# HP ProLiant Network Adapter Teaming

## White Paper

Part number: 4AA0-1139ENW

**Legal notices**

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.

Printed in the US.

Part number: 4AA0-1139ENW

# Contents

## Revision history

## 1 Introduction

## 2 An executive overview of teaming

## 3 Teaming fundamentals for the first-time user

## 4 The mechanics of teaming for the advanced user

# Figures

# Tables

# Revision history

## Revision tables

**Table 0-1** Teaming Whitepaper Revision History

| Date | Author | Major Revision | Minor Revision | Reference Driver Version |
|---|---|---|---|---|
| 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 |
|---|---|---|
| August 16, 2005 | 8.10.0.0 | Added support for the following:<br>• NC320i, NC324i, NC325i and NC326i LOMs<br>• Read-only web interface accessible through the System Management Homepage.<br>• The ability to create a detailed log of your system configuration for submission to third-level technical support personnel.<br>• Status of the active path and fast path states, along with member ID information and readability enhancements.<br>• Dynamic dual-channel load balancing, so user intervention is minimal and the NIC grouping happens automatically.<br>• 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:<br>• NC370x NICs.<br>• 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:<br>• 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:<br>• NIC Teaming Utilization Tab. This tab displays current and peak speed, throughput, and utilization information for the selected team and its members.<br>• 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.<br>Removed support for the following:<br>• HP NC3121 Fast Ethernet Adapter |
| September 11, 2004 | 7.71.0.0 | Added support for the following:<br>• NC320T PCI Express Gigabit Server Adapter.<br>Resolved the following issues:<br>• 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.<br>• NCU no longer fails to create teams on a system that has never had this software installed before. |

**Table 0-2 Teaming Driver Revision History**

| Date | Driver Version | Overview of Changes |
|---|---|---|
| August 18, 2004 | 7.70.0.0 | Added support for the following:<br>• large-send offload (LSO) and checksum offload advanced teaming features, which can reduce TCP/IP processing overhead.<br>• NC 320T PCI Express Gigabit Server Adapter, NC 150T PCI 4-port Gigabit Combo Switch Adapter, NC 310F PCI-X Gigabit Server Adapter.<br>• 802.3ad dynamic teaming which allows the user to create channels on the switch without configuring the switch.<br>• ProLiant Essentials Licensing support for the ProLiant Essentials Intelligent Networking Pack.<br>• Active Path Failover guarantees the team to use a port that has a path to a selected destination.<br>• Fast Path Failover allows the team to use the most expedient route to a destination.<br>• Dual channel teaming allows receive load balancing on a team connected to two switches.<br>• 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, it is recommended that any prior sections be read first. This is recommended since some information is not repeated in every section.

## Section layout

### 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 to an Executive seeking business justification for deploying HP ProLiant Network Adapter Teaming.

### 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/implementer. This section can also 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.

### 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, etc. 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 2000 and Windows Server 2003. For technical information on network adapter teaming for other operating systems, refer to **www.hp.com**.

Information in this document is specific to the set of drivers contained in release version NCDE 8.10. 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, etc.  One of the most often overlooked fault-tolerant devices is the server network adapter.  Many server administrators (SA) spend thousands of dollars on eliminating single points of failure in their servers.  They do this 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, it can be a frustrating experience – both 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, etc.) 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

  Network cable disconnects is one type of failure that is often out of the control of the SA.  This type of failure can be caused by someone tripping over the network cable, from 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 can't 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, do occur.  It is the prudent SA that 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 it 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 refer to 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 since all hard drive hardware is 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. When the SA has the budget to purchase newer technology (for example, Gigabit Ethernet – 1000 Mb), the team of Fast Ethernet adapters can be replaced 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) a solution 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 both 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 purposefully 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. For example, HP’s patent-pending advanced teaming mechanisms [refer to the ProLiant Essentials Intelligent Networking Pack (INP) upgrade for HP

ProLiant Network Adapter Teaming] provide fault-recovery and load-balancing intelligence to a server that's never been available in the industry.

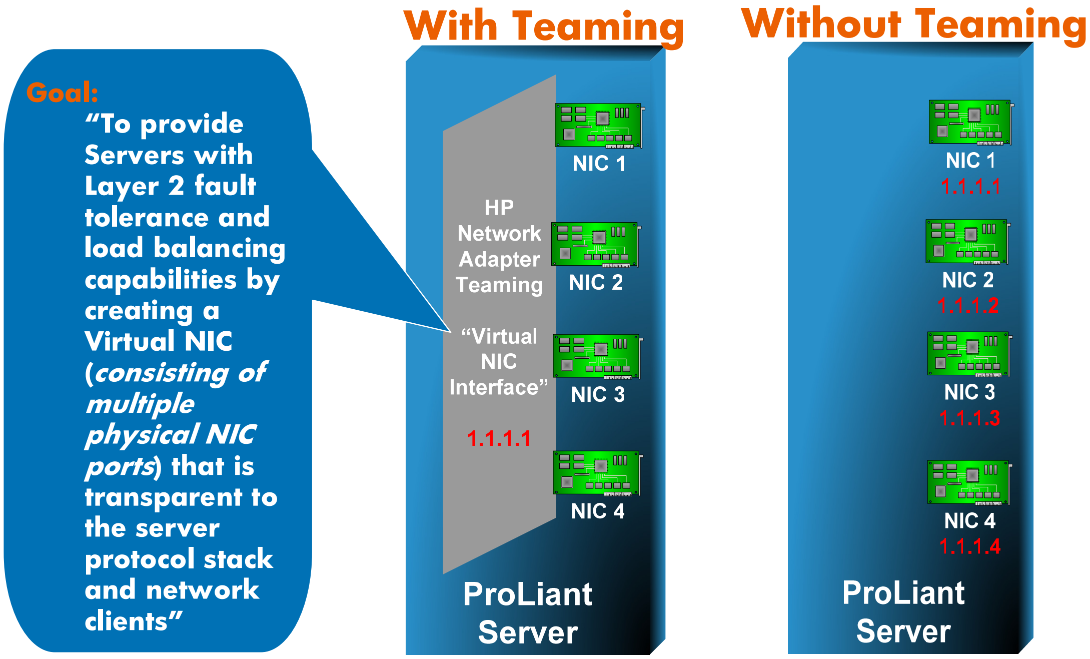
HP is committed to providing HP ProLiant server customers with consistent, high quality, and innovative server networking technology products.

# 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



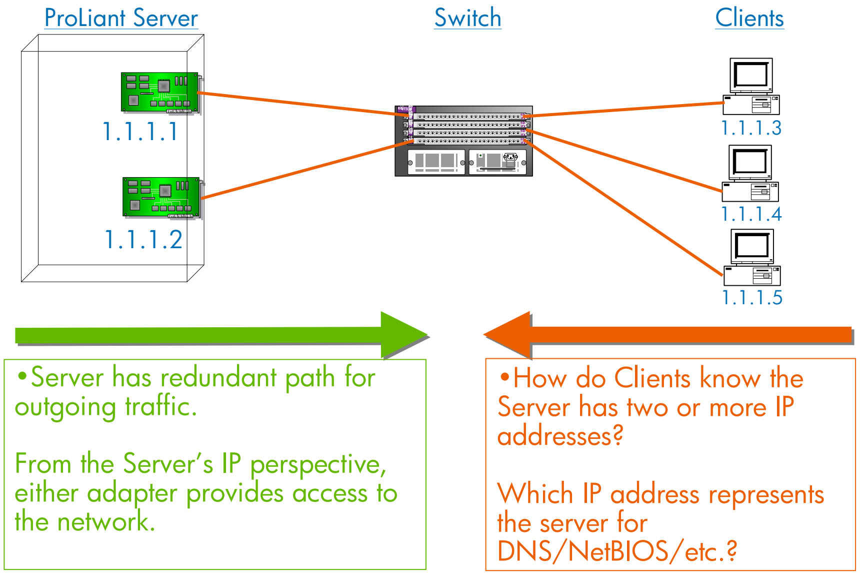
With Teaming
Without Teaming
Goal:
"To provide Servers with Layer 2 fault tolerance and load balancing capabilities by creating a Virtual NIC (*consisting of multiple physical NIC ports*) that is transparent to the server protocol stack and network clients"
HP Network Adapter Teaming
"Virtual NIC Interface"
1.1.1.1
NIC 1
NIC 2
NIC 3
NIC 4
ProLiant Server
NIC 1
1.1.1.1
NIC 2
1.1.1.2
NIC 3
1.1.1.3
NIC 4
1.1.1.4
ProLiant Server

### 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 (refer to Figure 3-2). Dual Homing does provide path redundancy for the server to the network. However, Dual Homing does not completely solve the network redundancy problem for a server. The biggest issue is that 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



- Server has redundant path for outgoing traffic.

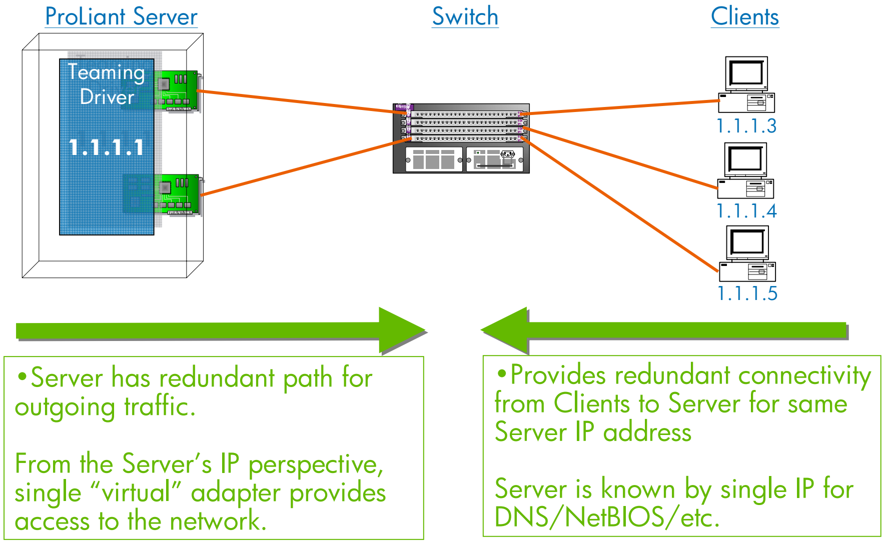
From the Server's IP perspective, either adapter provides access to the network.

- How do Clients know the Server has two or more IP addresses?

Which IP address represents the server for DNS/NetBIOS/etc.?

In contrast, Network Adapter Teaming provides transparent redundancy for both the server's outbound connectivity and all inbound client connectivity (refer to 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, etc.) 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



- Server has redundant path for outgoing traffic.

From the Server's IP perspective, single "virtual" adapter provides access to the network.

- Provides redundant connectivity from Clients to Server for same Server IP address

Server is known by single IP for DNS/NetBIOS/etc.

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

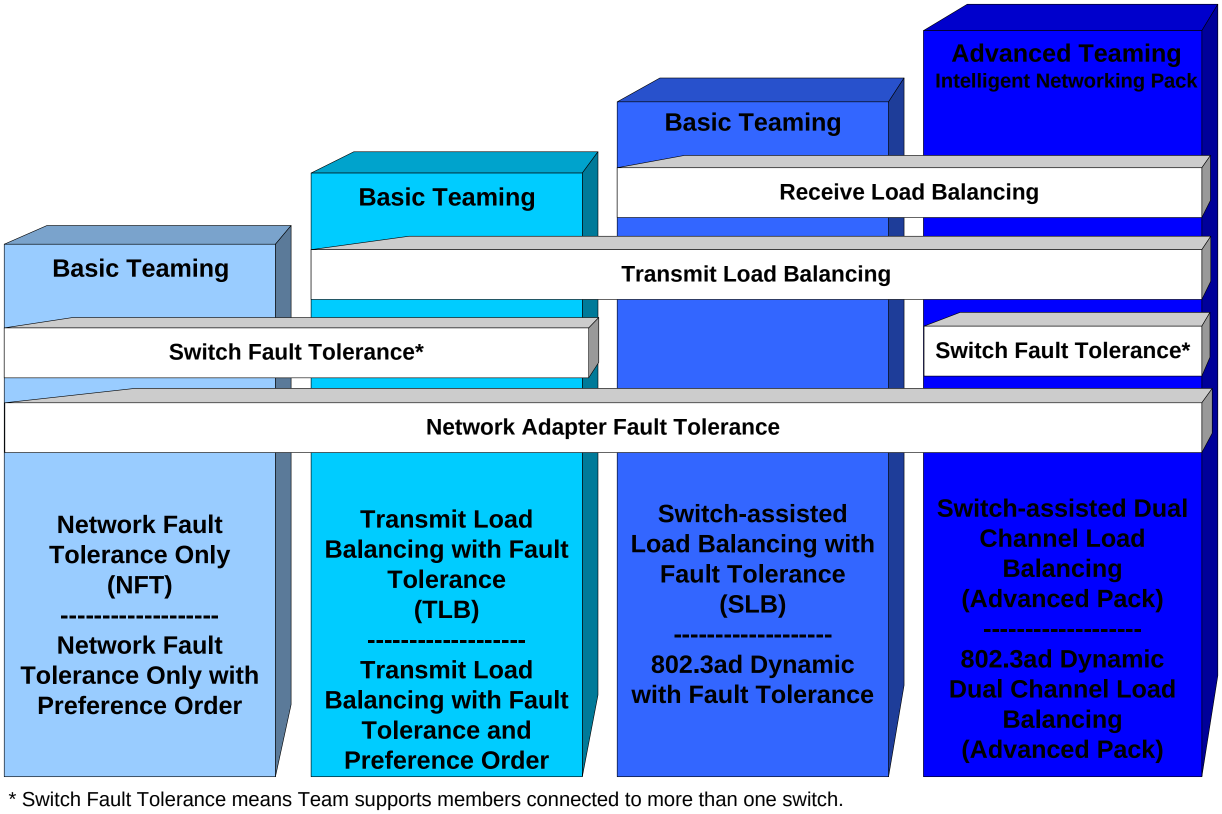
There are two widely used solutions for achieving load balancing for a server's network traffic. An SA can deploy a server-based routing protocol that supports load balancing (for example, OSPF, EIGRP, IS-IS, etc.) or the SA can deploy teaming. Server-based routing protocols are much more difficult to deploy and manage. In addition, they do not allow for load balancing between end devices that don't 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


Advanced Teaming
Intelligent Networking Pack
Basic Teaming
Basic Teaming
Basic Teaming
Receive Load Balancing
Transmit Load Balancing
Switch Fault Tolerance*
Switch Fault Tolerance*
Network Adapter Fault Tolerance
Network Fault Tolerance Only (NFT)
-------------------
Network Fault Tolerance Only with Preference Order
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
Switch-assisted Dual Channel Load Balancing (Advanced Pack)
-------------------
802.3ad Dynamic Dual Channel Load Balancing (Advanced Pack)

* 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. It can be said that the foundation of every team type supports NFT.

NFT is a standard feature of HP ProLiant Network Adapter Teaming and supports the features provided in the ProLiant Essentials Intelligent Networking Pack (INP) upgrade. Refer to "3-4 Basic teaming versus advanced teaming" for more information.

### 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. The only difference is 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, etc.), or preference for the team's Primary port to be located on a specific switch.

NFT with Preference Order is a standard feature of HP ProLiant Network Adapter Teaming and supports the features provided in the INP upgrade. Refer to "Basic teaming versus advanced teaming" for more information.

## 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.

TLB is a standard feature of HP ProLiant Network Adapter Teaming and supports the features provided in the INP upgrade. Refer to "Basic teaming versus advanced teaming" for more information.

## 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. The only difference is 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, etc.), or preference for the team's Primary port to be located on a specific switch.

Transmit Load Balancing with Fault Tolerance and Preference Order is a standard feature of HP ProLiant Network Adapter Teaming and supports the features provided in the INP upgrade. Refer to "Basic teaming versus advanced teaming" for more information.

## 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, etc.). SLB doesn't support switch redundancy since all ports in a team must be connected to the same switch. SLB is similar to the 802.3ad Dynamic team type discussed later.

SLB is a standard feature of HP ProLiant Network Adapter Teaming but doesn't support any of the advanced redundancy features provided in the Intelligent Networking Pack upgrade.

## 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 Switch-assisted Dual Channel Load Balancing (Advanced Pack)

Switch-assisted Dual Channel Load Balancing, referred to as Dual Channel or a "team of teams", is a special team type designed by HP to accomplish everything that NFT, TLB and SLB team types accomplish all in a single team type. Prior to Dual Channel, an HP ProLiant Network Adapter Teaming user had to choose between inbound load balancing (SLB or 802.3ad Dynamic) or switch redundancy (NFT, TLB). Dual Channel allows the user to create two teams, called groups, inside of a single team. Each group (A or B) is assigned one or more teamed ports. Also, each group can be connected to a different switch to provide switch fault tolerance. Full inbound and outbound load balancing is provided across both groups. Should all members in one group completely fail, the team remains available via the other group. If the groups are connected to different switches and one of the switches fail, the team remains available via the group attached to the functional switch.

Switch-assisted Dual Channel Load Balancing requires activation with a license key and is part of the INP upgrade for HP ProLiant Network Adapter Teaming. Refer to "Basic teaming versus advanced teaming" for more information.

## 3-2-8 802.3ad Dynamic Dual Channel Load Balancing (Advanced Pack)

802.3ad Dynamic Dual Channel Load Balancing, referred to as Dynamic Dual Channel, is identical to Dual Channel 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 benefits of Dynamic Dual Channel are that an SA will not have to manually configure the switch and will not have to manually assign teamed ports to individual groups (A or B). Dynamic Dual Channel utilizes the 802.3ad Link Aggregation Protocol to automatically determine which teamed ports are connected to common switches. Dynamic Dual Channel will then dynamically create Port Trunk (channels) for those "switch port to team port" combinations.

802.3ad Dynamic Dual Channel Load Balancing requires activation with a license key and is part of the INP upgrade for HP ProLiant Network Adapter Teaming. Refer to "Basic teaming versus advanced teaming" for more information.

## 3-2-9 Automatic

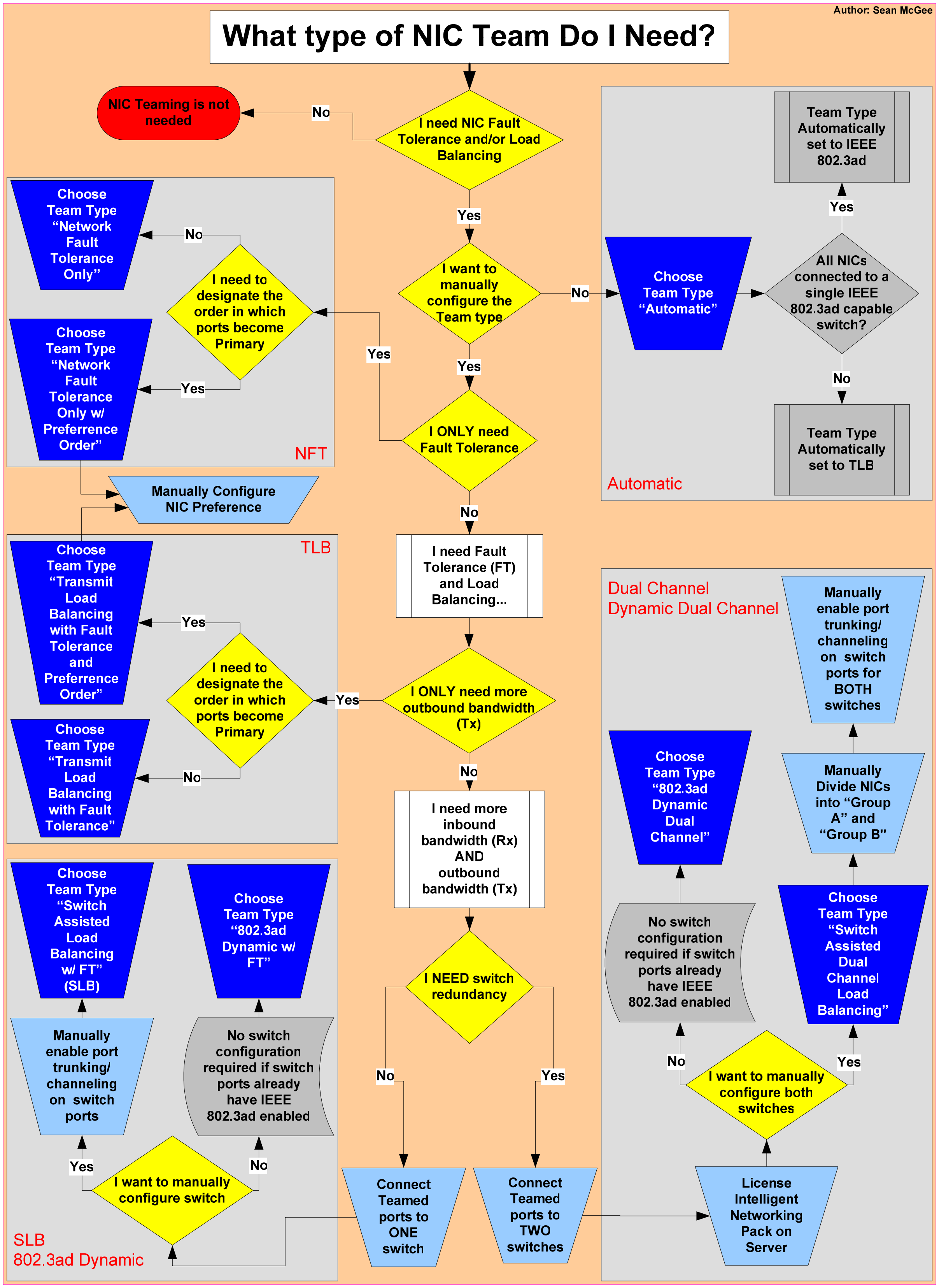
The Automatic team type is not really an individual team type. Automatic teams decide whether to operate as an NFT or a TLB team or as an 802.3ad Dynamic team. If all teamed ports are connected to a switch that supports the IEEE 802.3ad Link Aggregation Protocol (LACP) and all teamed ports are able to negotiate 802.3ad operation with the switch, then the team will choose to operate as an 802.3ad Dynamic team. However, if the switch doesn't support LACP or if any ports in the team don't 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.

Automatic is a standard feature of HP ProLiant Network Adapter Teaming and supports the features provided in the INP upgrade. Refer to "Basic teaming versus advanced teaming" for more information.

# 3-3 How to choose the best team type

There are several important decisions that need to be made when choosing the team type to use when deploying HP ProLiant Network Adapter Teaming (refer to Figure 3-5).

**Figure 3-5** How to choose the best team type to use



Author: Sean McGee
What type of NIC Team Do I Need?
I need NIC Fault Tolerance and/or Load Balancing
No
NIC Teaming is not needed
Yes
I want to manually configure the Team type
No
Choose Team Type "Automatic"
All NICs connected to a single IEEE 802.3ad capable switch?
Yes
Team Type Automatically set to IEEE 802.3ad
No
Team Type Automatically set to TLB
Automatic
Yes
I ONLY need Fault Tolerance
Yes
I need to designate the order in which ports become Primary
No
Choose Team Type "Network Fault Tolerance Only"
Yes
Choose Team Type "Network Fault Tolerance Only w/ Preferrence Order"
NFT
Manually Configure NIC Preference
No
I need Fault Tolerance (FT) and Load Balancing...
I ONLY need more outbound bandwidth (Tx)
Yes
I need to designate the order in which ports become Primary
Yes
Choose Team Type "Transmit Load Balancing with Fault Tolerance and Preferrence Order"
No
Choose Team Type "Transmit Load Balancing with Fault Tolerance"
TLB
No
I need more inbound bandwidth (Rx) AND outbound bandwidth (Tx)
I NEED switch redundancy
No
Connect Teamed ports to ONE switch
Yes
Connect Teamed ports to TWO switches
License Intelligent Networking Pack on Server
I want to manually configure switch
Yes
Manually enable port trunking/ channeling on switch ports
Choose Team Type "Switch Assisted Load Balancing w/ FT" (SLB)
No
No switch configuration required if switch ports already have IEEE 802.3ad enabled
Choose Team Type "802.3ad Dynamic w/ FT"
SLB
802.3ad Dynamic
I want to manually configure both switches
No
No switch configuration required if switch ports already have IEEE 802.3ad enabled
Choose Team Type "802.3ad Dynamic Dual Channel"
Yes
Choose Team Type "Switch Assisted Dual Channel Load Balancing"
Manually Divide NICs into "Group A" and "Group B"
Manually enable port trunking/ channeling on switch ports for BOTH switches
Dual Channel
Dynamic Dual Channel

# 3-4 Basic teaming versus advanced teaming

Basic HP ProLiant Network Adapter Teaming is provided as a standard on all supported HP ProLiant servers. Advanced teaming, called ProLiant Essentials Intelligent Networking Pack (INP), is provided as a license-enabled option for HP ProLiant customers requiring the advanced-redundancy and load-balancing features it provides.

**Table 3-1** Basic teaming versus advanced teaming capabilities comparison

| | Basic teaming | Advanced teaming (INP) |
|---|---|---|
| NFT | ✓ | ✓ |
| NFT with Preference | ✓ | ✓ |
| TLB | ✓ | ✓ |
| TLB with Preference | ✓ | ✓ |
| SLB | ✓ | ✓ |
| 802.3ad Dynamic | ✓ | ✓ |
| Link loss detection | ✓ | ✓ |
| Transmit validation heartbeats | ✓ | ✓ |
| Receive validation heartbeats | ✓ | ✓ |
| **Dual Channel** | | ✓ |
| **Dynamic Dual Channel** | | ✓ |
| **Active Path** | | ✓ |
| **Fast Path** | | ✓ |

The basic redundancy mechanisms that come standard with HP ProLiant Network Adapter Teaming are link loss, transmit validation heartbeats, and receive validation heartbeats. These basic redundancy mechanisms monitor each teamed port for link, the ability to transmit a frame, and the ability to receive a frame. These basic tests help HP ProLiant Network Adapter Teaming choose teamed ports that have been validated for “simple” network connectivity.

The INP upgrade for HP ProLiant Network Adapter Teaming is a set of advanced teaming features that provides advanced redundancy and advanced load balancing. The advanced redundancy features are called Active Path and Fast Path and the advanced load-balancing feature is called Dual Channel. These advanced features provide network-aware intelligence to HP ProLiant Network Adapter Teaming that was not previously available in any HP or non-HP server networking product.

- Active Path

  Active Path is a mechanism used by HP ProLiant Network Adapter Teaming to intelligently and proactively determine which teamed ports should or should not be used by the team based on “tested” connectivity with an external network device called an Echo Node. Each teamed port tests connectivity to the Echo Node by transmitting a frame to the Echo Node and monitoring for a response. Active Path can determine if teamed ports have become segregated onto different networks (for example, because of link loss on an upstream switch uplink, server network adapter ports connected to the wrong switch ports, VLAN misconfiguration on a switch, etc.). Teamed port segregation is determined based on successful communication with the Echo Node on a per teamed port basis.

  Active Path is extremely important for configurations with teamed ports connected to more than one switch.

- Fast Path

  Fast Path is a mechanism used by HP ProLiant Network Adapter Teaming to intelligently and proactively determine the best teamed ports to use to provide optimum network connectivity for the server. As with Active Path, Fast Path can determine if teamed ports have become segregated onto different broadcast domains (because of link loss on an upstream switch uplink, server network adapter ports are connected to the wrong switch ports, VLAN misconfiguration on a switch, etc.) Also, Fast Path is used by HP ProLiant Network Adapter Teaming to determine which teamed port is the optimal port to use as the team’s Primary port among one or more teamed ports with validated connectivity to the preferred Spanning Tree root switch.

  Fast Path is extremely important for configurations with teamed ports connected to more than one switch.

- Dual Channel

Switch-assisted Dual Channel Load Balancing, simply referred to as Dual Channel, is a special team type designed by HP Engineering to accomplish everything that NFT, TLB and SLB team types accomplish all in a single team type (refer to Table 3-2).

Dual Channel is extremely important for configurations requiring full load balancing AND switch redundancy at the same time.

**Table 3-2** Dual Channel capabilities comparison

| | Fault Tolerance | Transmit Load Balancing | Receive Load Balancing | Switch Redundancy |
|---|---|---|---|---|
| NFT | ✓ | | | ✓ |
| TLB | ✓ | ✓ | | ✓ |
| ****Dual Channel**** | ✓ | ✓ | ✓ | ✓ |
| ****Dynamic Dual Channel**** | ✓ | ✓ | ✓ | ✓ |
| SLB | ✓ | ✓ | ✓ | |
| 802.3ad | ✓ | ✓ | ✓ | |
| NFT | ✓ | | | ✓ |

While the basic redundancy mechanisms come standard with HP ProLiant Network Adapter Teaming, the advanced redundancy and load-balancing mechanisms provided by INP require the purchase of a per-server license through the ProLiant Essentials licensing mechanism. Refer to "3-5-3 License" for more information.

# 3-5 Basic deployment steps for HP ProLiant Network Adapter Teaming

## 3-5-1 Download

HP ProLiant Network Adapter Teaming is listed as a downloadable component on the drivers download page for every HP network adapter and HP ProLiant server that supports it. To find the download for HP ProLiant Network Adapter Teaming, go to http://h18004.www1.hp.com/support/files/networking/us/index.html and select the type of network adapter in the server. Next, select the operating system support needed. When the driver page loads, go to the "Driver – Network" section.

Select and download the driver labeled "HP Network Configuration Utility". The downloadable package is called a "Component Pack" and has a file name like CP*xxxxx*.exe.

The INP does not require separate installation. It is already installed with basic HP ProLiant Network Adapter Teaming but requires the purchase of a license key to activate it.

## 3-5-2 Install

The HP ProLiant Network Adapter Teaming driver is a self-installing executable. Launch the CP*xxxx*.exe file and follow any directions to complete the installation. If any 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.

## 3-5-3 License

To utilize the advanced redundancy mechanisms (Active Path and Fast Path) and the Dual Channel team mode, an INP license must be installed. If the advanced teaming features are not needed, proceed to the next step.

The license can be purchased online by going to http://h18004.www1.hp.com/products/servers/proliantessentials/inp/index.html and selecting **BUY ONLINE**.

The license is installed by launching the HP Network Configuration Utility (NCU) and clicking the **License Manager** button. The INP license can be manually entered into the License Manager or a key file can be used to import the license. Once a license is installed on a server, it cannot be removed until the operating system is reinstalled.

## 3-5-4 Launch

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

# 3-6 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.

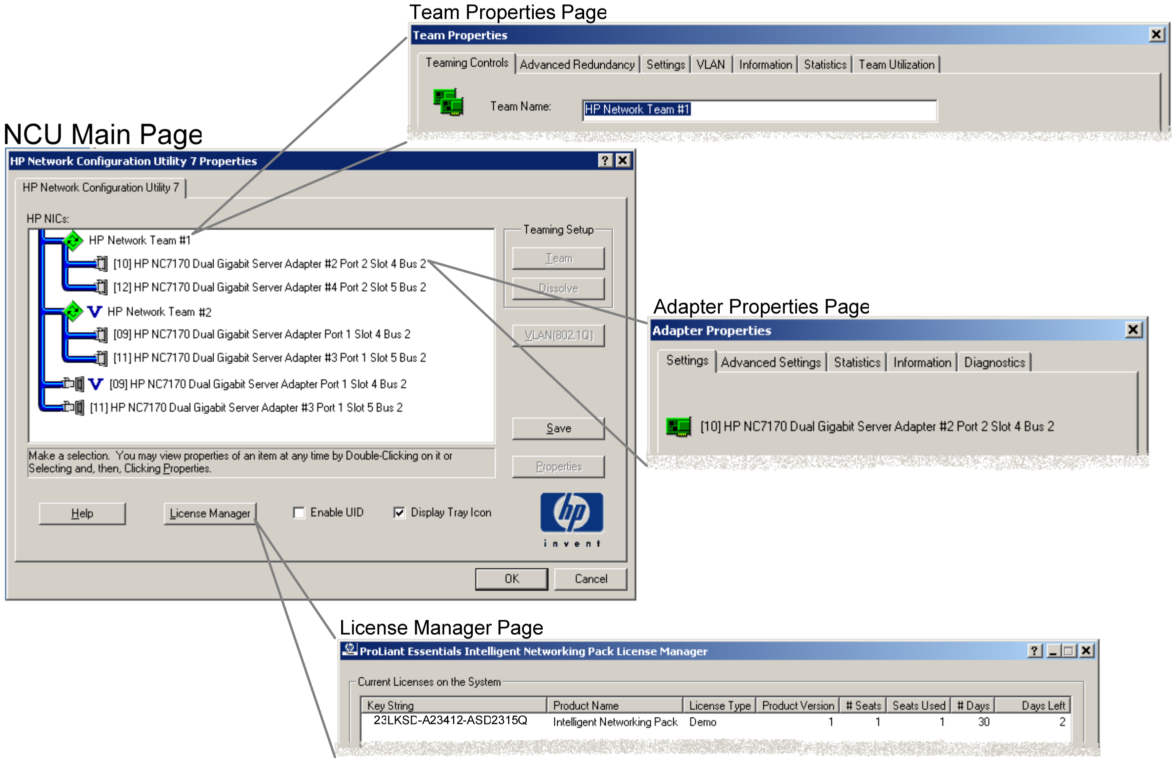
## 3-6-1 NCU main page

The NCU main page, labeled HP Network Configuration Utility 7 Properties, is used to create and dissolve teams, monitor the overall health of teams and individual network adapters, and to export team configuration settings to an XML file. The main page is also used to access the Properties pages for team settings, NIC settings, and the Licensing Manager (used for licensing the INP advanced features). Refer to Figure 3-6.

### Primary actions from the NCU main page

- Creating a team—Highlight two or more individual (non-teamed) network adapter ports and click the **Team** button.
- Adding a team member—Highlight an existing team and one or more (non-teamed) network adapter ports and click the **Add** button.
- Deleting a team—Highlight an existing team and click the **Dissolve** button.
- Dropping a team member—Highlight a member of an existing team and click the **Drop** button.
- Team settings—Highlight an existing team and click the **Properties** button on the right or simply double click the team.
- Adapter Properties—Highlight any teamed or non-teamed port and click the **Properties** button on the right or simply double click the port.
- Managing Licenses—To manage the installation of INP licenses, click the **License Manager** button.

Figure 3-6 NCU main page



Team Name: HP Network Team #1

HP Network Team #1
[10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2
[12] HP NC7170 Dual Gigabit Server Adapter #4 Port 2 Slot 5 Bus 2
V HP Network Team #2
[09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2
[11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2
V [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2
[11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2

Make a selection. You may view properties of an item at any time by Double-Clicking on it or Selecting and, then, Clicking Properties.

☐ Enable UID ☑ Display Tray Icon

[10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2

| Key String | Product Name | License Type | Product Version | # Seats | Seats Used | # Days | Days Left |
|---|---|---|---|---|---|---|---|
| 23LKSD-A23412-ASD2315Q | Intelligent Networking Pack | Demo | 1 | 1 | 1 | 30 | 2 |

## 3-6-2 Team Properties page

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

### Team Properties page tabs

- Teaming Controls—This tab is used to change the team name, select team type, select the Transmit Load Balancing algorithm, change port preference order for NFT with Preference teams, and to assign group membership for Dual Channel teams.
- Advanced Redundancy—This tab is only utilized when the server has a valid INP license. It is used to manage Active Path, Fast Path, and monitor Fast Path port cost and path cost.
- Settings—This tab is used to manually set 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, etc.)
- VLAN—This tab is used for assigning a team to one or more VLANs. It is also used to manually choose which VLANs to use for the default/native VLAN, VLAN for receive path validation (heartbeats), Active Path VLAN, and Fast Path VLAN (if PVST+ is selected for Fast path).
- Information—This tab is only used for monitoring 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—This tab is used to monitor 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, etc.
- Team Utilization—This tab is used for monitoring the effectiveness of load balancing. It provides a real-time numeric and graphical view of team-level utilization by percentage and bandwidth, as well as, individual teamed port-level usage information.

## 3-6-3 Adapter Properties page

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

### Adapter Properties page tabs

- Settings—This tab is used for setting the individual adapter's speed and duplex and for setting a non-teamed port's locally administered MAC address (LAA).
- Advanced Settings—This tab is used for managing per-adapter advanced settings such as max frame size, 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 page described in "3-6-2 Team Properties page".
- VLAN—This tab is used for managing VLAN membership for non-teamed ports. Teamed ports do not have this page since VLAN membership is managed at the team level. Refer to the VLAN tab described in "3-6-2 Team Properties page".
- Statistics—This tab is used to monitor real-time adapter statistics such as number of frames received/transmitted, bytes transmitted/received, heartbeats transmitted/received for teamed ports, etc. The Advanced button can be activated to monitor advanced statistics such as CRC errors, collisions, underruns, etc. on a per-Adapter basis.
- Information—This tab is only used for monitoring 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/slot/port information, etc.
- Diagnostics—This tab is used for running manual diagnostics on a per-adapter level. The diagnostic tests are manually selected and executed by clicking the **Run Tests** button. Running Diagnostic tests on an adapter will cause its normal network operation to be halted. Diagnostic tests are not recommended for adapters in production.

## 3-6-4 License Manager page

The License Manager page (refer to Figure 3-6) is used for adding and viewing INP licenses.

# 3-7 Getting started

## 3-7-1 Example deployment of basic teaming

1. Perform the initial installation.

   Complete the "Basic deployment steps for HP ProLiant Network Adapter Teaming" steps in section 3-5 above. Optionally, skip step "3-5-3 License".

2. Select the network adapter ports for a team.

   From the NCU main page, highlight two or more network adapter ports. Then, click the **Team** icon.

3. Set the team type **Automatic**.
   a. Click the **Properties** icon on the NCU main page 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.
   c. Note the name assigned to the team in the Team Name window (for example, HP Network Team #1).
   d. (Optional) Rename the team by changing the text in the Team Name window.
   e. Click the **OK** button on the Team Properties page when finished.

4. Apply all changes.
   a. Click the **OK** button on the NCU main page.
   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 (in other words, Network and Dial-up Connections) window from the Control Panel.
   b. Under Device Name (in the Details view), open the Properties page 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.

## 3-7-2 Example deployment of advanced teaming (INP)

1. Perform the initial installation.

   Complete the "Basic deployment steps for HP ProLiant Network Adapter Teaming" steps in section 3-5 above. Do not skip step "3-5-3 License".

2. Select the network adapter ports for a team.

   From the NCU main page, highlight two or more network adapter ports. Then click the **Team** icon on the main page.

3. Set the team type set to **Automatic**.
   a. Click the **Properties** icon on the NCU main page 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.
   c. Note the name assigned to the team in the Team Name window (for example, HP Network Team #1).
   d. (Optional) Rename the team by changing the text in the Team Name window.

4. Enable and configure advanced redundancy mechanisms.
   a. Select the **Advanced Redundancy** tab from the Team Properties page.
   b. Enable Active Path Failover by checking the box.
   c. Select **Community Address ARP** as the Echo Node Response Mechanism.
   d. Type the IP address of the designated Echo Node (for example, gateway router) in the Echo Node IP Address field.
   e. Select an unused IP address on the same subnet as the Echo Node and type it in the Community Probe IP Address field. (**Note**: The same Community Probe IP address can be used for all Active Path-enabled teams on the same subnet. HP recommends this type of configuration.)
   f. (Optional) If Spanning Tree is enabled on the switches connected to the server, enabled **Fast Path** by checking the box.
   g. (Optional) In the Mechanism Priority window, highlight **Active Path Failover** and move it above **Fast Path Failover** by using the up arrow on the right.

   h. Click the **OK** button at the bottom of the Team Properties page when finished.

5. Apply all changes.
   a. Click the **OK** button at the bottom of the NCU main page.
   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.
6. Assign an IP address to the team.
   a. Open the Network Connections (in other words, Network and Dial-up Connections) window from the Control Panel.
   b. Under Device Name (in the Details view), open the Properties page 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 2000 and Windows 2003, 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, Q57W2K.SYS, N1000NT5.SYS, N100NT5.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 (in other words, CPQTEAM.SYS). Its function is 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, etc.). NDIS is more than just a single layer since 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 2000, the Miniport driver used with the HP network adapter will be Q57W2K.SYS, N100NT5.SYS, or N1000NT5.SYS depending on the adapter in use.

For Microsoft Windows 2003, the Miniport driver used with the HP network adapter will be Q57XP32.SYS, N100325.SYS, or N1000325.SYS depending on the adapter in use.

### 4-2-2-2 Teaming intermediate driver

For Microsoft Windows 2000 and 2003, 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, visit http://h18004.www1.hp.com/support/files/networking/us/index.html.

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

One of the most important concepts to understand when implementing HP ProLiant Network Adapter Teaming is that of 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, etc.) 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.

In order for an HP network adapter team to appear as a single virtual network adapter, it is necessary for all networking devices to refer to 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, 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 other than Dual Channel, 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. With the exception of Dual Channel, the team always sends ARP replies using the same MAC address. This allows the team to be known by one MAC address to all network entities.

# 4-3 Types of HP ProLiant network adapter teams

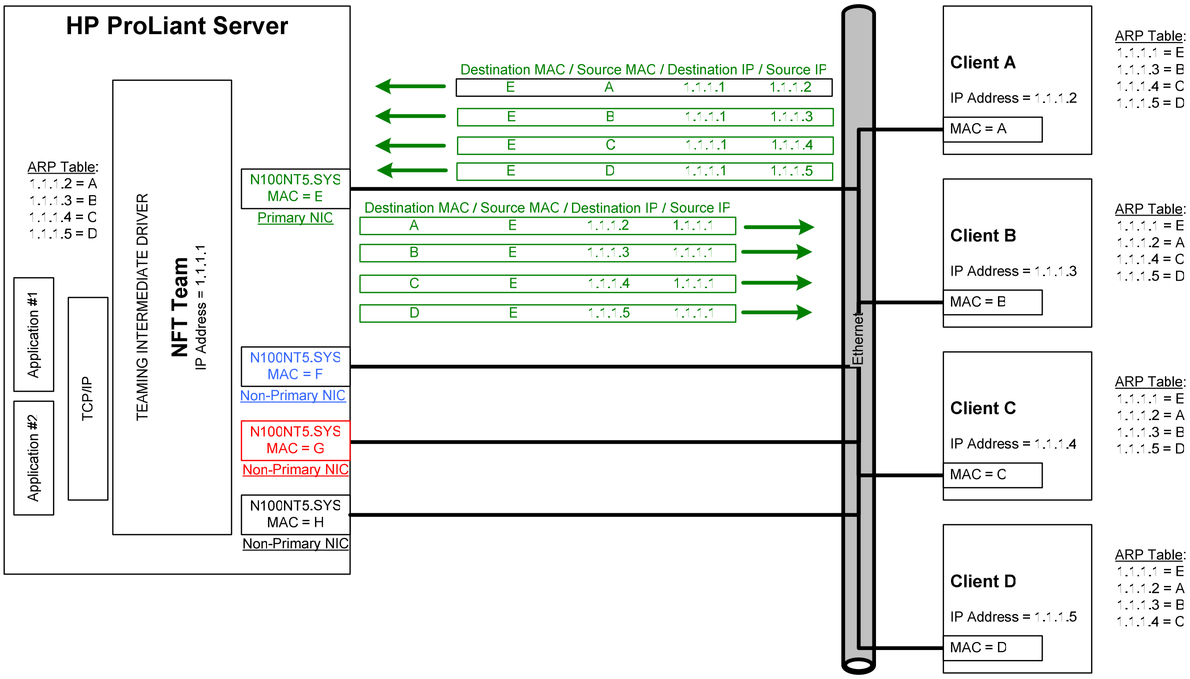
There are seven team types for HP network adapters: Network Fault Tolerance Only (NFT), Network Fault Tolerance Only with Preference, Transmit Load Balancing with Fault Tolerance (TLB), Switch-assisted Load Balancing with Fault Tolerance (SLB), Switch-assisted Dual Channel Load Balancing (Advanced Pack), 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, refer to Figure 3-4 on page 15.

## 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



| Destination MAC | Source MAC | Destination IP | Source IP |
|---|---|---|---|
| E | A | 1.1.1.1 | 1.1.1.2 |
| E | B | 1.1.1.1 | 1.1.1.3 |
| E | C | 1.1.1.1 | 1.1.1.4 |
| E | D | 1.1.1.1 | 1.1.1.5 |

| Destination MAC | Source MAC | Destination IP | Source IP |
|---|---|---|---|
| A | E | 1.1.1.2 | 1.1.1.1 |
| B | E | 1.1.1.3 | 1.1.1.1 |
| C | E | 1.1.1.4 | 1.1.1.1 |
| D | E | 1.1.1.5 | 1.1.1.1 |

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

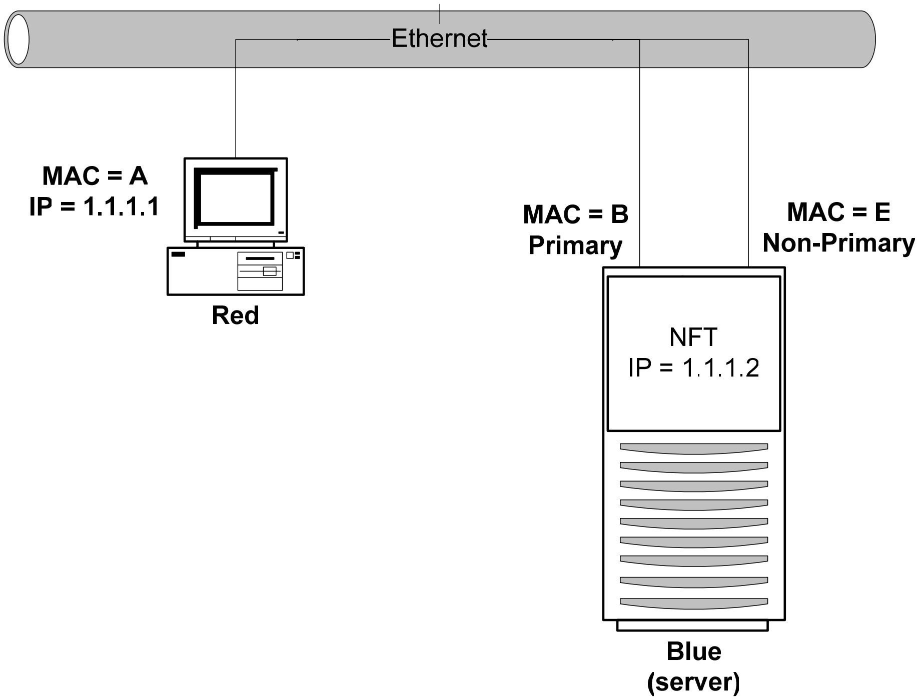
Before learning the specifics of NFT and how it communicates on the network, it is recommended that “4-2-3 HP teaming and Layer 2 versus Layer 3 addresses” be thoroughly reviewed and understood as well as “Appendix A: Overview of network addressing and communication.”

### 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 29, Blue has been modified to be a server utilizing an HP network adapter team in NFT mode with two network ports in a team (refer to 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



### 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. Once 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.

### b) 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) (refer to "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.

### c) 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.

### d) 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.

### e) 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.

**f) 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

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

**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.

**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 (refer to "Heartbeat functions" for a complete description of the different heartbeat types).

Heartbeat frames contain only Layer 2 addresses (refer to "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.

**Heartbeat functions**

There are three main functions of the heartbeat frame used for HP ProLiant Network Adapter Teaming: receive validation, transmit validation, and Switch MAC table updates. The same heartbeat 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*

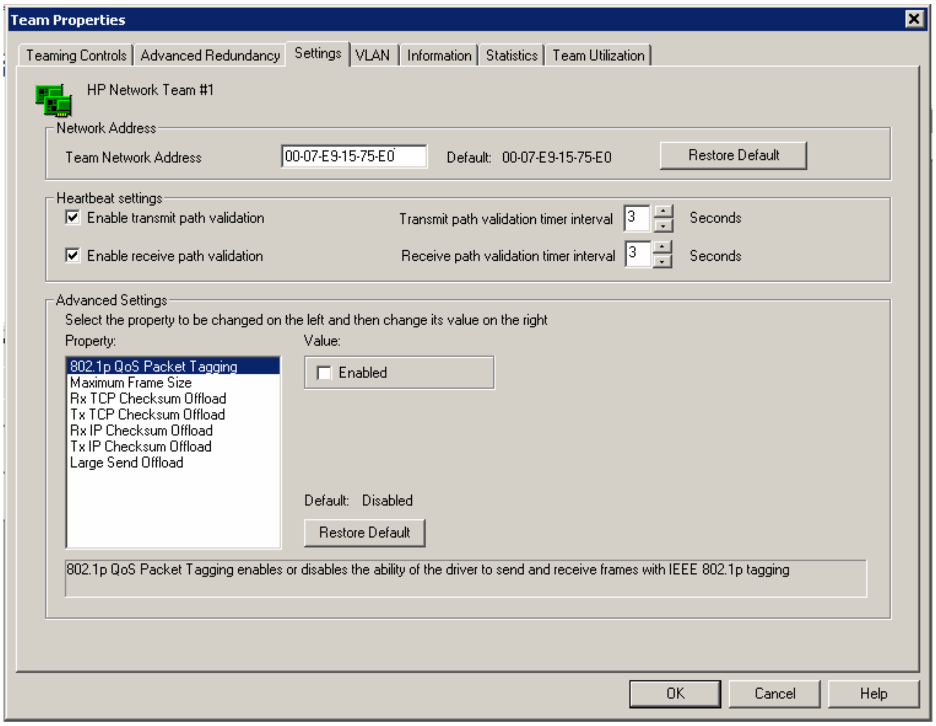
The Primary Port must ensure that the switch has the team's MAC address on the correct port, and that the switch keeps the team's MAC address in its MAC table [or Content Addressable Memory (CAM) table]**.** The Switch MAC table update heartbeat is used with most team types when a failover occurs, and 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.

**Heartbeat 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 page 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 page—Settings tam containing the heartbeat settings



HP Network Team #1

Network Address

Team Network Address: 00-07-E9-15-75-E0 Default: 00-07-E9-15-75-E0

Heartbeat settings

☑ Enable transmit path validation — Transmit path validation timer interval: 3 Seconds

☑ Enable receive path validation — Receive path validation timer interval: 3 Seconds

Advanced Settings

Select the property to be changed on the left and then change its value on the right

Property:

- 802.1p QoS Packet Tagging
- Maximum Frame Size
- Rx TCP Checksum Offload
- Tx TCP Checksum Offload
- Rx IP Checksum Offload
- Tx IP Checksum Offload
- Large Send Offload

Value:

□ Enabled

Default: Disabled

802.1p QoS Packet Tagging enables or disables the ability of the driver to send and receive frames with IEEE 802.1p tagging

*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 will place the port in a state called "Failed (TX Heartbeat)" and the team will failover 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 will 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 |

## 4-3-1-2-2 Advanced redundancy mechanisms

### Active Path

Active Path is a mechanism used by HP ProLiant Network Adapter Teaming to intelligently and proactively determine which teamed ports should or should not be used by the team based on "tested" connectivity with an external network device called an Echo Node. Each teamed port tests connectivity to the Echo Node by transmitting a frame to the Echo Node and monitoring for a response. Active Path can determine if teamed ports have become segregated onto different networks (for example,, because of link loss on an upstream switch uplink, server network adapter ports connected to the wrong switch ports, VLAN misconfiguration on a switch, etc.). Teamed port segregation is determined based on successful communication with the Echo Node on a per teamed port basis.

Active Path's goals:

- To determine if the teamed ports have become segregated into different broadcast domains (in other words, teamed ports pass or fail the connectivity test with the Echo Node)
- To only use teamed ports with validated connectivity to the Echo Node for server data communication

Active Path is considered an "active" mechanism because it actively transmits and receives a specific type of frame between teamed ports and an external network device called the Echo Node. Active Path uses the results of its connectivity tests to intelligently make decisions to meet the goals above.

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?"

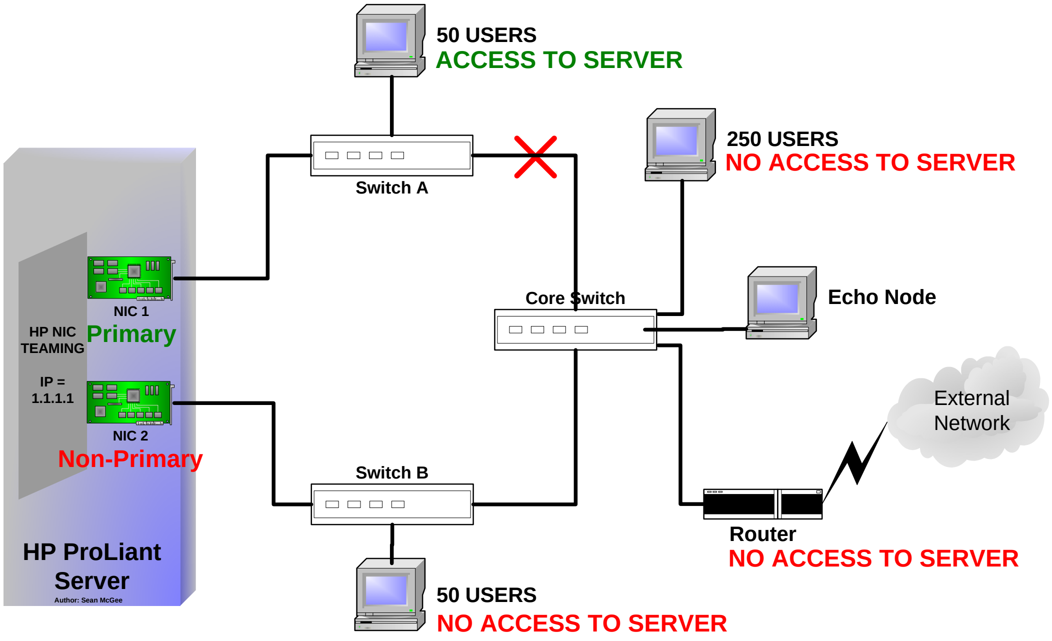
These basic mechanisms do not provide for redundancy beyond the network adapter port in the team. In other words, these basic redundancy mechanisms have no ability to detect network conditions that would prevent a particular team member port from being used by the team. If one teamed port is isolated 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 doesn't 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, refer to 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 don't have connectivity with the core segment to prevent transmitting on a teamed port that can't reach the intended target (for example, client, router, etc.).

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 can't 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 doesn't know which network segment is the best segment to choose.

**Figure 4-4** Upstream link failures cause server isolation



To solve this problem, HP Engineering developed the Active Path mechanism. The Active Path mechanism allows the teaming driver to actively monitor network connectivity on a per-team member port basis with an external network device (designated as the Echo Node). Team member ports that have connectivity with the Echo Node are considered to be "eligible for use by the team". Therefore, Active Path becomes another validation mechanism in addition to link loss, transmit path validation, and receive path validation.

**General description of Active Path operation**

The goal of Active Path is to verify that each individual member port in a team has connectivity to the designated Echo Node that is located on the "important" physical segment of the broadcast domain. The method that Active Path uses must be able to operate independently of team member roles (Primary, non-Primary). In other words, Active Path must be able to transmit Echo Node probes and receive responses on each individual member port. Additionally, Active Path must accomplish this without using the team's IP address. The team's IP address cannot be used since the Echo Node will always resolve the team's IP address to a single port in the team, which would prevent Echo Node probe responses from being sent to every port in the team.

Active Path uses a special frame as the Echo Node probe. This frame is identical to a normal ARP request frame except for one field – the Sender's IP address field. In an Echo Node probe, the sender's IP address field is intentionally left blank. This frame format accomplishes three goals: 1) the probe provides the MAC address that the Echo Node should respond to (in the sender's Hardware Address field), 2) the probe does not provide a sender's IP address so that the ARP cache on the Echo Node is not affected by the Echo Node probe, and 3) by using a modified ARP request frame, the Echo Node does not require any type of special configuration other than the assignment of an IP address. **Note**: Not all TCP/IP stacks will respond to the Directed ARP Echo Node probe.

When Active Path is enabled on a team, an Echo Node probe is transmitted from every member port in the team every X seconds (where X is equal to the setting of the Echo Node probe interval). The Active Path mechanism expects to receive a response on every member port every *Y* number of seconds (where *Y* is equal to the setting of the Echo Node probe timeout). Once a particular member port fails to receive an Echo Node probe response, the port is considered to be in a failed/disabled state. The port continues to transmit Echo Node probes every X number of seconds but will not be re-enabled for use until at least one Echo Node probe response is received on the port. The default values for Echo Node probe interval (X) and Echo Node probe timeout (Y) are both 3 seconds.

If the Echo Node is unreachable or unavailable for all ports in a team, the team considers all member ports equal and will not disable any ports based on the Active Path mechanism. Another mechanism would have to cause a member port failure. As long as at least one member port can reach the Echo Node, failure to communicate with the Echo Node for any other port will cause the port to be disabled for use by the team.

There are two Active Path mechanisms: Directed ARP and Community Address ARP.

*Active Path's Directed ARP mechanism operation*

Active Path's Directed ARP can be described as a "one-to-one" validation mechanism between each port in a team and the designated Echo Node. With Directed ARP, each and every teamed port transmits an Echo Node probe request to the designated Echo Node and expects an individual Echo Node reply in response. The Echo Node probe requests are sent at the interval configured in the teaming GUI in the Echo Node Probe Interval field. Each teamed port that transmits an Echo Node probe request will wait for the number of Echo Node probe timeout seconds (configured in the teaming GUI). If the Echo Node probe request times out without receiving a response from the Echo Node, the teaming driver will assume the teamed port does not have connectivity with the Echo Node and will place the port in Active Path failure. This prevents the port from being used as the Primary port in the team and will prevent the team from using it as a Secondary (non-Primary) in load balancing teams (for example, TLB). If the Primary port in a team loses connectivity with the Echo Node, another port in the team with confirmed connectivity with the Echo Node will assume the role of Primary (resulting in a failover).

Directed ARP's Echo Node probe

Special considerations were required when choosing a frame type to use for individually validating every teamed port's connectivity with an external device (Echo Node). First, the frame type used by Echo Node had to allow for a teamed port to predictably receive a response from the Echo Node to validate its individual connectivity. Since the team's IP address resolves to a specific MAC address that's associated with the team's Primary port, normal ICMP (in other words, PING) frames could not be used. If ICMP Requests were transmitted by each teamed port, the responses would all be transmitted to the Primary port (programmed with the team's MAC address). If the Primary port was isolated from the Echo Node, no responses would be received for any teamed ports. As a result, Active Path would be unable to determine which teamed ports did or didn't have connectivity with the Echo Node and a failure recovery would not be possible. Second, the frame type had to be supported by as many IP protocol stacks as possible, so no special configuration or modifications to the designated Echo Node were required. Third, a frame type had to be chosen that would not cause loss of communication for normal data traffic between the Echo Node and the team since, in many environments, the Echo Node could be another server, another client, a switch, or a router functioning as the team's gateway.

To comply with the above requirements, Directed ARP was designed to use a modified ARP request frame as the Echo Node probe. The format of the ARP request frame was chosen in order to prevent Echo Node probes from changing the team's ARP cache entry in the Echo Node's ARP table. By omitting the Source IP address in the ARP frame, the designated Echo Node doesn't have enough information to modify its ARP entry for the team's IP address. However, the Echo Node will respond back to the Source MAC address in the ARP frame with an ARP response. Since each teamed port inserts its own unique MAC address in the Source MAC Address field, each teamed port receives an individual response from the Echo Node. As a result, Directed ARP's Echo Node probe provides for validation to each teamed port, works with most (but not all) IP protocol stacks, and does not cause loss of communication with the team for normal data traffic.

**Table 4-3** Directed ARP Echo Node probe REQUEST frame format

| Ethernet V2 frame format | Value | 64 bytes total |
|---|---|---|
| Destination MAC address | *"Broadcast" unless specified in GUI* | 6 bytes |
| Source MAC address | *"MAC address of Teamed port"* | 6 bytes |
| 802.1Q/p tag (optional) | *"Value of VLAN ID and Priority"* | 4 bytes |
| Type | 0x0806 (ARP) | 2 bytes |
| Hardware type | 0x01 (Ethernet) | 2 byte |
| Protocol type | 0x800 (IP) | 2 byte |
| Hardware address length | 0x06 | 1 byte |
| Protocol address length | *0x04* | 1 byte |
| ARP operation | 0x001 (ARP request) | 2 bytes |
| Source MAC address | *"MAC address of Teamed port"* | 6 bytes |
| Source IP address | 0.0.0.0 | 4 bytes |
| Destination MAC address | 00:00:00:00:00:00 | 6 bytes |
| Destination IP address | *"Echo Node IP Address from GUI"* | 4 bytes |
| Checksum | *"varies"* | 4 bytes |

**Table 4-4** Directed ARP Echo Node probe REPLY frame format

| Ethernet V2 frame format | Value | 64 bytes total |
|---|---|---|
| Destination MAC address | *"MAC address of Teamed port"* | 6 bytes |
| Source MAC address | *"Echo Node's MAC Address"* | 6 bytes |
| 802.1Q/p tag (optional) | *"Value of VLAN ID and Priority"* | 4 bytes |
| Type | 0x0806 (ARP) | 2 bytes |
| Hardware type | 0x01 (Ethernet) | 2 byte |
| Protocol type | 0x800 (IP) | 2 byte |
| Hardware address length | 0x06 | 1 byte |
| Protocol address length | *0x04* | 1 byte |
| ARP operation | 0x002 (ARP response) | 2 bytes |
| Source MAC address | *"Echo Node's MAC Address"* | 6 bytes |
| Source IP address | *"Echo Node's IP Address"* | 4 bytes |
| Destination MAC address | *"MAC address of Teamed port"* | 6 bytes |
| Destination IP address | 0.0.0.0 | 4 bytes |
| Checksum | *"varies"* | 4 bytes |

Directed ARP's Designated Echo Node

The Echo Node is a network device that is "designated" by the implementer for use by a team (with Active Path enabled) to validate network connectivity on a per teamed port basis. No special software or configuration should be required for a device to function as the Echo Node. The designated Echo Node simply needs to have an IP address and be physically located on the same broadcast domain as the teamed ports.

**ATTENTION:** Only certain devices can be designated as the Echo Node for Directed ARP because of the frame format used for Directed ARP's Echo Node probe request. If a device will not function as the Echo Node when using Directed ARP, consider utilizing the Community Address ARP mechanism or using a different device as the designated Echo Node.

If the Echo Node is disconnected from the network or fails to respond to Echo Node probes for any reason, all teamed ports using the device as an Echo Node will not be able to receive Echo Node probe replies. When this happens, the teaming driver WILL NOT fail all teamed ports (in other words, a failed Echo Node will not cause loss of connectivity for the server). From the team's perspective, all teamed ports have become equal with regard to Active Path – none can reach the Echo Node. As a result, the teaming driver will indicate that the teamed ports are all in the Active Path degraded state (degraded indicates that the port is in use by the team but isn't functioning at its optimum level) and are still eligible to be used by the team. Any teamed port that was previously in the Active Path failed state due to lack of communication with the failed Echo Node will also be eligible for use and will be reinstated to the Active Path degraded state, equal to all other teamed ports in that team. Operationally, the team will continue to work as if Active Path was disabled since the Echo Node is not available to any teamed port and the Active Path mechanism can no longer be used to select teamed ports with better network connectivity. However, Active Path will continue to send Echo Node probe requests in case the designated Echo Node is restored. Future implementations of this mechanism may have an option for a backup Echo Node.

While the same Echo Node can be used for every server in the broadcast domain, it is not required. If an implementer chose to, they could designate a different Echo Node for every server using Active Path.

HP recommends choosing a designated Echo Node that is physically located on the most important physical segment of the broadcast domain. Since Active Path validates end to end network connectivity from the teamed port to the Echo Node, the farther the Echo Node is physically from the server the more network that is validated by Active Path. HP also recommends choosing an Echo Node that has reliable connectivity to the network. Whenever the Echo Node is disconnected from the network, Active Path is effectively disabled on any server configured to monitor the disconnected Echo Node. As stated above, a disconnected Echo Node will not cause the team to disable all teamed ports.

#### *Active Path's Community Address ARP mechanism operation*

Active Path's Community Address ARP can be described as a "one-to-many" validation mechanism between the designated Echo Node and all members in a particular team. This is accomplished by the Echo Node transmitting a single Echo Node probe reply that is received by all teamed ports configured to monitor the particular Echo Node. In contrast to Directed ARP, each individual teamed port is not necessarily required to transmit its own Echo Node request, yet all teamed ports individually receive an Echo Node probe reply.

With Community Address ARP, the team's Primary port transmits an Echo Node probe request to the designated Echo Node. The Echo Node then transmits an Echo Node reply that is received by every teamed port in that team. The Echo Node probe requests are sent at the interval configured in the teaming GUI in the Echo Node Probe Interval field. After the team's Primary port transmits an Echo Node probe request to the Echo Node, each teamed port will wait for the number of Echo Node probe timeout seconds configured in the teaming GUI. If for a particular teamed port, the Echo Node probe request times out without receiving a response from the Echo Node, the teaming driver will assume the teamed port does not have connectivity with the Echo Node and change the port state to Active Path Failed. This prevents the port from being used as the Primary port in the team and will prevent the team from using it as a Secondary (non-Primary) in load balancing teams (for example, TLB). In addition, when the team detects connectivity problems with the Echo Node for any teamed port, the team begins transmitting Echo Node requests on all ports in the team. If the Primary port in a team loses connectivity with the Echo Node, another port in the team with confirmed connectivity with the Echo Node will assume the role of Primary (resulting in a failover).

For Community Address ARP, the Echo Node probe replies from the Echo Node are received by every port in the team because of the special IP and MAC addresses configured in the GUI. There are two additional configuration fields – Community Probe IP Address and Community Probe MAC Address required for Community Address ARP. These addresses are ONLY used for Echo Node purposes and are not used for data communication for the team. These designated (by the server administrator) Community Address IP and MAC addresses are used as the Source IP and MAC addresses in the Echo Node probe request packets transmitted by teamed ports. The teamed ports transmit the Echo Node probe request with their individual MAC address as the Source MAC address in the Ethernet header, but use the Community Probe MAC address in the Source MAC Address ARP field. When the Echo Node receives a Community Address ARP Echo Node probe request, the Echo Node updates its ARP table with the Community Address IP and MAC addresses contained in the Source MAC and IP ARP fields, not the Source MAC address from the Ethernet header. The Echo Node then transmits a single Echo Node probe reply (in other words, ARP reply). This Echo Node reply is sent using the information from the Echo Node's ARP table (in other words, the Community Probe MAC address). All teamed ports have been configured in hardware to receive on at least two unicast MAC addresses: the team's MAC address and the configured Community Probe MAC address. Also, the teamed ports never transmit using the configured Community Probe MAC address. As a result, none of the switches in the LAN add the Community Probe MAC address to their MAC-to-port table (in other words, MAC table or CAM table). This means that when the Echo Node sends an Echo Node Reply destined to the MAC address defined as the Community Probe MAC address, all switches in the network give a copy of the frame to every port. The result is that all teamed ports configured to monitor connectivity with the Echo Node that transmitted the Echo Node probe reply will receive the frame and validate connectivity – either within a single team on a single server or across multiple teams on multiple servers if all are configured with the same Community Probe IP/MAC and Echo Node. Network devices that aren't configured to receive on the particular Community Probe MAC address will ignore the Echo Node probe reply frame (dropped in hardware).

As discussed earlier, Community Address ARP differs from Directed ARP in that only the Primary teamed port transmits the Echo Node probe requests, yet all teamed ports (with connectivity to the Echo Node) will receive an individual response from the Echo Node. The only scenario where Echo Node probes are transmitted on non-Primary ports with Community Address ARP is when any teamed port is in the Active Path failed state. If any teamed port is in an Active Path failed state, all teamed ports transmit Echo Node probe requests. This behavior is required in order to quickly determine the best teamed port(s) to use in case of a teamed port failure. For example, if the Primary port is isolated from the Echo Node, then the Echo Node probe requests will not be received by the Echo Node and Echo Node probe replies will not be transmitted (by the Echo Node) for reception by any teamed ports. In order to recover from this situation, all teamed ports will transmit Echo Node probe requests in an attempt to find at least one teamed port with connectivity to the Echo Node.

Community Address ARP's Echo Node probe

Much like Directed ARP's Echo Node probe, Community Address ARP also had to take into account the same considerations:

1. Validate individual connectivity for a teamed port.

2. There is no special configuration on the designated Echo Node.

3. The Echo Node probe frame can't cause loss of connectivity for the team.

Community Address ARP was designed to work with IP stacks that fail to respond to the Directed ARP probe packets. Community Address ARP allows practically any network device to function as the Echo Node. Although Directed ARP works with most IP stacks, certain routers fail to respond to the Directed ARP probes. The most likely cause of the failure was the Echo Node not creating an ARP entry because the Source IP in the ARP request is intentionally blank. In several IP stacks, this prevented the device designated as the Echo Node from responding to the Echo Node probe used by Directed ARP.

Due to the above considerations, Community Address ARP was designed to use a standard ARP request frame as the Echo Node probe, as opposed to Directed ARP's modified ARP request frame. Also, Community Address ARP utilizes a special "Echo Node Only" MAC and IP address called the Community Probe MAC address and Community Probe IP address (configured in the GUI). The Echo Node Only addresses are used to prevent Echo Node probes from changing the team's ARP entry in the Echo Node's ARP table and to allow a single response from the Echo Node to be received by multiple teamed ports. Each teamed port that transmits an Echo Node probe using the Community Address ARP mechanism inserts its individual MAC address in the Echo Node probe request's (ARP request) Ethernet header. However, every teamed port inserts the Community Probe IP address and Community Probe MAC address in the Source IP and Source MAC ARP header fields, respectively. Since the Echo Node receives a standard ARP request frame with a Source IP and Source MAC, any device with an IP address can be designated as the Echo Node (usually with no configuration required on the Echo Node end). As a result, Community Address ARP's Echo Node probe provides for validation to each individual teamed port, works with any IP protocol stack, and does not cause loss of communication with the team for normal data traffic.

**Table 4-5** Community Address ARP Echo Node probe REQUEST frame format

| Ethernet V2 frame format | Value | 128 bytes total |
|---|---|---|

Table 4-5 Community Address ARP Echo Node probe REQUEST frame format

| Ethernet V2 frame format | Value | 128 bytes total |
|---|---|---|
| Destination MAC address | *"Broadcast" unless specified in GUI* | 6 bytes |
| Source MAC address | *"MAC address of Teamed port"* | 6 bytes |
| 802.1Q/p tag (optional) | *"Value of VLAN ID and Priority"* | 4 bytes |
| Type | 0x0806 (ARP) | 2 bytes |
| Hardware type | 0x01 (Ethernet) | 2 byte |
| Protocol type | 0x800 (IP) | 2 byte |
| Hardware address length | 0x06 | 1 byte |

Table 4-6 Community Address ARP Echo Node probe REPLY frame format

| Ethernet V2 frame format | Value | 64 bytes total |
|---|---|---|
| Destination MAC address | *"Community Probe MAC from GUI"* | 6 bytes |
| Source MAC address | *"Echo Node's MAC Address"* | 6 bytes |
| 802.1Q/p tag (optional) | *"Value of VLAN ID and Priority"* | 4 bytes |
| Type | 0x0806 (ARP) | 2 bytes |
| Hardware type | 0x01 (Ethernet) | 2 byte |
| Protocol type | 0x800 (IP) | 2 byte |
| Hardware address length | 0x06 | 1 byte |
| Protocol address length | *0x04* | 1 byte |
| ARP operation | 0x002 (ARP response) | 2 bytes |
| Source MAC address | *"Echo Node's MAC Address"* | 6 bytes |
| Source IP address | *"Echo Node IP Address"* | 4 bytes |
| Destination MAC address | *"Community Probe MAC from GUI"* | 6 bytes |
| Destination IP address | *"Community Probe IP from GUI"* | 4 bytes |
| Checksum | *"varies"* | 4 bytes |

### Community Address ARP's Designated Echo Node

An Echo Node is any network device with an IP address that is "designated" by the implementer for use by a team (with Active Path's Community Address ARP mechanism enabled) to validate network connectivity on a per teamed port basis. No special software or configuration should be required for any device to function as the Echo Node. The designated Echo Node simply needs to have an IP address and be physically located on the same broadcast domain as the teamed ports.

If the Echo Node is disconnected from the network or fails to respond to Echo Node probes for any reason, all teamed ports using the device as an Echo Node will not be able to receive Echo Node probe replies. When this happens, the teaming driver WILL NOT fail all teamed ports (in other words, a failed Echo Node will not cause loss of connectivity for the server). From the team's perspective, all teamed ports have become equal in regards to Active Path – none can reach the Echo Node. As a result, the teaming driver will indicate that the teamed ports are all in the Active Path degraded state (degraded indicates that the port is in use by the team but isn't functioning at its optimum level) and are still eligible to be used by the team. Any teamed port that was previously in the Active Path Failed state due to lack of communication with the failed Echo Node will also be eligible for use and will be reinstated to the Active Path degraded state, equal to all other teamed ports in that team. Operationally, the team will continue to work as if Active Path was disabled on the team since the Echo Node is not available to any teamed port and the Active Path mechanism can no longer be used to select teamed ports with better network connectivity. However, Active Path will continue to send Echo Node probe requests in case the designated Echo Node is restored. Future implementations of this mechanism may have an option for a backup Echo Node.

While the same Echo Node can be used for every server in the broadcast domain, it is not required. It is possible to designate a different Echo Node for every Active Path team on every server on the network. With Community Address ARP, however, the most efficient implementation is to use a single Echo Node for all servers in the same broadcast domain. Because the Echo Node probe replies from the Echo Node are received by all teamed ports configured with the same Community Probe IP/MAC address and same Echo Node, an Echo Node reply generated from one server's request can be used to validate teamed ports for another server's teamed ports.

HP recommends choosing a designated Echo Node that is physically located on the most important physical segment of the broadcast domain. Since Active Path validates end-to-end network connectivity from the teamed port to the Echo Node, the farther the Echo Node is physically from the server the more network that is validated by Active Path. HP also recommends choosing an Echo Node that has reliable connectivity to the network. Whenever the Echo Node is disconnected from the network, Active Path is effectively disabled on any server configured to monitor the disconnected Echo Node. As mentioned above, a disconnected Echo Node will not cause the team to disable all teamed ports.

#### Active Path configuration

To use Active Path, an SA must enable it on a per-team basis from the Advanced Redundancy tab located on the HP ProLiant Network Adapter Teaming configuration GUI. Next, the SA should select which Active Path mechanism to use – Directed ARP or Community Address ARP from the Echo Node Response Mechanism drop-down box. "

For both Directed ARP and Community Address ARP, an IP address for an appropriate Echo Node must be provided in the Echo Node IP Address field. Optionally, a user can specify the MAC address of the Echo Node. If the Echo Node MAC address is configured, Echo Node probes will be transmitted using this MAC address as the destination MAC address in the Ethernet header. This may be desired by the implementer in order to prevent the Echo Node probes from being transmitted as broadcast frames. The downside is that this field must be manually updated should the MAC address of the Echo Node ever change.

An implementer may also manually adjust the timers for the Echo Node probe interval and Echo Node probe timeout. Echo Node probe interval determines how often an Echo Node probe is transmitted to the Echo Node. The default is 3 seconds. Echo Node timeout determines the number of seconds each team member will wait for the Echo Node probe reply before assuming the team member should be considered Active Path Failed. With the default values, six seconds is the longest it should take for a team to detect a fault between a teamed port and the Echo Node.

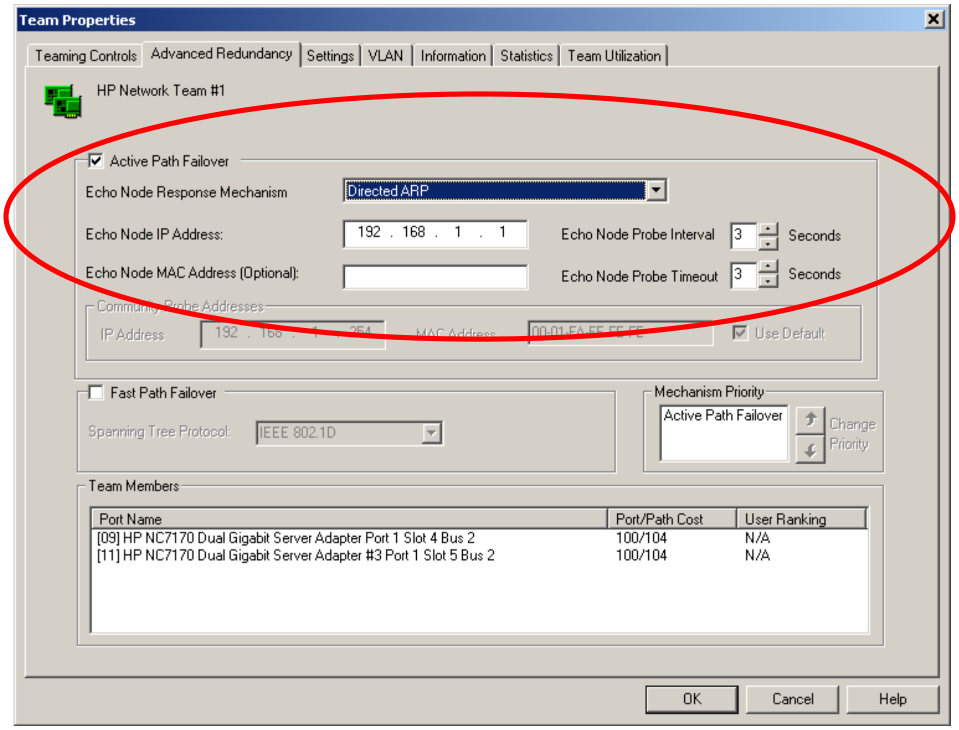
##### *Directed ARP configuration*

To enable Active Path's Directed ARP mechanism, select **Directed ARP** from the Echo Node Response Mechanism drop-down box and provide the IP address of the designated Echo Node.

Other than the configuration requirements described in the previous section, no other Directed ARP specific configuration is required.

Directed ARP guidelines—Not all network devices (for example, Cisco routers) will function as Echo Nodes. If this situation is encountered, HP recommends using Community Address ARP instead of Directed ARP.

Figure 4-5 Active Path failover configuration: Directed ARP



HP Network Team #1

☑ Active Path Failover

Echo Node Response Mechanism: Directed ARP

Echo Node IP Address: 192 . 168 . 1 . 1

Echo Node MAC Address (Optional):

Echo Node Probe Interval: 3 Seconds

Echo Node Probe Timeout: 3 Seconds

Community Probe Addresses

IP Address: 192 . 168 . 1 . 254

MAC Address: 00-01-FA-FF-FF-FE

☑ Use Default

□ Fast Path Failover

Spanning Tree Protocol: IEEE 802.1D

Mechanism Priority: Active Path Failover

Change Priority

Team Members

| Port Name | Port/Path Cost | User Ranking |
|---|---|---|
| [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2 | 100/104 | N/A |
| [11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2 | 100/104 | N/A |

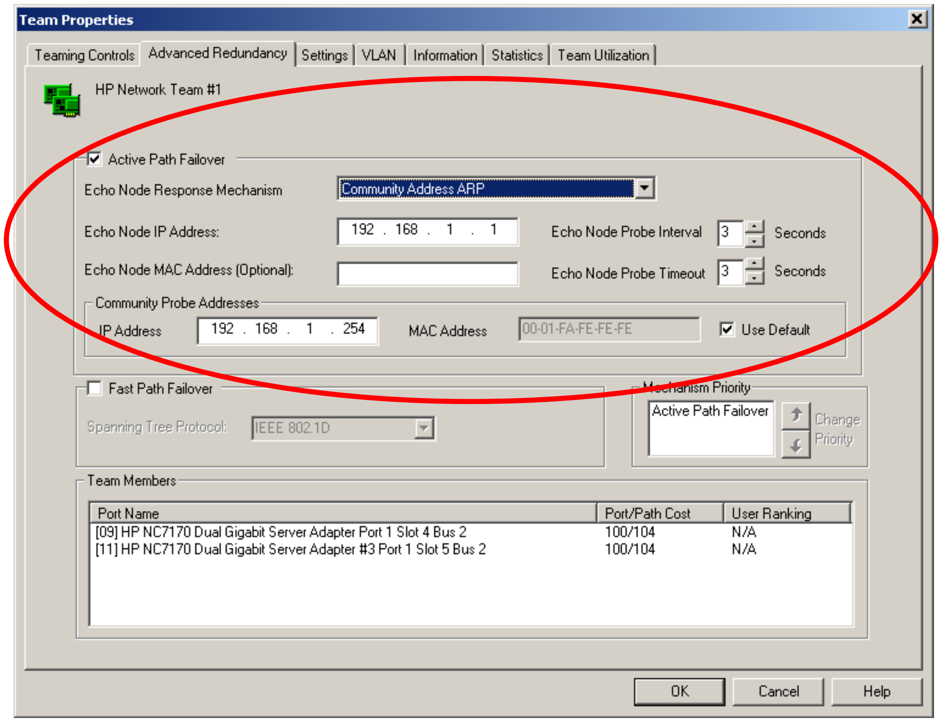
*Community Address ARP configuration*

To enable Active Path's Community Address ARP mechanism, select **Community Address ARP** from the Echo Node Response Mechanism drop-down box. In addition to supplying the Echo Node IP address, it is necessary to supply a Community Probe IP Address.

Community Address ARP guidelines

- Use the same Echo Node, Community Probe IP, and Community Probe MAC address for as many servers as possible on the same subnet. This allows a single Echo Node reply from the Echo Node to validate connectivity for all teamed ports on the network.
- Ensure the Community Probe IP address and Community Probe MAC address are not used by any other device in the network (other than another team for Community Address ARP).
- Ensure the Community Probe IP address and Community Probe MAC address are not used as the team's IP address and/or MAC address.
- If DHCP is in use, the IP address for the Community Probe IP address may need to be excluded from the DHCP scope on the DHCP server.

**Figure 4-6** Active Path failover configuration: Community Address ARP



HP Network Team #1

☑ Active Path Failover

Echo Node Response Mechanism: Community Address ARP

Echo Node IP Address: 192 . 168 . 1 . 1

Echo Node Probe Interval: 3 Seconds

Echo Node MAC Address (Optional):

Echo Node Probe Timeout: 3 Seconds

Community Probe Addresses

IP Address: 192 . 168 . 1 . 254

MAC Address: 00-01-FA-FE-FE-FE

☑ Use Default

☐ Fast Path Failover

Spanning Tree Protocol: IEEE 802.1D

Mechanism Priority

Active Path Failover

Team Members

| Port Name | Port/Path Cost | User Ranking |
|---|---|---|
| [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2 | 100/104 | N/A |
| [11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2 | 100/104 | N/A |

## Fast Path

Fast Path is a mechanism used by HP ProLiant Network Adapter Teaming to intelligently and proactively determine the best teamed ports to use to provide optimum network connectivity for the server. As with Active Path, Fast Path can determine if teamed ports have become segregated onto different broadcast domains (because of link loss on an upstream switch uplink, server network adapter ports that are connected to the wrong switch ports, VLAN misconfiguration on a switch, etc.) Also, Fast Path is used by HP ProLiant Network Adapter Teaming to determine which teamed port is the optimal port to use as the team's Primary port among one or more teamed ports with validated connectivity to the preferred Spanning Tree root switch.

Summary of Fast Path's goals

- To determine if the teamed ports have become segregated onto different broadcast domains (in other words, teamed ports see different Spanning Tree root switches).
- To only use teamed ports with connectivity to the preferred Spanning Tree root switch for server data communication.

To choose a single teamed port to be the team's Primary port (used for all receive traffic in NFT and TLB modes) that has access to the highest bandwidth links or least amount of switch hops from the team to the preferred Spanning Tree root switch.

Unlike Active Path, Fast Path is considered a "passive" mechanism because it does not actively transmit any frames; it only passively receives a specific type of frame, Spanning Tree Bridge Protocol Data Unit (BPDU), from the attached network(s) and uses the information to intelligently make decisions to meet its goals.

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—"Can I successfully transmit a frame onto the network?"
- Receive path validation—"Can I successfully receive any kind of frame from the network?"

These basic redundancy mechanisms do not provide for redundancy beyond the network adapter port in the team. In other words, these basic redundancy mechanisms have no intelligence 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 doesn't 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.

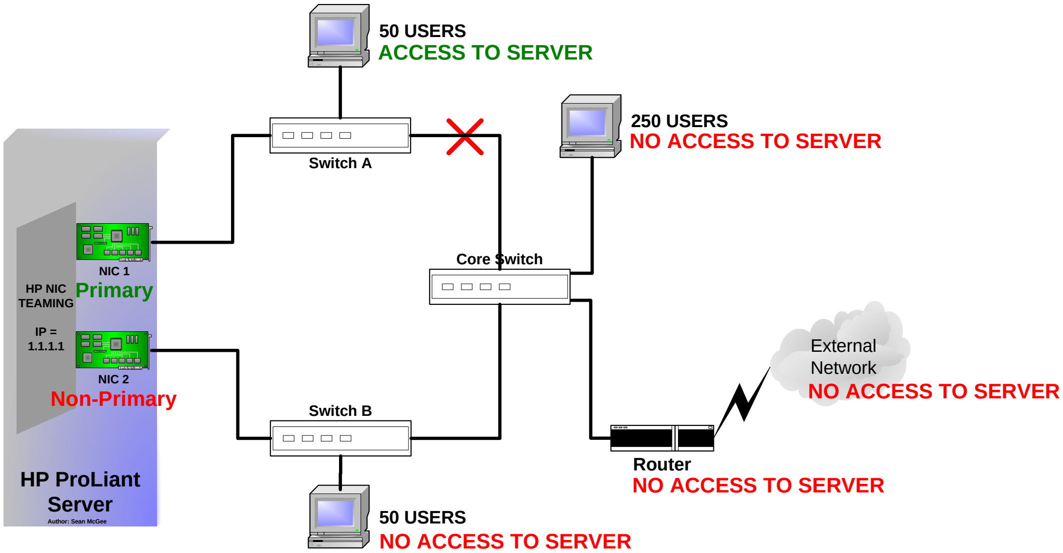
In addition to the basic redundancy mechanisms not being intelligent enough, use of Active Path alone doesn't always allow for optimal use of teamed resources. Active Path lacks the granularity to choose the BEST teamed port. Active Path determines if teamed ports are good or bad (in other words, binary result). Active Path can't determine good versus best. Fast Path, on the other hand, provides the granularity that Active Path lacks. Fast Path can make the same decisions that Active Path can make (in other words, good versus bad), however, Fast Path goes one step further and determines good versus best.

For example, refer to the diagram of an NFT team in Figure 4-7. 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 don't have connectivity with the core segment to prevent transmitting on a teamed port that can't reach the intended target (for example, client, router, etc.).

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 can't 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 doesn't know which network segment is the best segment to choose.

Figure 4-7 FP—Upstream link failures cause server isolation

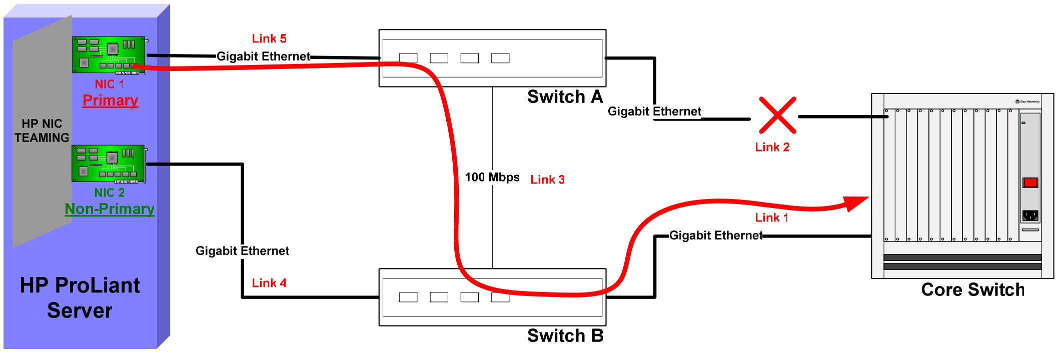


In this case, the Fast Path mechanism, designed and developed by HP Engineering is an alternative solution. The Fast Path mechanism allows the teaming driver to passively monitor network connectivity on a per-teamed port basis with the Core Switch (designated as the Spanning Tree root switch). Teamed ports that receive Spanning Tree BPDU frames from a root switch with the highest Spanning Tree bridge priority are considered "eligible for use by the team. Any teamed port that receives a BPDU from a different root switch with a lower Spanning Tree bridge priority is considered "ineligible for use by the team" since the teamed port has been segregated from the main network segment. Therefore, Fast Path becomes another validation mechanism in addition to link loss, transmit path validation, receive path validation, and Active Path.

There are also situations when upstream link failures can cause inefficient use of a server's teamed ports. For the first example, refer to Figure 4-8. In this example, Switch A has lost direct link with the Core Switch (in other words, Spanning Tree root switch). Spanning Tree eventually restores connectivity by unblocking the link between Switch A and Switch B. However, the link between Switch A and Switch B is only a 100 Mbps link instead of 1000 Mbps. As a result, any device connected to Switch A now only has access to 100 Mb worth of bandwidth to the core network. If the server's teamed port (in other words, NIC 1) connected to switch A is used as the team's Primary port, the maximum available receive bandwidth for the server is 100 Mb. However, if the same server used the teamed port (in other words, NIC 2) connected to Switch B as the team's Primary port, the maximum available receive bandwidth would be increased to 1000 Mb.

Redundancy mechanisms that only test connectivity (for example, heartbeats, Active Path) will not detect this problem since connectivity with the core network still exists. A more granular mechanism needs to be implemented that can detect connectivity problems (bad versus good), and detect slower versus faster paths (good versus best).

Figure 4-8 FP—Upstream link failures cause server receive bandwidth bottleneck

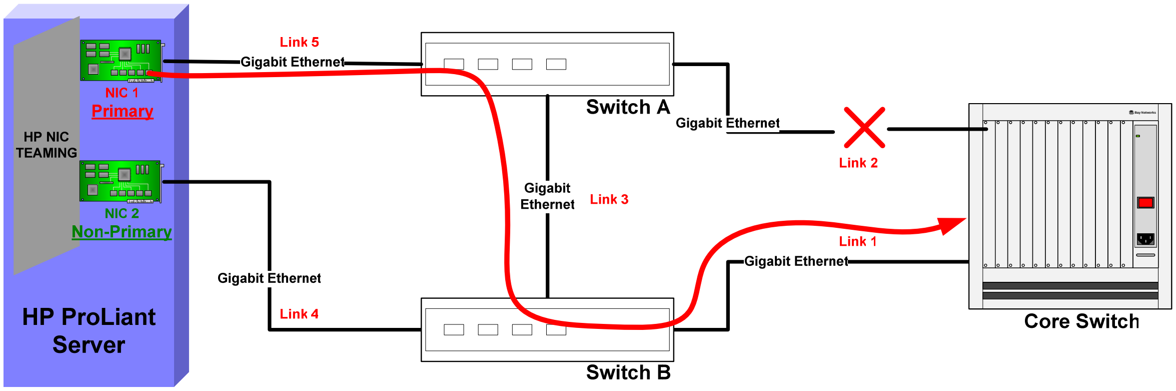


The Fast Path mechanism is a solution to this problem. In addition to Fast Path detecting full connectivity loss, as described in the first example, Fast Path can also detect when a server's teamed ports have different maximum bandwidth paths to the Core Switch (root switch). By examining the path cost information contained in Spanning Tree BPDU frames reported by directly attached switches, a team can detect which switch is the best one to use for the team's Primary port. The result is that Fast Path provides the intelligence for a team to detect, and proactively react to, network events to maximize the available bandwidth for server data traffic.

The second example of upstream link failures causing inefficient use of a server's teamed ports is found in Figure 4-9. In this example, Switch A has lost direct link with the Core Switch (in other words, Spanning Tree root switch). Spanning Tree eventually restores connectivity by unblocking the link between Switch A and Switch B. The link between switch A and Switch B is also 1000 Mbps link (like the failed link between Switch A and the Core Switch). From a bandwidth perspective, both Switch A and Switch B still have 1000 Mb access to the core switch. However, Switch A's connectivity to the Core Switch is slightly inferior to Switch B's connectivity because of hop count (in other words, latency). Since Switch A's traffic must be transmitted through Switch B, native Switch A traffic's hop count and latency to the core segment of the network is technically worse than native Switch B traffic. As a result, the most efficient use of the server's teamed ports is to use the teamed port (in other words, NIC 2) connected to Switch B as the team's Primary port instead of the teamed port (in other words, NIC 1) connected to Switch A.

Redundancy mechanisms that only test connectivity (for example, heartbeats, Active Path) will not detect this problem since connectivity with the core network still exists. A more granular mechanism needs to be implemented that can not only detect connectivity problems (bad versus good), but can also detect lower versus higher hop count on a per-teamed port "path-to-core segment" basis (good versus best).

Figure 4-9 FP—Upstream link failures add extra hop for server receive traffic



The solution to this problem is the Fast Path mechanism designed and developed by HP Engineering. In addition to Fast Path detecting full connectivity loss, as described in the first example, Fast Path can also detect when a server's teamed ports have different hop count paths to the Core Switch (root switch). By listening to the path cost information contained in Spanning Tree BPDU frames reported by directly attached switches, a team can detect which switch is the best one to use for the team's Primary port. The result is that Fast Path provides the intelligence for a team to detect, and proactively react to, network events to minimize network latency for server data traffic.

### General description of Fast Path operation

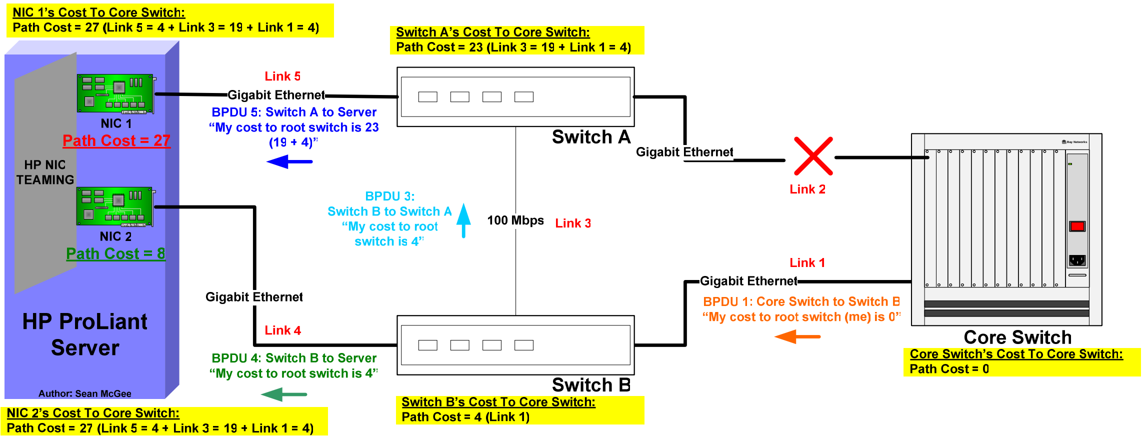
**Note**: This section assumes the reader has a technical understanding of the inner-workings of Spanning Tree – both IEEE 802.1D and Cisco's Per VLAN Spanning Tree Plus (PVST+). If additional information about Spanning Tree is required, refer to external sources (for example, "Interconnections: Bridges, Routers, Switches, and Internetworking Protocols" by Radia Perlman).

Fast Path's operation can be simply described as the intelligence for a server to passively make teamed port usage decisions as if the server was a switch (without interacting with or affecting the behavior of Spanning Tree on the LAN in any way). This is accomplished by the teaming driver "passively" listening to Spanning Tree BPDUs being transmitted by the switches in the local network. BPDUs exchanged between switches enable them to effectively run the Spanning Tree algorithm to detect loops and bypass broken links within the network topology. Each of these Spanning Tree configuration BPDUs carry such information as the bridge ID (switch ID) of the designated root switch, the cost to reach the root switch, the bridge ID of the switch that transmitted the BPDU, the port number on the switch that transmitted the BPDU, etc. When a switch transmits these frames on a switch port that is connected to a server's teamed port and Fast Path has been enabled on the team, these BPDUs (and their information about the network topology) can be passively received and utilized by the teaming driver.

In the example in Figure 4-10, an HP server with teamed ports is connected to two switches, Switch A and Switch B. The network administrator originally had both Switch A and Switch B connected directly to the Core Switch via Gigabit (1000 Mbps) links. This provided both Switch A and Switch B with equal access to the Core Switch and the rest of the network. As a backup, the network administrator directly connected Switch A and Switch B together via Link 3. Unfortunately, Switch A's uplink to the core switch was accidentally disconnected. This caused Switch A's effective bandwidth to the rest of the network to be reduced to 100 Mbps since all of Switch A's traffic must traverse Link 3 (100 Mbps).

Based on the network configuration described in the previous paragraph, the teaming driver will receive two BPDUs, one from Switch A (on NIC 1) and one from Switch B (on NIC 2). The BPDU transmitted by Switch A to NIC 1 indicates a path cost to the core switch of 23 (cost of 19 for Link 3 plus a cost of 4 for Link 1), while the BPDU transmitted by Switch B to NIC 2 indicates a path cost to the core switch of only 4 (cost of 4 for Link 1). HP teaming will determine that Switch B provides the best path to the core network and will use NIC 2 as the team's Primary port.

Figure 4-10 Fast Path operation overview


NIC 1's Cost To Core Switch:
Path Cost = 27 (Link 5 = 4 + Link 3 = 19 + Link 1 = 4)
Switch A's Cost To Core Switch:
Path Cost = 23 (Link 3 = 19 + Link 1 = 4)
Link 5
Gigabit Ethernet
BPDU 5: Switch A to Server
"My cost to root switch is 23 (19 + 4)"
Switch A
Gigabit Ethernet
Link 2
NIC 1
Path Cost = 27
HP NIC TEAMING
NIC 2
Path Cost = 8
BPDU 3:
Switch B to Switch A
"My cost to root switch is 4"
100 Mbps
Link 3
Link 1
Gigabit Ethernet
BPDU 1: Core Switch to Switch B
"My cost to root switch (me) is 0"
Gigabit Ethernet
Link 4
HP ProLiant Server
BPDU 4: Switch B to Server
"My cost to root switch is 4"
Switch B
Core Switch
Core Switch's Cost To Core Switch:
Path Cost = 0
Switch B's Cost To Core Switch:
Path Cost = 4 (Link 1)
Author: Sean McGee
NIC 2's Cost To Core Switch:
Path Cost = 27 (Link 5 = 4 + Link 3 = 19 + Link 1 = 4)

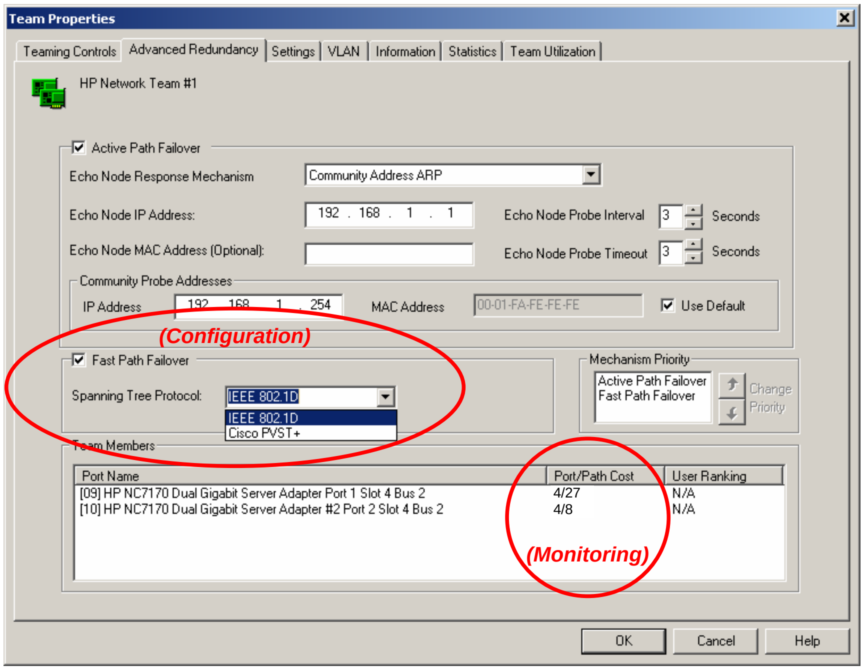
## Fast Path configuration

*Fast Path configuration and monitoring*

Fast Path is enabled from the Advanced Redundancy tab in NCU by checking the **Fast Path Failover** checkbox (refer to Figure 4-11).

Fast Path can also be monitored from the Advanced Redundancy tab. Under the Team Members section, the Port/Path Cost column provides real-time monitoring information. The Port Cost value (in other words, 4 in Figure 4-11) is the Spanning Tree cost associated with the current speed of the teamed port. Fast Path uses IEEE 802.1D guidelines for assigning costs based on the operating speed of teamed ports. For example, 4 would indicate the individual teamed port is running at 1000 Mbps, 19 would indicate 100 Mbps, and 100 would indicate 10 Mbps. The Path Cost (in other words, 27 or 8 in Figure 4-11) indicates the individual teamed port's full Path Cost from the teamed port itself all the way to the root switch. The monitoring information displayed in Figure 4-11 would be the results of the network configuration described in Figure 4-10.

**Figure 4-11** Fast Path configuration and monitoring



HP Network Team #1

☑ Active Path Failover

Echo Node Response Mechanism: Community Address ARP

Echo Node IP Address: 192 . 168 . 1 . 1

Echo Node Probe Interval: 3 Seconds

Echo Node MAC Address (Optional):

Echo Node Probe Timeout: 3 Seconds

Community Probe Addresses

IP Address: 192 . 168 . 1 . 254

MAC Address: 00-01-FA-FE-FE-FE

☑ Use Default

☑ Fast Path Failover

Spanning Tree Protocol: IEEE 802.1D

IEEE 802.1D<br>Cisco PVST+

Mechanism Priority

Active Path Failover<br>Fast Path Failover

Team Members

| Port Name | Port/Path Cost | User Ranking |
|---|---|---|
| [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2 | 4/27 | N/A |
| [10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2 | 4/8 | N/A |

*Spanning Tree protocol selection for Fast Path*

Fast Path only has two configuration options:

- Enabled or disabled, as discussed in the previous section
- Spanning Tree Protocol (STP) type

Both of these settings are located on the Advanced Redundancy tab in NCU. Refer to Figure 4-11.

There are two predominate Spanning Tree Protocols (STP) in use today. They are IEEE 802.1D Spanning Tree Protocol and Cisco® Per VLAN Spanning Tree Plus (PVST+) Protocol. The Fast Path implementer should choose which version of Spanning Tree is in use on the attached network.

If IEEE 802.1D is chosen under the Fast Path configuration section, the teaming driver configures the teamed ports to listen for IEEE 802.1D BPDUs. This involves registering, in hardware, a specific multicast MAC address (01-80-C2-00-00-00) on all teamed ports and monitoring for Spanning Tree frames in the IEEE 802.1D frame format. Since the IEEE 802.1D STP is VLAN unaware, it is not necessary to configure the VLAN for Fast Path to listen to.

If Cisco PVST+ is chosen under the Fast Path configuration section, the teaming driver configures the teamed ports to listen for Cisco PVST+ BPDUs. This involves registering, in hardware, a specific multicast MAC address (01-00-0C-CC-CC-CD) on all teamed ports and monitoring for Spanning Tree frames in Cisco PVST+ frame format. Since Cisco PVST+ is VLAN aware, configuring the VLAN for PVST+ to monitor may be necessary. By default, Fast Path using Cisco PVST+ will monitor the first VLAN configured on the team. The VLAN monitored by Fast Path can be changed on the VLAN configuration tab in NCU.

### Recommended configurations for Fast Path environments

HP recommends the following:

- Switch ports connected to teamed ports should be configured to bypass Spanning Tree's block/listen/learn stages. These stages are not needed for ports connected to end node devices (for example, servers, PCs, etc.)
- Switch ports connected to teamed ports MUST transmit BPDUs. Without BPDUs, Fast Path will not have the information it needs to make decisions.
- On most Cisco switches, such as the HP BladeSystem Cisco Gigabit Ethernet Module (CGESM), PortFast is the best setting to use for switch ports connected to teamed ports. This setting accomplishes both recommendations (a) and (b) above.
- On HP BladeSystem GbE blade switches, bypass is the best setting to use for switch ports connected to teamed ports. This setting accomplishes both recommendations (a) and (b) above.
- On HP BladeSystem GbE2 blade switches, switch ports connected to teamed ports should have Spanning Tree enabled. This setting accomplishes recommendation (b) above. Currently the GbE2 does not provide

a feature that accomplishes recommendation (a) above. As a result, the switch port will not allow communication for up to one minute after link is established because of Spanning Tree loop prevention stages (listen, learn). This can cause problems when using protocols like PXE to deploy a server image. An upcoming version of firmware for the GbE2, version 2.2, may include the feature called Fast Forward. This feature accomplishes the same thing as PortFast and should be enabled for all BladeSystem server ports.

### 4-3-1-2-3 Redundancy mechanism priority

Refer to "4-4-2 Mechanism priority" on page 67.

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

**Table 4-7** Redundancy mechanism comparison chart

| | Link loss | TX path validation | RX path validation | Active Path | Fast Path |
|---|---|---|---|---|---|
| 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) | ✓ | ✓ | ✓ | | |
| Advanced redundancy (for fee) | | | | ✓ | ✓ |
| Detects bad ports versus good ports | ✓ | ✓ | ✓ | ✓ | ✓ |
| Detects good ports versus best ports | | | | | ✓ |
| Mechanism can be manually disabled | | ✓ | ✓ | ✓ | ✓ |

### 4-3-1-2-5 Failover events

#### Link loss

When a network port is a member of a team and loses physical link (in other words, 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 port's 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.

#### 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.

### Active Path

As an Advanced Redundancy mechanism, Active Path is used to choose which teamed ports have validated connectivity with the important network segment (by testing communication with the designated Echo Node). Whenever a teamed port loses connectivity with the Echo Node while other teamed ports maintain connectivity with the Echo Node, the teamed port is considered to be in an Active Path Failed state. If the teamed port is the team's Primary port, a failover will occur and another teamed port with validated connectivity to the Echo Node will assume the role of Primary port.

If all teamed ports lose connectivity with the Echo Node, no failover will occur since no teamed port is better than another in regards to the Active Path mechanism. The team will operate as if Active Path was not enabled except that all teamed ports will be labeled as Active Path Degraded and the team will be in a degraded state.

### Fast Path

Fast Path, another advanced redundancy mechanism, is used to validate that all teamed ports have connectivity with the important network segment by monitoring for connectivity to the best root switch. The best root switch is based on the IEEE 802.1D specification that says the bridge (in other words, switch) with the lowest bridge ID is the preferred root Fast Path also can choose the best teamed port from a set of teamed ports that all have validated connectivity to the main network segment. As a result, Fast Path can initiate a failover from one teamed port operating as the Primary port to another non-Primary teamed port for either of two reasons. The first failover is caused by a teamed port losing connectivity with the main network segment, in which case the team will label the teamed port as Split LAN and will disable the teamed port for use by the team. The second failover occurs when the team discovers that a non-Primary port has better connectivity to the root switch than the current Primary port. In this case, the roles are switched, and the original Primary port remains a functioning member of the team in a non-Primary role.

If all teamed ports lose connectivity with all root switches (in other words, no team members are receiving BPDUs), no failover will occur since one teamed port is no better than another with regard to the Fast Path mechanism. The team will operate as if Fast Path was disabled except that all teamed ports will be labeled as Fast Path Degraded and the team will be in a degraded state.

## 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, the Manual Mode button 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 will initiate a failover from the Primary port to an operational Non-Primary port whenever a failover event occurs (refer to "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:

- 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. 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 (in other words, 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 page.  In addition, the User Ranking mechanism is listed in the Mechanism Priority list box on the Advanced Redundancy tab on the Team Properties page.  The User Ranking mechanism is always ranked lower than any other advanced redundancy mechanisms (for example, Active Path or Fast Path).  Otherwise, if User Ranking had a higher priority, the other advanced redundancy mechanisms would be effectively disabled.

For all other aspects of this team type operation, refer to 4-3-1 “Network Fault Tolerance Only (NFT)” on page 28.

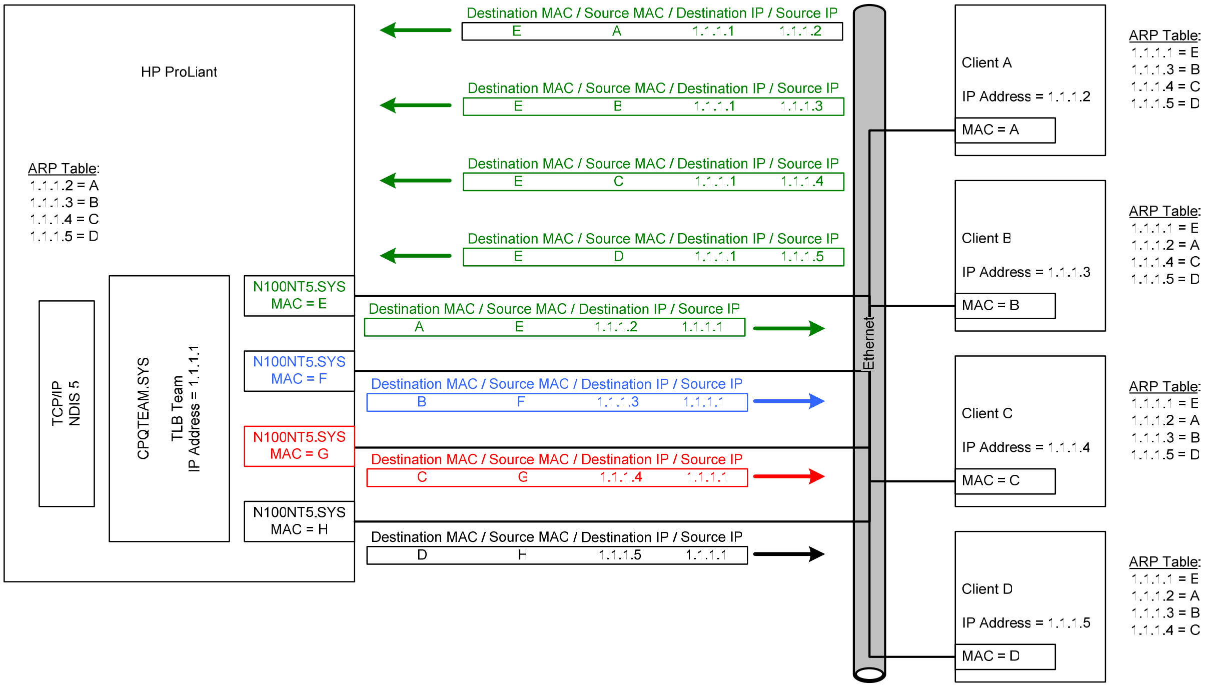
### 4-3-2-1 NFT with Preference Order applications

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 (refer to Figure 4-12). 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 4-12** Overview of TLB communication



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

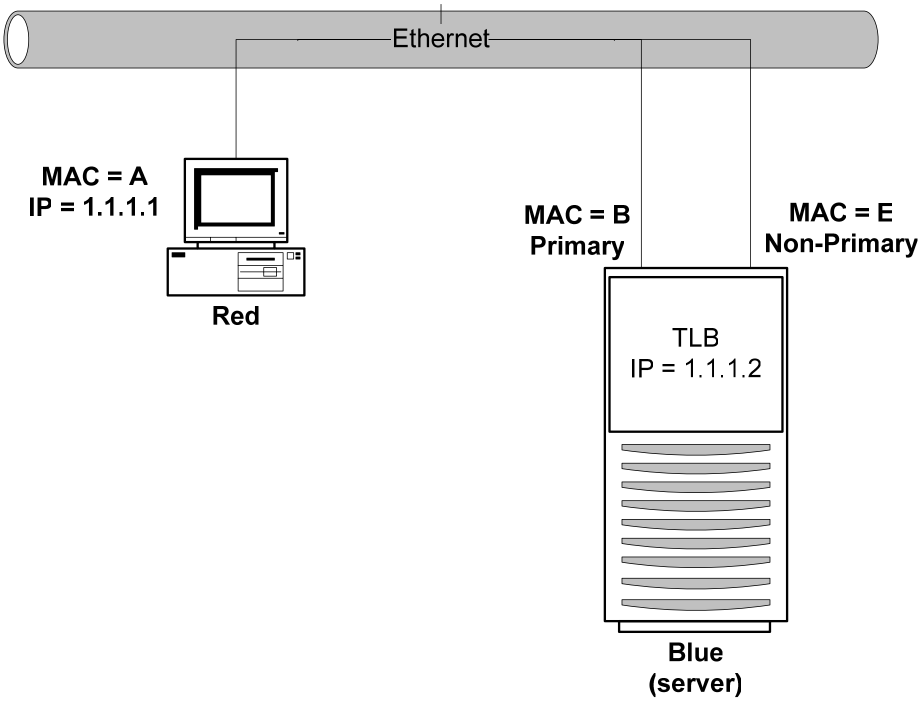
Before learning the specifics of TLB and how it communicates on the network, it is recommended that 4-2-3 "HP teaming and Layer 2 versus Layer 3 addresses" be thoroughly reviewed and understood, as well as "Appendix A: Overview of network addressing and communication."

### 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 75, 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 (refer to Figure 4-13). 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 4-13 Scenario 1-C: A device PINGs a TLB team on the same Layer 2 network



### 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.

Once 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.

### b) 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) (refer to "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.

### c) 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.

### d) 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.

### e) 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.

**f) 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 either Red's MAC address or Red's IP address. Once 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 (refer to "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, 802.3ad Dynamic, and Dual Channel) 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.

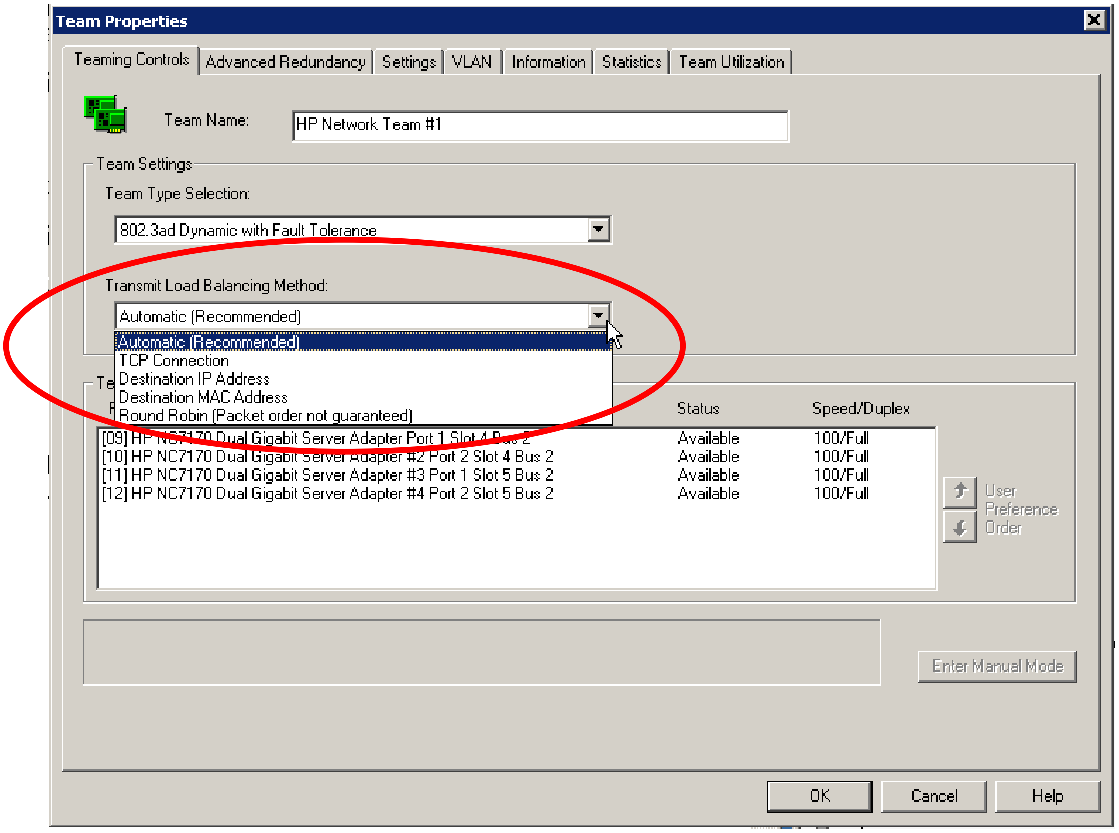
Frame order preservation is important for several reasons – to prevent frame retransmission because frames arrive out of order and to prevent performance-decreasing frame reordering within OS protocol stacks. In order to avoid frames from being transmitted out of order when communicating with a target 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.

**Figure 4-14** Transmit Load Balancing method configuration



Team Name: HP Network Team #1

Team Type Selection: 802.3ad Dynamic with Fault Tolerance

Transmit Load Balancing Method: Automatic (Recommended)

Automatic (Recommended)
TCP Connection
Destination IP Address
Destination MAC Address
Round Robin (Packet order not guaranteed)

```
                                                              Status      Speed/Duplex
[09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2   Available   100/Full
[10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2   Available   100/Full
[11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2   Available   100/Full
[12] HP NC7170 Dual Gigabit Server Adapter #4 Port 2 Slot 5 Bus 2   Available   100/Full
```

**Table 4-8** 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 (refer to “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 (refer to “TLB Destination IP Address method” below). If the frame doesn’t have an IP header, the frame is load balanced by destination MAC address (refer to “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 and IP address but not a TCP header, then the frame is load balanced by destination IP address (refer to “TLB Destination IP Address method” below). If the frame doesn’t have an IP header, the frame is load balanced by destination MAC address (refer to “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 will consecutively assign 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 (refer to Table 4-9).

**Table 4-9** 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 |
| 111 | network port 4 | 111 | network port 3 |

If the Destination IP Address algorithm is chosen, and the frame doesn’t have an IP header, the frame is load balanced by destination MAC address (refer to “TLB Destination MAC Address method” below).

#### Scenario 2-C: a TLB team using IP address-based load balancing

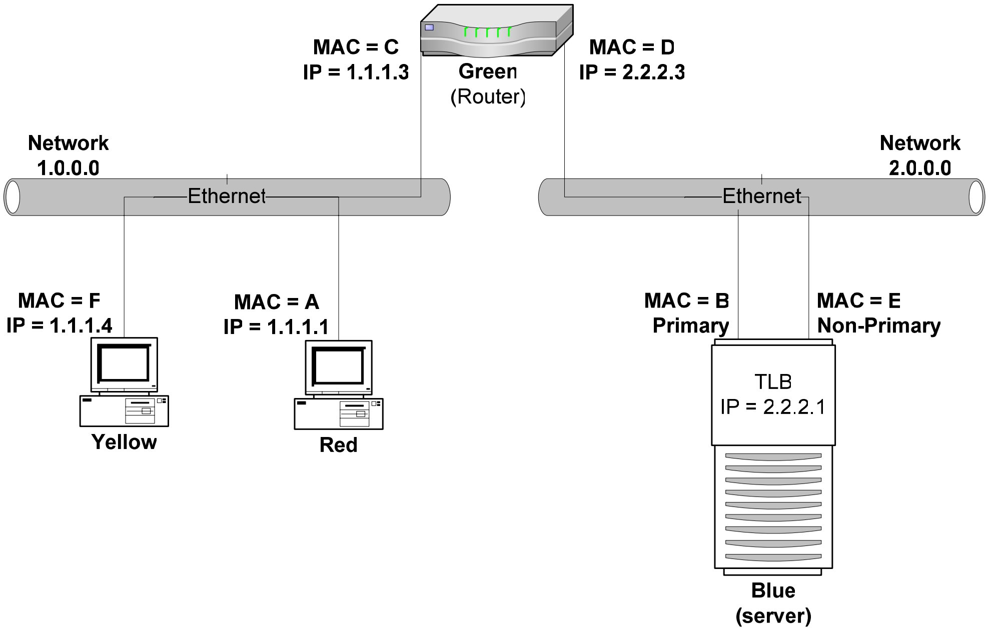
Using the concepts reviewed in Scenario 2-A of “Example scenarios of network addressing and communication” in Appendix A and Figure 4-15, this section describes how TLB IP addressed- based load-balancing functions.

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 (refer to Table 4-9 – 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 4-9 – 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 4-15, 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 4-15** Scenario 2-C: TLB team using IP address for load balancing algorithm



## 4-3-3-2-4TLB 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 (refer to Table 4-10), 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 4-10** 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 |

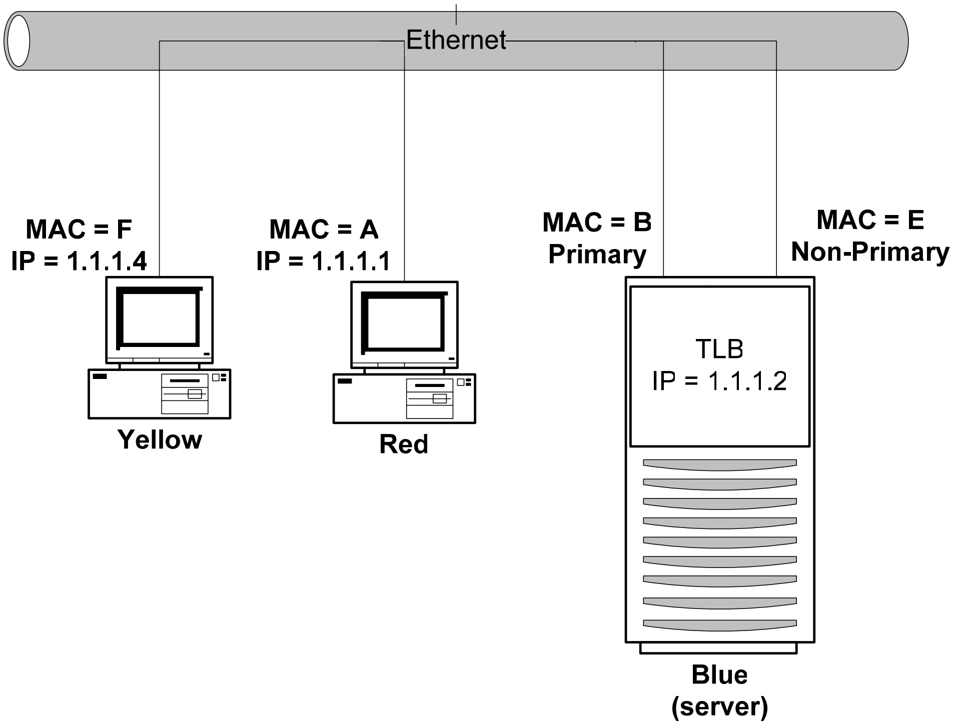
### Scenario 1-D: A TLB team using Destination MAC Address load-balancing method

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 (refer to Table 4-10 – 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 4-10, 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 4-16** 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 (refer to "4-3-3-2-1 TLB Automatic method"), TCP Connection method (refer to "4-3-3-2-2 TLB TCP Connection method"), or Destination IP Connection method (refer to "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. Refer to "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:

- 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 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 page.  In addition, the User Ranking mechanism is listed in the Mechanism Priority list box on the Advanced Redundancy tab on the Team Properties page.  The User Ranking mechanism is always ranked lower than any other advanced redundancy mechanisms (for example, Active Path or Fast Path).  Otherwise, if User Ranking had a higher priority, the other advanced redundancy mechanisms would be effectively disabled.

For all other aspects of this team type operation, refer to "4-3-3 Transmit Load Balancing with Fault Tolerance" on page 48.

### 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.

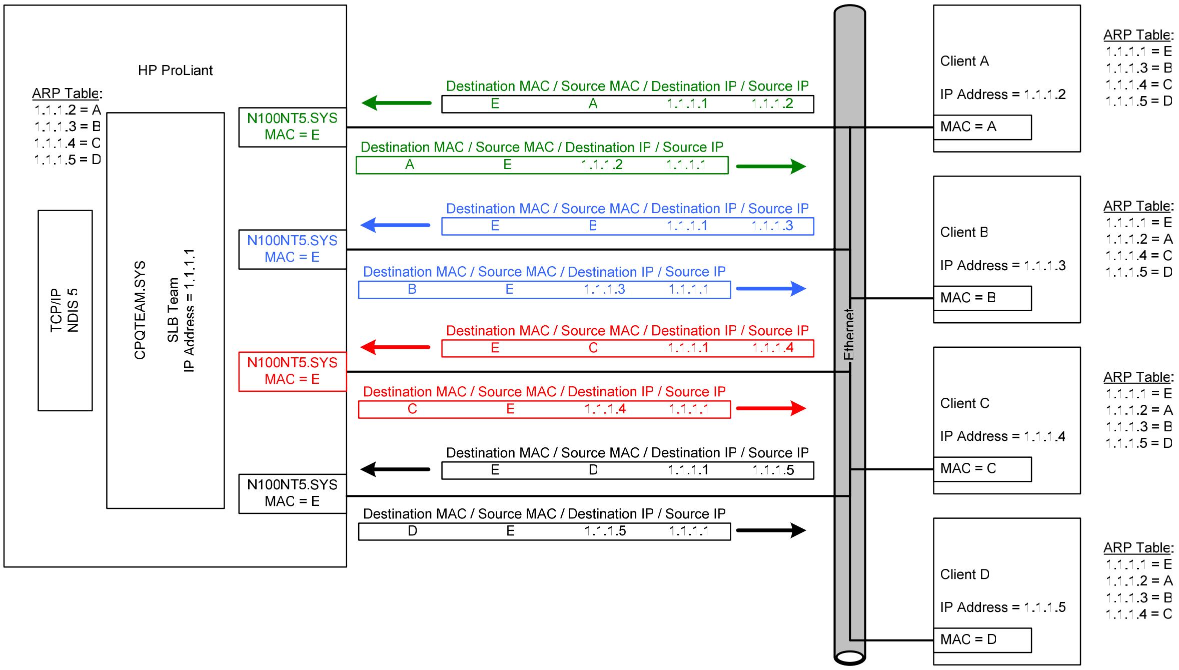
## 4-3-5 Switch-assisted Load Balancing with Fault Tolerance (SLB)

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 (refer to Figure 4-17).

**Figure 4-17** 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 (refer to 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 [refer to 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 (refer to Table 4-11). 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. Because of this, receive traffic is not evenly distributed across all ports in the FEC/GEC group.

**Table 4-11** Example of load-balancing algorithms

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

| | |
|---|---|
| Packet 1 - Source MAC address:00-00-00-00-00-03<br>Last binary bit = 0000 001**1**<br>Packet 1 - Destination MAC address:00-00-00-00-00-01<br>Last binary bit = 0000 000**1**<br>*XOR result of binary bits **1** & **1** = **0*** | *frame is transmitted out port 1* |
| Packet 2 - Source MAC address:00-00-00-00-00-02<br>Last binary bit = 0000 001**0**<br>Packet 2 - Destination MAC address:00-00-00-00-00-01<br>Last binary bit = 0000 000**1**<br>*XOR result of binary bits **0** & **1** = **1*** | *frame is transmitted out port 2* |

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. None of the advanced redundancy mechanisms (for example, Active Path and Fast Path) are supported by SLB. The advanced redundancy mechanisms require the ability to predict and control which ports in a team will receive a particular frame. Since the switch connected to an SLB team controls how the team receives frames, the advanced redundancy mechanisms would not function reliably.

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 [refer to 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. Refer to the SLB sections for more details.

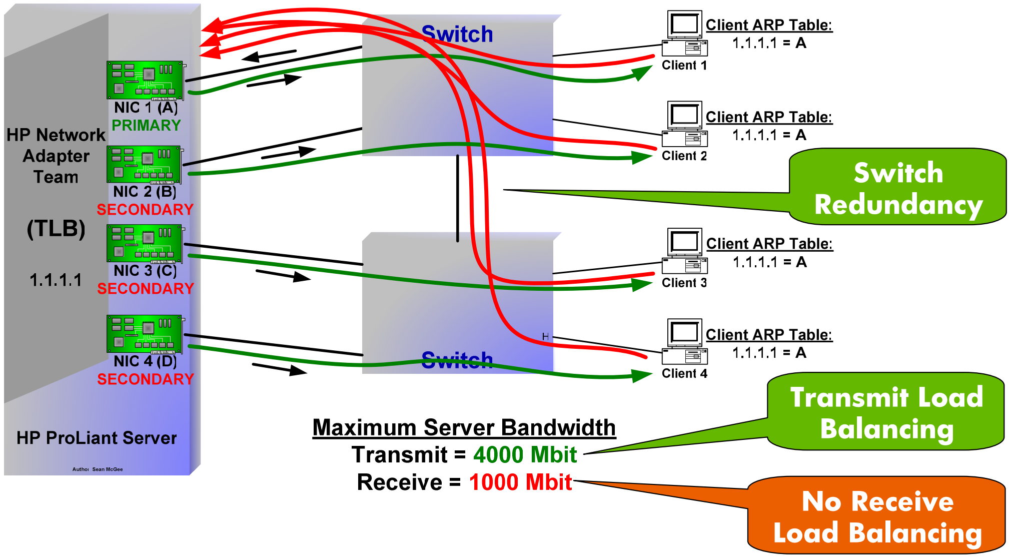
## 4-3-7 Switch-assisted Dual Channel Load Balancing (Advanced Pack)

Switch-assisted Dual Channel Load Balancing, simply referred to as Dual Channel, is a special team type designed by HP to accomplish everything that NFT, TLB and SLB team types accomplish all in a single team type.

Prior to Dual Channel, an HP ProLiant Network Adapter Teaming user had to choose between inbound load balancing or switch redundancy. Inbound load balancing is defined as the ability to receive server traffic on more than one teamed port in the same team. Switch redundancy is defined as connecting teamed ports to more than one switch to recover from a switch failure. Prior to the introduction of Dual Channel, HP ProLiant Network Adapter Teaming did not provide a team type that would accomplish both needs:

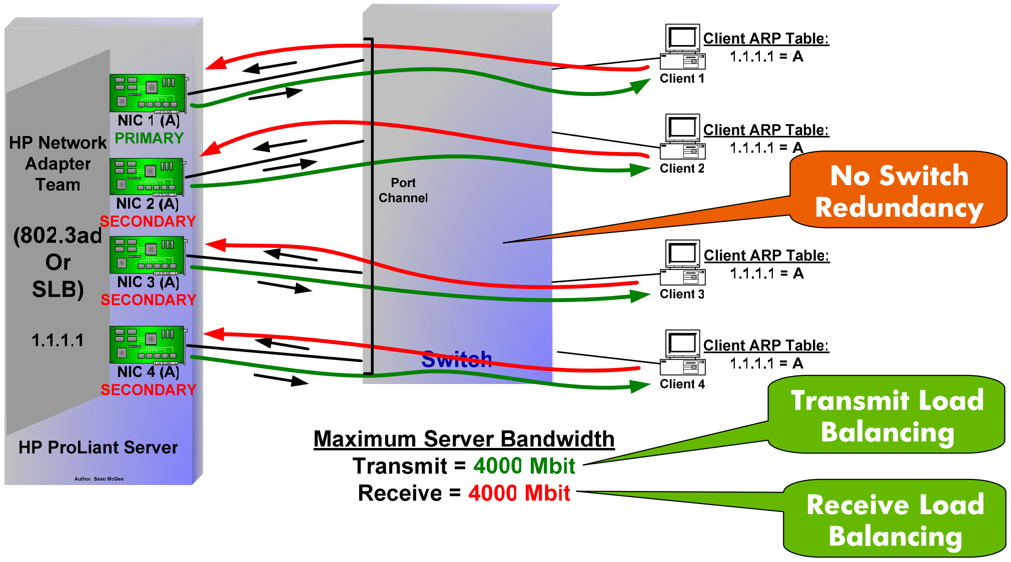
- TLB can provide switch redundancy since it allows you to connect teamed ports from the same team to one or more switches (refer to Figure 4-18). TLB also supports transmit load balancing. However, TLB does not support receive load balancing.

**Figure 4-18** TLB doesn't provide receive load balancing



- SLB supports transmit load balancing. SLB also supports receive load balancing due to assistance from the port trunking technology on the switch. However, SLB requires all teamed ports be connected to the same switch (refer to Figure 4-19). As a result, SLB doesn't provide switch redundancy.

**Figure 4-19** SLB doesn't provide switch redundancy



Dual Channel and Dynamic Dual Channel were developed to address the predicament of having to choose inbound load balancing (SLB or 802.3ad) versus switch redundancy (NFT or TLB). With Dual Channel, a server can have full transmit and receive load balancing with switch redundancy (refer to Figure 4-20).

**Figure 4-20** Dual Channel & Dynamic Dual Channel provide full load balancing and switch redundancy

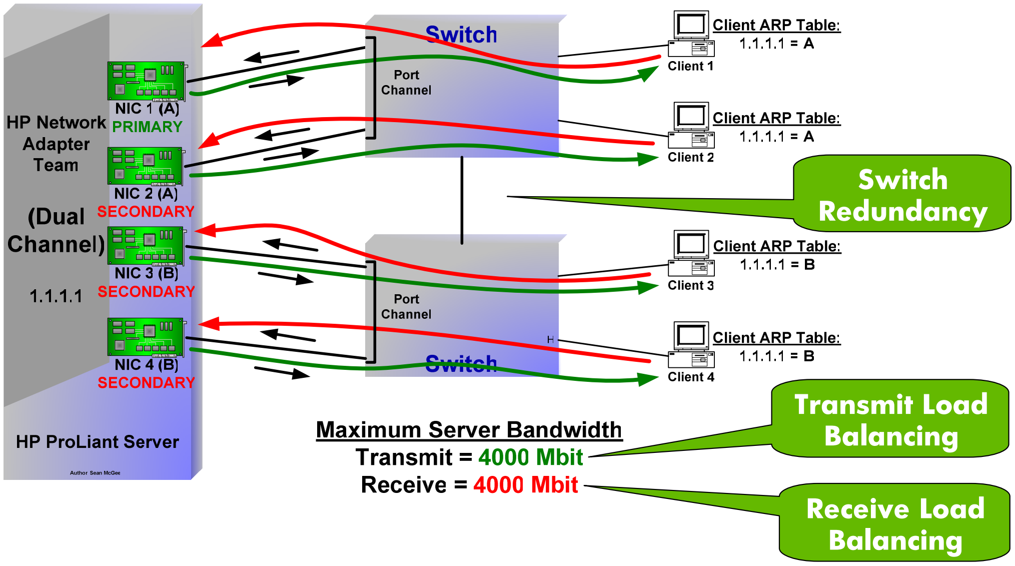
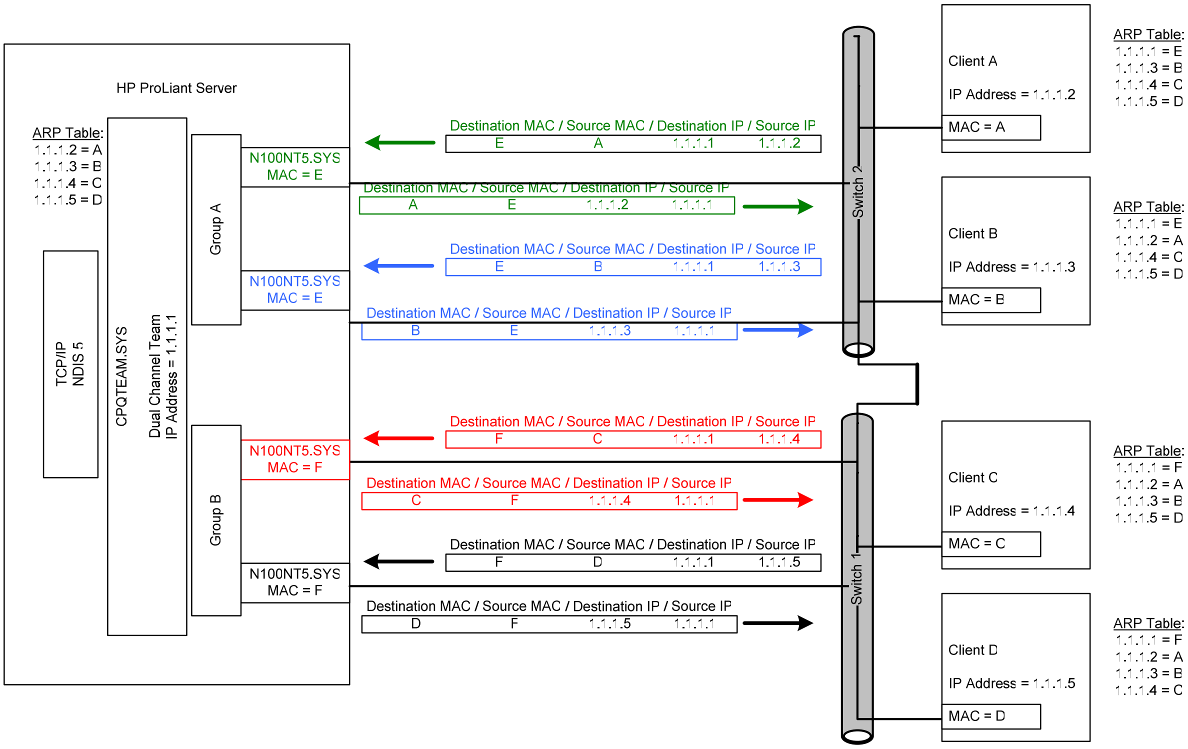


**Table 4-12** Dual Channel/Dynamic Dual Channel capabilities comparison to other team types

| | Fault Tolerance | Transmit Load Balancing | Receive Load Balancing | Switch Redundancy |
|---|---|---|---|---|
| NFT | ✓ | | | ✓ |
| TLB | ✓ | ✓ | | ✓ |
| ***Dual Channel*** | ✓ | ✓ | ✓ | ✓ |
| ***Dynamic Dual Channel*** | ✓ | ✓ | ✓ | ✓ |
| SLB | ✓ | ✓ | ✓ | |
| 802.3ad | ✓ | ✓ | ✓ | |

Dual Channel works by dividing the teamed ports into two groups – Group A and Group B. Each group functions much like an individual team. Because of this, Dual Channel is sometimes described as a team of teams because it is a team comprising two team-like entities called Group A and Group B. The teamed ports within each group behave exactly like an SLB team. This means that all teamed ports within a particular group must be connected to the same switch. Also, if there is more than one teamed port in the same group, the switch must be configured for Port Trunking (just like SLB). Within each group, any teamed port can fail and the group is still functional (just like SLB). If all of the teamed ports in a group fail, the team is still available via the other group. Since the IP address is "owned" by the team instead of either group (A or B), the server maintains network connectivity (with the same IP address) as long as at least one teamed port in at least one group is functional.

**Figure 4-21** Overview of Dual Channel communication



## 4-3-7-1 Dual Channel configuration

Dual Channel configuration is more complicated than other team types because of the need to separate teamed ports into two groups.

To create a Dual Channel team, select at least two network ports on the main NCU page and click the **Team** icon. Next, select **Switch-assisted Dual Channel Load Balancing (Advanced Pack)** from the Team Type Selection drop-down box on the Teaming Controls tab. Refer to Figure 4-22.

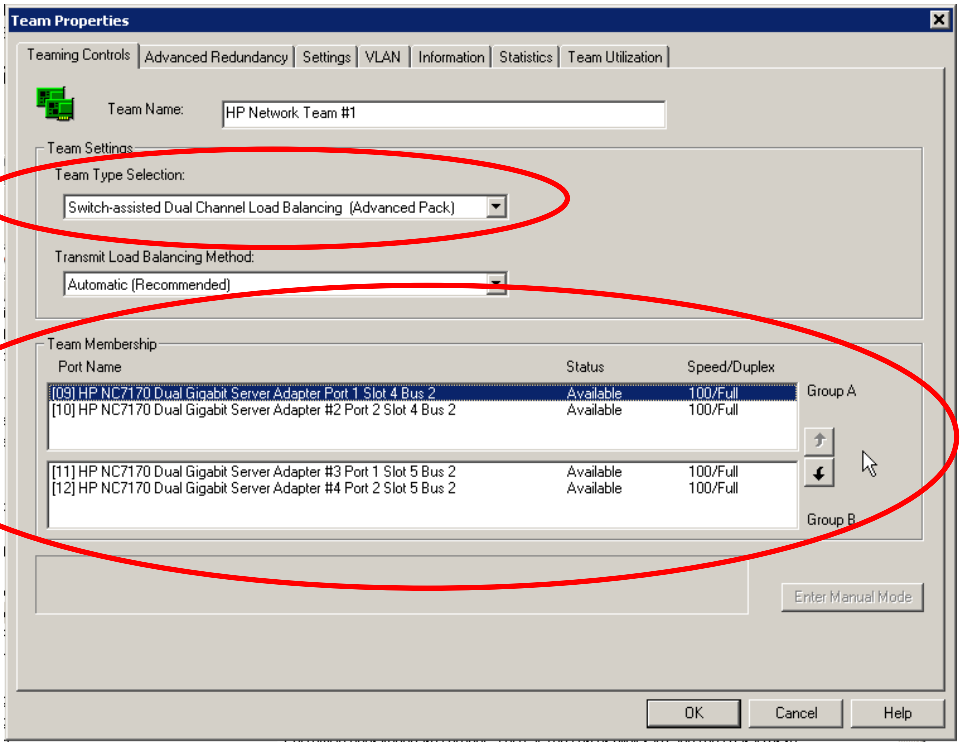
**Note**: If a valid INP license is not installed on the server, the Dual Channel team type will not be listed in the Team Type Selection drop-down box.

After Dual Channel is selected as the team type, the Team Membership section of the tab will change to show the Group configuration. At least one teamed port must be assigned to each group. To move teamed ports between groups, highlight the teamed port and click the up or down arrow on the right hand side of the Team Membership panel.

A group in Dual Channel ONLY requires one teamed port. If only a single teamed port is put in a group, the switch does not need to be configured for Port Trunking. However, if a group has two or more teamed ports, the switch MUST be configured for Port Trunking (like SLB). Also, groups do not have to have the same number of teamed ports. For instance, Group A could have two teamed ports and Group B could have four teamed ports. Or Group A could have one teamed port and Group B could have five teamed ports. However, HP recommends that whenever a Dual Channel team is created with an even number of teamed ports, the ports should be evenly distributed between both groups. Since both Group A and Group B are used for receiving traffic, having an even amount of ports in each group ensures the maximum amount of receive throughput for the team.

The minimal configuration for Dual Channel requires two teamed ports; one port in Group A and one port in Group B. Port Trunking doesn’t need to be configured on the switch for either Group. This configuration provides the flexibility of a TLB team with the throughput of an SLB team.

Figure 4-22 Dual Channel configuration



Team Name: HP Network Team #1

Team Type Selection: Switch-assisted Dual Channel Load Balancing (Advanced Pack)

Transmit Load Balancing Method: Automatic (Recommended)

| Port Name | Status | Speed/Duplex |
|---|---|---|
| [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2 | Available | 100/Full |
| [10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2 | Available | 100/Full |
| [11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2 | Available | 100/Full |
| [12] HP NC7170 Dual Gigabit Server Adapter #4 Port 2 Slot 5 Bus 2 | Available | 100/Full |

## 4-3-7-2 Dual Channel transmit balancing algorithm

The transmit load-balancing algorithms used by Dual Channel are the same methods used by TLB.

Refer to 3-2-3 "Transmit Load Balancing with Fault Tolerance" for a detailed discussion.

## 4-3-7-3 Dual Channel receive load-balancing algorithm

### 4-3-7-3-1 Receive load balancing on a per-port trunk basis

The receive load-balancing algorithm used for a group of teamed ports in a Dual Channel team is identical to the receive load-balancing algorithms used for SLB. Basically, the switch controls the receive load balancing for a particular group in a Dual Channel team (just like an SLB or 802.3ad Dynamic team).

Refer to the SLB section on receive load balancing for a detailed discussion.

### 4-3-7-3-2 Receive load balancing on multiple trunks

**ARP Intercept**

Dual Channel is unique from SLB and 802.3ad Dynamic in that it achieves receive load balancing within a single team that's connected to two switches. As discussed above, the switch controls the receive load balancing with a particular group in a Dual Channel team.

However, another mechanism within the Dual Channel team is needed to achieve receive load balancing across both groups. This mechanism is called ARP Intercept. The teaming driver intercepts ARP responses sent by the IP stack to clients. The teaming driver replaces the Source MAC Address in the ARP data of the frame so that the client is provided with either the MAC address of Group A or Group B, depending on which group the teaming driver chooses to receive the client's data. Since the teaming driver intercepts the ARP frame originating within the server's protocol stack, the feature is referred to as ARP Intercept.

Refer to Figure 4-21 on page 64. Notice that two of the four clients have the MAC address of Group A (MAC E) as the server's IP address in their ARP tables The other two clients have the MAC address of Group B (MAC F) as the server's IP address in their ARP tables.

The client's ARP table entries expire after a few minutes; as a result, the clients will re-ARP for the server's IP address. The clients may or may not get the same IP address for the server depending on the results from the load-balancing algorithm.

**Periodic non-cross trunk load balancing when server transmits ARP request**

Periodically, the server must broadcast an ARP request for the MAC address of an IP address to which it needs to talk. Whenever the server broadcasts the ARP request, all clients within the broadcast domain receive it. Any client that has the server's IP address in its ARP table may use the Source MAC Address (one of the server's MAC addresses from one of the groups) in the ARP data to update its ARP table. If most, or all, of the clients do this, then all receive traffic to the server will be received in a single group instead of in both groups. This happens because all clients have the same MAC address in their ARP table for the server. This condition will eventually correct itself as soon as client ARP table entries begin to expire. As the client's re-ARP for the server's MAC address, receive load balancing will slowly spread across both groups in the Dual Channel team.

It is possible to speed up the recovery process by reducing the ARP table timeout value in the clients IP stack. Refer to the OS manufacturer for information on how to accomplish this.

### 4-3-7-4 Dual Channel redundancy mechanisms

### 4-3-7-5 Basic redundancy mechanisms

The basic redundancy mechanisms on Dual Channel work exactly like that of NFT or TLB with one exception – the heartbeat mechanisms treat all teamed ports within a group as a single teamed port. In other words, when using receive validation heartbeats, the mechanism need only receive a frame on any teamed port in the group to validate the entire group for the mechanism.

This is done since teaming cannot predict how the switch load-balancing algorithm will deliver frames to an individual teamed port within a group.

#### 4-3-7-5-1 Advanced redundancy mechanisms

The advanced redundancy mechanisms on Dual Channel work exactly like that of NFT or TLB with one exception – the Active Path and Fast Path mechanisms treat all teamed ports within a group as a single teamed port. In other words, when using Active Path, the mechanism need only receive an Echo Node response frame on any teamed port in the group to validate the entire group for the Active Path mechanism. In addition, Fast Path needs only to receive a Spanning Tree configuration BPDU on at least one teamed port in the group to validate connectivity with the preferred root switch and to choose the group with the higher bandwidth path to the root switch.

This is done since teaming cannot predict how the switch load-balancing algorithm will deliver frames to an individual teamed port within a group.

### 4-3-7-6 Dual Channel network adapter failure recovery

Network adapter failure recovery for Dual Channel is identical to SLB within a particular group. Failure recovery between groups is identical to TLB.

### 4-3-7-7 Dual Channel applications

Dual Channel is ideal for environments that need NIC fault tolerance, switch fault tolerance, transmit load balancing and receive load balancing, all in a single team type.

### 4-3-7-8 Recommended configurations for Dual Channel environments

HP recommends the following configurations for Dual Channel teams:

- Make sure that each group in a Dual Channel team is connected to a different switch. If both groups are attached to the same switch, switch redundancy is lost.
- Whenever a Dual Channel team is created with an even number of teamed ports, evenly distribute the ports between both groups.
- Ensure all ports in the same group are connected to the same switch.

## 4-3-8 802.3ad Dynamic Dual Channel Load Balancing (Advanced Pack)

802.3ad Dynamic Dual Channel Load Balancing (Dynamic Dual Channel) functions exactly like Dual Channel. The differences between Dual Channel and Dynamic Dual Channel are in the configuration of the team. With Dual Channel, teamed ports must be manually assigned to either Group A or Group B and the switch ports connected to the teamed ports must have Port Trunking manually enabled. With Dynamic Dual Channel, teamed ports do not need to be manually assigned to Groups (A or B) and switch configuration is done automatically with 802.3ad Link Aggregation Control Protocol (LACP). Dynamic Dual Channel uses information learned from the LACP negotiation process to automatically group the teamed ports from each switch.

All other aspects of Dynamic Dual Channel are identical to Dual Channel. Refer to the Dual Channel sections for more details.

#### 4-3-8-1 802.3ad Dynamic Dual Channel configuration

Dynamic Dual Channel configuration is easier than Dual Channel configuration. Dual Channel requires the user to separate teamed ports into two groups in the NCU and requires manual switch configuration on both switches. Dynamic Dual Channel configuration is identical to 802.3ad Dynamic except that two switches are used instead of one.

To create a Dual Channel team, select at least two network ports on the main NCU page and click the **Team** icon. Next, select **802.3ad Dynamic Dual Channel Load Balancing (Advanced Pack)** from the Team Type Selection drop-down box on the Teaming Controls tab. Refer to Figure 4-22.

**Note**: If a valid INP license is not installed on the server, the Dynamic Dual Channel team type will not be listed in the Team Type Selection drop-down box.

Lastly, confirm that all switch ports connected to Dynamic Dual Channel network adapter ports have 802.3ad Link Aggregation Control Protocol (LACP) enabled.

### 4-3-9 Automatic

The Automatic team type is not really a separate team type. Automatic teams decide whether to operate as 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 switch ports have LACP enabled, the team will operate as an 802.3ad Dynamic team. However, if the switch doesn't support LACP or if any ports in the team don't have successful LACP communication with the switch, the team will operate as a TLB team. As network and server configurations change, the Automatic team type ensures that HP ProLiant servers connect to the network in the most efficient team type possible.

To monitor which team type Automatic mode uses, refer to the Information tab on the Team Properties page discussed in 3-6-2 "Team Properties page".

# 4-4 Team types and redundancy mechanisms interoperation

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

**Table 4-13** Team type and redundancy mechanism compatibility

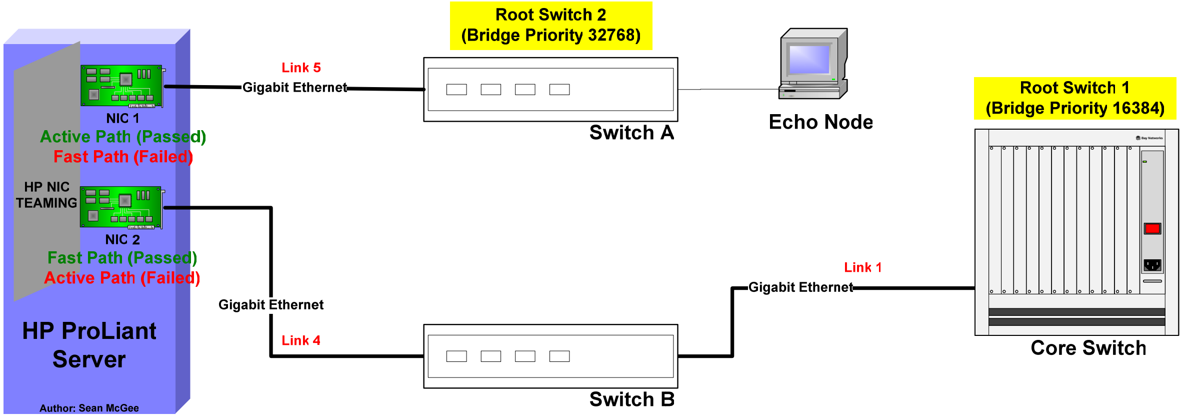
| | NFT | TLB | SLB | 802.3ad Dynamic | Dual Channel | Dynamic Dual Channel |
|---|---|---|---|---|---|---|
| Transmit validation heartbeats | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Receive validation heartbeats | ✓ | ✓ | | | ✓ | ✓ |
| Active Path | ✓ | ✓ | | | ✓ | ✓ |
| Fast Path | ✓ | ✓ | | | ✓ | ✓ |

## 4-4-2 Mechanism priority

Most team types can support multiple redundancy mechanisms simultaneously. This allows an implementer to deploy any combination of redundancy mechanisms depending on what the network environment will support. All basic redundancy mechanisms (in other words, link loss, transmit path validation heartbeats, and receive path validation heartbeats) and all advanced redundancy mechanisms (in other words, Active Path and Fast Path) can be used together. As a result, an implementer may need to prioritize the mechanisms in case of a conflict.

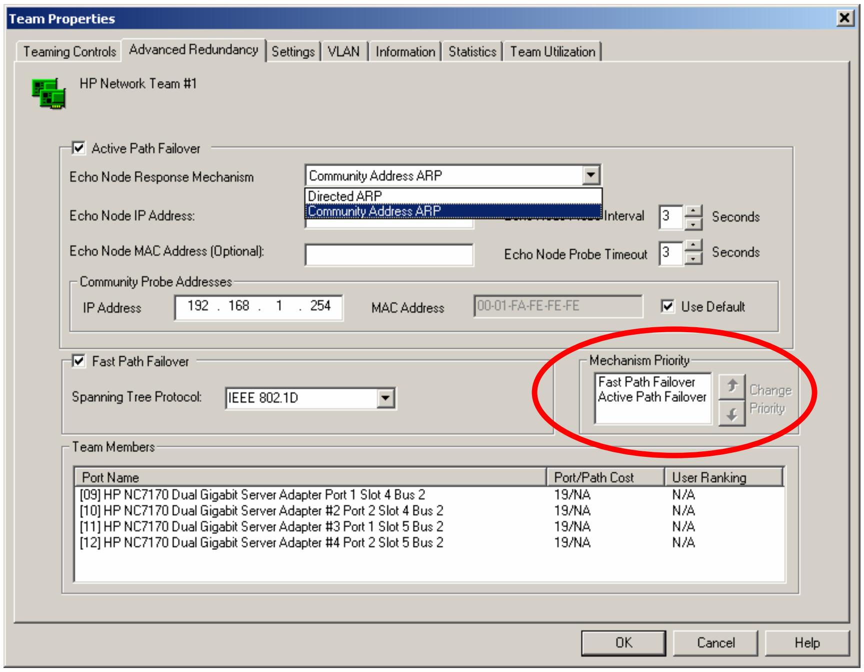
When multiple mechanisms are in use on the same team, certain network scenarios may result in one mechanism in use by the team choosing a different resolution to the problem than another mechanism in use by the same team. For instance, there are situations when Active Path and Fast Path may disagree on which teamed port should be the team's Primary port (refer to Figure 4-23). In this scenario, a team has been deployed with both Active Path and Fast Path. In this situation, Active Path prefers NIC 1 while Fast Path prefers NIC 2. Active Path prefers NIC 1 because only NIC 1 can reach the Echo Node connected to Switch A. Fast Path prefers NIC 2 because NIC2 has access to the best root switch (root switch 1 – Core Switch). NIC 1 connected to Root Switch 2 (Switch A) is inferior because of its Spanning Tree bridge priority. If Switch A and the Core Switch could communicate, Switch A would agree that the Core Switch is the root switch since the Core Switch's Spanning Tree Bridge Priority is better than its own (Switch A). Remember, Spanning Tree calculations always prefer a lower number – the lowest Path Cost is the best, the lowest Bridge ID is the best, the lowest Bridge Priority is the best, etc.

**Figure 4-23** Active Path and Fast Path disagree on Primary teamed port



Because Active Path and Fast Path disagree, the implementer needs to designate which mechanism has a higher priority. On the Advanced Redundancy tab in the NCU (refer to Figure 4-24), the implementer may highlight the individual mechanism and use the arrows on the right to move a mechanism higher or lower on the priority list. If Fast Path is selected as the higher priority mechanism, then NIC 2 will be chosen as the team's Primary port in the scenario above and NIC 1 will be disabled for use by the team. If Active Path is selected as the higher priority mechanism, the result is the opposite – NIC 1 will be chosen as the team's Primary port and NIC 2 will be disabled for use by the team.

**Figure 4-24** Mechanisms priority configuration: Fast Path prioritized higher than Active Path



HP Network Team #1

☑ Active Path Failover

Echo Node Response Mechanism: Community Address ARP
Directed ARP
Community Address ARP

Echo Node IP Address:

Interval: 3 Seconds

Echo Node MAC Address (Optional):

Echo Node Probe Timeout: 3 Seconds

Community Probe Addresses

IP Address: 192 . 168 . 1 . 254

MAC Address: 00-01-FA-FE-FE-FE

☑ Use Default

☑ Fast Path Failover

Spanning Tree Protocol: IEEE 802.1D

Mechanism Priority
Fast Path Failover
Active Path Failover

Team Members

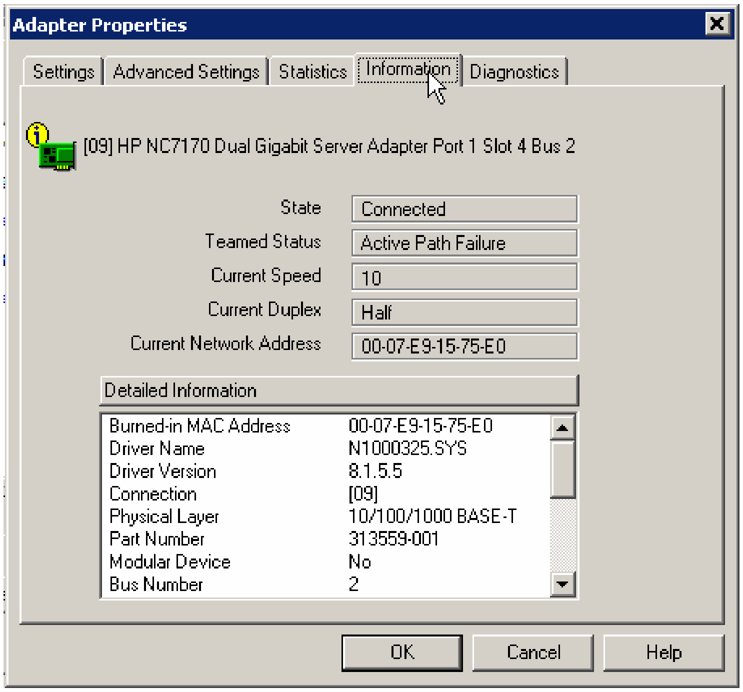
| Port Name | Port/Path Cost | User Ranking |
|---|---|---|
| [09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2 | 19/NA | N/A |
| [10] HP NC7170 Dual Gigabit Server Adapter #2 Port 2 Slot 4 Bus 2 | 19/NA | N/A |
| [11] HP NC7170 Dual Gigabit Server Adapter #3 Port 1 Slot 5 Bus 2 | 19/NA | N/A |
| [12] HP NC7170 Dual Gigabit Server Adapter #4 Port 2 Slot 5 Bus 2 | 19/NA | N/A |

# 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 4-25 Teamed port Information tab



[09] HP NC7170 Dual Gigabit Server Adapter Port 1 Slot 4 Bus 2

State: Connected
Teamed Status: Active Path Failure
Current Speed: 10
Current Duplex: Half
Current Network Address: 00-07-E9-15-75-E0

Detailed Information

| | |
|---|---|
| Burned-in MAC Address | 00-07-E9-15-75-E0 |
| Driver Name | N1000325.SYS |
| Driver Version | 8.1.5.5 |
| Connection | [09] |
| Physical Layer | 10/100/1000 BASE-T |
| Part Number | 313559-001 |
| Modular Device | No |
| Bus Number | 2 |

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. The most likely cause is 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. The most likely cause is changing a parameter for a team member using the Local Area Connection property sheet.
- **Degraded (Fast Path)**—The team member is not receiving BPDUs, and no other team member is receiving BPDUs. Because all team members have equal Fast Path status, the team member is still in use by the team.
- **Degraded (Active Path)**—The team member cannot reach the Echo Node and no other team member can reach it. Because all team members have equal Active Path status, the team member is still in use by the team.
- **Degraded (Rx Heartbeat)**—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 (Active Path)**—The member is not receiving replies from the configured Echo Node.
- **Failed (Split LAN)**—The team member is receiving BPDUs from an inferior Root Bridge. Teaming has determined that team members are connected to different networks.
- **Failed (Fast Path)**—The team member is not receiving BPDUs.
- **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 Heartbeat)**—The team member is not receiving frames.
- **Failed (Tx Heartbeat)**—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.

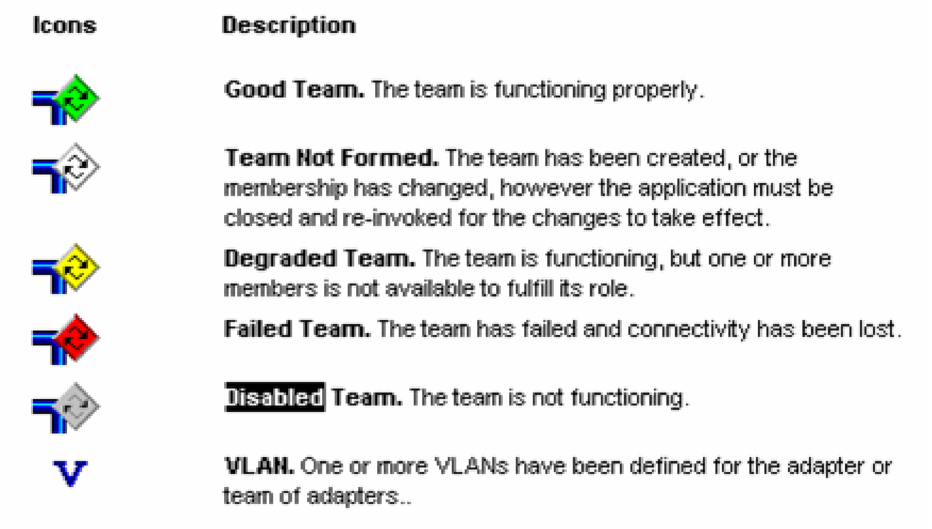
Listed below are the possible team state conditions with definitions:

- **Ok**—The team is functioning properly. (green team icon)
- **Degraded**—The team is functioning, but one or more member(s) is not available to fulfill its role. (yellow team icon)
- **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).

## 4-5-3 Team icons

Team icons are shown on the main page to allow the user to easily recognize status and conditions of teams. Refer to Figure 4-26 for team state icons. Refer to the HELP file for HP ProLiant Network Adapter Teaming for additional information.

Figure 4-26 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. |
| V | **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.

## 4-6-1 Checksum offloading

There are up to four features that are supported on certain HP NC Series network adapters–Receive TCP Checksum Offloading, Transmit TCP Checksum Offloading, Transmit IP Checksum Offloading, and Receive IP Checksum Offloading. HP ProLiant Network Adapter Teaming supports the offloading advanced feature when it is supported by all the teamed ports in the same team. If one or more teamed ports in the same team have the feature enabled and the feature is supported by all teamed ports in the same team, the feature is promoted to the team's Advanced Settings tab as ENABLED and the teaming driver automatically enables this feature on all ports. If all teamed ports in the same team have the feature disabled and the feature is supported by all teamed ports in the same team, the feature is promoted to the team's Advanced Settings tab as DISABLED.

Although some adapters' implementation of checksum offloading is not 100% compatible with other adapters, the team handles the incompatibilities internally and the features are available.

## 4-6-2 802.1p QoS tagging

HP ProLiant Network Adapter Teaming supports the advanced feature 802.1p QoS, when supported by all teamed ports in the same team. This feature is used to mark frames with a priority level for transmission across an 802.1p QOS-aware network. If all teamed ports in the same team have the feature enabled, the feature is promoted to the team's Advanced Settings tab as ENABLED. If one or more teamed ports in the same team have the feature disabled, and the feature is supported by all teamed ports in the same team, the feature is promoted to the team's Advanced Settings tab as DISABLED. If any one team member does not support 802.1p QOS, then the feature will not be promoted to the team's Advanced Settings tab. Most HP NC Series network adapters support this advanced feature.

## 4-6-3 Large Send Offload (LSO)

HP ProLiant Network Adapter Teaming supports the advanced feature Large Send Offload (LSO) (sometimes referred to as TCP Segmentation Offload) when it is supported by all the teamed ports in the same team. This feature is only available in Windows Server 2003 operating systems (server, enterprise, etc) and allows teamed ports in the same team to offload large TCP packets for segmentation in order to improve performance. If one or more teamed ports in the team have the feature enabled and the feature is supported by all teamed ports in the same team, the feature is promoted to the team's Advanced Settings tab as ENABLED and the teaming driver automatically enables this feature on all adapters. If all teamed ports in the same team have the feature disabled and the feature is supported by all teamed ports in the same team, the feature is promoted to the team's Advanced Settings tab as DISABLED.

Although some adapters' implementation of LSO is not 100% compatible with other adapters, the team handles the incompatibilities internally and the features are available.

LSO is available on Windows 2000 but Microsoft doesn't recommend its use in end-user environments. Refer to **http://www.microsoft.com/whdc/device/network/taskoffload.mspx** for more information.

## 4-6-4 Maximum frame size (jumbo frames)

HP ProLiant Network Adapter Teaming supports the advanced feature Maximum Frame Size (Jumbo Frames) when supported by all teamed ports in the same team. This feature allows teamed ports to increase maximum frame size for TCP/IP packets transmitted and received in order to improve performance. If all teamed ports in the same team support jumbo frames, the feature is promoted to the team's Advanced Settings tab with the lowest size configured for the team members. For example, if there are two adapters, one configured for 4088 bytes and the other for 9014 bytes, and they are teamed, the Maximum Frame Size feature is set to 4088 bytes. A setting of 1514 bytes is equal to Jumbo Frames being disabled. Maximum Frame Size greater than 1514 is equal to Jumbo Frames being enabled. Jumbo Frames are supported on NC Series gigabit adapters only. In addition, the specified Maximum Frame Size in HP ProLiant Network Adapter Teaming does not include the four-byte Cyclic Redundancy Check (CRC) portion of an Ethernet frame. Some switch settings do include the CRC portion in their Jumbo Frame Size configuration. Therefore, it may be necessary to increase the switch's Jumbo Frame Size setting by four bytes in order for Jumbo Frames to work correctly.

## 4-6-5 802.1Q Virtual Local Area Networks (VLAN)

The NCU supports the configuration of VLANs on standalone HP ProLiant Network adapters and on HP ProLiant Network adapter teams. This allows a network adapter or a team to belong to more than one VLAN at the same time. When multiple VLANs are configured on the same network adapter or team, 802.1Q VLAN tagging is enabled and used to mark every transmitted frame with a VLAN identifier (number between 1 and 4094). The use of VLAN tagging requires the support of the network infrastructure and/or the receiving network device.

For teams configured to use VLANs, receive validation heartbeats are only transmitted on a single VLAN. This means that if four VLANs are configured on the team and the VLAN configured for receive validation heartbeats is 20, the teaming driver will use VLAN 20 for transmitting heartbeats between team members.

Much like deploying SLB, the use of VLANs requires the switch or switches to be configured properly. Every team member's switch port must be configured with the same VLAN configuration. This means that if a team is to operate on four different VLANs, every team member must have all four VLANs configured on their respective switch port.

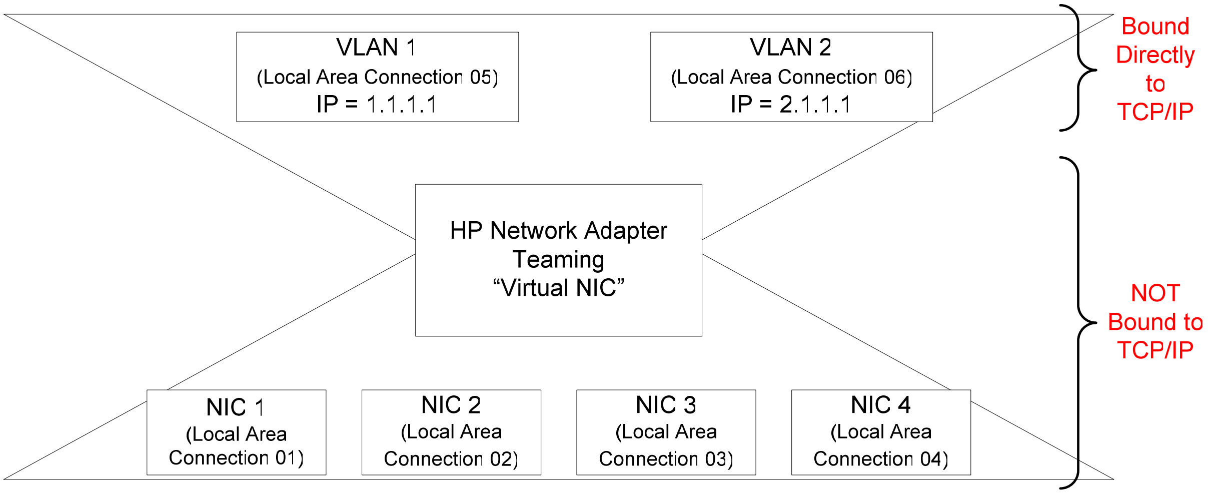
The maximum configurable VLANs per team is 64. The valid VLAN number range per VLAN is 1 to 4094.

Example of VLAN tagging used with HP ProLiant Network Adapter Teaming (refer to Figure 4-27):

- Four NICs teamed together as a single virtual NIC using HP ProLiant Network Adapter Teaming
- Two VLANs configured on top of the virtual NIC to create two virtual interfaces

- Provides the same functionality to the OS as having two NICs installed but provides for fault tolerance and load balancing across four NICs

**Figure 4-27** VLAN tagging used with HP ProLiant Network Adapter Teaming


VLAN 1
(Local Area Connection 05)
IP = 1.1.1.1
VLAN 2
(Local Area Connection 06)
IP = 2.1.1.1
Bound Directly to TCP/IP
HP Network Adapter Teaming "Virtual NIC"
NOT Bound to TCP/IP
NIC 1 (Local Area Connection 01)
NIC 2 (Local Area Connection 02)
NIC 3 (Local Area Connection 03)
NIC 4 (Local Area Connection 04)

# 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, refer to **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** button 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\cqniccmd.exe /c %windir%\temp\hpteam.xml).

For more information on the cqniccmd.exe tool, refer to 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 cqniccmd.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 once the new team is created. You can overcome this by using netsh.exe included with Windows 2003 Server (it may be downloaded from Microsoft for Windows 2000 servers).

# 5 Teaming feature matrix

**Table 5-1** Teaming feature matrix

| | NFT &<br>NFT with Preference | TLB &<br>TLB with Preference | SLB &<br>802.3ad Dynamic | Dual Channel &<br>Dynamic Dual Channel |
|---|---|---|---|---|
| Number of ports supported per team | 2-8 | 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 | 800/800 |
| Maximum theoretical transmit/receive throughput (in Mbps) with maximum number of 1000 Mbps ports | 1000/1000 | 8000/1000 | 8000/8000 | 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) | ✓(Dynamic Dual Channel 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 address(es) on the network | ✓ | ✓ | ✓ | ✓ |

* Dual Channel and Dynamic Dual Channel transmit frames with the same source MAC address from all teamed ports in the same group only.

# 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 (refer to "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 (refer to "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.

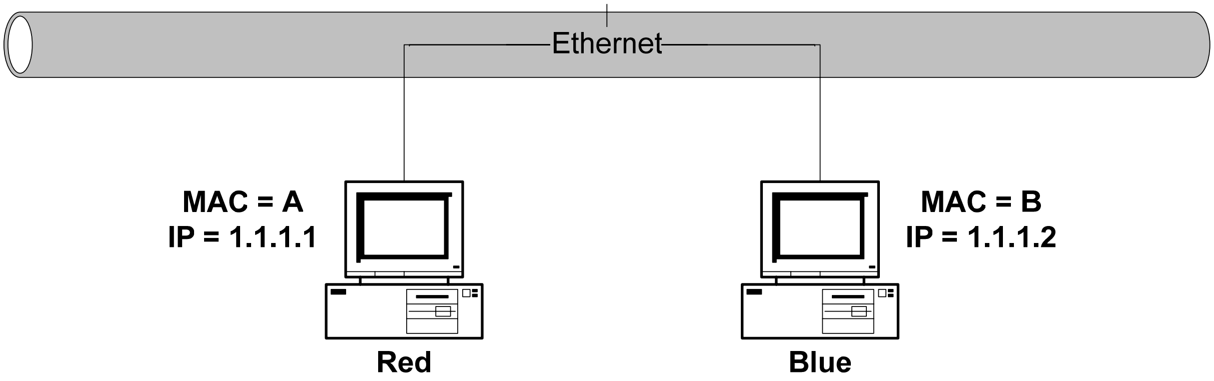
## Example scenarios of network addressing and communication

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.

## 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



### 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.

Once 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.

### 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.

### 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.

### 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) 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.

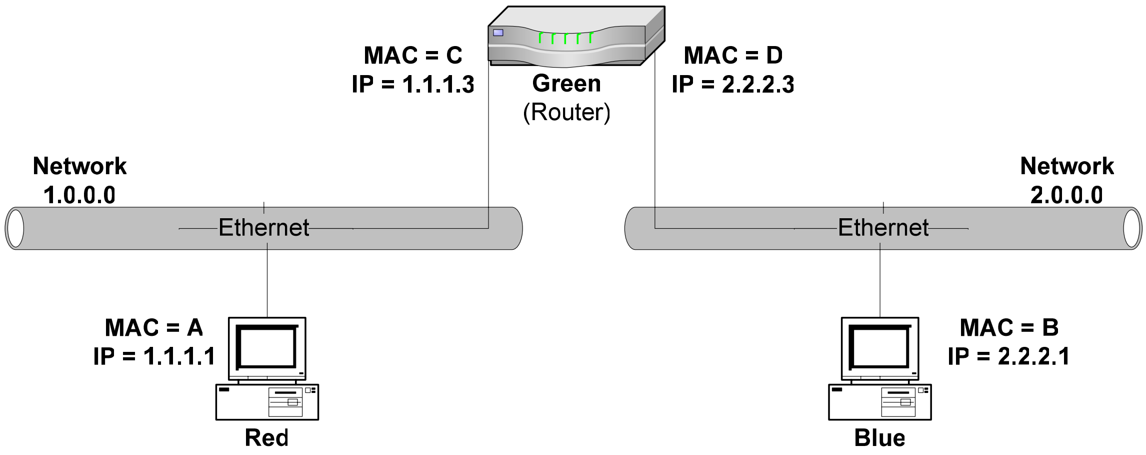
Red sees the ARP request and transmits a unicast ARP reply directly to Blue providing its MAC address (A). Blue receives the ARP reply and puts Red's MAC address (A) and IP address (1.1.1.1) in its ARP cache.

### 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.

## 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



### 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.

### 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.

### 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).

### 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) 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.

### 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.

### 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.

### 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.

### 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).

### 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.

### 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.

### 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, 802.3ad Dynamic or Dual Channel 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 don't 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. The exception would be the NC150T, which does not support SLB or Dual Channel.

**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 page of any network interface under Network and Dial-up Connections (Microsoft UI). Select **HP Network Teaming and Configuration** and click the **UNINSTALL** button. 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.

For Dual Channel, they need to know which sets of teamed ports constitute Group A versus Group B. All the teamed ports in Group A should be configured on the switch as one port trunk and all the teamed ports in Group B should be configured on the switch as another port trunk.

**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 address/es and will randomly transmit frames for any of the NICs down any of the links for any of the NICs. If the NICs aren't 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 2000 Service Pack level that is required for HP ProLiant Network Adapter Teaming to work correctly?**

**A22** No, not for Windows 2000 or Windows 2003. Windows NT4 requires Service Pack 5.

**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 an 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 page 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.  Refer to NICSCRPT.PDF, which is located in the c:\compaq\network directory, if teaming has been installed on the system.

**Q27 I need NIC redundancy, switch redundancy and RX/TX load balancing. What teaming type should I use?**

**A27** When NIC redundancy, switch redundancy, and load balancing are all required, the only option is Dual Channel or Dynamic Dual Channel. An alternative is to use switch vendor redundancy mechanisms to make a single switch highly redundant. For example, Cisco provides an option called High Availability on some switches. This option allows a Cisco switch to have redundant Supervisor modules.

An HP customer using a Cisco switch with redundant Supervisor modules and redundant power supplies can create an SLB team, a Dual Channel team, or a Dynamic Dual Channel team of several ports, connect the team to two modules inside the same Cisco switch, and enable High Availability. This provides transmit/receive load balancing, network port fault tolerance, switch power supply fault tolerance, Supervisor fault tolerance, and switch module fault tolerance. This option is usually more expensive than purchasing a license to use Dual Channel or Dynamic Dual Channel.

**Q28 Which switch Port Trunking/Port Bonding technologies is SLB (Switch-assisted Load Balancing) compatible with?**

**A28** FEC/GEC, Load Sharing, MLT, IEEE 802.3ad (Static mode – no LACP), etc.

**Q29 Which switch Port Trunking/Port Bonding technologies is Dual Channel compatible with?**

**A29** FEC/GEC, Load Sharing, MLT, IEEE 802.3ad (Static mode – no LACP), etc.

**Q30 Which switch Port Trunking/Port Bonding technologies is 802.3 Dynamic compatible with?**

**A30** IEEE 802.3ad (Dynamic mode – using LACP)

**Q31 Can I connect teamed ports to more than one switch?**

**A31** Yes, with NFT, TLB, and Dual Channel 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.

**Q32 Who in HP is responsible for development and support of HP ProLiant Network Adapter Teaming?**

**A32** 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, etc.). NSP is one of many groups that constitute HP's Industry Standard Servers division.

**Q33 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?**

**A33** 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.

**Q34 How do I upgrade HP ProLiant Network Adapter Teaming drivers?**

**A34** HP provides the HP ProLiant Network Adapter Teaming driver with an installation utility. The same component completes installation and upgrades. Download the appropriate HP ProLiant Network Adapter Teaming driver from http://h18004.www1.hp.com/support/files/networking/us/index.html.

**Q35 Where do I find drivers for HP's legacy network adapters?**

**A35** http://h18000.www1.hp.com/support/files/networking/nics/index.html

**Q36 Can I obtain an evaluation license for the ProLiant Essentials Intelligent Networking Pack?**

**A36** A one-day and a 30-day evaluation license are available upon request by visiting http://www.hp.com/go/tryessentials.

**Q37 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?**

**A37** 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 (CP*xxxx*.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. Consult the latest HP ProLiant Network Adapter Teaming Component Pack for the latest utilities and associated documentation.

- Cqniccmd.exe

  CQNICCMD is a Windows 2000 and Windows Server 2003 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, refer to the NICSCRPT.PDF file in the HP ProLiant Network Adapter Teaming Component Pack.

- Hpnetinfo.exe

  HPNETINFO is a Windows Server 2003 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, etc. 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 add a ProLiant Essentials Intelligent Networking Pack license to the system or display licenses previously installed on the system.

  For more information on NALicense.exe, refer to 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, refer to 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://h18004.www1.hp.com/products/servers/networking/index.html

**ProLiant Ethernet adapters**
http://h18004.www1.hp.com/products/servers/networking/index-nic.html

**ProLiant switch adapters**
http://h18004.www1.hp.com/products/servers/networking/nc150t/index.html

**Intelligent Networking Pack**
http://h18004.www1.hp.com/products/servers/proliantessentials/inp/index.html

## Whitepapers

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

## Network adapter and network adapter teaming drivers

Software/drivers by NIC: http://h18004.www1.hp.com/support/files/networking/us/index.html

Software/drivers by server:http://h18004.www1.hp.com/support/files/server/us/index.html

**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 Network Configuration Utility" 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.

**Active Path** An advanced redundancy mechanism used by HP ProLiant Network Adapter Teaming that allows the team to monitor per-teamed port connectivity with an external device (Echo Node). Requires the team to have an Intelligent Networking Pack license.

**Advanced Pack** Same as Intelligent Networking Pack (INP).

**ALB Adaptive Load Balancing** Refer to 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's 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 CP*xxxxx*.exe.

**Degraded** A team or teamed port status that indicates it is still in use but not completely whole.

**Dual Channel** A special team type used by HP ProLiant Network Adapter Teaming that allows for NIC port redundancy, switch redundancy, outbound load balancing, and receive load balancing all on the same team. Requires the team to have an Intelligent Networking Pack license.

**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. Refer to 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. Refer to SLB.

**GUI**
Graphical User Interface.

**HP Teaming and Configuration GUI**
Now referred to as the HP Network Configuration Utility (NCU). Refer to 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.

**Intelligent Networking Pack**
Refer to INP.

**INP**
Intelligent Networking Pack. A set of licensed features that allow for advanced redundancy for HP ProLiant Network adapter teams. Features include Fast Path, Active Path, Dual Channel, and 802.3ad Dynamic Dual Channel. Also referred to as "Advanced Pack".

**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 refer to an actual hardware device (NIC) or to a software device (for example, team, VLAN, virtual NIC interface, Loopback, etc.)

**LOM**
LAN On Motherboard. The acronym refers to a network adapter that is built directly onto the computer's motherboard. LOM's 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 refer to a package of drivers for ProLiant server network adapters, teaming, firmware, etc. 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 made available on CD or the individual drivers are available for download from www.hp.com.

**Network adapter** A physical network adapter card. Network adapter cards may contain one of 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. Refer to 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 refer to 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.