Network Load Balancing Software Administrator's Guide

Document Number 007-3895-001

CONTRIBUTORS

Written by Carolyn Curtis, with material by Carlin Otto and Susan Thomas Illustrated by Dan Young

Production by Linda Rae Sande

Engineering contributions by James Yarbrough, John Becker, and Ken Robinson St. Peter's Basilica image courtesy of ENEL SpA and InfoByte SpA. Disk Thrower image courtesy of Xavier Berenguer, Animatica.

© 1999, Silicon Graphics, Inc.— All Rights Reserved

The contents of this document may not be copied or duplicated in any form, in whole or in part, without the prior written permission of Silicon Graphics, Inc.

RESTRICTED RIGHTS LEGEND

Use, duplication, or disclosure of the technical data contained in this document by the Government is subject to restrictions as set forth in subdivision (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS 52.227-7013 and/or in similar or successor clauses in the FAR, or in the DOD or NASA FAR Supplement. Unpublished rights reserved under the Copyright Laws of the United States. Contractor/manufacturer is Silicon Graphics, Inc., 2011 N. Shoreline Blvd., Mountain View, CA 94043-1389.

Silicon Graphics, CHALLENGE, Indy, IRIS, and IRIX are registered trademarks and Indigo2, IRIS InSight O2, OCTANE, Onyx2, Origin2000, Origin2000, and Origin2000 GIGAchannel are trademarks of Silicon Graphics, Inc. ACEswitch is a trademark of Alteon Networks, Inc. EtherChannel and Catalyst are registered trademarks of Cisco Systems, Inc. Macintosh is a registered trademark of Apple Computer, Inc. Windows NT is a registered trademark of Microsoft Corporation.

Network Load Balancing Software Administrator's Guide Document Number 007-3895-001

Contents

List of Figures v **List of Tables** vii

About This Guide ix

	Audience ix
	Structure of This Document ix
	Other Required Documentation x
	Conventions x
	Product Support xi
1.	Network Load Balancing Software and the Ethernet 1
	Network Load Balancing Software Features 1
	Compatibility 2
	Supported Platforms 3
	Supported Switches and Ethernets 3
	Network Load Balancing Driver and System Resources 5
2.	Network Load Balancing Software Installation and Configuration 7
	Installing the Software 7
	IP Addresses and the Network Load Balancing Software 8
	Preparing for Network Load Balancing Software Configuration 9
	Network Load Balancing Software and the Shell Command File /etc/init.d/network 11
	Configuring the Network Load Balancing Software and Driver 12
	Adding Network Load Balancing to a System 13
	Making the Network Load Balancing Interface the Primary Interface 16
	Configuring Multiple Network Load Balancing Software Interfaces 19
	Verifying the Network Load Balancing Software Connection 20

3. **Performance Tuning** 23

Tuning Parameters 23

Changing Input Load Balancing 24

Adjusting the Balance Period 25

Adjusting the Balance Threshold Value 25

Adjusting the Average Weight 26

4. Troubleshooting 27

Initial Remedies 27

Connection Has Not Been Functional Since Last Boot 28

Connection Has Been Functional in the Recent Past 29

High Rate of Packet Loss Occurs 30

System Cannot Communicate With Other Systems 31

System Does Not Load Miniroot or Boot From the Network 31

A. Validated Ethernet Switches 33

B. Error Messages 35

List of Figures

Figure 1-1	Network Load Balancing Software Example 4
Figure 2-1	Network Load Balancing Device as Secondary Interface Example (netstat -in) 16
Figure 2-2	Network Load Balancing Device as Primary Interface Example (netstat -in) 19
Figure 4-1	Displaying Available Interfaces With netstat -in 29

List of Tables

Table 1-1	Network Load Balancing Software Subsystems	5
Table 2-1	Default Network Interface Parameters 10	
Table A-1	Validated Ethernet Switches 33	

About This Guide

This guide describes the Network Load Balancing Software from Silicon Graphics, Inc., which lets you consolidate Ethernet ports into a single IP address. It also optimizes performance by determining and setting the load balance for input and output packets. The Network Load Balancing Software runs on CHALLENGE, Origin200, Origin2000, and Onyx2 systems running IRIX 6.5 or later.

Audience

This manual is your guide to configuring, testing, and monitoring your Network Load Balancing Software network connection. This guide has been written so you can perform all the basic Network Load Balancing Software administration tasks.

This guide is written for network system administrators. It presumes general knowledge of the Ethernet and of the Silicon Graphics system in which it is installed.

Note: This guide is not an in-depth network administration guide; it does not provide information for planning, managing, and maintaining an Ethernet network.

Structure of This Document

This guide consists of the following chapters and appendixes:

- Chapter 1, "Network Load Balancing Software and the Ethernet," describes the software's features.
- Chapter 2, "Network Load Balancing Software Installation and Configuration," explains how to install the software and how to associate Ethernet ports with it.
- Chapter 3, "Performance Tuning," explains how to tune parameters to control how the Network Load Balancing Software distributes outbound packets, and how it balances inbound packets.

- Chapter 4, "Troubleshooting," describes what to do when your network connection has problems.
- Appendix A, "Validated Ethernet Switches," summarizes Ethernet switches for which the Network Load Balancing Software has been validated.
- Appendix B, "Error Messages," lists Network Load Balancing Software error messages and gives suggestions for remedies.

Other Required Documentation

For instructions on configuring a system for networking, see the latest version of *IRIX Admin: Networking and Mail.*

For general information about your Silicon Graphics system, see its owner's guide.

If you do not have your owner's guide handy, you can get it and other Silicon Graphics documentation online in the following locations:

- If you have installed the guide on your system, or if it is installed on a server on the network, you can use the IRIS InSight Library: from the Toolchest, choose Help > Online Books > SGI EndUser or SGI Admin, and select the applicable owner's or hardware guide.
- If you have access to the Internet, you can use the Technical Publications Library. Enter the following URL in your Web browser location window: http://techpubs.sgi.com/library/

Once you are in the library, choose Catalogs > Hardware Catalog > and look under the Owner's Guides for the applicable owner's guide.

Conventions

In command syntax descriptions and examples, square brackets ([]) surrounding an argument indicate an optional argument. Variable parameters are in *italics*. If the variable appears in an italics context, angle brackets (< >) are used around the variable to differentiate it from the literal. Replace these variables with the appropriate string or value.

Command-line flags and switches are in **boldface regular type**; these are preceded with dashes or with a plus or minus, for example, **-e**.

Commands, IRIX filenames, and document titles are in *italics*.

Helvetica Bold font is used for labels on hardware, such as the names of LEDs.

Messages and prompts that appear onscreen are shown in fixed-width type. Entries that are to be typed exactly as shown are in **boldface fixed-width type**.

Product Support

Silicon Graphics, Inc., provides a comprehensive product support and maintenance program for its products. If you are in North America, and would like support for your Silicon Graphics-supported products, contact the Technical Assistance Center at 1-800-800-4SGI or your authorized service provider. If you are outside North America, contact the Silicon Graphics subsidiary or authorized distributor in your country.

Network Load Balancing Software and the Ethernet

The Network Load Balancing Software lets you consolidate as many as 40 Ethernet ports into a single IP address, and balance the load on the ports.

This chapter consists of these sections:

- "Network Load Balancing Software Features" on page 1
- "Compatibility" on page 2
- "Network Load Balancing Driver and System Resources" on page 5

Network Load Balancing Software Features

The Network Load Balancing Software acts as a virtual network device, with its own IP address.

You associate real devices—the actual Ethernet ports—with the Network Load Balancing Software; instead of using and advertising the ports' IP addresses, you use the Network Load Balancing Software's IP address. The Network Load Balancing Software and its ports look like a single device to the clients.

The method of assigning a single IP address to a group of Ethernet devices is to collect the Ethernets under a new network device, the *load-balancing device*. Each Ethernet device must have its own IP address, and this address cannot be identical to that of any other network device on the system. Only the IP address of the load-balancing device is advertised, typically through the name assigned to the IP address; the IP addresses of the individual Ethernets are not.

Besides offering the efficiency of a single IP address for several ports, the Network Load Balancing Software automatically optimizes load sharing:

- The Network Load Balancing Software distributes outbound packets on a
 per-connection basis. (Packets with the same source and destination IP addresses
 and port numbers are considered to belong to the same connection.) It guarantees
 that all packets for a single connection are sent out on the same Ethernet device.
- The Network Load Balancing Software balances input on a per-client basis. Input balancing is dynamic and requires no special configuration.

Each client is given a MAC address to use for the IP address belonging to the load-balancing device. The clients do not all have the same IP-to-MAC address mapping; thus they send all of the packets they transmit to the MAC address they were given without regard to connection.

You can adjust parameters to fine-tune the load balancing; this process is explained in Chapter 3, "Performance Tuning."

The Network Load Balancing Software can detect link failure and can reconfigure the load-balancing device in real time to adapt for the failed link. The software can detect interface and cable failure between the server and the switch; it automatically redirects inbound and outbound traffic to an active adapter port, thus eliminating a single point of network failure.

Compatibility

This section discusses compatibility in the following subsections:

- "Supported Platforms" on page 3
- "Supported Switches and Ethernets" on page 3

The Network Load Balancing Software provides IP alias support and dynamic session shifting. As you can with any network device, you can set up aliases for the Network Load Balancing Software; follow instructions in the latest version of *IRIX Admin: Networking and Mail.* However, you cannot assign aliases to the physical devices.

Supported Platforms

The Network Load Balancing Software runs on the following Silicon Graphics systems:

- Challenge DM, L, and XL system with at least one two-port Fast Ethernet VME adapter board installed
- Origin200 system with up to three PCI Fast Ethernet adapter boards per module
- Origin200 GIGAchannel two-module system with two Gigabit Ethernet PCI adapter boards or with up to five 4-port XIO Fast Ethernet adapter boards
- Origin2000 or Onyx2 system with up to six 4-port Fast Ethernet adapter boards per module, or with two to six Gigabit Ethernet adapter boards per module

In addition, the Network Load Balancing Software is interoperable with OCTANE, O2, Indy, Indigo2, and any other client that supports gratuitous ARP.

Supported Switches and Ethernets

The Network Load Balancing Software supports load distribution and bandwidth aggregation for mixed IP traffic, that is, multiple connections, or connections to multiple hosts.

A network switch is required as an interface between the client Ethernet ports and the host system running the Network Load Balancing Software. The Ethernets used with the load-balancing device must be connected to a Layer 2 switch and must be on the same virtual local area network (VLAN). There are no additional requirements for the switch, nor are there any requirements for any other systems connected to the switch to send packets to the load-balancing device. No special switch configuration is required for Layer 2 devices.

The software is compatible with Ethernet 10-Base-T, 100-Base-T, and gigabit Ethernet. The Network Load Balancing Software supports 2 to 40 Ethernet ports per subnet, and 1 to 20 load-balancing devices per system, assuming two ports minimum per group. Appendix A, "Validated Ethernet Switches," summarizes Ethernet switches with which the Network Load Balancing Software has been validated.

Note: No speed information is made available by the Ethernet driver, so the device does not distinguish between 10-Base-T, 100-Base-T, and gigabit Ethernet. Because the load-balancing device does not take maximum line bandwidth into account (that is, bandwidth differences between attached devices), mixing 10-Base-T, 100-Base-T, and gigabit Ethernet might not give the expected results.

Figure 1-1 diagrams an example of Network Load Balancing Software and switch topology.

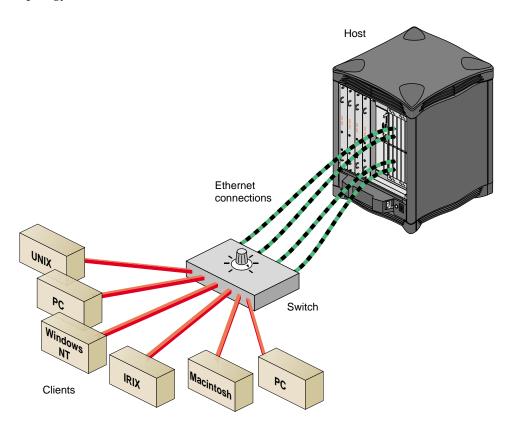


Figure 1-1 Network Load Balancing Software Example

Note: For more information on networks on Silicon Graphics systems, see the latest version of *IRIX Admin: Networking and Mail*.

Network Load Balancing Driver and System Resources

The Network Load Balancing driver includes the subsystems summarized in Table 1-1.

 Table 1-1
 Network Load Balancing Software Subsystems

Subsystem	Description	Approximate Size (512-Byte Blocks)
nlb_eoe.sw.nlb	Network Load Balancing driver software	116
nlb_eoe.man.nlb	Man pages	12
nlb_eoe.man.relnotes	Release notes	32

The Network Load Balancing software places an additional memory usage load of 40 bytes per client on the kernel: each client communicating with the server via the server's load-balancing interface causes the server to allocate 40 bytes of memory. This amount remains allocated as long as the route entry exists. When the client is no longer sending packets, the route entry is eventually removed due to an ARP entry expiration.

The worst-case cost of balancing the input load is three ARP packets (one request, one reply, and one gratuitous request) per balance period. By default, this period is one second, so the cost for the default would be three packets per second. This overhead does not appear to have any measurable effect.

Network Load Balancing Software Installation and Configuration

This chapter describes how to configure the Network Load Balancing Software, in these sections:

- "Installing the Software" on page 7
- "IP Addresses and the Network Load Balancing Software" on page 8
- "Preparing for Network Load Balancing Software Configuration" on page 9
- "Network Load Balancing Software and the Shell Command File /etc/init.d/network" on page 11
- "Configuring the Network Load Balancing Software and Driver" on page 12
- "Adding Network Load Balancing to a System" on page 13
- "Making the Network Load Balancing Interface the Primary Interface" on page 16
- "Configuring Multiple Network Load Balancing Software Interfaces" on page 19
- "Verifying the Network Load Balancing Software Connection" on page 20

Installing the Software

Use Inst to install the Network Load Balancing Software. (For information on using Inst, see *IRIX Admin: Software Installation and Licensing* in the Technical Publications Library.)

If you are installing the Network Load Balancing Software for the first time, the subsystems marked default are the ones that are installed if you use the *go* command from the Inst menu.

To install a different set of subsystems, use the *install*, *remove*, *keep*, and *step* commands in Inst to customize the list of subsystems to be installed, then choose *go*.

IP Addresses and the Network Load Balancing Software

Your system must have an *IP address* and a *network connection name* for each Network Load Balancing Software (NLBS) network interface and for each device to be assigned to the NLBS.

- IP addresses are usually assigned by a site's network administrator, who obtains them from the Network Information Center.
- The network connection name is entered in the /etc/hosts file*. This name is paired with an IP address. Each network interface must have one network connection name that is unique to the domain and one globally unique IP address. The following example line that might appear in an /etc/hosts file shows an IP address (from the test network, 192.2.0) and a network connection name, followed by its short form (or alias).

```
192.0.2.5 sys2.mrktg.group1.com sys2
```

Generally, network interface IP addresses must always be unique and cannot share a subnet on the same system. (For more information, see information on IP addresses in the latest version of *IRIX Admin: Networking and Mail.*)

However, network devices assigned to a Network Load Balancing Software device can share a subnet address with the Network Load Balancing Software device or the other network devices assigned to it. For Network Load Balancing Software, these addresses must be unique only within the subnet.

The IP addresses for network devices assigned to a Network Load Balancing Software device must be addresses that can legally appear on the wire. Although these addresses are not used or advertised, they appear in ARP packets.

These addresses need not be on the same network or subnet as the Network Load Balancing Software; you can reuse addresses currently in use on other subnets. The IP addresses of the devices configured for load balancing and that of the Network Load Balancing Software need not be related.

8

^{*} For more information about the /etc/hosts file, see the hosts(4) man page and IRIX Admin: Networking and Mail

The network portion of the IP address for each local area network must be unique; you cannot use the same network address for your system's Ethernet and Network Load Balancing connections.

To get IP addresses for the Network Load Balancing Software and the devices you are associating with it, follow these steps:

- Determine which real devices you are associating with the Network Load Balancing Software.
- 2. Obtain an IP address for the Network Load Balancing Software. This address must be unique.
- 3. Obtain IP addresses for the real (Ethernet) devices if they have not already been assigned.
- 4. If necessary, configure the Ethernet devices with *ifconfig*. For information on this command, see its man page, ifconfig(1M).

Note: For specifics, see the sections on planning and setting up a network in the latest version of *IRIX Admin: Networking and Mail*.

Preparing for Network Load Balancing Software Configuration

Before starting Network Load Balancing Software configuration, collect the necessary information and perform housekeeping so that the process goes smoothly. Follow these steps:

- 1. Determine the number of networks to which your system will be connected after the Network Load Balancing Software is installed.
- 2. If the system is to have more than one network connection, decide which will be the *primary network*. The primary network interface should be the one where all or most of the system's network services or clients reside.

Note: If your system must boot over the network, the Network Load Balancing Software connection cannot be the primary network interface.

3. For each network connection, select a network connection name and IP address.

The network connection name of the primary network connection must be the same as the system's hostname. You can display your system's hostname by using the *hostname* command within a shell window:

% /usr/bsd/hostname

You can display the current IP address associated with the network connection name *hostname* by typing one of the following commands in a shell window:

- % /sbin/grep hostname /etc/hosts
- % /usr/bin/ypmatch hostname hosts

The names you create for non-primary network interfaces can be anything you want. To facilitate recognition, the names usually include both the hostname and an indication of the protocol (for example, *nlb-mars* or *nlb2-mars*).

4. Determine if any of your system's network interfaces require special configuration for the *subnetwork mask* (*netmask*), *broadcast address*, and *route metric*.

Table 2-1 summarizes the default operational parameters for the network interfaces.

 Table 2-1
 Default Network Interface Parameters

Parameter	Default	Description
Netmask	No subnet. That is, the bits in the standard network portion of the Internet address are set to 1; the bits in the standard host portion of the Internet address are set to 0 (for class B addresses, 0xFFFF0000; for class C, 0xFFFFFF00).	32-bit value used to create two or more subnetworks from a single Internet address, by increasing the number of bits used as the network portion and decreasing the number of bits used as the host portion. When creating the mask, assign 1 to each network bit and 0 to each host bit.
Broadcast address	For the Internet address family, the host portion of the IP address is set to 1s (for class B addresses, <i>x.x.</i> 255.255; for class C addresses, <i>x.x.x.</i> 255).	Address used by this interface for contacting all systems on the local area network.
Route metric	0	Hop count value advertised by the routing daemon (<i>routed</i>) to other routers. Higher numbers make the route less desirable and less likely to be selected as a route. Settings range from 0 (most favorable) to 16 (least favorable, infinite).
arp	Address Resolution Protocol (ARP) is enabled and used by the interface.	Address Resolution Protocol translates IP addresses to link-layer (hardware) addresses. When this parameter is disabled, the interface does not use ARP.
debug	Disabled	When debugging is enabled, a wider variety of error messages is displayed when errors occur.

If any of these operational parameters needs special configuration, you must create or edit an <code>/etc/config/ifconfig-#.options</code> file, where the pound sign (#) matches the network interface's order in the <code>netif.options</code> file. For example, for the <code>netif.options</code> line <code>if3name=1b0</code>, create or edit the file <code>/etc/config/ifconfig-3.options</code>.

The following example contents of an <code>/etc/config/ifconfig-#.options</code> file specifies a subnetwork mask for the associated network.

netmask 0xFFFFFF80

Complete instructions for configuring operational parameters are provided in the device configuration instructions in *IRIX Admin: Networking and Mail*.

5. Update the site's hosts and ethers databases to include the correct information about this system.

Network Load Balancing Software and the Shell Command File /etc/init.d/network

During system startup and any time it is invoked specifically, the shell command file /etc/init.d/network configures and initializes the network interfaces and software. Some of the script's procedures are accomplished by calling other utilities and reading configuration files. Among other tasks, /etc/init.d/network performs the following:

- Determines the system's hostname, as defined in the /etc/sys_id file.
- Determines the network hardware and interfaces available in the operating system.
 This information can be viewed with the *hinv* command.
- Determines the ordering for the network interfaces.
 - This information is defined in the /etc/config/netif.options file. If the netif.options file has not been altered, the default ordering is configured (as defined in the network script).
- Determines the network connection name for each network interface.
 - This information is defined by the if#name lines in the /etc/config/netif.options file. If the netif.options file has not been altered, the default names (as defined in the network script) are used.
- Determines the IP address for each interface by looking up each network connection name in the /etc/hosts file.

- Determines the settings for each network interface's operational parameters.
 This information is defined in the /etc/config/ifconfig-#.options. If an ifconfig-#.options file does not exist for the interface, the default settings are assigned.
- Configures and starts (enables) the number of network interfaces specified by the *if_num* variable in the network script.
- Invokes /etc/init.d/network.ls, which handles the device-specific configuration of the load-balancing devices.

The results of the *network* script's configuration can be viewed with the */usr/etc/netstat -i* and */usr/etc/ifconfig* commands.

Configuring the Network Load Balancing Software and Driver

This section describes how to configure a new network load-balancing network connection for a system with multiple functional Ethernet connections. Configuration consists of assigning Ethernet devices to the network load-balancing device and performing the normal network device configuration for it. Follow these steps:

1. Edit /etc/config/nlb.options to assign the Ethernet devices for which load balancing is to be performed. This file initially contains the following lines:

```
lblname=lb0
lbldevs=
lb num=1
```

2. To configure one network load-balancing device on the system, set the shell variable *lb1devs* to a comma-separated list of Ethernet devices to be attached to the network load-balancing device lb0. For example, if two Ethernet devices in the system, ef1 and ef2, are to be assigned to the network load-balancing device, *nlb.options* looks like the following:

```
lb1name=lb0
lb1devs=ef1,ef2
lb_num=1
```

Adding Network Load Balancing to a System

This section provides instructions for adding a Network Load Balancing device to a system with previously configured Ethernet devices. In these instructions, one of the existing Ethernet interfaces serves as the primary network interface; the Network Load Balancing interface serves as the secondary network interface.

Note: If your site uses an NIS service, the changes described in this section must also be made to the database on the NIS server.

Follow these steps:

- 1. Make sure you have completed preliminary steps:
 - Obtain network connection names and IP addresses for each Network Load Balancing Software network interface, as described in "IP Addresses and the Network Load Balancing Software" on page 8.
 - Prepare for configuration, as described in "Preparing for Network Load Balancing Software Configuration" on page 9.
 - Configure the software and driver, as described in "Network Load Balancing Software and the Shell Command File /etc/init.d/network" on page 11.
- 2. If necessary, determine your system's hostname:
 - # /usr/bsd/hostname
- 3. As superuser, open the /etc/hosts file and find the line containing your system's hostname. This line configures your system's Ethernet network interface.

If you find the Network Load Balancing Software name or address, verify that the IP address and the name are correct for the Network Load Balancing Software connection.

If you do not find an entry for this name, search for the Network Load Balancing Software IP address. Searching for all instances of the system's hostname usually identifies all the network connection names for the system.

If the line is not correct or is missing, edit the file so that there is a line containing the IP address and network connection name for the Network Load Balancing Software network interface.

A typical format for an entry in /etc/hosts file is as shown:

IPaddress full_network_connectionname alias

For example, a portion of an /etc/hosts file might look like this, where the host sys2 has two entries and sys3 has one:

```
192.0.2.1 sys3.mrktg.group1.com sys3
192.0.2.5 sys2.mrktg.group1.com sys2
192.0.2.8 nlb-sys2.engr.group1.com nlb-sys2
```

- 4. If your system uses more than one network connection, for each one (in addition to Network Load Balancing Software), verify that the name and IP address are the correct; follow instructions in step 3.
- 5. Copy the line containing your system's hostname and place the copy immediately below the original in the /etc/hosts file. This new line configures your system's Network Load Balancing network interface.
- 6. On the new line, change each instance of your system's hostname to nlb-hostname. That is, precede your system's hostname with nlb-.
- 7. On the same (new) line, change the address (digits on the left) to the Network Load Balancing IP address.

For example, the lines for a system with a hostname of sys3 in the domain group1.com looks like this:

```
x.x.x.x sys3.group1.com sys3 #Ether primary x.x.x.x nlb-sys3.group1.com nlb-sys3 #Load Bal. secondary
```

In these entries, each *x* represents one to three digits.

- 8. Save and close /etc/hosts.
- 9. Edit /etc/config/netif.options so that the network startup script can configure the load-balancing devices. Find the first two pairs of lines showing interface names that are not currently occupied.

In the following example, the first two pairs are in use:

```
iflname= ef0
ifladdr=$HOSTNAME
if2name= ef1
if2addr=gate-$HOSTNAME
: if3name=
: if3addr=gate2-$HOSTNAME
: if4name=
: if4addr=gate3-$HOSTNAME
```

In this case, you change

```
: if3name=
```

: if3addr=gate2-\$HOSTNAME

to

```
if3name= lb0
if3addr=nlb-$HOSTNAME
```

Note: Make sure the names or name formats you enter correspond to entries in the */etc/hosts* file.

If this file has, for example, five pairs of entries for five Ethernets already configured on the system, you would add the following two lines below those entries:

```
if6name=lb0
if6addr=nlb-$HOSTNAME
```

Note: You do not need to edit the entries for the existing Ethernet interfaces. Each Ethernet to be attached to the Network Load Balancing device must have an IP address assigned, even though these addresses are not actually used. You might need to change the addresses of client systems previously on the subnets accessed via the Ethernets that are being assigned to the Network Load Balancing interface.

- 10. Save and close /etc/config/netif.options.
- 11. If applicable, edit /etc/config/nlb.options to attach additional Ethernets to the Network Load Balancing device. (See "Configuring the Network Load Balancing Software and Driver" on page 12 for information on this file.)
- 12. Build your changes into the operating system:

```
# /etc/autoconfig
```

Automatically reconfiguring the operating system.

- 13. Turn on the Network Load Balancing Software using *chkconfig*.
- 14. Reboot the system to start using the reconfigured kernel.

Once the system is rebooted, output from *netstat -in* might look like that in Figure 2-2.

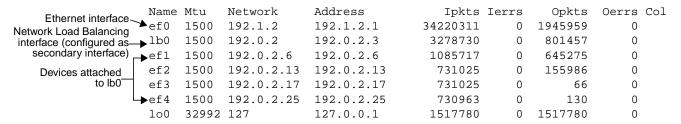


Figure 2-1 Network Load Balancing Device as Secondary Interface Example (netstat -in)

Making the Network Load Balancing Interface the Primary Interface

This section provides instructions for configuring a system so that the Network Load Balancing connection serves as the primary network interface and an Ethernet interface serves as the secondary one.

Follow these steps:

- 1. Open a shell window and log in as superuser.
- 2. If necessary, determine your system's hostname:
 - # /usr/bsd/hostname
- 3. Open the /etc/hosts file and find the line containing your system's hostname. If the file does not contain a line for your hostname, enter it; for example: 192.0.2.1 sys3.mrktg.group1.com sys3
- 4. Copy the line and place the copy immediately below the original.
- 5. On the original line, change the address (the digits on the left) to the IP address for the Network Load Balancing network. This line configures your system's Network Load Balancing network interface. Do not change the hostname.
- 6. On the new line, change the hostname to the new connection name, which is used for the original primary Ethernet (such as gate5-sys3 in the example below).
 - On this new line, do not change the IP address (which represents the Ethernet connection).

The following example shows entries in /etc/hosts for a system with a hostname of sys3 in the domain group1.com:

```
x.x.x.x sys3.group1.com sys3 #NLBS primary x.x.x.x gate5-sys3.group1.com gate5-sys3 #Ether secondary x.x.x.x gate1-sys3.group1.com gate1-sys3 #NLBS attached to 1b0 x.x.x.x gate2-sys3.group1.com gate2-sys3 #NLBS attached to 1b0 x.x.x.x gate3-sys3.group1.com gate3-sys3 #NLBS attached to 1b0 x.x.x.x gate4-sys3.group1.com gate4-sys3 #NLBS attached to 1b0
```

In this example, each x represents one to three decimal digits.

7. Save and close /etc/hosts.

If your site uses an NIS service, the changes described above must also be made to the database on the NIS server.

- 8. Open the /etc/config/netif.options file.
- 9. Edit the file as follows:

iflname= ef0
ifladdr=\$HOSTNAME
if2name= ef1

• Find the first two lines showing interface names that are not currently occupied and edit them to make the existing primary interface into a secondary interface. In the following example, iflname and iflname are currently used:

```
if2addr=gate-$HOSTNAME
: if3name=
: if3addr=gate2-$HOSTNAME
: if4name=
: if4addr=gate3-$HOSTNAME
In this case, you would change
: if3name=
: if3addr=gate2-$HOSTNAME
to
if3name= ef0
```

If this file has, for example, five pairs of entries for five Ethernets already configured on the system, you would make the existing primary interface into a secondary interface by adding the following two lines below those entries:

```
if6name=ef0
if6addr=gate5-$HOSTNAME
```

if3addr=gate2-\$HOSTNAME

■ Edit the entries for the existing primary interface to make the load-balancing device the primary interface: change

```
iflname= ef0
ifladdr=$HOSTNAME

to
iflname=NLBinterface
ifladdr=nlb-$HOSTNAME
```

where *NLBinterface* is the name of the Network Load Balancing interface. For example:

```
if1name=lb0
if1addr=nlb-$HOSTNAME
```

- 10. Save and close /etc/config/netif.options.
- 11. If applicable, edit /etc/config/nlb.options to attach additional Ethernets to the Network Load Balancing device. (See "Configuring the Network Load Balancing Software and Driver" on page 12 for information on this file.)
- 12. Build your changes into the operating system:

```
# /etc/autoconfig
Automatically reconfiguing the operating system.
```

- 13. Turn on the Network Load Balancing Software using *chkconfig*.
- 14. Reboot the system to start using the reconfigured kernel.

Figure 2-2 shows *netstat -in* output with a load-balancing device as the primary Ethernet interface.

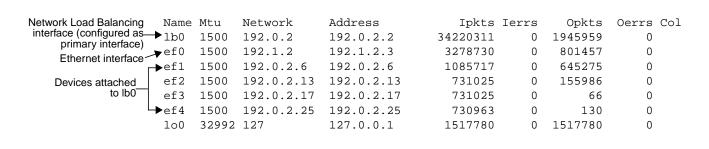


Figure 2-2 Network Load Balancing Device as Primary Interface Example (netstat -in)

Configuring Multiple Network Load Balancing Software Interfaces

To configure multiple network load-balancing devices, follow these steps:

- 1. Follow instructions for configuring and adding the load-balancing device to the system, as described in this chapter.
- 2. Use *systune* to change the tunable parameter, *lb_devices*. Enter

```
systume lb_devices number of devices
```

For example, this command sets the number of load-balancing devices to 2:

```
systune lb_devices 2
```

For information on this command, see its man page, systune(1M).

- 3. Edit /etc/config/nlb_options:
 - Set *lb1name* and *lb2name* to the load-balancing device names you are using.
 - Set the shell variable *lb1devs* to a comma-separated list of Ethernet devices to be attached to the network load-balancing device lb0.
 - Change the shell variable *lb_num* from the default of 1 to the number of network load-balancing devices configurable on the system.
 - Save and close the file.

The following example shows entries for two network load-balancing devices:

```
lb1name=lb0
lb1devs=ef1,ef2
lb2name=lb1
lb2devs=ef3,ef4
lb num=2
```

4. For changes in the *systune lb_devices* parameter to take effect, build your changes into the operating system:

```
# /etc/autoconfig
```

Automatically reconfiguring the operating system.

- 5. Turn on the Network Load Balancing Software using *chkconfig*.
- 6. Reboot the system to start using the reconfigured kernel.

Verifying the Network Load Balancing Software Connection

To verify that the system's Network Load Balancing Software connection is working, follow these steps:

1. List all network interfaces on the system. A Network Load Balancing Software interface, such as *lb*#, must be listed.

```
% /usr/etc/netstat -in
```

If the listing does not include a Network Load Balancing Software interface, the Network Load Balancing Software may not have been installed, may not be configured, or may not be built into the operating system. Fix the problem before continuing.

2. Check that the load-balancing configuration is correct by entering

```
% /usr/etc/lbconfig interfacename
```

where *interfacename* is the name of a load-balancing interface, such as lb0. For example:

```
# lbconfig lb0
Interfaces configured for lb0: ef1,ef2,ef3,ef4
```

Compare the output of this command to the configuration in /etc/config/nlb.options. If it does not correspond, reboot and check again. If it still does not correspond, contact your service provider.

- 3. Use *ifconfig* to get the broadcast address:
 - % /etc/ifconfig interfacename
- 4. Use the broadcast address in a *ping* command, as follows:
 - % /usr/etc/ping broadcastaddress
- 5. Verify that all clients on the subnet respond.
- 6. You can use lbstat -c to further evaluate the state of the load-balancing device(s) on the system. In particular, the state for each Ethernet port attached to the load-balancing device should be LINK_GOOD. If a port does not show this status, check its cabling.

Performance Tuning

Tunable parameters control how well the Network Load Balancing Software distributes outbound packets and how it balances inbound packets. This chapter describes these parameters in these sections:

- "Tuning Parameters" on page 23
- "Changing Input Load Balancing" on page 24

Tuning Parameters

Network Load Balancing Software interface names are as follows:

```
nlb# lb0 lb1 ...
```

The Network Load Balancing Software takes a rolling average of source and destination IP addresses and port numbers: after taking the first average, subsequent averages can be tuned, as explained in "Adjusting the Average Weight" on page 26. The software then applies a hash function and collapses the result to a 32-bit number so that it can send related packets out on the same interface.

The Network Load Balancing Software has tunable parameters for adjusting the hash function multiplier, number of devices to configure, and load balancing. Reset these parameters with *systune*; for information, see its man page, systune (1M).

The Network Load Balancing Software tunable parameters are as follows:

lb_hash_multiplier

Hash function multiplier, which selects the output interface. The default multiplier is 612542337. For example:

systume lb_hash_multiplier 612542337

lb_devices

Maximum number of load-balancing devices to be configured when the system is booted. The default is 1; range is 1 to 10. This example configures the software for three active devices:

systune lb_devices 3

For the new setting to take effect, reconfigure the kernel and reboot.

See "Network Load Balancing Software and the Shell Command File /etc/init.d/network" on page 11 in Chapter 2 for information on setting the number of load-balancing devices.

lb_balance_period

Period in milliseconds used for input load balancing. This variable determines how often the input-balancing algorithm is executed. The default is 1000 ms (that is, one second); the range is 100 to 600,000 ms. See "Adjusting the Balance Period" on page 25 for instructions on adjusting this parameter.

lb_balance_thresh

Threshold for mean deviation in input byte counts above which input load balancing is triggered. The balance period is set in milliseconds. See "Adjusting the Balance Threshold Value" on page 25 for instructions on adjusting this parameter.

lb_average_weight

Weighting in units of 100 used to calculate the rolling load averages for the interfaces. See "Adjusting the Average Weight" on page 26 for instructions on adjusting this parameter.

Changing Input Load Balancing

The Network Load Balancing Software monitors packet count on the physical interfaces and reassigns clients depending on the load. The averages are calculated from the input packet counts for each interface. You can control this process by adjusting these parameters:

- the balance period, which determines when the snapshot is taken for the rolling average
- the threshold for mean deviation in input byte counts: this parameter determines when load balancing is triggered
- the average weight used for calculating the rolling load averages for the interfaces

The rest of this section explains how to adjust these parameters.

Adjusting the Balance Period

The balance period determines how often the input-balancing algorithm is executed. More frequent execution increases overhead proportionally. The default is 1000 ms (that is, one second).

To reset the balance period, set *lb_balance_period* to the desired number of milliseconds (100 to 600,000); for example:

systume lb_balance_period 100000

Adjusting the Balance Threshold Value

The balance threshold value determines when load balancing is triggered; it is the threshold for mean deviation in input packet counts above which input load balancing is triggered. The Network Load Balancing Software calculates mean deviation, which should be less than the balance threshold.

When the mean absolute deviation, expressed as a percentage of the mean input load, exceeds this threshold, the Network Load Balancing Software attempts to balance the load by assigning a client another MAC address.

Setting the balance threshold value to a low number can result in a better input balance. However, such input balance might be impossible to achieve because the software continually moves clients from one MAC address to another.

The default balance threshold is 1%; the range is 1 to 100,000%. For example: systume lb_balance_period 100

Adjusting the Average Weight

The Network Load Balancing Software uses the following formula to calculate the rolling average:

$$S(n) = \frac{k}{100} \times S(n-1) + \left(1 - \frac{k}{100}\right) \times l(n)$$

where

$$0 \le \frac{k}{100} < 0$$

and

S(n) The approximate average of n samples.

l(n) The nth sample.

k The tunable parameter, an integer that is the average weight. Default is 50; range is 0 to 99.

For example:

systume lb_average_weight 35

If k is 0, S(n) is always l(n), the instantaneous load. As k increases, l(n) provides less and S(n-1) provides more of the value of S(n). The calculation uses k% of S(n-1) and 100-k% of l(n).

Troubleshooting

This chapter describes what to do when your load-balancing network connection has problems. The chapter describes the following topics:

- "Initial Remedies" on page 27
- "Connection Has Not Been Functional Since Last Boot" on page 28
- "Connection Has Been Functional in the Recent Past" on page 29
- "High Rate of Packet Loss Occurs" on page 30
- "System Cannot Communicate With Other Systems" on page 31
- "System Does Not Load Miniroot or Boot From the Network" on page 31

Initial Remedies

When you experience difficulty with the load-balancing network connection at a particular system, you can do the following:

- 1. Check the physical connections at the system; consult the owner's guides for the equipment.
- 2. Search or read the /var/adm/SYSLOG file and console window for error messages. If you find any Network Load Balancing Software driver messages, read about them in Appendix B.
- 3. Use **lbstat** -C to evaluate the state of the load-balancing device(s) on the system. In particular, the state for each Ethernet port attached to the load-balancing device should be LINK_GOOD. If a port does not show this status, check its cabling.

Connection Has Not Been Functional Since Last Boot

If the Network Load Balancing Software connection has not been working since the last time the system was booted, or if this is an initial installation of a Network Load Balancing Software product, one or more of the following could be occurring:

- The operating system has not been rebuilt to include the driver.
- The Network Load Balancing Software interface or the aggregated Ethernets have not been configured properly.

Follow the instructions below to determine the cause of the problem:

- 1. Verify that the installed IRIX operating system and Network Load Balancing Software software are the correct versions by doing the following:
 - Determine the correct versions of IRIX and of the Network Load Balancing Software. The Network Load Balancing Software release notes indicate the correct IRIX and Network Load Balancing Software versions.
 - Use the *versions* command to display the installed release identifications:

```
% /usr/sbin/versions eoe.sw.base
eoel date Execution Only Environment 1, 6.5
```

If the version is not correct, install the correct version.

Note: The *eoe.sw.base* subsystem must be 6.5 or later for the Network Load Balancing Software to work.

- 2. View /etc/config/nlb.options; use lbconfig to check for Ethernets attached to the load-balancing devices.
- 3. Use *netstat*, as shown below, to display the currently configured network interfaces.

```
% /usr/etc/netstat -in
```

Figure 4-1 shows example output.

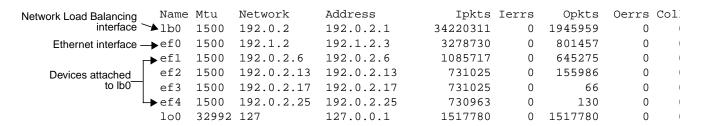


Figure 4-1 Displaying Available Interfaces With netstat -in

If the Network Load Balancing Software interface is not displayed, continue to the next step. If the interface is displayed, but the configuration is incorrect, follow the instructions in Chapter 2 to reconfigure it.

- 4. Verify the Network Load Balancing Software entries in the /etc/config/netif.options file. For example, check the spelling of the network interface.
- 5. Use /etc/autoconfig to rebuild the operating system to include the Network Load Balancing Software driver. Reboot the system to start using the new operating system.
- 6. Invoke *netstat -in* again. If the Network Load Balancing Software interface is still missing, contact your service provider.

Connection Has Been Functional in the Recent Past

Follow these steps to resolve the problem:

- 1. Ensure that the system is using the operating system that was built most recently. Use /etc/autoconfig to rebuild the operating system, and then reboot to start using it.
- 2. Watch the messages on the terminal during restart to verify that each network interface is configured correctly. The messages should look similar to these examples:

```
Configuring 1b0 as sys3
Configuring ec0 as gate-sys3
```

If the load-balancing driver is not mentioned on the terminal during startup, there is a problem with the software.

3. Use *netstat* to display the currently configured network interfaces.

% /usr/etc/netstat -in

If the interface is displayed, but the configuration is incorrect, follow the instructions in Chapter 2 to reconfigure it.

The /etc/config/netif.options file may have an incorrect entry (for example, a misspelled network interface); verify all file contents carefully.

If the interface does not display, the software may be dysfunctional. Contact your service provider.

High Rate of Packet Loss Occurs

If high packet loss is displayed by the *ping* command, follow these steps:

- 1. Confirm that the source and destination IP addresses are on the same subnet. If there was any doubt that they are, enter ping -R, or traceroute to check the route. (For information on the latter command, see its manpage, traceroute(1M).)
- 2. If the source and destination IP addresses are attached to the same switch and *ping* still drops packets:
 - Enter netstat -C on the source and destination systems.
 - Enter ping -f on the source and destination systems to make the problem more obvious.

In the *netstat* output, inspect the packet counts and indications of input errors or output errors while *ping* is running.

If the output shows (approximately) one packet dropped per second, check for defective hardware. Make sure all cables are seated properly.

- 3. If the *netstat* output shows no such errors:
 - Consult the documentation for the switch and its management software.
 - Use statistics tools specific to the protocol you are using (IP, UDP, TCP/IP) to determine drops.
 - Check for defective hardware such as switch ports and problems with the cabling.

System Cannot Communicate With Other Systems

If *ping* does not elicit a response from any system, the /etc/hosts and /etc/config/netif.options files may not have been set up properly. For example, the files may be configuring the Network Load Balancing Software interface with the Ethernet IP address.

Verify that the IP (inet) addresses for all network interfaces are correct. To display the currently configured IP addresses, enter

```
% /usr/etc/netstat -in
```

If the addresses are not correct, follow the instructions in Chapter 2 to reconfigure the Network Load Balancing Software.

If the problem persists, identify other problems, as described in "Verifying the Network Load Balancing Software Connection" on page 20.

System Does Not Load Miniroot or Boot From the Network

Silicon Graphics workstations and servers are capable of loading a small-sized version of the operating system (the miniroot) and booting themselves over the network; however, they are capable of doing this only over Ethernet local area networks (they cannot boot over Network Load Balancing Software networks) that are configured as the primary network interface.

If your system is unable to load the miniroot (or boot over the network), verify that its primary network interface is an Ethernet connection by following these instructions:

- 1. Restart the system from the System Maintenance menu. Do not rebuild the operating system during this restart.
- 2. Log on and open a shell window.
- 3. Use these commands to display the ordering of the network interfaces:

- 4. If the primary interface is an Ethernet (for example, *ef0* or *et0*), the Ethernet network connection may be dysfunctional. See *IRIX Admin: Networking and Mail* for information about Ethernet network connections.
 - If the primary interface is not an Ethernet, go to step 5.
- 5. Configure an Ethernet connection as the primary interface.
- 6. Reboot the system. When the system is up and running, it should be able to load the miniroot over the network and boot from it.

Validated Ethernet Switches

Table A-1 summarizes Ethernet switches that have been specifically validated for the Network Load Balancing Software.

 Table A-1
 Validated Ethernet Switches

Manufacturer	Switch	
Cisco Systems, Inc.	Catalyst 5000 Switching System	
	Catalyst 2900 Fast Ethernet Switching System	
Alteon Networks, Inc.	ACEswitch 110	
	ACEswitch 180	

The Network Load Balancing Software is compatible with any Layer 2 switch.

Error Messages

This appendix lists and explains the error messages that the Network Load Balancing Software prints to the console. Most of the messages refer to SIOCSLSCONF, the ioctl used to configure load-balancing devices.

1bn: link failure detected on attached interface if.

The Network Load Balancing device *n* has detected that the link *if* has failed.

lbn: All links have failed or been disabled.

The Network Load Balancing device *n* has detected the failure of the last configured link, or the last link has been administratively disabled.

lbn: link recovery for attached interface if.

The Network Load Balancing device *n* has detected that the link *if* has recovered (packets have been received on it).

Network load balancing failure detection daemon exiting.

The Network Load Balancing failure detection daemon is exiting, as it normally does when the system is shut down or rebooted. Occurrences of this message at times other than system shutdown or reboot indicate a problem.

 ${\tt SIOCSLSCONF}$ enable/disable: ${\tt lb}n$ not configured.

An attempt has been made to enable or disable an interface on a load-balancing device that has not been configured.

SIOCSLSCONF enable/disable lbn: invalid device count k.

The number of interfaces specified when enabling or disabling an interface on load-balancing device n was less than 1 or greater than MAX_INTERFACES as specified in /usr/include/sys/if_lb.h. The value given was k.

SIOCSLSCONF enable/disable lbn: bad device list, length k, error err.

The device list passed to the load-balancing configuration ioctl SIOCSLSCONF for enabling or disabling an interface on load-balancing device *n* could not be copied from user space. The length of the list is *k*. The error number for the error that occurred during the copy is *err*.

SIOCSLSCONF enable/disable lbn: NULL device name, device k.

The number of entries in the list of devices to be enabled or disabled for the load-balancing device n was less than the number specified in the configuration structure supplied to the SIOCSLSCONF ioctl.

SIOCSLSCONF enable/disable lbn: unit name not found.

The network device whose name is *name* could not be enabled or disabled on load-balancing device *n* because it does not exist.

SIOCSLSCONF enable/disable lbn: unit name not configured.

The unit *name* could not be enabled or disabled on load-balancing device *n* because it is not attached to that device.

 ${\tt SIOCSLSCONF:}\ {\tt lb}{\it n}\ {\tt must}\ {\tt be}\ {\tt down}\ {\tt for}\ {\tt reconfiguration.}$

An attempt was made to reconfigure the load-balancing device *n* before the device was taken down via *ifconfig*.

SIOCSLSCONF: lbn: invalid device count k

The number of interfaces specified when load-balancing device *n* was configured was less than 1 or greater than MAX_INTERFACES as specified in /usr/include/sys/if_lb.h. The value given was *k*.

SIOCSLSCONF: lbn: bad device list, length k, error err

The device list passed to the load-balancing configuration ioctl SIOCSLSCONF for load-balancing device n could not be copied from user space. The length of the list is k. The error number for the error that occurred during the copy is err.

SIOCSLSCONF: lbn: NULL device name, device k

The number of entries in the list of devices to be attached to the load-balancing device n was less than the number specified in the configuration structure supplied to the SIOCSLSCONF ioctl.

SIOCSLSCONF: lbn: unit name not found.

The network device whose name is *name* could not be attached to load-balancing device *n* because it does not exist.

SIOCSLSCONF: lbn: unit name has no primary address.

Every network device must have a primary IP address, which must be assigned before the load-balancing device is configured.

n indicates the load-balancing device being configured when the error was detected. The offending device's name is *name*.

SIOCSLSCONF: lbn: unit name has an invalid subnet mask x.

All network devices to be attached to a load-balancing device must have the subnet mask 0xffffffff, which must be set before the load -balancing device is configured.

n indicates the load-balancing device being configured when the error was detected. The offending device's name is *name*.

SIOCSLSCONF: lbn: unit name error err setting flags.

This error was encountered when attempting to set the device flags for a device being attached to a load-balancing device.

n indicates the load-balancing device being configured when the error was detected. The offending device's name is *name*.

SIOCSLSCONF: lbn: the units are not all of the same type.

All network devices attached to a load-balancing device must be of the same type: 10-Base-T, 100-Base-T, or gigabit Ethernet.

n indicates the load-balancing device being configured when the error was detected.

SIOCSLSCONF: unable to start lbn, error err.

Error err was encountered when attempting to start load-balancing device n.

Unable to allocate load balancing device data.

The kernel had insufficient memory available to allocate the data structures required by the load-balancing devices.

Tell Us About This Manual

As a user of Silicon Graphics products, you can help us to better understand your needs and to improve the quality of our documentation.

Any information that you provide will be useful. Here is a list of suggested topics:

- General impression of the document
- Omission of material that you expected to find
- Technical errors
- Relevance of the material to the job you had to do
- Quality of the printing and binding

Please send the title and part number of the document with your comments. The part number for this document is 007-3895-001.

Thank you!

Three Ways to Reach Us

- To send your comments by **electronic mail**, use either of these addresses:
 - On the Internet: techpubs@sgi.com
 - For UUCP mail (through any backbone site): [your_site]!sgi!techpubs
- To **fax** your comments (or annotated copies of manual pages), use this fax number: 650-932-0801
- To send your comments by **traditional mail**, use this address:

Technical Publications
Silicon Graphics, Inc.
2011 North Shoreline Boulevard, M/S 535
Mountain View, California 94043-1389