- TLB teams that communicate with TCP/IP network devices via a router should use the Automatic, TCP port or, IP address-based load balancing algorithm (configured via the NCU).

4-3-4 Transmit Load Balancing with Fault Tolerance and Preference Order

Transmit Load Balancing with Fault Tolerance and Preference Order is exactly like Transmit Load Balancing with Fault Tolerance except that it allows the SA to rank teamed ports with a user ranking. This user ranking (or preference) is used by the teaming driver as a criterion in deciding which teamed port should be the primary port. Teamed ports are ranked higher or lower than each on the Team Controls tab on the Team Properties window.

For all other aspects of this team type operation, see "4-3-3 Transmit Load Balancing with Fault Tolerance" on page 29.

4-3-4-1 Transmit Load Balancing with Fault Tolerance and Preference Order applications

Transmit Load Balancing with Fault Tolerance and Preference Order is mainly used in teaming configurations where one or more teamed ports are better than others. For example, a TLB team of a Gigabit network adapter and a Fast Ethernet adapter could utilize Preference order to rank the Gigabit adapter higher than the Fast Ethernet adapter. As long as the Gigabit adapter is in a good state, it will rank higher than the Fast Ethernet adapter and will be the team's primary port.

4-3-5 Switch-assisted Load Balancing with Fault Tolerance (SLB)

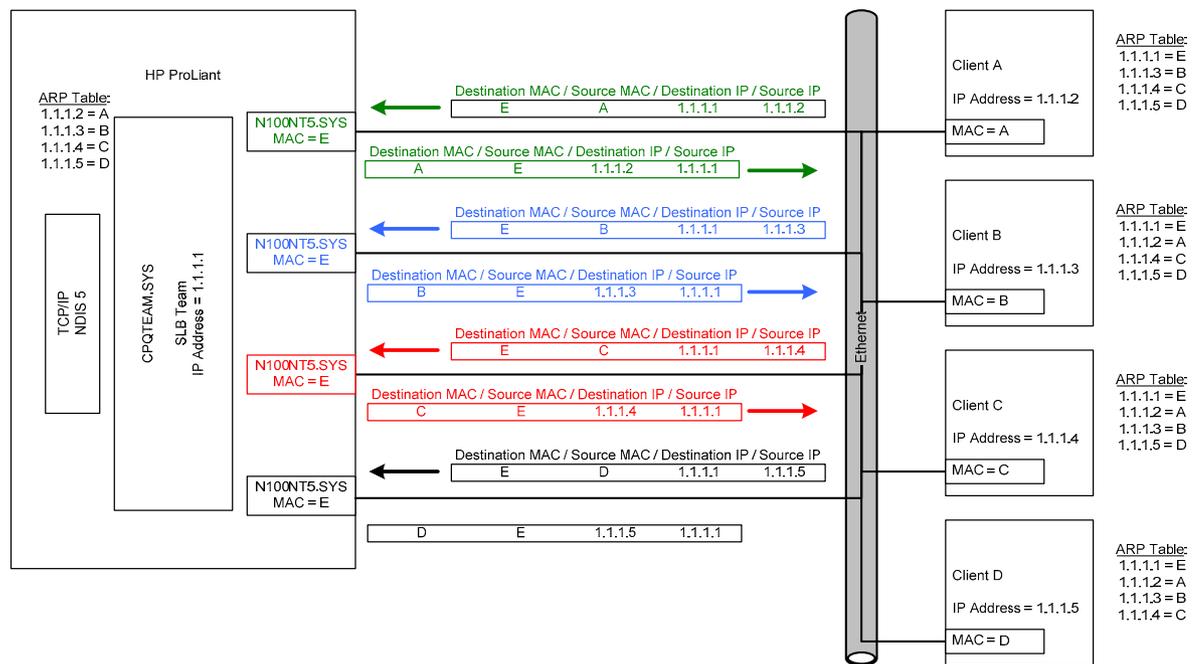
Switch-assisted Load Balancing (SLB) is an HP term that refers to an industry-standard technology for grouping multiple network ports into one virtual network port and multiple switch ports into one virtual switch port. HP's SLB technology works with multiple switch vendors' technologies. Compatible technologies include: HP ProCurve Port

Trunking, Cisco Fast EtherChannel (FEC)/Gigabit EtherChannel (GEC) (Static Mode Only – no PAGP), IEEE 802.3ad Link Aggregation (Static Mode only – no LACP), Nortel’s Multilink Trunking (MLT), and Extreme Network® Load Sharing. Therefore, SLB mode is sometimes referred to as Fast EtherChannel mode (FEC), Gigabit EtherChannel mode (GEC), 802.3ad Static mode, or Port Trunking. Note that SLB is not the same thing as Server Load Balancing (SLB) as used by some switch vendors (for example, Cisco). Switch-assisted Load Balancing operates independently of, and in conjunction with, Server Load Balancing.

SLB incorporates most of the features of NFT and TLB, and adds the feature of load balancing received traffic. The major feature that TLB, NFT, and Dual Channel have and the SLB lacks is switch fault tolerance since all teamed ports in SLB teams (and 802.3ad Dynamic teams) are required to be connected to the same switch.

In SLB mode, two to eight ports may be teamed together as a single virtual network port. The load-balancing algorithm used in SLB allows for the load balancing of the server’s transmit and receive traffic (see Figure 4-9).

Figure 4-9 Overview of SLB communication



Unlike NFT and TLB, SLB does not utilize the concepts of primary and non-primary ports within a team. All ports within a team are considered equal and perform identical functions as long as the particular port is in a functioning state. The algorithm for load-balancing transmit traffic used by SLB is identical to the algorithm used by TLB. Unlike TLB, SLB load balances all traffic regardless of the protocol being used.

4-3-5-1 Network addressing and communication with SLB

SLB functions identically to TLB (see 4-3-1-1 “Network addressing and communication with TLB” and the scenarios described in that section) except in its use of MAC addresses. SLB requires a switch capable of grouping multiple switch ports as a single switch port and SLB teaming uses the same source MAC address for all frames transmitted on all teamed ports in the same team. This does not violate IEEE standards because the switch is fully aware of the port groupings and expects that all teamed ports may transmit using the same source MAC address.

4-3-5-2 SLB outbound load-balancing algorithms

The algorithms for load balancing transmit traffic used by SLB is identical to the algorithms used by TLB [see 4-3-3-2 “Transmit Load Balancing Methods (algorithms)”].

4-3-5-3 SLB inbound load-balancing algorithm

The switch connected to the SLB team determines which load-balancing algorithm is used to load balance receive traffic for the team. An SLB team does not control which teamed port in the team receives the incoming traffic. Only the switch can choose which teamed port to use to send the traffic to the server. Therefore, consult the switch manufacturer to determine the algorithm the switch uses.

As an example, a detailed discussion of Cisco’s EtherChannel load-balancing algorithm is described below. Most of the information about how Cisco switches load-balance traffic on a port trunk is applicable to other switch vendors’ Port Trunking technology.

4-3-5-3-1 Cisco EtherChannel

Cisco's Fast EtherChannel (FEC) and Gigabit EtherChannel (GEC) technology is a MAC layer (Layer 2) load-balancing technology using two or more switch ports grouped together as one logical switch port. Depending on the specific load-balancing algorithm used, FEC/GEC may not efficiently load balance traffic to an SLB team.

FEC/GEC was originally designed as a switch-to-switch technology allowing two switches to increase the bandwidth between each other by aggregating multiple ports together as a single logical port for both transmits and receives. This is in contrast to TLB that only balances transmit traffic. An algorithm had to be used that could statistically divide the traffic over each port in the FEC/GEC group in an attempt to divide it evenly.

There are at least three algorithms that have been developed: source-based, destination-based, and XOR (see Table 4-6). The source-based algorithm utilizes the last one or two bits (depending on the number of ports in the FEC/GEC group) of the source address in the packet. If the bit is 0, the first port is used. If the bit is 1, the second port is used. The destination-based algorithm utilizes the last one or two bits (depending on the number of ports in the FEC/GEC group) of the destination address in the packet. If the bit is 0, the first port is used. If the bit is 1, the second port is used. The XOR algorithm utilizes the last one or two bits (depending on the number of ports in the FEC/GEC group) of the destination AND source addresses in the packet. The algorithm "XORs" the bits. If the result is 0, then the first port is used. If the result is 1, then the second port is used.

FEC/GEC has developed into not only a switch-to-switch technology but also a switch-to-node technology. In most cases, the node is a Multi-homed server with network port drivers that support FEC/GEC. Problems can arise with switches using the destination-based algorithm when switch-to-node FEC/GEC is used. Because the destination address of the FEC/GEC node is always the same, the switch always sends traffic to that server on the same port and receive traffic is not evenly distributed across all ports in the FEC/GEC group.

Table 4-6 Example of load-balancing algorithms

Preliminary information:		
MAC address of a two port FEC/GEC group Last byte (01) conversion to binary	MAC address of a two port FEC/GEC group Last byte (01) conversion to binary	MAC address of a two port FEC/GEC group Last byte (01) conversion to binary
MAC address of client1 Last byte (02) conversion to binary	MAC address of client1 Last byte (02) conversion to binary	MAC address of client1 Last byte (02) conversion to binary
MAC address of client2 Last byte (03) conversion to binary	MAC address of client2 Last byte (03) conversion to binary	MAC address of client2 Last byte (03) conversion to binary
Packet 1 is a frame transmitted from client 2 to the two-port FEC/GEC group. Packet 2 is a frame transmitted from client 1 to the two-port FEC/GEC group.		
Method 1: DESTINATION-BASED ALGORITHM (method used by HP teaming)		
Packet 1 - Destination MAC address:00-00-00-00-00-01 Last binary bit = 0000 000 1		<i>frame is transmitted out port 2</i>
Packet 2 - Destination MAC address:00-00-00-00-00-01 Last binary bit = 0000 000 1		<i>frame is transmitted out port 2</i>
Method 2: SOURCE-BASED ALGORITHM		
Packet 1 - Source MAC address:00-00-00-00-00-03 Last binary bit = 0000 001 1		<i>frame is transmitted out port 2</i>
Packet 2 - Source MAC address:00-00-00-00-00-02 Last binary bit = 0000 001 0		<i>frame is transmitted out port 1</i>
Method 3: XOR ALGORITHM (best method to use on switch)		
Packet 1 - Source MAC address:00-00-00-00-00-03 Last binary bit = 0000 001 1 Packet 1 - Destination MAC address:00-00-00-00-00-01 Last binary bit = 0000 000 1 XOR result of binary bits 1 & 1 = 0		<i>frame is transmitted out port 1</i>

Packet 2 - Source MAC address:00-00-00-00-00-02

frame is transmitted out port 2

Last binary bit = 0000 001**0**

Packet 2 - Destination MAC address:00-00-00-00-00-01

Last binary bit = 0000 000**1**

XOR result of binary bits **0** & **1** = **1**

The effects of the destination-based algorithm do not indicate a fault in the network port drivers nor on the switch. Destination-based load balancing is considered a functional FEC/GEC algorithm because packets between switches may not always use the same destination or source addresses. Only single-node-to-switch FEC/GEC uses the same destination address (for example, server team to switch).

The algorithm used for load balancing has no effect on fault tolerance; fault tolerance will function the same in any implementation.

Note: Some switches have the option to change the load-balancing algorithm. In such cases, HP advises using the algorithms in this order of preference: XOR (both), source-based, and destination-based.

4-3-5-4 Switch vendor Port Trunking technology supported by SLB

SLB is designed to support most switch vendors' implementation of Port Trunking technology (even if the technology is called by another name). The most important requirement is that SLB only supports Port Trunking technology that has been manually enabled. SLB does not support any kind of Port Trunking auto-configuration protocols (for example, LACP, PAgP). If automatic Port Trunking is required, 802.3ad Dynamic team type should be used with an IEEE 802.3ad Dynamic capable switch.

Examples of switch vendor Port Trunking technologies:

- Any switch that supports Static 802.3ad (no LACP)
- Extreme network's load sharing
- HP ProCurve Trunking (or Port Trunking)
- HP GbE & GbE2 Port Trunking
- Nortel MultiLink Trunking (MLT), but not Split MultiLink Trunking (SMLT)

4-3-5-5 SLB redundancy mechanisms

The only redundancy mechanisms used by SLB are link loss and transmit validation heartbeats. Any SLB teamed port that loses link or fails the transmit validation heartbeat test will be set to a failed state and will not be used by the team until the issue is resolved.

4-3-5-6 SLB network adapter failure recovery

With SLB, the recovery mechanism is somewhat different than those discussed previously. All members of the team transmit and receive frames with the same source MAC address in the Ethernet header and there is no concept of a Primary or Non-Primary port as there is in NFT and TLB. With SLB, there are no receive validation heartbeat frames, and consequently no receive validation heartbeat failures. In a two-port SLB team, all members are capable of receiving data frames (based on the switch's load-balancing algorithm), and transmitting data frames (based on the teaming driver's load balancing algorithm). In the event of a failover, all transmit traffic is redistributed among the working teamed ports. After a teamed port failure occurs in a two-port team, all transmit traffic is sent using a single teamed port since only one port is currently working. All receive traffic is determined by the switch, which should detect that only one teamed port is functional. If a failed teamed port is restored, all transmit and receive traffic is once again load balanced among all teamed ports.

A failure on any teamed port within the same SLB team has the same effect as a failure on any other teamed port because all teamed ports are considered "equal" (except for link loss or transmit validation heartbeat failures).

4-3-5-7 SLB applications

Switch-assisted Load Balancing is deployed in environments that require fault tolerance and additional transmit and receive throughput greater than the capacity of the Primary port. SLB environments also require a switch manually configured to provide Port Trunking.

4-3-5-8 Recommended configurations for SLB environments

HP recommends the following:

- Transmit validation heartbeats should be enabled (default).
- MAC addresses should not be manually set to a locally administered address (LAA) via the Microsoft Adapter Properties User Interface. A user should not implement LAAs on individual network ports that are

members of a team; otherwise teaming may not function correctly. Setting an LAA for the team is permitted via the NCU.

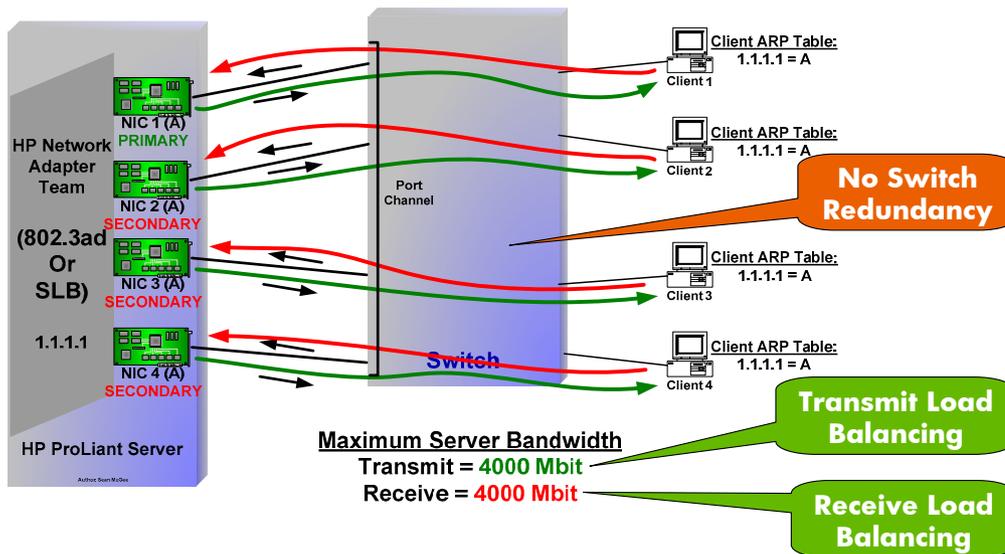
- Spanning Tree’s blocking, listening, and learning stages should be disabled or bypassed on all switch ports to which an HP ProLiant Network adapter team port is attached. These stages are not needed when a non-switch networking device (for example, server) is attached to the switch port. HP ProCurve switches have a feature called STP Fast Mode that is used to disable these Spanning Tree stages on a port-by-port basis. Cisco switches have an equivalent feature called PortFast.
- SLB teams that communicate with TCP/IP network devices via a router should use the Automatic, TCP port or the IP address-based load-balancing algorithm (configured via the NCU).
- Implementers thoroughly understand the configuration guidelines set by the switch vendor because SLB is dependent on the switch being configured in a compatible mode. HP’s SLB technology has been designed to allow for flexibility. Therefore, the NCU may allow configuration of an SLB team that will not work correctly with a particular vendor’s switch.
- The switch’s load-balancing algorithm should be set to XOR or SOURCE-BASED but not DESTINATION-BASED [see 3-2-5 “Switch-assisted Load Balancing with Fault Tolerance” and 4-3-5-3-1 “Cisco EtherChannel”]

The switch’s load-balancing algorithm should be set to balance by IP address if most traffic destined for the server originates on a different network and must traverse a router.

4-3-6 802.3ad Dynamic with Fault Tolerance

802.3ad Dynamic is exactly like SLB in every way except for one – switch configuration of Port Trunking. With SLB, the switch ports connected to the team must have Port Trunking manually enabled in order for SLB to work correctly. With 802.3ad Dynamic, the Link Aggregation Control Protocol (LACP) is used by the team to communicate with the switch and automatically configure Port Trunking. This eliminates the need for manual configuration on the switch.

All other aspects of 802.3ad Dynamic are identical to SLB. See the SLB sections for more details.



4-4 Team types and redundancy mechanisms interoperation

4-4-1 Team type and redundancy mechanisms compatibility chart

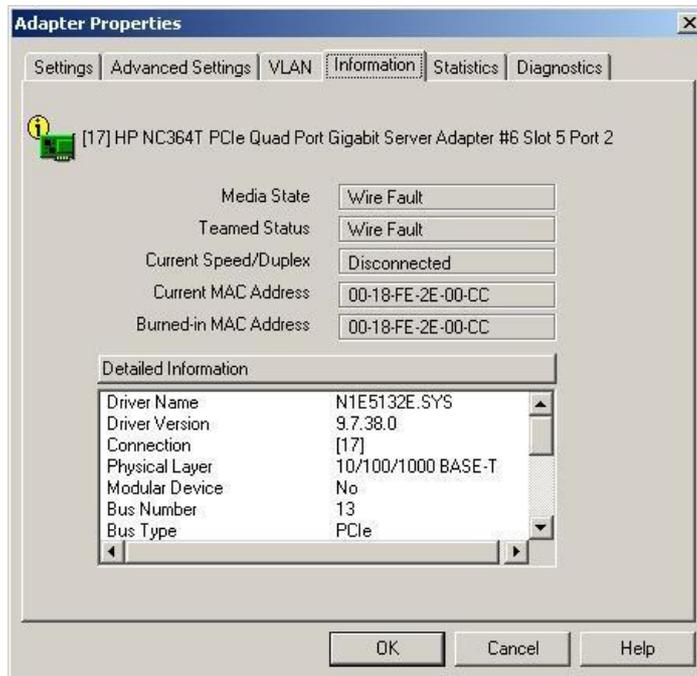
	NFT	TLB	SLB	802.3ad Dynamic
Transmit path validation	✓	✓	✓	✓
Receive path validation	✓	✓		

4-5 Team status and icons

4-5-1 Network adapter teamed status

The NCU reports a “Teamed Status” on the Information tab of each teamed port in the server. This status allows the user to determine the condition of the teamed port’s teamed status and take corrective action if necessary.

Figure 4-10 Teamed port Information tab



Listed below are the possible status conditions with definitions:

- **Available**—The team member is functioning normally.
- **Not Teamed**—The adapter is not part of the team. This is most likely caused by adding an adapter to the team but not applying the change.
- **Unknown**—The team member's status could not be determined.
- **Wire Fault**—The member does not have link.
- **Not Joined**—The member cannot be joined in the team because it has an incompatible setting. This is most likely caused by changing a parameter for a team member using the Local Area Connection property sheet.
- **Degraded (Rx Path)**—The team member is not receiving frames and no other team member is receiving frames. Because all team members are equal, the team member is still in use by the team.
- **Degraded (Multiple)**—The team member has multiple degraded conditions.
- **Failed (LACP)**—The team member failed to establish a LACP channel.
- **Failed (LACP Standby)**—The team member has failed because the team has more members than the switch supports in the LACP protocol. The port is blocked by the switch.
- **Failed (Rx Path)**—The team member is not receiving frames.
- **Failed (Tx Path)**—A failure occurred while attempting to send a frame on the team member.
- **Failed (Multiple)**—The team member has multiple failed conditions.

4-5-2 Team state

The NCU reports a Team State on the Information tab of each configured team. This status allows the user to determine the overall condition of the team and take corrective action if necessary.

Listed below are the possible team state conditions with definitions:

- **Ok**—The team is functioning properly (green team icon).
- **Disabled**—The team has been disabled.

- **Redundancy Reduced**—The team has one or more failed members but at least two members are still functional. Redundancy has been reduced since not all members are functional but redundancy is still available since more than one member is functional (yellow team icon).
- **Redundancy Lost**—The team has one or more failed members and only a single member is functional. Redundancy has been lost since only one member is functional (yellow team icon).
- **Failed**—The team has failed and connectivity has been lost. (red team icon).
- **Unknown**—The state of the team cannot be determined.

4-5-3 Team icons

Team icons are shown on the main window to allow the user to easily recognize status and conditions of teams. See Figure 4-11 for team state icons. See the HELP file for HP ProLiant Network Adapter Teaming for additional information.

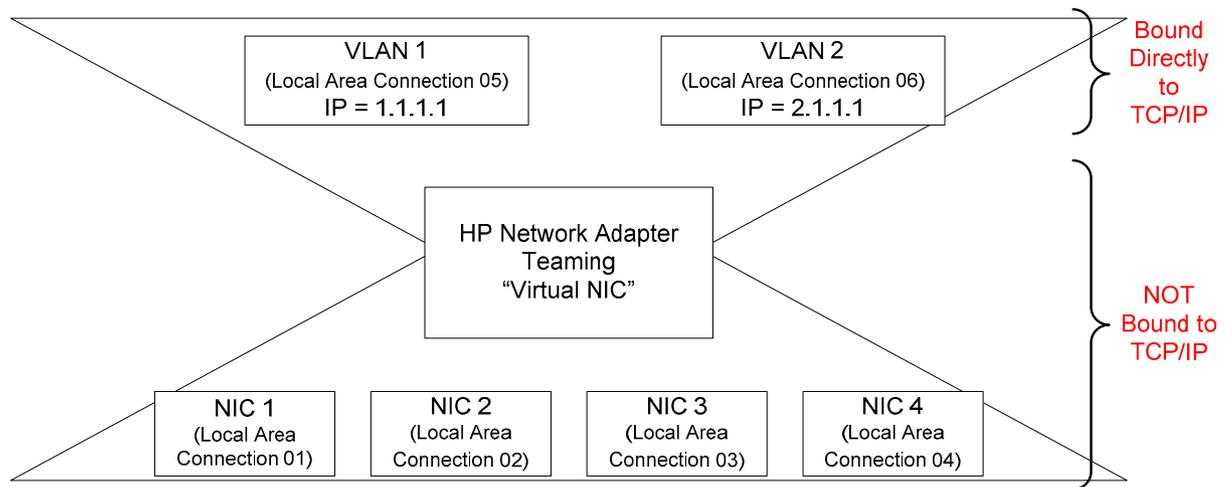
Figure 4-11 Team state icons

Icons	Description
	Good Team. The team is functioning properly.
	Team Not Formed. The team has been created, or the membership has changed, however the application must be closed and re-invoked for the changes to take effect.
	Degraded Team. The team is functioning, but one or more members is not available to fulfill its role.
	Failed Team. The team has failed and connectivity has been lost.
	Disabled Team. The team is not functioning.
	VLAN. One or more VLANs have been defined for the adapter or team of adapters..

4-6 HP ProLiant Network Adapter Teaming and advanced networking features

When adapter ports are members of a team, some advanced features are no longer configurable on the port's Advanced Settings tab. These advanced features are either promoted to the team's Advanced Settings tab because the features are compatible and supported by all team members, or not promoted because one or more ports do not support the features. There are also special cases where a feature is supported by all team members but the implementation is not compatible between the team members because they use different miniport drivers. In this case, it is not promoted to the team's Advanced Settings tab.

Figure 4-12 VLAN tagging used with HP ProLiant Network Adapter Teaming



4-7 HP ProLiant Network Adapter Teaming configuration and deployment with ProLiant Essentials Rapid Deployment Pack (RDP)

One of the methods of deployment for HP ProLiant Network Adapter Teaming is to use HP ProLiant Essential's Rapid Deployment Pack (RDP) remote scripting feature. This feature allows an SA to deploy a specific teaming configuration to any number of target servers in a single action. RDP is a server deployment solution that facilitates the installation, configuration, and deployment of high-volumes of servers through either a GUI-based or a web-based console, using either scripting or imaging technology. For additional information on RDP, see <http://h18004.www1.hp.com/products/servers/management/rdp.html>.

Example steps for HP ProLiant Network Adapter Teaming deployment using RDP:

- 1) Create the desired team on one server manually. When satisfied with the teaming configuration, click the **Save** in the NCU tool and save the file as hpteam.xml to a location on the RDP server.
- 2) Create a new job in the RDP console called "NIC Teaming Job."
- 3) Edit the job properties and include a copy file task that copies the latest NCU (for example, cp123456.exe) to the target server's %windir%\temp folder (if you do not need to upgrade the NCU tool, you may omit this step).
- 4) Add an additional copy file task that copies the Hpteam.xml file (from step 1) to the same location on the server as the NCU tool copied in step 3.
- 5) If needed, add a Run Script task that executes the NCU installer (for example, %windir%\temp\cp123456.exe /s).
- 6) Add an additional Run Script task that executes the teaming CLI tool and uses the Hpteam.xml configuration file (for example, c:\windows\system32\cqnccmd.exe /c %windir%\temp\hpteam.xml).

For more information on the cqnccmd.exe tool, see nicscript.pdf located in c:\compaq\network on any server with Teaming installed.

- 7) Test the new job on a target server with no teaming installed to ensure the desired result.

Note: If any of the target servers already have a team in place and your intention is to overwrite that team, consider the cqnccmd.exe /d parameter to delete existing teams. Note that in doing so, your target may lose the connection to the RDP server and may be unable to complete the job. This can be reconciled by putting all of the above scripts into one batch file that is sent to the server via the rdp console and then told to execute locally. Lastly, if your network does not use DHCP on the subnet where the team is created, you will lose access to the target server after the new team is created. You can overcome this by using netsh.exe included with Windows operating systems.

5 Teaming feature matrix

	NFT & NFT with Preference	TLB & TLB with Preference	SLB & 802.3ad Dynamic
Number of ports supported per team	2-8	2-8	2-8
Maximum theoretical transmit/receive throughput (in Mbps) with maximum number of 100 Mbps ports	100/100	800/100	800/800
Maximum theoretical transmit/receive throughput (in Mbps) with maximum number of 1000 Mbps ports	1000/1000	8000/1000	8000/8000
Supports network port fault tolerance	✓	✓	✓
Supports Transmit Load Balancing		✓	✓
Supports Receive Load Balancing			✓
Requires a switch that supports a compatible form of load balancing. (in other words, requires configuration on the switch)			✓
Teamed ports can be connected to more than one switch for switch fault tolerance (all ports must be in the same broadcast domain)	✓	✓	
Requires switch to support 802.3ad LACP			✓ (802.3ad Dynamic only)
Can utilize heartbeats for network integrity checks	✓	✓	
Can team ports that do not support a common speed	✓		
Can team ports operating at different speeds as long as the ports support a common speed	✓	✓	✓
Can team ports of different media	✓	✓	✓
Load balances TCP/IP		✓	✓
Load balances non-IP traffic		✓	✓
Supports load balancing by destination IP address, destination MAC address, TCP port, and Round Robin		✓	✓
All teamed ports within a team transmit frames with the same MAC address			✓
All ports within a team utilize the same IP addresses on the network	✓	✓	✓

Appendix A: Overview of network addressing and communication

Understanding the concepts of network addressing is the key to understanding how HP's Network Adapter Teaming works. This section provides a brief overview of network addressing as a baseline for explaining how HP's Network Adapter Teaming can create one logical network port from a team of two or more ports.

Layer 2 versus Layer 3

Devices on a computer network use unique addresses, much like telephone numbers, to communicate with each other. Each device, depending on its function, will use one or more of these unique addresses. The addresses correspond to one or more layers of the OSI model. Most often, network devices use an address at Layer 2 (Data Link Layer) called a MAC address, and an address at Layer 3 (Network Layer) called a protocol address (for example, IP, IPX, AppleTalk). One could say that a MAC address is one that is assigned to the hardware, whereas a protocol address is one that is assigned to the software.

MAC addresses are in the format of 00-00-00-00-00-00 (hexadecimal), IP addresses are in the format of 0.0.0.0 (dotted decimal), and IPX addresses are in the format of 000000.000000000000 (hexadecimal). Because multiple protocols can reside on the same network device, it is not uncommon for a single network device to use one MAC address and one or more protocol addresses.

Ethernet devices communicate directly using the MAC address, not the protocol address. For instance, when a PING is initiated for the address 1.1.1.1, the network device must find a corresponding MAC address for the IP address of 1.1.1.1. A frame is then built using the MAC address that corresponds to 1.1.1.1 and sent to the destination computer. The frame carries the sender's protocol address in its payload, which is how the destination network device knows to which device to respond. This means that protocol addresses must be resolved to MAC addresses. For IP, this is done using ARP (see "Example scenarios of network addressing and communication" in Appendix A). For IPX, the MAC address is part of the IPX address, so no special mechanism is needed.

Address – unicast versus broadcast versus multicast

There are three types of Layer 2 and Layer 3 addresses: unicast, broadcast, and multicast. A unicast address is one that corresponds to a single network device, either a single MAC address, or a single IP address. A broadcast address is one that corresponds to all network devices. A multicast address is one that corresponds to multiple network devices, but not necessarily all network devices. When a station transmits a frame to a unicast address, the transmitting device intends for only a single network device to receive the frame. When a station transmits a frame to a broadcast MAC address or IP address, the station intends for all devices on a particular network to receive the frame. When a station transmits a frame to a multicast MAC or IP address, the station intends for a predefined group of network devices to receive the frame. A group, as used here, can be defined as more than one network device, but less than all the network devices on a particular network.

A multicast MAC address is used in HP ProLiant Network Adapter Teaming for the purpose of transmitting and receiving heartbeat frames (see "Heartbeats"). Because the heartbeat frames are Layer 2 only frames (only use MAC addresses), HP ProLiant Network adapter teams do not need a protocol address assigned to them (for example, IP address) for heartbeat frames to function.

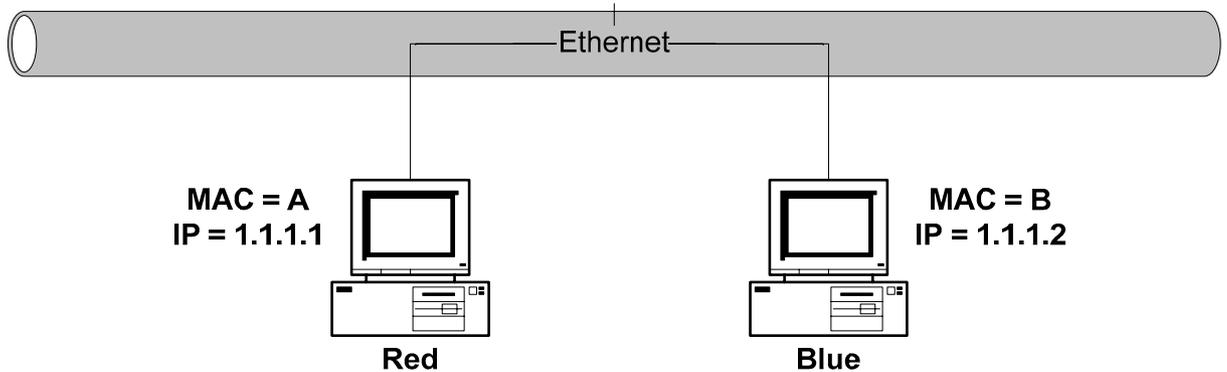
Example scenarios of network addressing and communication

As discussed earlier, protocol addresses (for example, IP, IPX) must be resolved to hardware addresses (MAC) for network devices to communicate. What follows are two simple scenarios with one network device (named Red) PINGing another network device (named Blue). The first scenario cites one device PINGing another on the same Layer 2 network. The second scenario cites one device PINGing another on a different Layer 2 network, which requires the use of a router to effect communication.

These scenarios provide a baseline of typical network addressing and communication using IP. This baseline will be referred to later in this document to differentiate how HP ProLiant Network Adapter Teaming functions in these same scenarios. By understanding the differences in simple examples such as these (without HP's Network Adapter Teaming technology involved), implementers will have a better understanding of how HP's Network Adapter Teaming technology may work in their environment.

5-1-1 Scenario 1-A: one device PINGs another on the same Layer 2 network

Figure A-1 Scenario 1-A: One device PINGs another on the same Layer 2 network



5-1-1-1 1) Red transmits a broadcast ARP request asking for Blue's MAC address.

A user on Red issues the command `ping 1.1.1.2` to initiate a PING to Blue. The number 1.1.1.2 refers to Blue's IP address, or protocol address. First, Red determines whether or not Blue is on the same Layer 2 network by running an algorithm (details of this algorithm are beyond the scope of this document) using its own IP address of 1.1.1.1, its own subnet mask (not shown), and Blue's IP address of 1.1.1.2. If Blue is on a different Layer 2 network, then Red will need to use its gateway, or router, to get to Blue.

After Red has determined that Blue is on the same Layer 2 network, Red must find out what Blue's MAC address is. First, Red checks its own ARP cache for a MAC address entry matching the IP address of 1.1.1.2. ARP is used to map protocol addresses to hardware addresses. If Red does not have a static entry or an entry cached from a previous conversation with Blue, then it must broadcast an ARP Request frame containing the IP address of Blue on the network asking Blue to respond and provide its MAC address. Red must broadcast this ARP request because without knowing Blue's unique MAC address, it has no way of sending a frame directly (unicast) to Blue.

5-1-1-2 2) Blue transmits a unicast ARP reply to Red, providing its MAC address (B).

Blue sees the ARP request containing its own IP address and responds with a unicast ARP reply directly to Red. Blue also notes Red's MAC address (A) and IP address of 1.1.1.1, and enters them into its ARP cache. Red receives the ARP reply and enters Blue's MAC address (B) and IP address (1.1.1.2) into its own ARP cache.

5-1-1-3 3) Red transmits a unicast PING request to Blue using Blue's MAC address (B).

Red can now create a PING request frame using Blue's MAC address (B). Red sends the PING request to Blue using Blue's MAC address (B). Blue receives the PING request frame and notices that a station with an IP address of 1.1.1.1 is requesting that it respond.

5-1-1-4 4) Blue transmits a broadcast ARP request asking for Red's MAC address.

Note: This step may not occur if Blue's ARP table still contains an entry for Red as a result of steps 1 and 2.

Blue checks its ARP cache for a MAC address entry that corresponds to 1.1.1.1. If Blue does not find one (in other words, ARP cache timed out since last communication with Red), then Blue broadcasts an ARP request asking for Red's MAC address.

5-1-1-5 5) Red transmits a unicast ARP reply to Blue providing its MAC address (A).

Note: This step will not occur if step 4 does not take place.

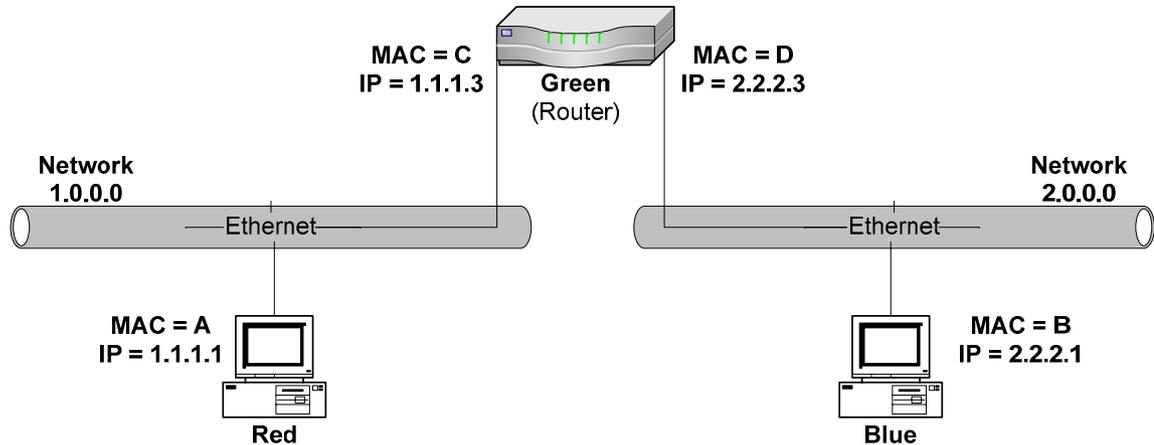
Red sees the ARP request and transmits a unicast ARP reply directly to Blue providing its MAC address (A). Blue receives the ARP reply and puts Red's MAC address (A) and IP address (1.1.1.1) in its ARP cache.

5-1-1-6 6) Blue transmits a unicast PING reply to Red using Red's destination MAC address (A).

Blue transmits a unicast PING reply to Red using Red's MAC address (A) and the user sees the PING REPLY message printed on the screen. This completes the entire conversation.

5-1-2 Scenario 2-A: one device PINGs another on a different Layer 2 network

Figure A-2 Scenario 2-A: one device PINGs another on a different Layer 2 network



5-1-2-1 1) Red transmits a broadcast ARP request on Network 1.0.0.0 asking for Green's MAC address.

A user on Red issues the command `ping 2.2.2.1` to initiate a PING to Blue. The number 2.2.2.1 refers to Blue's IP address, or protocol address. First, Red determines whether or not Blue is on the same Layer 2 network by running an algorithm (details of this algorithm are beyond the scope of this document) using its own IP address of 1.1.1.1, its own subnet mask (not shown), and Blue's IP address of 2.2.2.1. If Blue is on a different Layer 2 network, then Red will need to use its gateway or router (Green) to get to Blue.

Once Red has determined that Blue is on a different Layer 2 network, Red must use Green as a gateway to get to Blue. Red communicates directly with Green at Layer 2 but communicates directly with Blue at Layer 3. This means that Red must transmit a frame with the Layer 2 address (MAC) of Green, but the same frame will have Blue's Layer 3 address (IP) in it. When Green receives the frame, it sees the Layer 3 data destined for Blue and forwards the frame onto Blue via Green's interface that is attached to Blue's Layer 2 network (Network 2.0.0.0). This means that Red must find out what Green's MAC address is. First, Red checks its own ARP cache for an entry that matches 1.1.1.3. If Red does not have an entry cached, then it must broadcast an ARP request frame on network 1.0.0.0 asking Green to respond and provide its MAC address.

5-1-2-2 2) Green transmits a unicast ARP reply to Red providing its MAC address (C).

Green sees the ARP request and responds with a unicast ARP reply to Red. Also, Green enters Red's MAC address and IP address into its ARP cache. Red receives Green's ARP reply and enters Green's MAC address (C) and IP address (1.1.1.3) into its ARP cache.

5-1-2-3 3) Red transmits a PING request to Blue (2.2.2.1) using the destination MAC address (C) of Green's 1.1.1.3 interface, because Green is Red's gateway to Blue.

Red can now create a PING request frame using Green's MAC address and Blue's IP address. Red sends the PING request. Green receives the PING request and determines that the frame is meant for Blue because of the Layer 3 address (IP).

5-1-2-4 4) Green transmits a broadcast ARP request on Network 2.0.0.0 asking for Blue's MAC address.

Green looks in its ARP cache for a MAC address for Blue. If one is not found, Green broadcasts an ARP request frame on Blue's Layer 2 network asking for Blue's MAC address.

5-1-2-5 5) Blue transmits a unicast ARP reply to Green providing its MAC address (B).

Blue sees the ARP request frame and responds with a unicast ARP reply frame to Green. Also, Blue enters Green's MAC address and IP address into its ARP cache. Green receives the ARP reply from Blue and enters Blue's MAC address (B) and IP address (2.2.2.1) into its ARP cache.

5-1-2-6 6) Green forwards Red's PING request to Blue using Blue's destination MAC address (B).

Green now transmits Red's original PING request frame onto Blue's network using Blue's MAC address and Blue's IP address as the destination MAC and destination IP address. The source MAC address is Green's MAC address (D) and the source IP address is Red's IP address (1.1.1.1). Blue receives the frame and notices that a station with an IP address of 1.1.1.1 is asking for it to respond to a PING. Before Blue can respond with a PING reply, it must determine whether or not 1.1.1.1 is on the same layer 2 network. Blue runs an algorithm (details of this algorithm are beyond the scope of this document) using its own IP address (2.2.2.1), its own subnet mask (not shown) and the IP address of Red (1.1.1.1). Blue then determines that Red is on a different network. Because of this, Blue must use its gateway (Green) to get the PING reply back to Red.

5-1-2-7 7) Blue transmits a broadcast ARP request on Network 2.0.0.0 asking for Green's MAC address.

Note: This step may not occur if Blue's ARP table still contains an entry for Green resulting from steps 4 and 5.

Blue checks its ARP cache for the MAC address that corresponds to the IP address of 2.2.2.3 (Blue's gateway). If an entry is not found, Blue must broadcast an ARP request asking for Green's MAC address.

5-1-2-8 8) Green transmits a broadcast ARP reply to Blue providing its MAC address (D).

Note: This step will not occur if step 7 does not take place.

Green sees the ARP request and responds with a unicast ARP reply directly to Blue. Also, Green enters Blue's MAC address and IP address into its ARP cache. Blue receives the ARP reply and puts Green's MAC address (D) and IP address (2.2.2.3) in its ARP cache. Blue now has all the information it needs to send a PING reply to Red.

5-1-2-9 9) Blue transmits a unicast PING reply to Red (1.1.1.1) using the MAC address of Green's 2.2.2.3 interface (D).

Blue transmits a unicast PING reply to Red through Green by using Green's MAC address as the destination MAC address, Red's IP address as the destination IP address, Blue's MAC address as the source MAC address and Blue's IP address as the source IP address. Green receives the PING reply and determines that the frame is meant for Red because of the Layer 3 address (IP).

5-1-2-10 10) Green transmits a broadcast ARP request on Network 1.0.0.0 asking for Red's MAC address.

Note: This step will not occur if Green's ARP table still contains an entry for Red resulting from steps 1 and 2.

Green looks in its ARP cache for a MAC address for Red. If one is not found, Green broadcasts an ARP request frame on network 1.0.0.0 asking for Red's MAC address.

5-1-2-11 11) Red transmits a unicast ARP reply to Green providing its MAC address (A).

Note: This step will not occur if step 10 does not take place.

Red sees the ARP request frame and responds with a unicast ARP reply frame to Green. Also, Red enters Green's MAC address and IP address into its ARP cache. Green receives the ARP reply from Red and enters Red's MAC address (A) and IP address (1.1.1.1) into its ARP cache.

5-1-2-12 12) Green forwards Blue's PING reply to Red using the destination MAC address of Red (A).

Green transmits Blue's PING reply frame onto Red's network using Red's MAC address (A) and Red's IP address (1.1.1.1) as the destination MAC and destination IP address. The source MAC address is Green's MAC address (C) and the source IP address is Blue's IP address (2.2.2.1). The user sees the PING REPLY message printed on the screen. This completes the entire conversation.

Appendix B: Frequently asked questions

HP ProLiant Network Adapter Teaming Frequently Asked Questions

Q1 Why is traffic not being load balanced out of my server?

- or -

Why is traffic not being load balanced during backups?

A1 Either TLB or SLB is required for load balancing of transmit traffic. NFT will not provide for any type of load balancing. HP ProLiant Network Adapter Teaming uses either the MAC or IP address of the destination, or the TCP port information in the frame to make its load-balancing decisions. If the destination always has the same MAC and IP address (another server or client), no load balancing will result. If the destination has the same MAC address but the IP address is different (for example, several clients on the other side of a router), then HP ProLiant Network Adapter Teaming needs to be configured to load balance by TCP port or IP address instead of by MAC address. HP recommends using the Automatic method for load balancing.

Q2 Why is traffic not being load balanced into my server?

- or -

Why isn't traffic being load balanced during backups?

A2 A team type of SLB or 802.3ad Dynamic and a supporting switch are needed to achieve receive load balancing. Receive load balancing is determined by the switch connected to the HP team in SLB mode. If the HP NIC team is configured for SLB, then receive load balancing should occur if the switch is configured properly. Consult the technical resources provided by the switch manufacturer.

Q3 I am trying to team two NICs for load balancing but the system will not let me. Why?

- or -

I have an HP NC series Fast Ethernet adapter and an HP NC series copper Gigabit Ethernet adapter in a TLB or SLB team but I can not add an HP NC series fiber Gigabit Ethernet adapter to the team. Why?

A3 To team multiple ports together for load balancing, all ports must be capable of supporting a common speed. For instance, any 10/100 port can be teamed for load balancing with a 100/1000 port because both ports support a common speed. The ports do not have to be operating at the common speed. Teaming a 1000 fiber port with a 10/100 port is not supported for load balancing because the ports don't support a common speed.

Q4 Can I team HP NC Series fiber Gigabit Ethernet adapters with HP NC Series copper Gigabit Ethernet adapters?

A4 Yes, any team type.

Q5 What is the difference between HP's load-balancing teams and Microsoft's Network Load Balancing (NLB) or Window's Load Balancing Service (WLBS) features?

A5 HP teaming provides for fault tolerance and load balancing across network ports and is aimed at server resilience. Microsoft's NLB and WLBS are for fault tolerance and load balancing across servers and are aimed at application resilience.

Q6 Can I use HP ProLiant Network Adapter Teaming with Microsoft's NLB or WLBS features?

A6 Yes, however, some special configuration may be required. Support is limited to NLB and WLBS in Multicast mode only, not Unicast mode.

Q7 What effect will teaming have on the use of Cisco's Hot Swap Router Protocol (HSRP) or the IETF's Virtual Router Redundancy Protocol (VRRP) in my environment?

A7 None. HSRP and VRRP operate independently of teaming.

Q8 Can I use HP ProLiant Network Adapter Teaming in conjunction with Cisco Local Director?

A8 Yes, teaming will work correctly with Cisco Local Director.

Q9 I want to force a locally administered MAC address on my HP ProLiant Network adapter team. How should I do it?

A9 Open the HP Network Teaming and Configuration Properties GUI (NCU). Click the appropriate team in the GUI interface and select **Properties**. Go to the Settings tab and type the LAA address in the Team Network **Address field**.

Q10 How do I uninstall HP ProLiant Network Adapter Teaming?

A10 HP ProLiant Network Adapter Teaming can be uninstalled by opening the properties window of any network interface under Network and Dial-up Connections (Microsoft UI). Select **HP Network Teaming and Configuration** and click the **UNINSTALL**. Always dissolve any existing teams using the NCU before attempting an uninstall of HP ProLiant Network Adapter Teaming.

Q11 Is teaming multiple Fast Ethernet network adapters better than upgrading to a Gigabit Ethernet network adapter?

A11 Teaming multiple ports does provide for additional fault tolerance over using a single port. However, the throughput of several fast Ethernet ports will not usually be better than a single gigabit port.

Q12 Why does having Spanning Tree turned on for the HP ProLiant Network adapter team switch ports cause a problem sometimes?

A12 When link is lost on a port that has Spanning Tree enabled on it, Spanning Tree will isolate the port from communicating with the rest of the network for a certain time period. This time period can sometimes exceed a minute. This isolation period can cause communication loss, heartbeat failures, and undesired teaming failovers under certain conditions. Spanning Tree timeouts can also cause PXE boot failures. HP recommends bypassing the block, listen, and learn stages of Spanning Tree on switch ports connected to a server (for example, Cisco's PortFast).

Q13 Is HP ProLiant Network Adapter Teaming an industry-standard technology?

A13 The core of HP ProLiant Network Adapter Teaming technology is an industry-standard technology used for grouping network ports together for fault tolerance and load balancing. However, some of the special mechanisms that HP uses to enhance network port teaming are unique to HP teaming technology.

Q14 Can I use third party/non-HP network adapters with HP ProLiant Network Adapter Teaming?

A14 No, only HP branded network adapters may be used.

Q15 What does the Network Infrastructure Group need to do to help me deploy HP ProLiant Network Adapter Teaming correctly?

A15 For all team types, the Network Infrastructure Group needs to know the following:

- The VLAN IDs used on a particular team.
- Which group of ports constitutes a team. For each group, the following must be done:
- All ports must be configured for the same VLANs, if any.
- All ports must belong to the same broadcast domain(s).

For SLB teams, they also need to know which group of ports constitutes a team. For each group, all ports in each team must be configured as a single-port trunk/Multilink Trunk/EtherChannel group.

For 802.3ad Dynamic teams, they should only need to verify that LACP is enabled on the switch ports connecting to the server's teamed ports.

Q16 Can I use HP ProLiant Network Adapter Teaming in conjunction with Microsoft Cluster Server?

A16 Yes, however, Microsoft may request that teaming be disabled before technical assistance is provided. If teaming can be disabled and the problem still occurs, it may be assumed that the problem is not affected by teaming. You may always contact HP for support and HP Support can help resolve the issue.

Q17 My switch is set up for Switch-assisted Load Balancing (Port Trunking) but my network adapters are not. Why am I having communication problems?

A17 The switch assumes that all ports are capable of receiving on the same MAC addresses and will randomly transmit frames for any of the NICs down any of the links for any of the NICs. If the NICs are not configured for SLB teaming, they will drop any frame meant for another NIC. Because of this, the switch should only be configured for Port Trunking after the SLB team has been created.

If a single server with multiple NICs is connected to a switch configured for Port Trunking and PXE is being used to deploy the server, communication problems will most likely prevent PXE from completing a successful install on the server. To avoid such problems, disconnect all NICs except for the NIC providing PXE support or remove the Port-Trunking configuration on the switch.

Q18 What is the maximum number of network ports that can be in a single team?

A18 8 ports

Q19 What is the maximum number of teams that can be configured on a single HP server?

A19 16 teams

Q20 What is the maximum number of VLANs that can be configured on a single network port or a single team?

A20 64 VLANs

Q21 Why does my team lose connectivity for the first 30 to 90 seconds after the preferred Primary port's link is restored?

A21 This may be caused by Spanning Tree. Disable Spanning Tree on the port or enable the Spanning Tree bypass feature if available (for example, PortFast, bypass).

Q22 Is there a particular Windows Server 2003 Service Pack level that is required for HP ProLiant Network Adapter Teaming to work correctly?

A22 Service Pack 2 is recommended when using TOE, RSS, or SNP.

Q23 If I make an Altiris image of a server with a team, can I deploy that image onto other servers?

A23 Yes, but the team's MAC address registry entry will have to be restored individually on all servers the image was deployed on.

A better solution is to use the Windows Command Line utility called CQNICCMD.EXE (delivered in the NCU driver download) to import a team XML script to restore a teaming configuration.

Q24 Why do I still see heartbeat frames after disabling them?

A24 Even when heartbeats are disabled, an HP team must make sure that the team's MAC address is known by the switches in the network to prevent flooding of traffic destined for the team. To achieve this, the team needs to transmit a frame periodically. The frame used for this purpose is a heartbeat frame. Heartbeat frames are also used during a failover to notify the switch of MAC address changes on the teamed ports.

The purpose for disabling the transmit and receive validation heartbeats on the Settings tab of the Team Properties window is to either reduce the number of heartbeat frames on the network or to disable the Path Validation redundancy mechanism. Disabling transmit and receive validation do not completely disable the transmission of heartbeats on the network.

Q25 When should I increase the heartbeat timers for a team?

A25 The heartbeat timers should be increased when heartbeat failures are caused by latency in the network infrastructure. This is an extremely rare problem.

Q26 Is Unattended Installation of HP ProLiant Network Adapter Teaming supported?

A26 Yes. See NICSCRPT.PDF, which is located in the c:\compaq\network directory, if teaming has been installed on the system.

Q27 Which switch Port Trunking/Port Bonding technologies is SLB (Switch-assisted Load Balancing) compatible with?

A27 FEC/GEC, Load Sharing, MLT, IEEE 802.3ad (Static mode – no LACP), etc.

Q28 Which switch Port Trunking/Port Bonding technologies is 802.3 Dynamic compatible with?

A28 IEEE 802.3ad (Dynamic mode – using LACP)

Q29 Can I connect teamed ports to more than one switch?

A29 Yes, with NFT and TLB teams only. Also, all switch ports that have teamed ports connected to them must belong to the same broadcast domain. This means that the broadcast domain must span between the switches.

Q30 Who in HP is responsible for development and support of HP ProLiant Network Adapter Teaming?

A30 The Network Server Products (NSP) group in Austin, Texas provides all development and Level 3 support for HP ProLiant Network Adapter Teaming technology, in addition to almost all ProLiant networking products (for example, BL switches, ProLiant network adapters, Infiniband, iSCSI, and so on). NSP is one of many groups that constitute HP's Industry Standard Servers division.

Q31 What is the limit for the number of teams I can create in one server and what is the limit for the number of network ports that I can include in one team?

A31 The theoretical limit is 16 teams of 8 network adapter ports per server. This is defined as a "theoretical" limit because very few servers will allow the installation of enough network ports to create 16 teams of 8 network adapter ports.

Q32 How do I upgrade HP ProLiant Network Adapter Teaming drivers?

A32 HP provides the HP ProLiant Networking Software SoftPak with an installation utility. The same component completes installation and upgrades. Download the latest softpak from <http://www.hp.com/cgi-bin/hpsupport/index.pl>

Q33 Which HP ProLiant Network Adapters can I team together?

- or -

Can I team together embedded network adapter ports with standup network adapter ports?

- or -

Can I team together ports from more than one standup network adapter?

- or -

Can I team together Intel-based HP ProLiant Network Adapters with Broadcom-based HP ProLiant Network Adapters?

A33 One of the main benefits of HP ProLiant Network Adapter Teaming for ProLiant customers is its flexibility to support teaming together ports from almost all HP ProLiant NC series Network Adapters. HP ProLiant Network Adapter Teaming supports teaming ports together from HP ProLiant Network Adapters with Intel or Broadcom ASICs, whether the ports are embedded ports (LAN on Motherboard – LOM) or ports from standup network adapters. Support for a specific team type or advanced feature (for example, offloading) may vary depending on the combination of HP ProLiant Network Adapter ports.

Appendix C: Overview of utilities included with HP ProLiant Network Adapter Teaming

In the Component Pack (CPxxxx.exe) for HP ProLiant Network Adapter Teaming there are several command line utilities provided for special teaming-related management functions. Below is a list of the utilities with a brief description of each utility. Consult the latest HP ProLiant Network Adapter Teaming Component Pack for the latest utilities and associated documentation.

- **Cqniccmd.exe**
CQNICCMD is a Windows Server 2003 and Windows Server 2008 Enterprise utility that processes a network adapter configuration script file to duplicate the HP ProLiant Network Adapter Teaming configuration of a source ProLiant ML/DL/BL server on a target server. The utility can be run from the command line in a Command Prompt window, from the Run option on the Windows Start menu, or from a Windows command file. This utility can also be used during unattended deployment of the operating system.

For more information, see the NICSCRIPT.PDF file in the HP ProLiant Network Adapter Teaming Component Pack.
- **Hpnetinfo.exe**
HPNETINFO is a Windows Server 2003 and Windows Server 2008 Enterprise utility that allows for querying and retrieving network adapter related information from a server. For example, find the number of NICs in a server, display network adapter model numbers, retrieve Local Area Connection (LAC) information, and so on. The utility is normally used in conjunction with several sample batch files: getlacinfo.bat, getmodel.bat, getnetov.bat, getnicnumber.bat, and getnumlacs.bat. All of these batch files are delivered in the HP ProLiant Network Adapter Teaming Component Pack.
- **NALicense.exe**
The Network Adapter License (nalicense.exe) utility can be used to view currently installed licenses on the system.

For more information on NALicense.exe, see the NALICNSE.PDF file in the HP ProLiant Network Adapter Teaming Component Pack.
- **SetLACState.exe**
If a server contains more than one server adapter, and at least one of those adapters is not connected, the server is displayed as failed on some network management consoles, and SNMP traps may be triggered. The SetLACState utility allows the user to disable the Local Area Connection containing the unused adapter and, as a result, the server does not display as failed, and SNMP traps are avoided. This utility can also be used during unattended deployment of the operating system to disable the adapters that will not be used.

For more information on SetLACState.exe, see the SetLAC.PDF file in the HP ProLiant Network Adapter Teaming Component Pack.
- **Hpnetsvy.exe []**
HPNetSvy is a command line utility used to gather information about HP network products in a ProLiant server. This utility provides a snapshot of configuration information and statistics for HP ProLiant Ethernet Network adapters and HP ProLiant Network Adapter Teaming. The information provided by this tools is mainly used by HP Support and Engineering, but can be used by customers.

Usage: HpNetSvy [/f<filename> /a] [/n<iterations> /t<interval>] [/help | /?]

/f	File to write output to instead of displaying on screen. Default is to replace file contents unless /a is used.
/a	Appends output to file.
/n	Number of iterations to run the utility.
/t	Number of seconds to wait between iterations.

Appendix D: ProLiant networking information resources

Product information

<http://www.hp.com/go/ProLiantNICs>

Whitepapers

<http://h18004.www1.hp.com/products/servers/networking/whitepapers.html>

Network adapter and network adapter teaming drivers

Software and drivers: <http://www.hp.com/cgi-bin/hpsupport/index.pl>

Note: The HP ProLiant Network Adapter Teaming driver is a separate downloadable component listed under the network adapter or server that supports it. Choose the network adapter or server by OS and then select "HP ProLiant Networking Software " to download and install HP ProLiant Network Adapter Teaming.

Appendix E: Technical Support

To contact an HP Technical Support engineer regarding issues with HP ProLiant Network Adapter Teaming, call 1-800-652-6672. To speak with the appropriate support group, select the call routing options for HP Server Networking, or HP ProLiant Networking.

For online assistance, HP's ProLiant web-based Support Forum is located at

<http://forums.itrc.hp.com/cm/CategoryHome/1,,264,00.html>

Appendix F: Glossary

802.1D	Refers to the IEEE 802.1D specification. This is the original Spanning Tree specification.
802.3ad Dynamic	Refers to the IEEE 802.3ad specification. This specification provides for manual and automatic (dynamic) port grouping for fault tolerance and load balancing between two network devices.
ALB Adaptive Load Balancing	See Transmit Load Balancing (TLB).
ARP	Address Resolution Protocol. A Layer 2 protocol used to determine a MAC address from an IP address. A network device broadcasts an ARP request to ask for the MAC address for a given IP address. The network device with the given IP address responds back with an ARP reply providing its MAC address to the original requester.
BIA	Burned In Address. The Layer 2 address that is permanently assigned to a piece of hardware by the vendor. Referred to as a MAC address. Can be overridden by a Locally Administered Address (LAA).
Bit	The smallest value in binary. A bit is a single value that is equal to either 1 or 0. The collection of 8 bits is called a byte.
BPDU	Bridge Protocol Data Unit. A special configuration frame used by the Spanning Tree Protocol.
Broadcast domain	Set of all devices that will receive Layer 2 broadcast frames originating from any device within the set. Broadcast domains are typically bounded by routers because routers do not typically forward Layer 2 broadcast frames.
Byte	Eight bits
Collision domain	A single Ethernet network segment in which there will be a collision if two computers attached to the system transmit simultaneously.
Component Pack	An HP-specific term used for a self extractable package for installing drivers on a server. The file name syntax is CPxxxx.exe.
Degraded	A team or teamed port status that indicates it is still in use but not completely whole.
Echo Node	A special device used by the Active Path mechanism. This is a device that is simply "designated" for use. The device doesn't need any special software or configuration other than an IP address.
Echo Node probe	A special frame used by Active Path to validate connectivity with the Echo Node.
Failed	A team or teamed port status that indicates it is not in use because of a failure important enough to cause it to be disabled.
Failover	A term used to describe any event that causes the role of primary port in a team to change from one teamed port to another.
Fast Path	An advanced redundancy mechanism used by HP ProLiant Network Adapter Teaming that allows the team to monitor per-teamed port connectivity with the Spanning Tree root switch and monitor network bandwidth. Requires the team to have an Intelligent Networking Pack license.
FEC	Fast EtherChannel. A method of load balancing that both transmits and receives traffic across multiple Fast Ethernet connections (100 Mbps) between two devices. Developed by Cisco Systems. See SLB.
GEC	Gigabit EtherChannel. A method of load balancing that both transmits and receives traffic across multiple Gigabit Ethernet (1000 Mbps) connections between two devices. Developed by Cisco Systems. See SLB.

GUI	Graphical User Interface.
HP Teaming and Configuration GUI	Now referred to as the HP Network Configuration Utility (NCU). See NCU.
IEEE	Institute of Electrical and Electronics Engineers. A standards body for, among other things, network communications and protocols.
Intermediate driver	A special kind of networking driver used in Microsoft operating systems. This special network driver operates in between other networking drivers to perform a specialized function.
LAA	Locally Administered Address. A temporary Layer 2 address that is manually assigned to a hardware or software network device.
LACP	Link Aggregation Control Protocol. The protocol used by 802.3ad to dynamically create port trunk groups between two network devices that both support LACP.
LAN	Local Area Network.
Layer 2	The second layer of the OSI model, the Data Link Layer. A Layer 2 address is the same as a MAC (Media Access Control) address.
Layer 3	The third layer of the OSI model, the Network Layer. A Layer 3 address refers to a protocol address such as an IP or IPX address.
Local Area Connection	Also referred to as LAC. A LAC is one type of connection found in "Network Connections" on Microsoft operating systems. A LAC typically represents a device that connects the PC/Server to a local area network (LAN). A LAC can see an actual hardware device (NIC) or to a software device (for example, team, VLAN, virtual NIC interface, Loopback, and so on).
LOM	LAN On Motherboard. The acronym refers to a network adapter that is built directly onto the computer's motherboard. LOMs are not removable and do not use up slots in the computer (for example, PCI).
MAC address	Media Access Control address. With Ethernet, this refers to the 6-byte (48-bit) address that is unique to every Ethernet device.
Multi-homed	A device that is redundantly attached to a network or networks.
NCU	Network Configuration Utility. This is the GUI used by HP ProLiant Network Adapter Teaming for all configuration options related to NIC teaming and VLAN creation on all teamed and non-teamed ports.
NDIS	Network Driver Interface Specification. Simplified, it is the interface between a network adapter and Microsoft's protocol stack.
NCDE	Network Controller Drivers for Ethernet. Product name used by HP Engineering to see a package of drivers for ProLiant server network adapters, teaming, firmware, and so on. NCDE packages are released periodically and always have an associated revision level for the entire package. Individual driver versions inside an NCDE package may differ. NCDE packages are available for download from www.hp.com .
Network adapter	A physical network adapter card. Network adapter cards may contain one or more network adapter ports. Synonymous with NIC.
Network adapter port	A physical network adapter port or NIC port. Multiple network adapter ports can exist on the same physical network adapter card.
Network Configuration Utility	The GUI used to configure and monitor HP ProLiant Network Adapter Teaming.
NFT	Network Fault Tolerance. A team of network ports that transmits and receives on only one port with all other ports in standby.
NIC	Network Interface Card. Synonymous with network adapter.

NIC teaming	A phrase referring to HP ProLiant Network Adapter Teaming.
OSI Model	Open Systems Interconnect Model. The seven-layer model developed by the International Standards Organization (ISO) that outlines the mechanisms used by networked devices to communicate with each other.
PING	A type of packet used to validate connectivity with another network device. The packet asks another network device to respond to validate connectivity, a kind of “echo.” PING packets for IP are accomplished using the ICMP protocol.
PortFast	A special setting used mainly by Cisco switches and HP ProCurve switches to bypass the block, listen, and learn stages of Spanning Tree on a particular port. Port Fast ports still transmit BPDUs.
ProLiant Essentials	A mechanism used for licensing many different types of options for HP ProLiant Servers. See http://h71028.www7.hp.com/enterprise/cache/43768-0-0-225-121.aspx for more information.
PVST+	Cisco’s Per VLAN Spanning Tree Plus. A Cisco proprietary Spanning Tree protocol that runs Spanning Tree on each individual VLAN.
RAID	Redundant Array of Independent Disks. Disk drive technology used for redundancy, load balancing, or a combination of both.
SLB	Switch-assisted Load Balancing. Also known as FEC/GEC. A team of network ports that load balances transmits and receives on all ports.
STA	Spanning Tree Algorithm (IEEE 802.1D).
Switch MAC table	A list of MAC addresses and associated ports that are used by a switch to transfer frames between attached devices. Also referred to as a CAM table.
LLC	Logical Link Control.
Teamed port	Any network adapter port that has been added to a team.
TLB	Transmit Load Balancing. Was known as Adaptive Load Balancing (ALB). A team of network ports that receives on one port but load balances transmitted IP traffic on all ports. Other protocol traffic is transmitted by a single port.
Virtual Local Area Connection	All of these terms see a team. A team, comprising two or more real ports, represents itself to the host operating system as a single virtual network adapter/NIC interface/NIC port.
Virtual network adapter	
Virtual NIC interface	
Virtual NIC port	
XML	eXtensible Markup Language. Used by teaming to import or export team configurations.