Network Device Naming

Matt Domsch

Technology Strategist, Office of the CTO

Which is eth0?

Office of the CTO

Chassis Labels

Office of the CTO
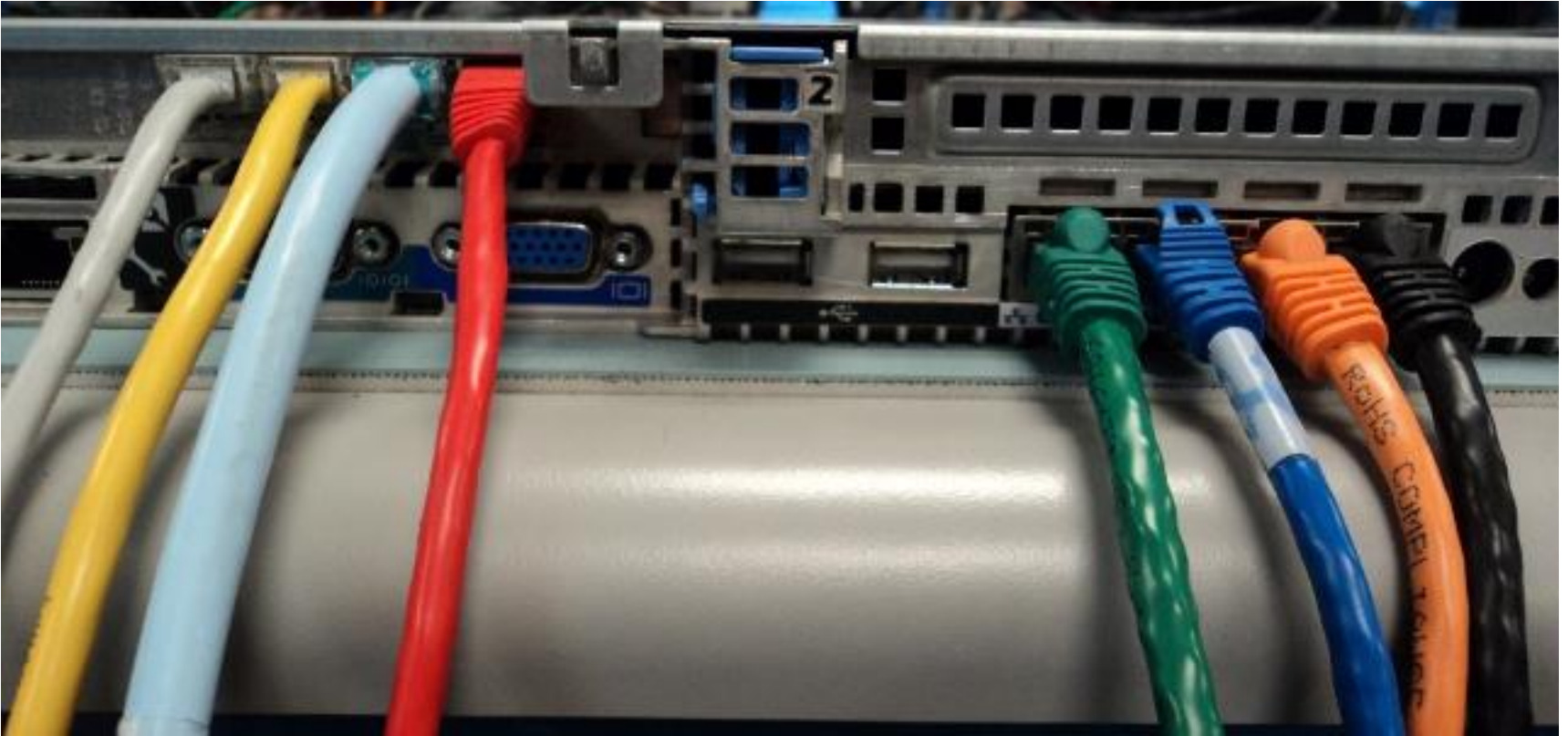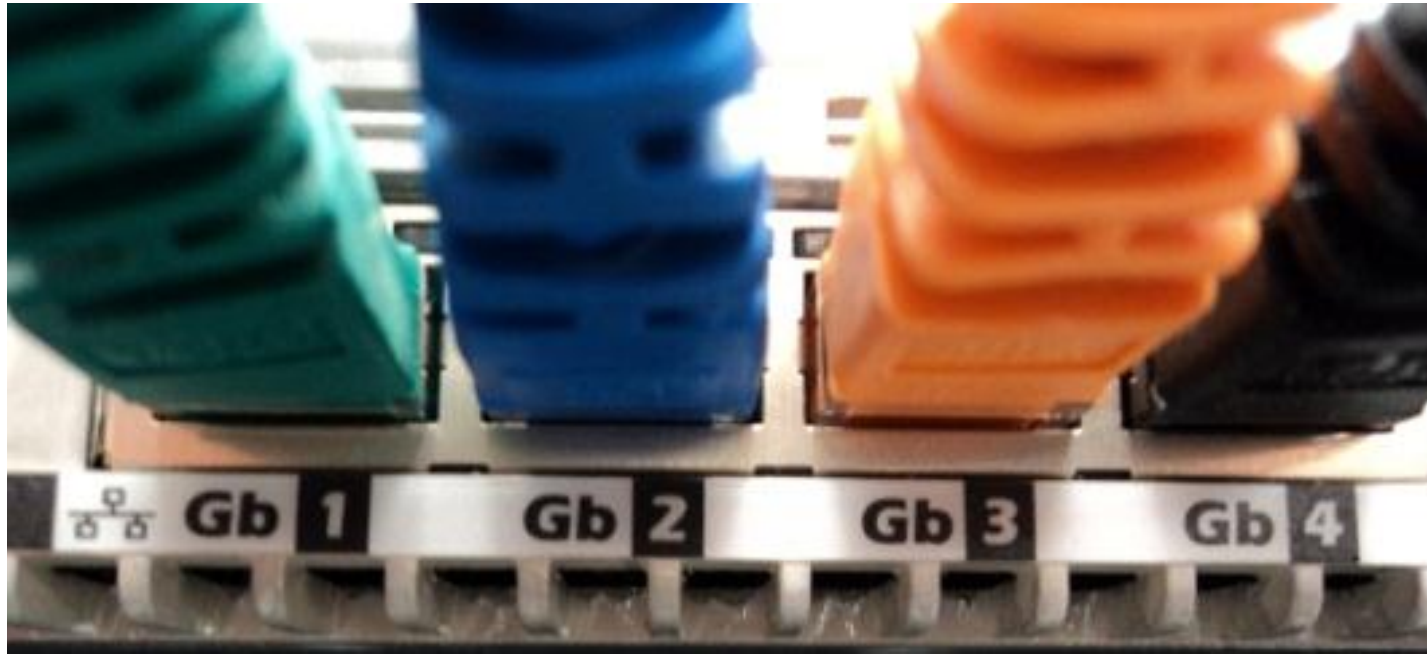
Problem: Enumeration != Naming

- System Administers expect that names are deterministic. Enumeration is non-deterministic.
  - Racing modprobes (hack: serialized by udev in RHEL6)
  - Racing firmware loading
  - Racing udev renames (claimed fixed Sept'09, but failures still reported)
  - PCI depth-first discovery varies based on motherboard layout

- Undefined relationship between chassis label and kernel name

- Other problems we've tried to tackle:
  - Names limited to 15 characters
  - Only one name for each device allowed
  - Existing methods add state (udev MAC->name mappings)
  - Works for all devices, embedded, add-in, USB, …

Current Solutions^WHacks

- ifcfg-eth0 and 70-persistent-net.rules
  - introduces state to stateless systems

- pci=bfsort
  - Whitelists, assumes PCI bus topology is meaningful

- Force sane PCI Bus layout
  - Breadth-first vs. depth-first list
  - PCI Express is switched, not a bus hierarchy anymore
  - Adds significant hardware cost

- PXELINUX + ks=bootif
  - Only works for 1 NIC, and only if you use PXE

We need a better solution.

Standards!

- HP added NIC ordering information to SMBIOS vendor fields
  - Displayed by dmidecode, but not otherwise used

- Dell added labels to device objects in SMBIOS 2.6 standard
  - PowerEdge 10G, 11G implement this

- Dell proposed labels and index in ACPI _DSM in PCI Firmware Spec (to be released soon)

Great, so now we have PCI device to BIOS label mappings. Let's use them!

Office of the CTO

Proposed (and rejected) ideas

1. Rename devices using biosdevname udev rules (userspace)

2. /dev/net/eth* device nodes for network devices (kernel + userspace)

3. Make OS installers assign names (userspace)

4. Libnetdevname (userspace)

5. Reserve the first N ethN names for LOMs (kernel)

Which brings us back full circle...

Current Proposal

- Use BIOS-provided indexes in udev rules to rename LOMs
  - lom[1-N] as BIOS reports on-board NICs (# matches chassis labels)
  - ethX for all others (no change for PCI slot/USB/others)

- Implementation needs some help
  - udev rules
  - Use SMBIOS 2.6 info on kernels >= 2.6.36, PCI DSM info later
  - biosdevname on kernels < 2.6.36, and for all IBM, HP and older Dell servers
  - methods to ask user if they want it (installers, post-install)

- NetworkManager to display BIOS labels as informational text

"The perfect is the enemy of the good" – Voltaire

Thanks

- Greg Kroah-Hartman
- David Miller
- Alan Cox
- Stephen Hemminger
- Kay Sievers
- Narendra K
- Jordan Hargrave
- Peter Jones
- Jon Masters
- Bill Nottingham
- Ben Hutchings
- Dann Frazier

- Marco d'Itri
- Bryan Kadzban
- Scott James Remnant
- Dan Williams
- John Linville
- Sujit K M
- David Zeuthen
- Bill Fink
- Rob Townley
- Kurt Van Dijck
- Hannes Reinecke
- Karl O. Pinc

- Alex Chiang
- Bjorn Helgaas
- Neil Horman
- Thomas Cantwell
- Chip Webb
- Mukund Khatri
- Charles Rose
- Richard Chan
- Vijay Nijhawan
- Vinod Kutty
- Jesse Barnes

Office of the CTO