Linux NIC Enumeration on HP ProLiant Gen8 Servers

Red Hat Enterprise Linux 6.1

Technical white paper

Table of contents

Abstract

Linux customers expect network device naming to be deterministic and persistent on their server platforms. Persistent device naming means the ordering of network devices remains the same across reboots. Deterministic naming of devices ensures that the network device naming under Linux matches the way the ports are labeled on the server chassis. Ideally, network devices under Linux should be ordered in such a way that the first device label belongs to the embedded network controller even when an add-on network adapter is inserted into the system. This document outlines the HP support for an industry-standard solution implemented in the platform firmware on ProLiant Gen8 servers, and how it can be used under Red Hat Enterprise Linux v6.1 (RHEL6.1) or later.

Introduction

LAN-on-motherboard (LOM) refers to the embedded network controller on a server. On ProLiant Gen8 servers, the LOM ports are typically labeled, bottom to top, right to left, on the back of the server chassis, as shown in Figure 1.

**Figure 1** Quad-port LOM on a ProLiant DL360p Gen8 serve. The ports are labeled right to left

Customers expect network devices to be enumerated in such a way that `eth0` is assigned to the LOM. The assignment of Ethernet names to network devices is not always consistent and this can be usually noticed with the addition of a stand-up NIC to a ProLiant server. For instance, consider a ProLiant ML350p Gen8 server that has been configured with a quad-port 1GbE LOM. Figure 2 displays the Ethernet labels `eth0` to `eth3` corresponding to the four network ports, and their respective MAC addresses.

Figure 2: LOM Ethernet labels on an ML350p Gen8 server
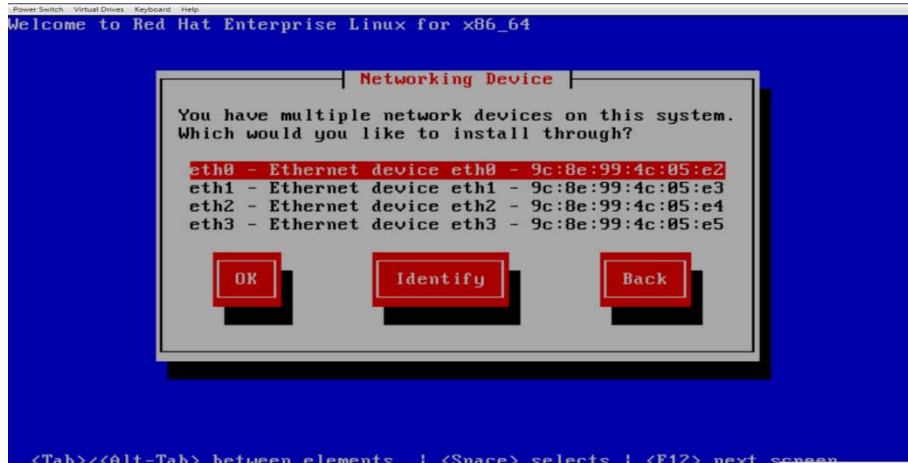
When a dual-port NIC card is inserted into the PCI Express Slot 2 of the server, `eth0` no longer corresponds to the LOM. Instead, as shown in Figure 3, `eth0` is assigned to one of the ports on the add-on NIC card.

Figure 3: The eth0 Ethernet device is no longer associated with the LOM
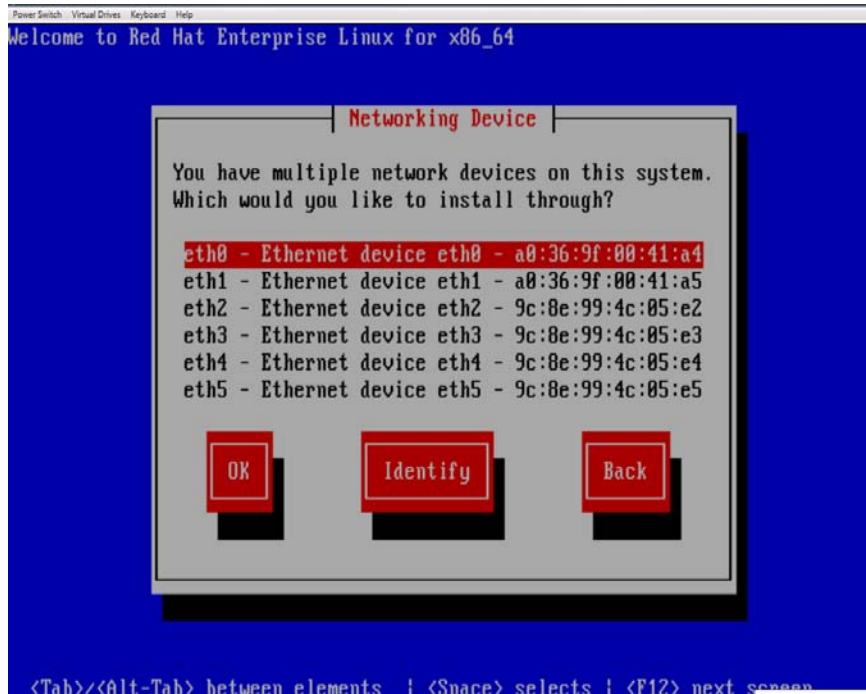
To resolve this issue, HP ProLiant Gen8 servers support the industry-standard SMBIOS-based solution that enumerates the LOM devices in a specific order and labels them suitably so an operating system such as RHEL6.1 can take advantage of it.

SMBIOS-based solution on Gen8

The `type 9` record in the SMBIOS table that is included in the platform firmware on ProLiant Gen8 platforms can help you determine the bus address of a given PCI Express slot in the system.

Example 1: A typical type 9 record in the SMBIOS table

```
# dmidecode –t 9
            Handle 0x0902, DMI type 9, 17 bytes
            System Slot Information
             Designation: PCI-E Slot 2
             Type: x4 PCI Express 3 x8
             Current Usage: In Use
             Length: Long
             Characteristics:
                    3.3 V is provided
                    PME signal is supported
            Bus Address: 0000:07:00.0
```

Example 1 shows that the PCI Express slot 2 belongs to bus address 7, device 0, and function 0. The `In Use` state indicates that an add-on card has been inserted in the slot.

The `type 41` record in the SMBIOS table indicates the ordering of embedded devices on a server platform. Such devices could include network, storage, and video controllers.

Example 2: Snippet of type 41 record entries in the SMBIOS table

```
# dmidecode –t 41

Handle 0x2901, DMI type 41, 11 bytes
Onboard Device
      Reference Designation: NIC Port 1
      Type: Ethernet
      Status: Enabled
      Type Instance: 1
      Bus Address: 0000:02:00.0

Handle 0x2902, DMI type 41, 11 bytes
Onboard Device
      Reference Designation: NIC Port 2
      Type: Ethernet
      Status: Enabled
      Type Instance: 2
      Bus Address: 0000:02:00.1
```

Example 2 shows that LOM ports located in the specific bus addresses are ordered as 1 and 2 as indicated in the `Type Instance` field.

Together the `type 9` and `type 41` SMBIOS records provide a mechanism for the OS to order the LOM devices and the add-on NIC devices consistently and label them suitably. This not only enables you to identify LOMs from add-on controllers but also ensures that the first Ethernet device is always assigned to the LOM. However, for the functionality to exist, the `biosdevname` helper utility must be operational.
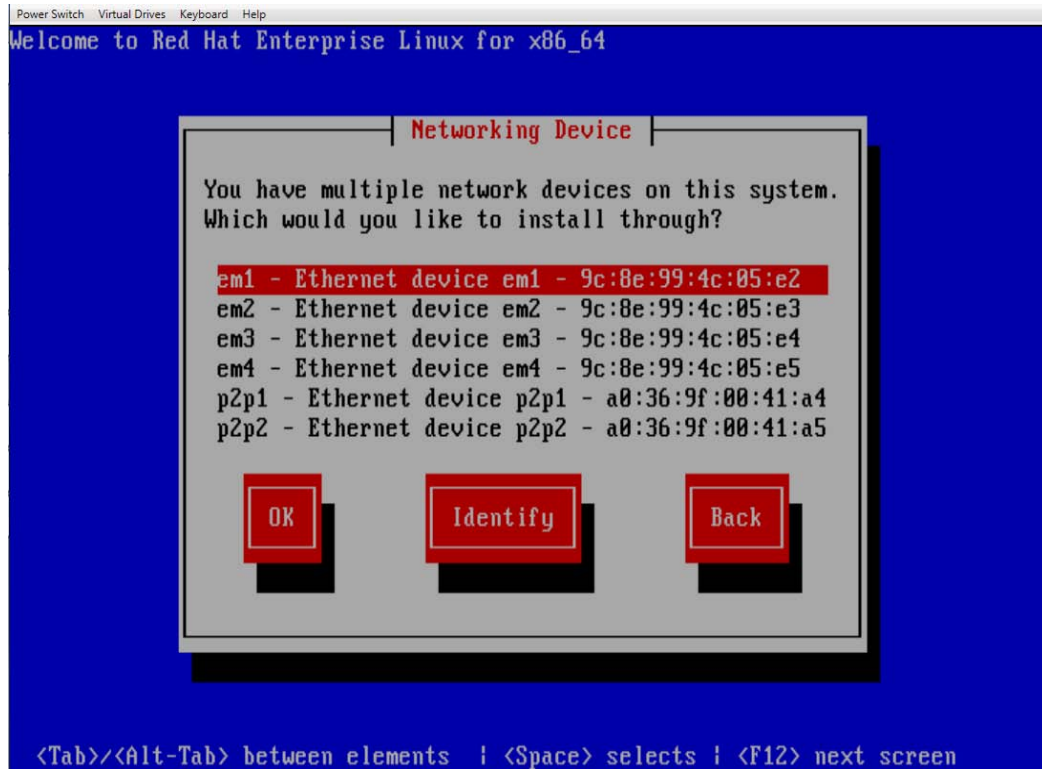
RHEL 6.1 and biosdevname

RHEL 6.1 introduces support for the `biosdevname` utility that brings together the SMBIOS information discussed in the previous section and enables consistent device naming for LOM and add-on NIC devices. This utility is packaged with RHEL and you must select it during the installation. All options except the `Minimal Install` option will install this package by default. You must specify the boot parameter `biosdevname=1` to enable this functionality during the install, as shown in Figure 4.

Figure 4: Enabling biosdevname during install

On Gen8 servers, utilizing `biosdevname` will result in a change in nomenclature for network devices. For example, LOMs will no longer be called `eth0` and `eth1`. Instead, they will be named `em1`and `em2` as shown in Figure 5. The `em` string denotes embedded. Unlike the `eth` names that start from 0, the `em` devices are ordered starting from 1. This matches the chassis labeling of ports that also start at 1 (see Figure 1).

```
Power Switch  Virtual Drives  Keyboard  Help
Welcome to Red Hat Enterprise Linux for x86_64

                 ┤ Networking Device ├

    You have multiple network devices on this system.
    Which would you like to install through?

    em1 - Ethernet device em1 - 9c:8e:99:4c:05:e2
    em2 - Ethernet device em2 - 9c:8e:99:4c:05:e3
    em3 - Ethernet device em3 - 9c:8e:99:4c:05:e4
    em4 - Ethernet device em4 - 9c:8e:99:4c:05:e5
    p2p1 - Ethernet device p2p1 - a0:36:9f:00:41:a4
    p2p2 - Ethernet device p2p2 - a0:36:9f:00:41:a5


         OK          Identify          Back


<Tab>/<Alt-Tab> between elements  ¦  <Space> selects  ¦  <F12> next screen
```

The name `p2p1` denotes port 1 of the add-on NIC card in slot 2. Similarly, `p2p2` denotes port 2 of the card in slot 2. Comparing the MAC address information from Figure 5 and Figure 3, you can see that the LOM devices are consistently enumerated ahead of the add-on ports in Figure 7 when `biosdevname` is in effect. The table below provides the traditional nomenclature and the new naming convention for the LOM and add-on network devices.

Table 1: Notice the old and new method of naming network devices

NIC Type	MAC address	Traditional method	Biosdevname method
LOM	9c:8e:99:4c:05:e2	eth2	em1
LOM	9c:8e:99:4c:05:e3	eth3	em2
LOM	9c:8e:99:4c:05:e4	eth4	em3
LOM	9c:8e:99:4c:05:e5	eth5	em4
ADD ON	A0:36:9f:00:41:a4	eth0	p2p1
ADD ON	A0:36:9f:00:41:a5	eth1	p2p2

The LOM devices get listed based on the ordering provided by the platform in the `type 41` SMBIOS record (see Example 2). The add-on card is enumerated next and is labeled based on the slot in which it is present. This information is gathered from the `type 9` SMBIOS record (see Example 1).

The string used in the `Reference Designation` field in the `type 41` SMBIOS record (see Example 2) is also exposed under `/sys` as `label`. For instance, the label for the device `em1` is shown in the following example:

```
# cat /sys/class/net/em1/device/label
NIC Port 1
```

NOTE: RHEL6.0 does not support the `biosdevname` utility. ProLiant G7 servers and earlier do not support the `type 41` SMBIOS record. Therefore, the above nomenclature is applicable only to RHEL6.1 and later on ProLiant Gen8 servers and only when `biosdevname` is in effect.

CAUTION: Using `biosdevname` results in a fundamental change in how Ethernet devices are named. This change in nomenclature may result in pervasive changes to deployment scripts and associated files in a data center. To avoid this issue, adapt your `kickstart` files suitably. Modify the `ksdevice` directive from `eth0` to `em1`. Several other configuration files, system services, scripts and special user-space tools that typically rely on the `ethX` naming scheme need to be reviewed to ensure that they understand the `emX` nomenclature. For upgrades from RHEL6.0 to RHEL6.1 or later, the older device naming scheme using `ethX` is preserved. If you want to use the new nomenclature outlined in this document, then implementing it for a fresh install is the cleanest solution and is therefore recommended.

Summary

HP ProLiant Gen8 servers are enabled for consistent device naming of network devices. RHEL 6.1 and later releases provide an industry-standard SMBIOS-based device naming solution that interoperates well with the latest generation of HP ProLiant servers. This results in deterministic naming of network devices even when an add-on NIC card is inserted into the system. Prior to enabling `biosdevname` and adopting the new naming convention, you should carefully evaluate your needs and the affect that the new network device nomenclature will have on your existing environments.

For more information

To read more about:

- HP ProLiant Gen8, go to:
  hp.com/go/proliantgen8

- Linux documentation, go to:
  hp.com/go/linux-docs
  Select HP Linux Server Management Software.

Share with colleagues

Get connected
www.hp.com/go/getconnected
Current HP driver, support, and security alerts
delivered directly to your desktop

Become a fan on f »   Follow on twitter »

5900-2216, March 2012