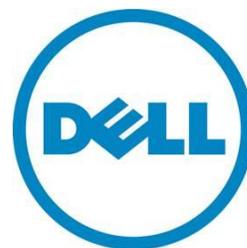

Consistent Network Device Naming in Linux

Narendra K (Narendra_K@Dell.com)

Dell Linux Engineering



This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

© 2012 Dell Inc. All rights reserved. Dell and its affiliates cannot be responsible for errors or omissions in typography or photography. Dell, the Dell logo, and PowerEdge are trademarks of Dell Inc. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell disclaims proprietary interest in the marks and names of others.

July 2012 | Rev 1.4

Contents

Introduction	5
The new naming scheme	5
Firmware information that “biosdevname” utilizes to suggest new names	5
Dell Systems that support the new network naming	6
RHEL 6.1 notes	7
Getting the new naming scheme to work on your system	8
Prerequisites	8
Building and installing biosdevname from source tar ball	8
Installing biosdevname from RPMs	9
Building biosdevname from source rpm	9
New naming built into operating system	9
Enable/disable the new naming scheme	9
During install time	9
Reverting to the traditional ‘ethN’ naming scheme	10
Applications/Scripts using hardcoded ‘ethN’ names	10
Examples:	10
Naming scheme for network interfaces from SRIOV virtual functions	11
Naming scheme for network interfaces from NIC partitions (NPAR)	11
Contribution	12
More Information.....	12
Mailing Lists	12
References	12

Tables

Table 1. Dell systems that support the new network naming	6
---	---

Introduction

Modern x86-based Servers support an increasing number of network interface ports on the motherboard in addition to add-in network adapters. Linux-based OSes name these interfaces as ethN. The naming of network interfaces is currently non-deterministic and not governed by any standard in terms of their relationship to the way the ports are wired on the system. Common user expectations such as 'eth0' representing the first network port on the motherboard as labeled on the server chassis cannot be fulfilled in many cases.

Ensuring that the Ethernet interface names follow the order of the devices as intended by the system designer might not be sufficient. The “ethN” names currently in use do not suggest the Ethernet interface’s physical location, whether it is on the system’s motherboard or if it is on an add-in card; or if it is on an add-in card with multiple ports, which port on the card it is on.

Consequently, a naming mechanism that can impart meaning to the network interface’s name based on the physical location of a network port in concordance to the intended system design is necessary. To achieve this, the system firmware has the ability to communicate the intended order for network devices on the motherboard to the Operating System via standard mechanisms such as SMBIOS and ACPI.

The new naming scheme, using the 'biosdevname' udev helper utility developed by Dell and released under GPL, suggests new names based on the location of the network adapters on the system as suggested by system BIOS.

The new naming scheme

Lan-On-Motherboard interfaces

em<port number>_< virtual function instance / NPAR Index>

(ethernet-on-motherboard <1,2 ..>)

PCI add-in interfaces

p<slot number>p<port number>_<virtual function instance / NPAR Index>

Firmware information that “biosdevname” utilizes to suggest new names

The system BIOS indicates the order of Onboard Network Devices to the OS via SMBIOS type 41 records. The system BIOS provides ‘system slot information’ to the OS via type 9 records. The biosdevname makes use of SMBIOS type 41 to suggest names to the onboard network devices and type 9 records to suggest new names for PCI add-in adapters.

The biosdevname falls back on PCI IRQ Routing Table in the absence of the above mentioned information.

Dell Systems that support the new network naming

Table 1. Dell systems that support the new network naming

System Model	RHEL 6.1	SLES 11 SP2
Dell PowerEdge		
R720	X ²	X ²
R720xd	X ²	X ²
R620	X ²	X ²
T620	X	X
M620	X ²	X ²
M610	X	X
M610X	X	X
M710	X	X
M710HD	X	X
M910	X	X
R210	X	X
R210 II	X	X
R310	X	X
R410	X	X
R415	X	X
R510	X	X
R515	X	X
R610	X	X
R710	X	X
R715	X	X
R810	X	X
R815	X	X
R910	X	X
T110	X	X
T110 II	X	X
T310	X	X
T410	X	X
T610	X	X
T710	X	X
M600	-	-
M605	-	-
M805	-	-
M905	-	-
R200	-	-
R300	-	-
R805	-	-
R900	-	-
R905	-	-
T100	-	-
T105	-	-
T300	-	-
T605	-	-
1955	-	-
1950	-	-
2950	-	-
2970	-	-
6950	-	-
1900	-	-
2900	-	-
Dell Precision		
R5400	-	-
T3500	-	-

Consistent Network Device Naming in Linux

T5500	-	-
T7500	-	-
T1600	X	X
R5500	-	-
M4600	X	X
M6600	X	X
T7600	-	-
T5600	-	-
T3600	-	-
PowerEdge C Series		
C1100	-	-
C2100	-	-
C6100	-	-
C6105	-	-
C6145	-	-

Notes:

1. Systems with '-' under RHEL 6.1 have their system BIOS implementing SMBIOS version 2.5 and 2.4, hence biosdevname will not suggest new names in RHEL 6.1
2. Naming SRIOV virtual functions which belong to the Lan-On-Motherboard devices is not supported.

RHEL 6.1 notes

- In RHEL 6.1, system BIOS should implement SMBIOS Version 2.6 Type 41 and Type 9 records for the biosdevname to suggest new names. biosdevname will not fall back to PCI IRQ Routing Table and new names will not be suggested in the absence of SMBIOS Version 2.6 compliant system BIOS.
- NPAR device naming is not supported in RHEL 6.1. The support for naming NPAR devices is included in RHEL 6.2.

Getting the new naming scheme to work on your system

Prerequisites

The system BIOS should implement SMBIOS Version 2.6 type 41 and type 9 tables. If they are not available, the system BIOS should expose a PCI IRQ Routing Table.

To check if BIOS implements SMBIOS type 41:

```
# /sbin/dmidecode -t 41
```

To check if BIOS implements SMBIOS type 9:

```
# /sbin/dmidecode -t 9
```

To check if BIOS implements PCI IRQ Routing Table:

```
# /sbin/biosdecode
```

Note: RHEL 6.1 requires SMBIOS version 2.6 to use the new naming scheme.

Building and installing biosdevname from source tar ball

biosdevname can be built from sources available for download here:

<http://linux.dell.com/files/biosdevname/permalink/>

git://linux.dell.com/biosdevname.git

1. Prerequisites: pciutils-devel and zlib-devel
2. Build and install biosdevname

```
$ tar zxvf biosdevname-<latest version>.tar.gz  
$ cd biosdevname-<latest version>  
$ ./configure  
$ sudo make && make install
```

3. Remove /etc/udev/rules.d/70-persistent-net.rules

This file contains existing interface names for persistency across reboots. If this file exists, the biosdevname will not kick in.

4. Rename ifcfg-ethN files to ifcfg-emN and ifcfg-p*. The content of these files is distribution specific. Please modify the fields that contain the interface name to reflect the new names

As an example ifcfg-ethN files in RHEL 6.1 contain a 'DEVICE=' field which will be assigned with the ethN name. Please modify it to suit the new naming scheme such as 'DEVICE=em1'

If the ifcfg-ethN file embeds interface hardware address, please comment it as biosdevname suggests a name based on the location of the adapter on the mother board

5. Reboot the system

Installing biosdevname from RPMs

The biosdevname can be installed using RPMs built for the distribution you are using. For example:

Fedora 15:

```
# rpm -ivh biosdevname-0.3.8-1.fc15.x86_64.rpm
```

RHEL 6.1:

```
# rpm -ivh biosdevname-0.3.6-11.el6.x86_64.rpm
```

Building biosdevname from source rpm

The biosdevname rpm can be built from the source rpm from:

<http://linux.dell.com/biosdevname/>

Pre-requisites: pciutils-devel and zlib-devel

```
# rpm -ivh biosdevname-<latest version>-1.src.rpm
```

```
# rpmbuild -bb <path-to-biosdevname.spec file>
```

On Fedora-based distributions, this is under ~/rpmbuild/SPECS. On SuSE based distributions, this is /usr/src/packages/SPECS

```
# rpm -ivh <path-to-biosdevname-binary-rpm>
```

New naming built into operating system

If the OS has biosdevname support built in, then the new naming scheme is enforced by default during install time and run time on supported Dell systems

For a list of Linux Distributions where biosdevname is integrated into the installer, see Table 1.

Enable/disable the new naming scheme

If you choose not to use the new naming scheme, it can be turned off during install time or post-install time.

During install time

To disable the use of the new naming scheme, during installation (attended or automated), pass the kernel command line parameter biosdevname=0 on the boot command line. The parameter should be passed on the boot command line after installation to ensure that a new network adapter plugged in post installation has a traditional 'eth' name.

Note: On SLES 11 SP2 biosdevname=0 needs to be passed to boot command line only during installation.

Certain distributions (see Table 1) have enabled the new naming scheme on all Dell systems by default and disabled for others. If you choose to use the new naming scheme on non-Dell systems with such distributions, set biosdevname=1 in the kernel command line during install.

Reverting to the traditional ‘ethN’ naming scheme

In case you have a system with the new naming scheme and would like to revert to the traditional naming scheme:

1. Rename all ifcfg-emN and ifcfg-p* configuration files and modify the contents of the files accordingly. The content of these files is distribution specific (see /usr/share/doc/initscripts-<version> for details)

As an example ifcfg-ethN files in RHEL 6.1 contain a ‘DEVICE=’ field which will be assigned with the emN name. Modify it to suit the new naming scheme such as ‘DEVICE=eth0’.

2. Comment the HWADDR variable in the ifcfg-eth* files if present as it is not possible to predict here which of the network devices will be named eth0, eth1 etc.
3. Reboot the system
4. Login and you should see the ethN names.

The **70-persistent-net.rules** file will be created under **/etc/udev/rules.d/** to make the names persistent across reboots.

Applications/Scripts using hardcoded ‘ethN’ names

If an application expects network interface names to be of the form ethN, then it might fail. Please report it via the distribution’s bug tracking system.

Examples:

On a PowerEdge R710 with

- Four Broadcom BCM 5709 Lan-On-Motherboard ports
- One Intel 82572EI single port network adapter on PCI slot 4
- One Intel 82576 dual port network adapter on PCI slot 3

The naming of the network interfaces would be:

```
[root@per710~]# ls /sys/class/net/  
em1 em2 em3 em4 lo p3p1 p3p2 p4p1
```

Using the new naming scheme

```
[root@per710 ~]# ifconfig em1
```

```
em1    Link encap:Ethernet  HWaddr 00:24:E8:2E:DE:FF
       inet addr:172.16.66.166  Bcast:172.16.127.255  Mask:255.255.192.0
```

...

```
[root@per710 ~]# ethtool -i p3p2_0
```

```
driver: igbvf
version: 1.0.8-k0
firmware-version: N/A
bus-info: 0000:07:10.1
```

Name for interfaces of Intel 82572EI single port network adapter on PCI slot 4

p4p1

This conveys that the interface is from the network adapter on PCI slot 4 and port 1

Naming scheme for network interfaces from SRIOV virtual functions

The interface names for virtual functions exposed by SRIOV adapters, like the Intel 82576 network adapter, would be as follows (assuming that we enable two virtual functions per physical function):

```
[root@per710 ~]# ls /sys/class/net/
```

```
em1 em2 em3 em4 lo p3p1 p3p1_0 p3p1_1 p3p2 p3p2_0 p3p2_1 p4p1
```

Physical function interface name	p3p1	p<slot 3>p<port number 1>
Virtual function interface names exposed by 'p3p1'	p3p1_0	p<slot 3>p<port number 1>_<virtual function instance 0> Virtual function instance 0 of port 1 of PCI add-in network adapter on SLOT 3
	p3p1_1	p<slot 3>p<port number 1>_<virtual function instance 1> (Virtual function instance 1 of port 1 of PCI add-in network adapter on SLOT 3)

Naming scheme for network interfaces from NIC partitions (NPAR)

The interface names for NIC partitions exposed by NPAR-capable devices, like the BCM 57800 NDC (Network Daughter Card) with two NPAR-capable 10G ports, on a PowerEdge R620 would be:

```
[root@PowerEdge-R620 ~]# ls /sys/class/net/
```

```
em1_1 em1_2 em2_1 em2_2 em3 em4 lo p3p1 p3p2
```

NPAR disabled	NPAR enabled
em1	em1_1 - em<port 1>_<NPAR Index 1> em1_2 - em<port 1>_<NPAR Index 2>
em2	em2_1 - em<port 2>_<NPAR Index 1>

em2_2- em<port 2>_<NPAR Index 2>

Contribution

Your contribution in making biosdevname achieve its purpose comprehensively would be invaluable.

- Test biosdevname and report bugs on the bug tracking systems corresponding to the distribution
- “biosdevname -d” provides extensive debugging information that can assist with bug resolution
- Contribute to enabling 'biosdevname' by fixing applications that depend on hardcoded eth names

More Information

- Biosdevname manpage
- Migrating from Conventional Device Names to Consistent Device Names for Ethernet Devices in RHEL6.
- http://linux.dell.com/files/whitepapers/consistent_network_device_naming_migration_guide_rhel_6.pdf

Mailing Lists

Queries relating to ‘biosdevname’ can be discussed on <https://lists.us.dell.com/mailman/listinfo/linux-poweredge>

linux-hotplug mailing list

linux-hotplug@vger.kernel.org

References

Presentation at Texas Linux Fest - April 2011

<http://marc.info/?l=linux-hotplug&m=128892593821639&w=2>

Presentation at Linux Plumbers Conference - 2010

<http://marc.info/?l=linux-hotplug&m=128892593821639&w=2>

Fedora Feature Description

<http://fedoraproject.org/wiki/Features/ConsistentNetworkDeviceNaming>

Fedora Test Day - January 2011

https://fedoraproject.org/wiki/Test_Day:2011-01-27_Network_Device_Naming_With_Biosdevname

SMBIOS v2.6 Reference Specification

Consistent Network Device Naming in Linux

http://www.dmtf.org/sites/default/files/standards/documents/DSP0134_2.6.0.pdf

ACPI DSM for Naming a PCI or PCI Express Device Under Operating Systems

http://www.pcisig.com/members/downloads/pcifw_r3_1_13Dec10.pdf