

Green IT Technical Survey

MIRACLE LINUX CORPORATION

Naohiro Ooiwa / Ichiro Suzuki

2009/04/08

Contents



- Introduction
- Green IT standards / organizations
- Functionalities on Linux
- Real-world measurement
- Future ideas
- Summary

Who are we? What do we do?



- An IPA project: Green IT on Linux
 - Japanese government sponsored project
 - Miracle Linux + LF-jp
- Information-technology Promotion Agency
 - Providing technology and resources for IT in Japan
 - IT Security / Software engineering / Open source etc.
- Miracle Linux Corp.
 - Linux distributor (Asianux)
- Linux Foundation Japan
 - you know...

Project Goals



- Finding problems in Green-IT on Linux/OSS
 - The report will be used as a base-line for IPA's planning.
- Sharing the findings within the community
 - Encourage developers
 - Make users aware of Green IT

Four Steps



- **Survey of Green-IT standards and organizations**
- **Survey of Linux functionalities to reduce power consumption**
- **Evaluation of these functionalities**
- **Improvement ideas**
 - Driver improvements to reduce power consumption of the e1000 NIC

In this session...



- We want to ...
 - Tell you our findings
 - Get your feed-backs
 - Discuss solutions / ideas

Standards and Organizations



- Standards
 - ENERGY STAR
 - EPEAT
- Organizations
 - The Green Grid
 - The Climate Savers Computing Initiative

ENERGY STAR



- A program to reduce power consumption for personal desktops/laptops, not for servers.
- Required power saving modes
 - Off (ACPI S5)
 - Sleep (ACPI S3)
 - Idle
- Power management is also required
 - WOL



Linux has already **satisfied** the requirements

EPEAT



- Environmental standard in U.S. to evaluate desktops, laptops, and monitors
- It has 51 evaluation items.
 - material / energy saving / packaging etc.
- The result is categorized into Gold, Silver, and Bronze.
- Requiring ENERGY STAR on software part.



Linux has also satisfied this standard

The Green Grid



- Global consortium for data center and information center
- Provides wide range of white papers
 - Analysis of power distribution, power efficiency metrics, chassis configuration, recommend ways to save power ...



PUE is good, but hard to compare without detailed conditions
Microscopic indicator is useful for software developers

The Climate Savers



- NPO that consists of companies and consumers who have high environmental consciousness.
- Detailed instructions to reduce power consumption for Windows and Mac.
 - How to power off display / stand-by / hibernate when a computer is idle



Comparable documents for Linux would be nice

Linux Functionalities



Battery Life Toolkit

Lm_sensors

IPMI tools

Powertop

Hibernation

(swsusp, uswsusp, TuxOnIce)

Tickless idle

Power Policy Manager

mc/smt power saving policy

Virtualization

Power QoS

Device Bus Power

Management

Display and Graphics Power
Saving

Wake On Lan

WiFi

Many of these found in LessWatts

Our impressions (1/2)



- Major functions have been implemented on Linux

Functions / OSS	Linux	Windows
Hibernation	○	○
Suspend	○	○
Power Management	○	○
Monitoring by IPMI	○	○
Monitoring by dedicated controllers	○	○
Virtualization	○	○

Our impressions (2/2)



- But,
 - There is no de-facto standard benchmark tools in OSS.
 - Interface to enable functions are not always easily accessible for users. Can't control uniformly.
 - Open source management software to govern entire functions to reduce power?

Bench Mark Tool



- De-facto bench mark software is important for H/W vendor, S/W developer, and Users.
- Indicator is necessary to improve power consumption.
- Commercial software: SPECpower.
- The OSS like SPECpower is needed?
 - Or other kind of bench mark is needed?

Uniform Control



- Some functions to reduce power are implemented, but the interface is mainly CUI-base such as `/proc/xxx`, `/sys/xxx`
- This is not easy for users
- We think it is better if GUI tools are provided.
 - And it's available from Desktop Environment (gnome, KDE, ...)

Power Consumption Management



- For administrators, set and monitor configuration and status of a lot of nodes are important.
 - JP1 can check status of these functionalities.
- Integrated with OSS cluster monitoring software such as nagios, Zabbix, Hinemos etc.?

Measurements



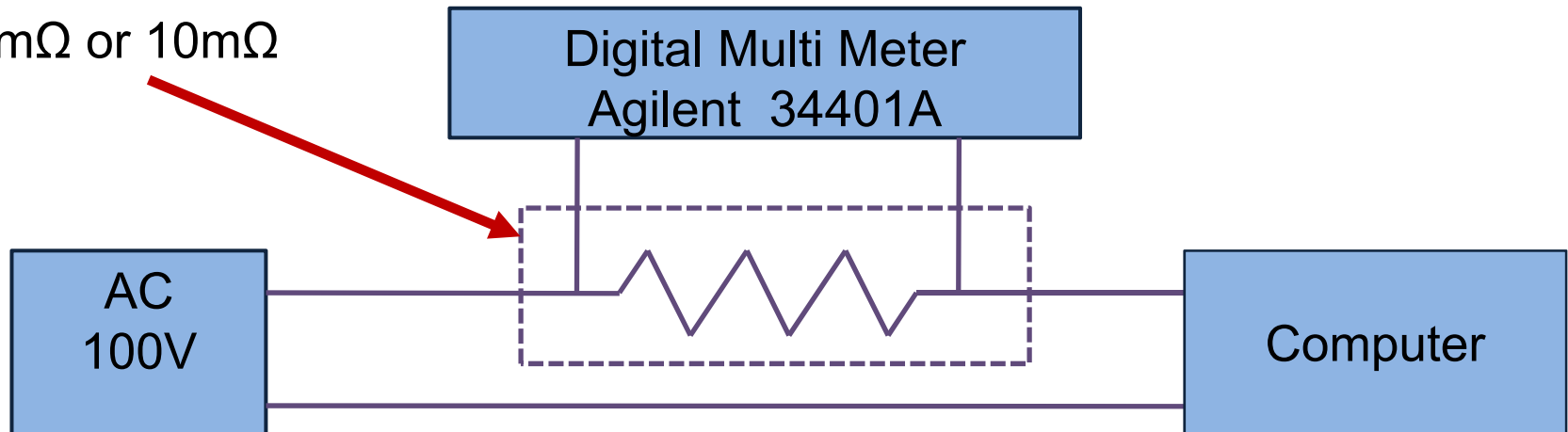
- Measured the effects of the functions to reduce power consumption
 - to see the effects quantitatively
 - not to get the precise reduction value that depends on the machines and H/W configurations.

Setup

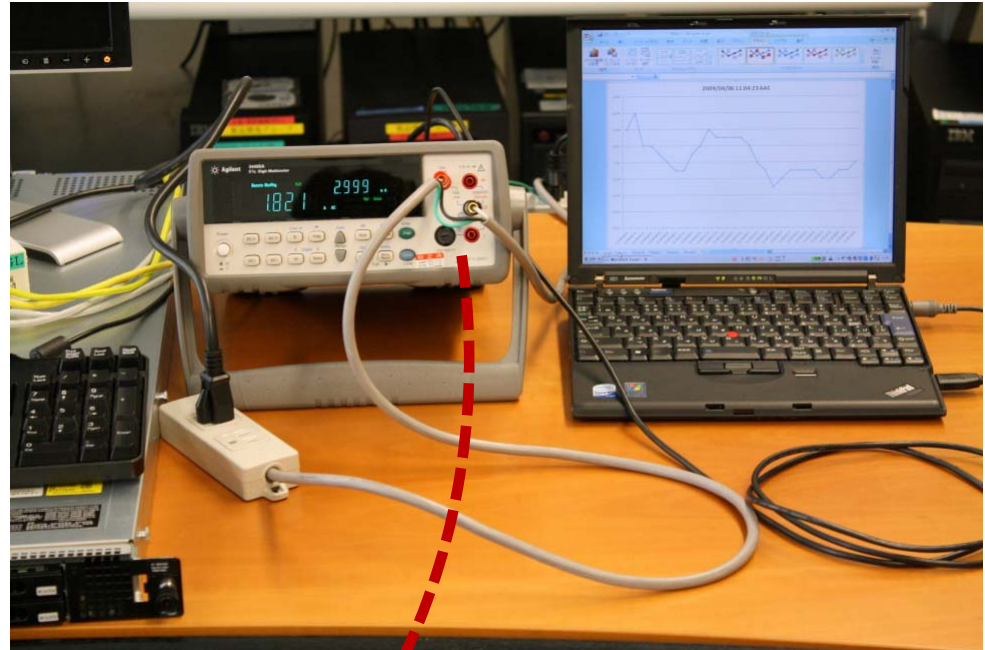


- We measured **total current** by a shunt resistor.
 - Not exactly the same as power consumption, but a good indicator. ($\cos\theta$ ignored)
 - Easy to measure

Shunt resistor
1m Ω or 10m Ω



Setup (cont'd)



Do the Next, Open your Window

MIRACLE

Target Computers



- **Server**

- DELL PowerEdge1950

- Intel Xeon 5460 (3.16GHz) x 2
 - Memory 2GB
 - HDD(SAS) 73GB (Seagate ST973402SS) x 2

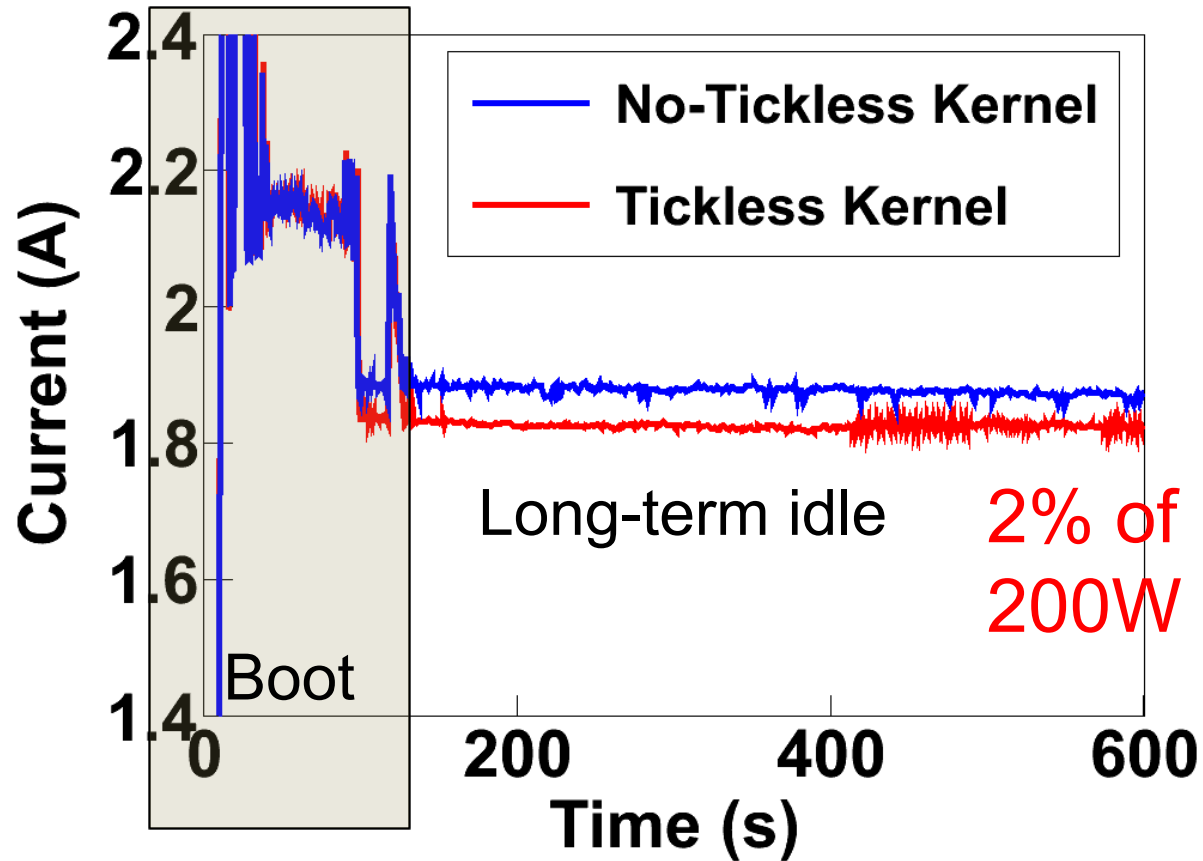
- **Laptop**

- Panasonic CF-W4

- Intel Pentium-M 1.2GHz x 1
 - Memory 768MB
 - HDD(ATA) 40GB (TOSHIBA MK4025GASL) x 1

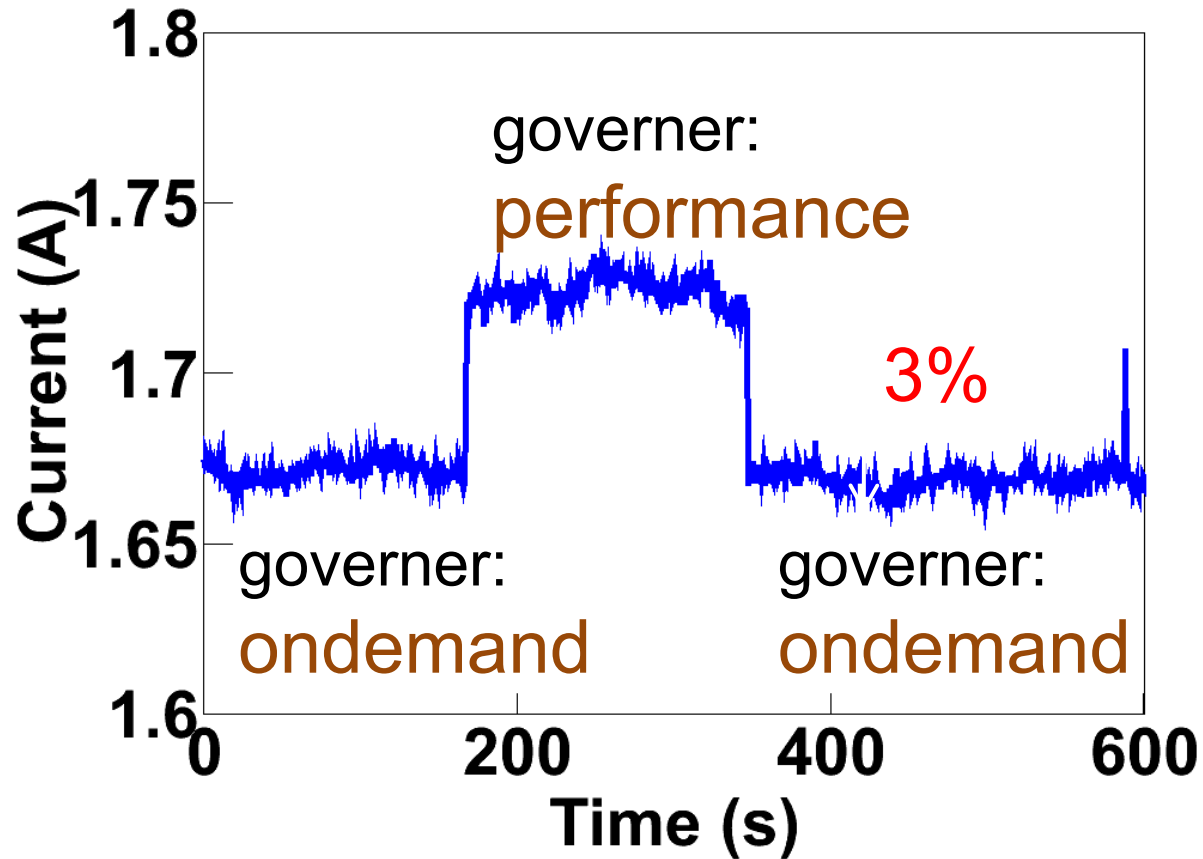
- **Focusing on idle state**

Tickless Idle



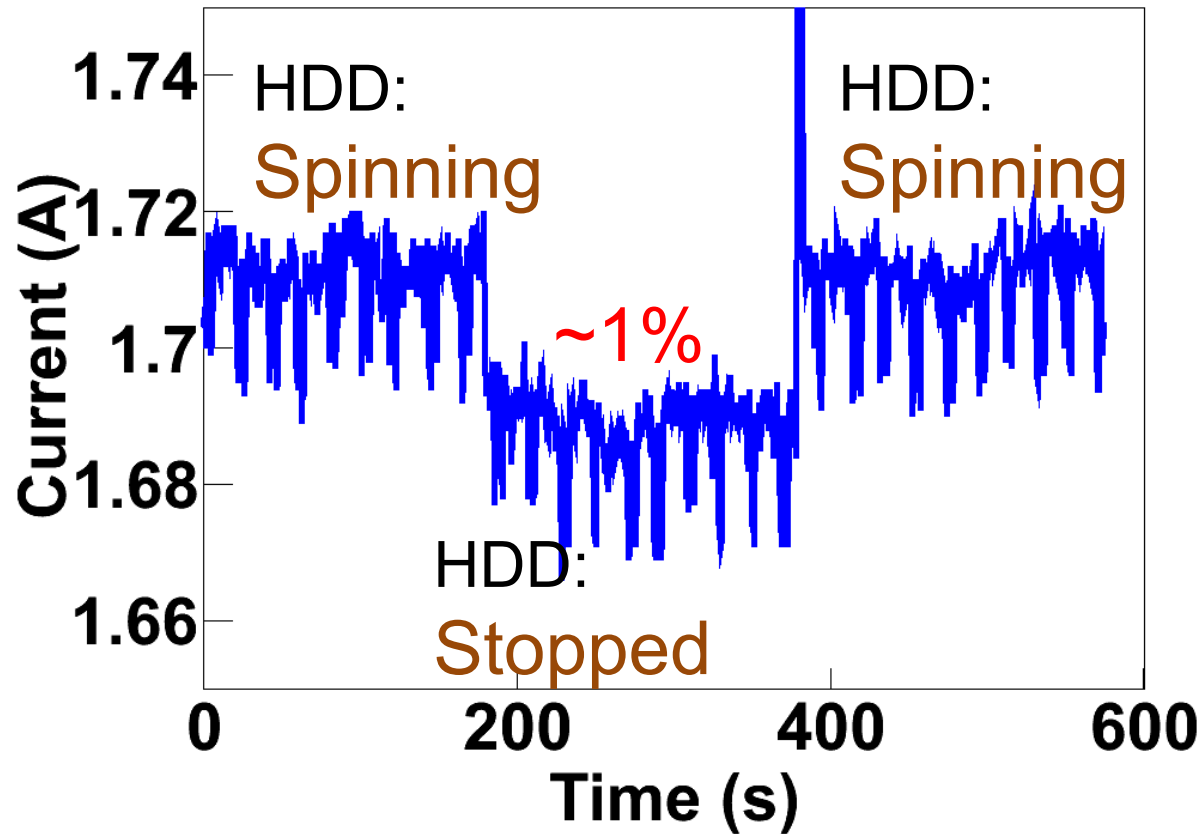
PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1

Processor Power Management (cpufreq)



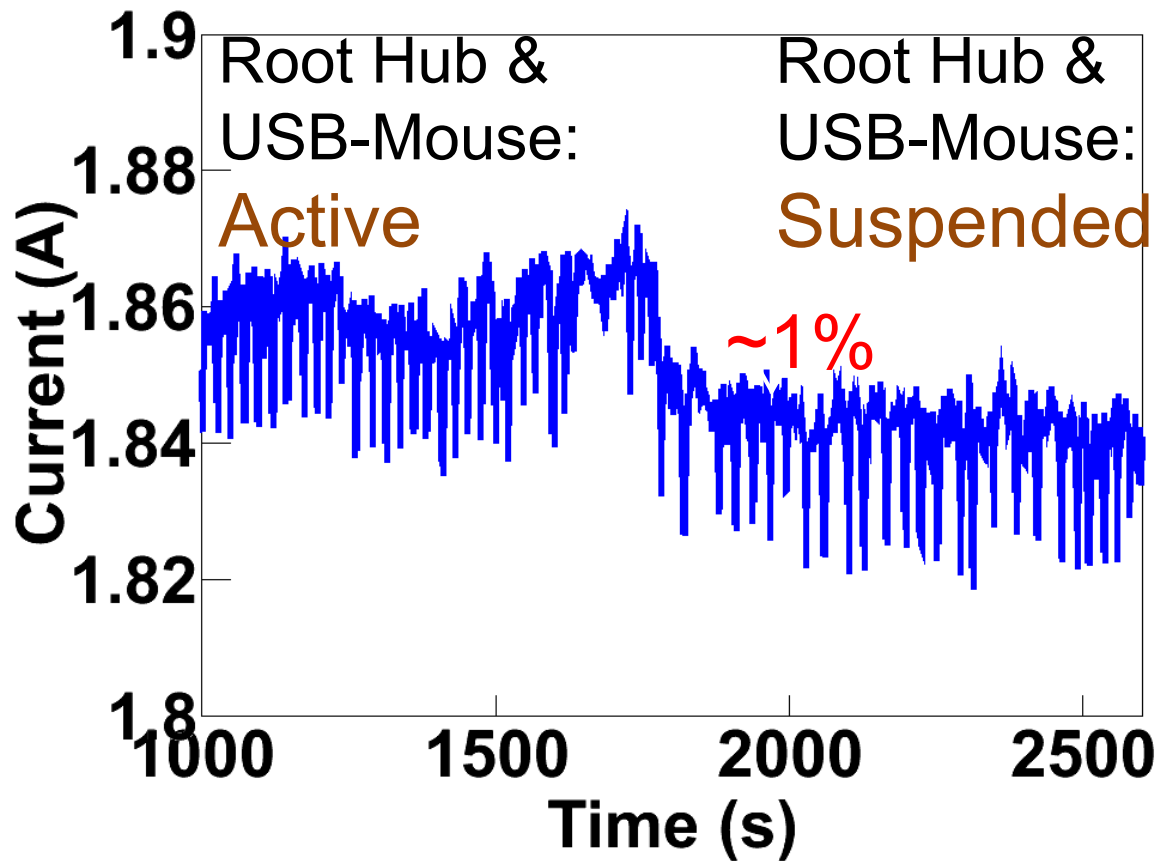
PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1

HDD spin down (sdparm/hdparm)



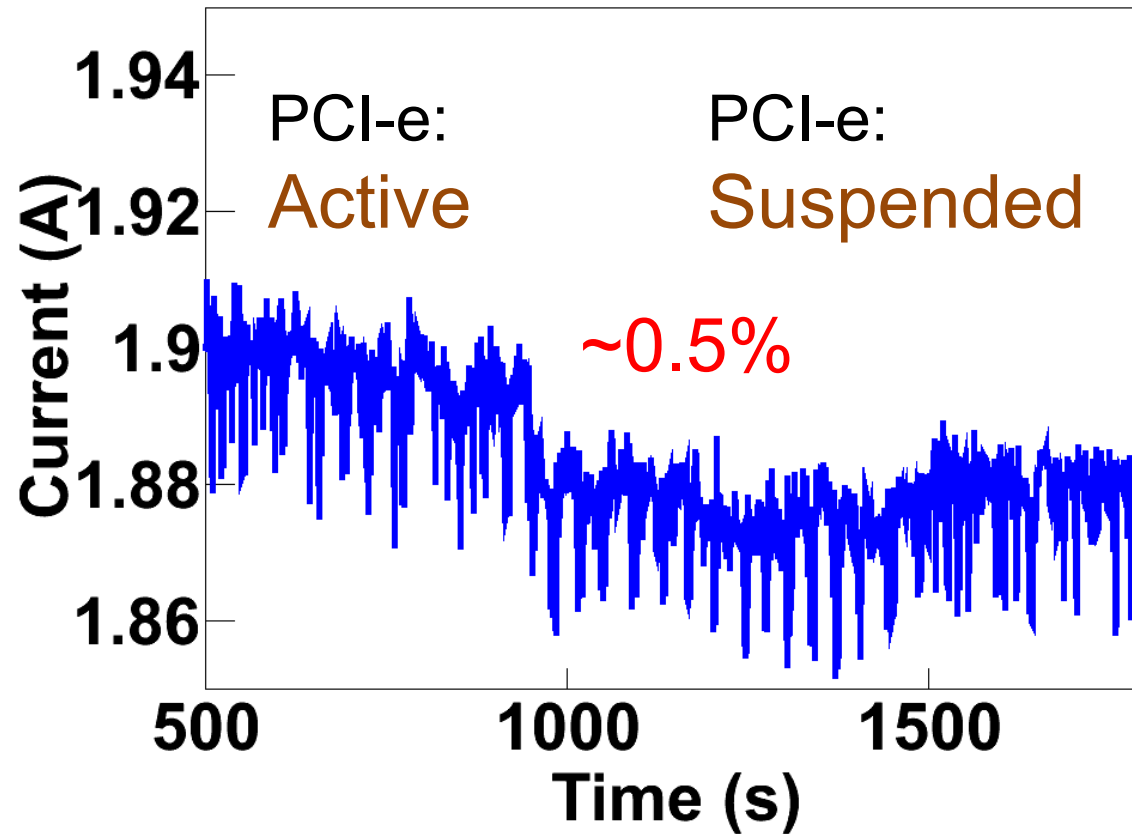
PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1

Bus Power Management (USB)



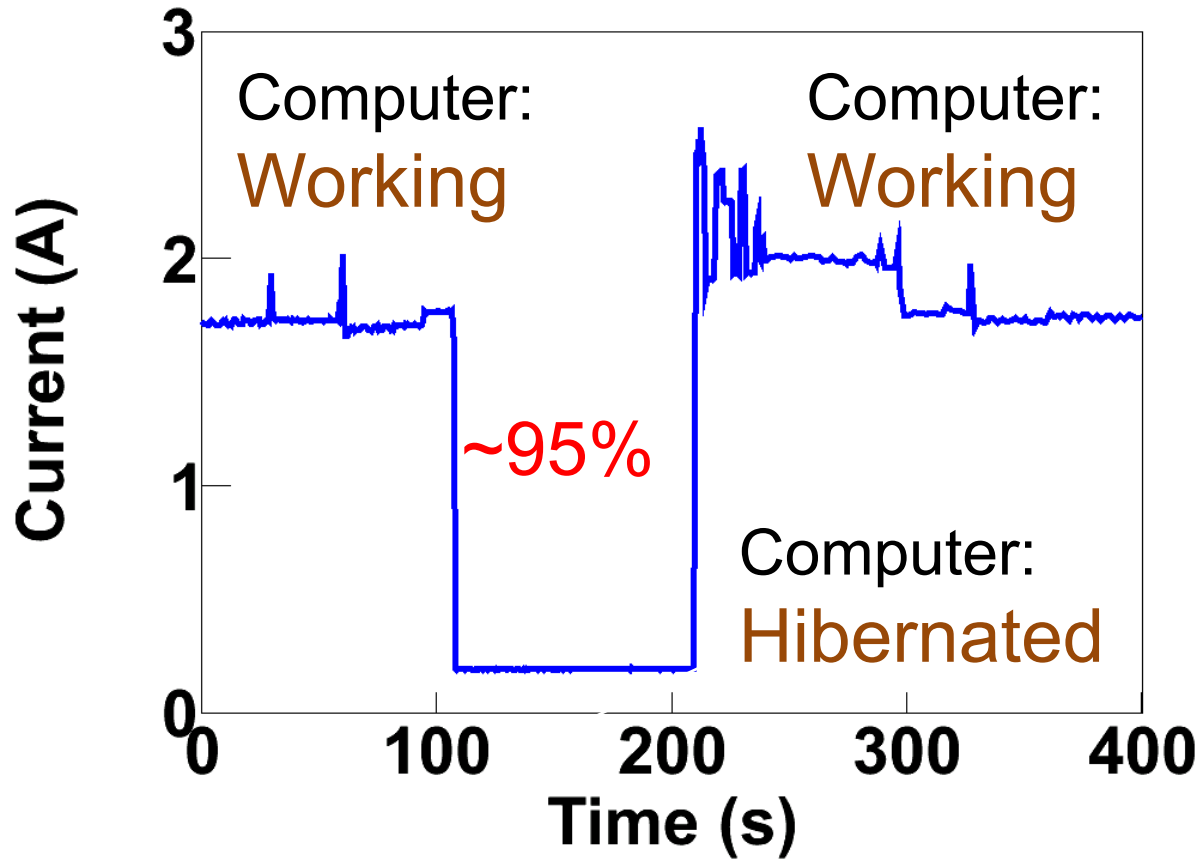
PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1

Bus Power Management (PCIe)

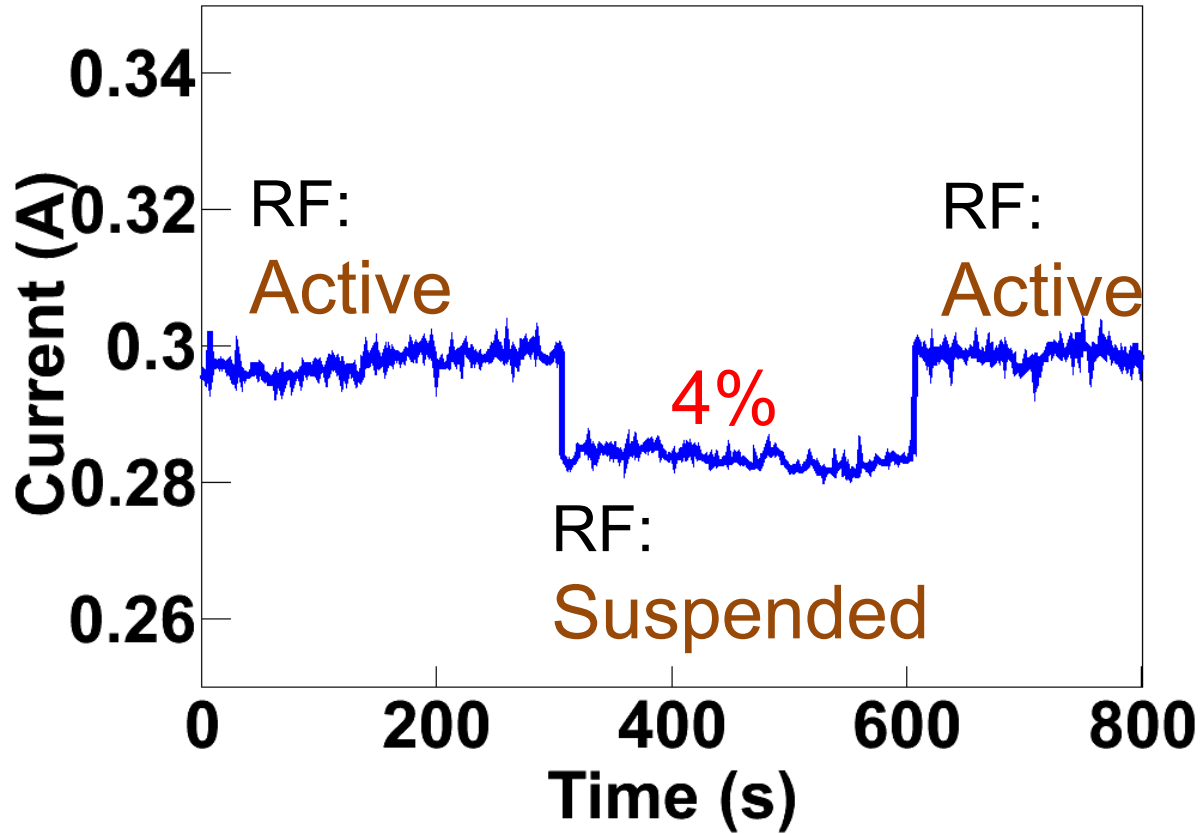


PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1

Hibernation(swsusp)

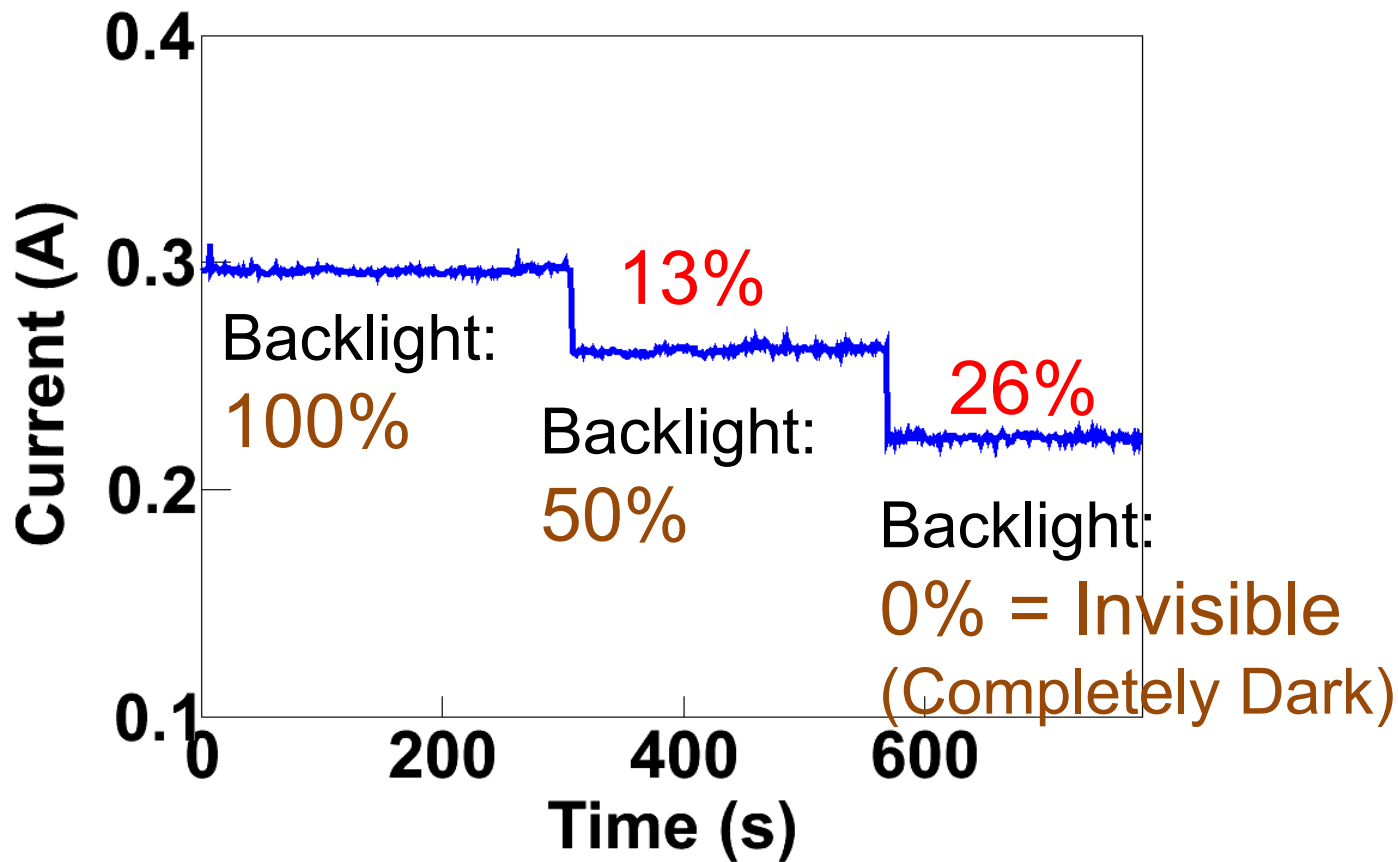


PowerEdge 1950 / Cent5.2 / Kernel-2.6.27.1



Panasonic CF-W4 / Ubuntu 8.10 / Kernel-2.6.24

xbacklight



Panasonic CF-W4 / Ubuntu 8.10 / Kernel-2.6.24

Summary of the Measurements

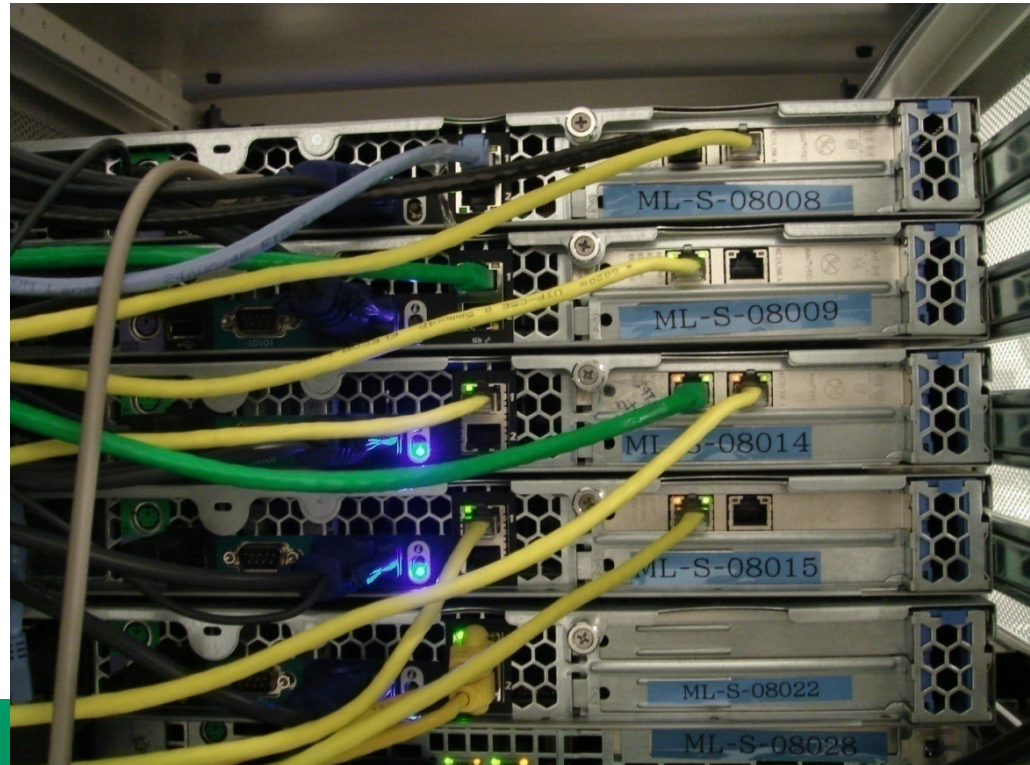


- Hibernation
 - Very effective as expected
- Other functions that reduces power when idle
 - 0.5 - 4% of the current is reduced
 - Total sum: ~10%
 - Cutting out peripheral devices is effective
 - About 10% (times system idle fraction) of power should be reduced by present Linux/OSS.

Peripheral Driver – e1000 NIC



- Disabling peripheral devices is very effective
- Some PCs has network devices even not used
 - unplug the network cable
 - run ifdown command



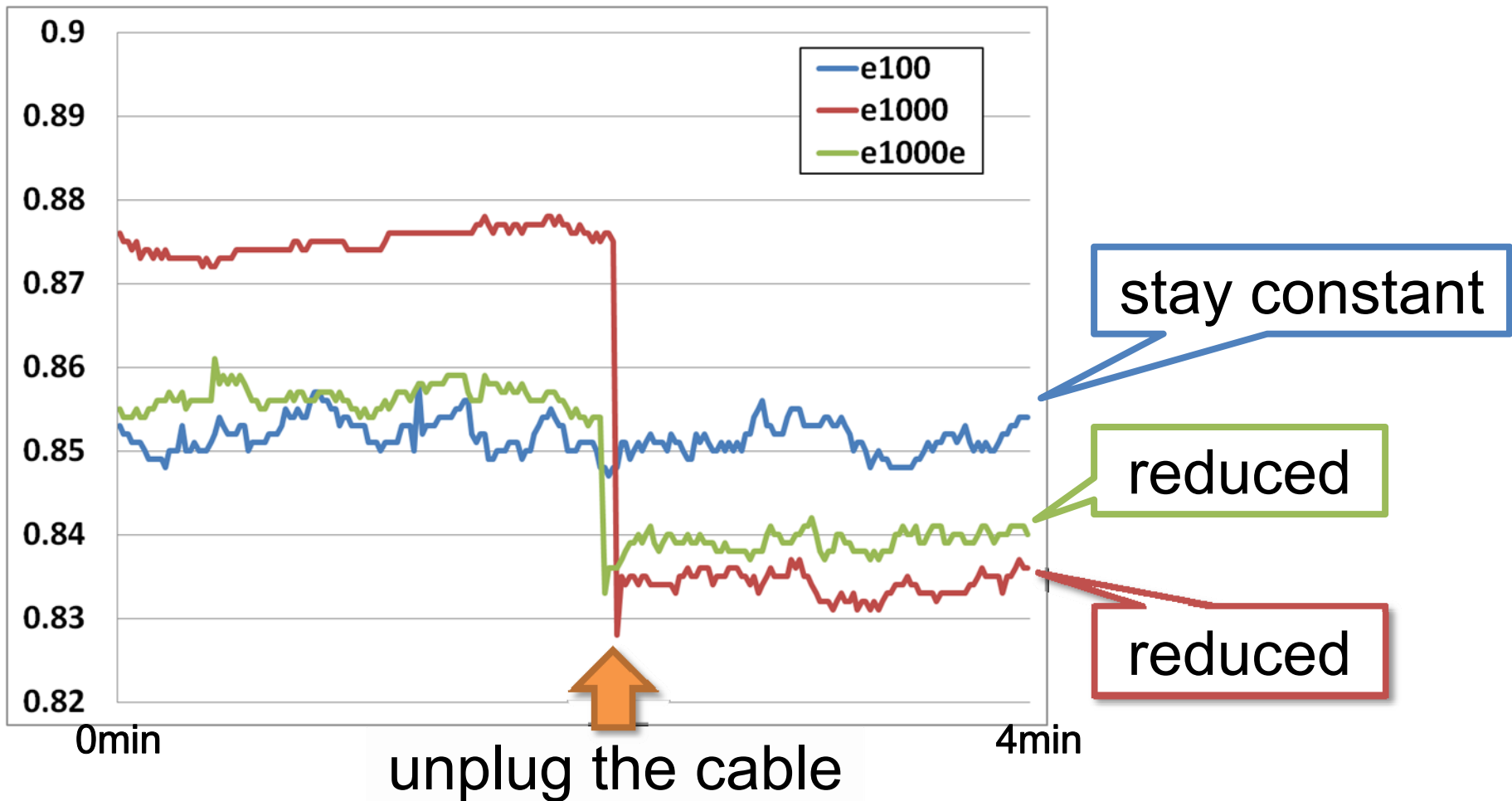
Do the Next, Open your Window

Environment



- Machine/OS
 - DELL PowerEdge SC440
 - Linux 2.6.18 kernel (runlevel 1)
- e100/e1000/e1000e driver
 - e100: 3.5.10-k2-NAPI
 - e1000: 7.3.20-k2-NAPI
 - e1000e: 0.2.9.5-NAPI
- e100/e1000/e1000e NIC
 - e100: Intel 82557/8/9/0/1 Ethernet Pro 100
 - e1000: Intel 82541PI Gigabit Ethernet Controller
 - e1000e: Intel 82572EI Gigabit Ethernet Controller

Link Down of e100/e1000/e1000e



- The e100 device doesn't reduce the power.

Modification of e100 Driver

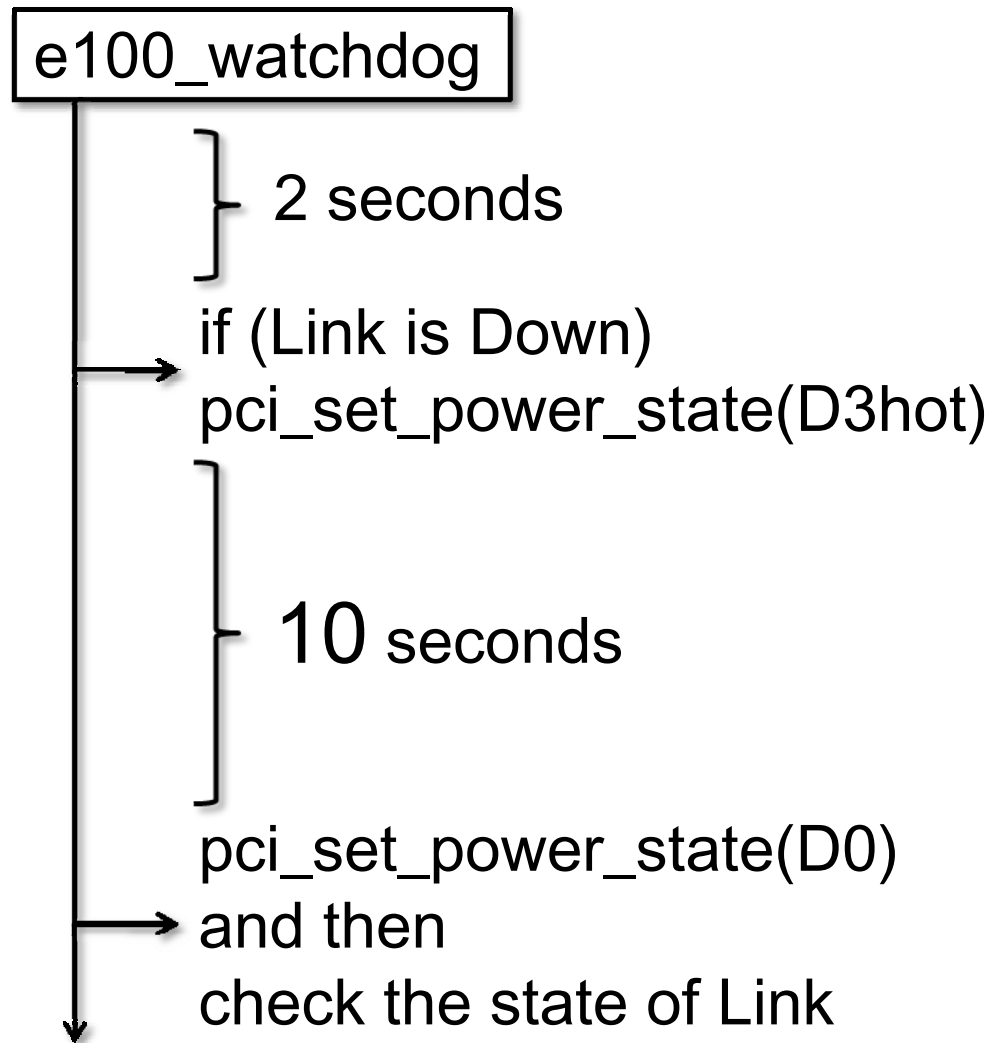


[problem]

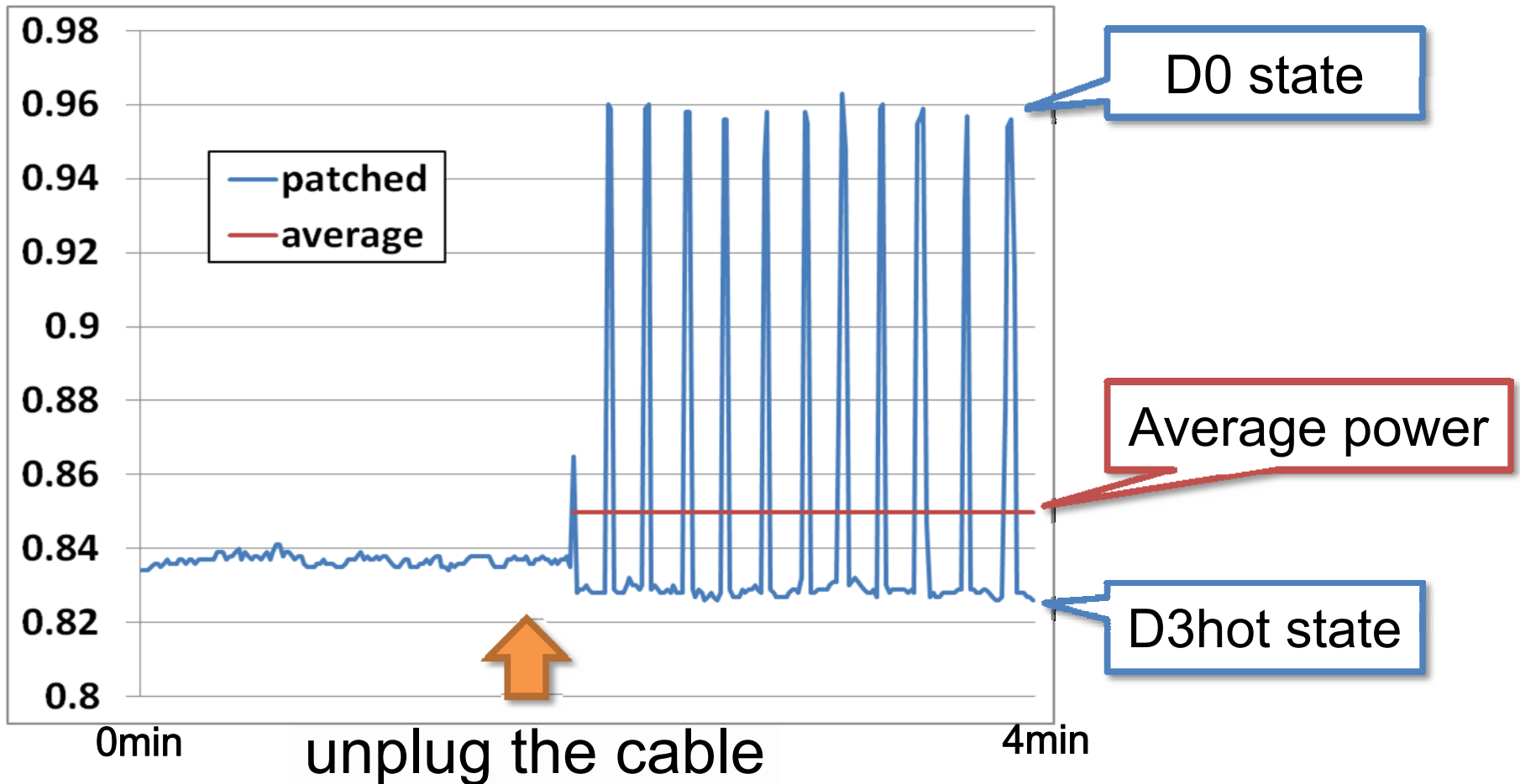
- doesn't let the power down when unplugging a LAN cable

[solution]

- set power state of a PCI device to D3hot at the check point of watchdog

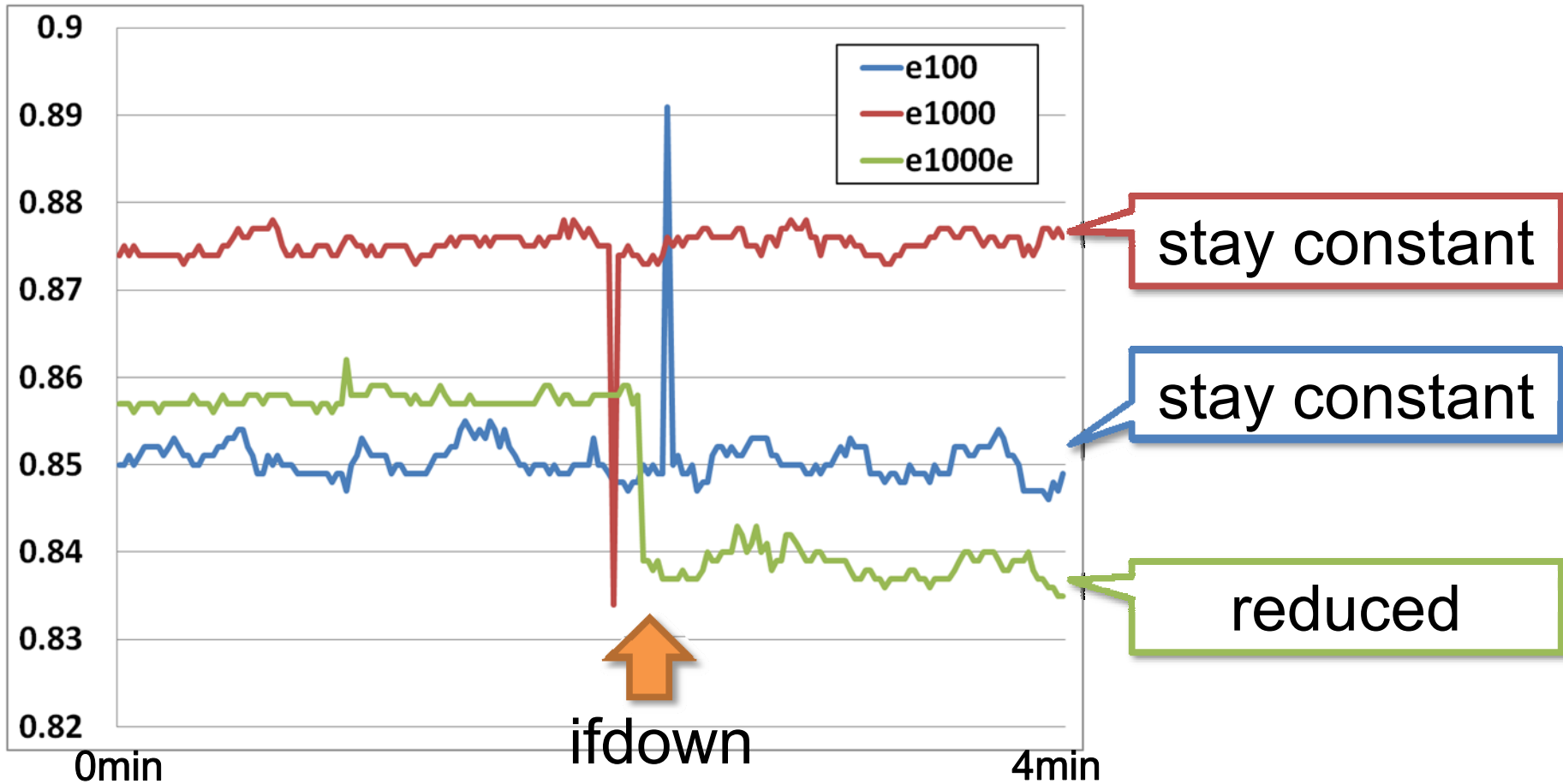


Modified e100 Driver



- In transitions to D0, there were big spikes.

Changes by ifdown Command



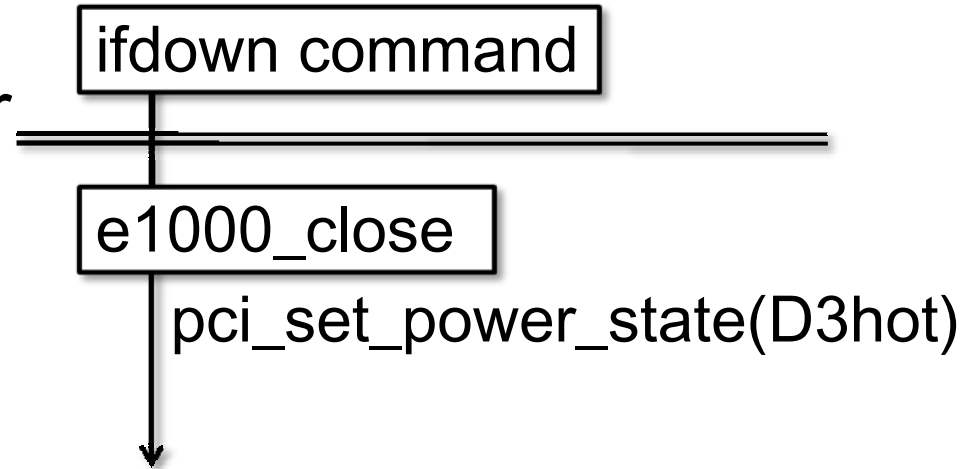
- Only e1000e device is smart.

Modification of e1000 Driver



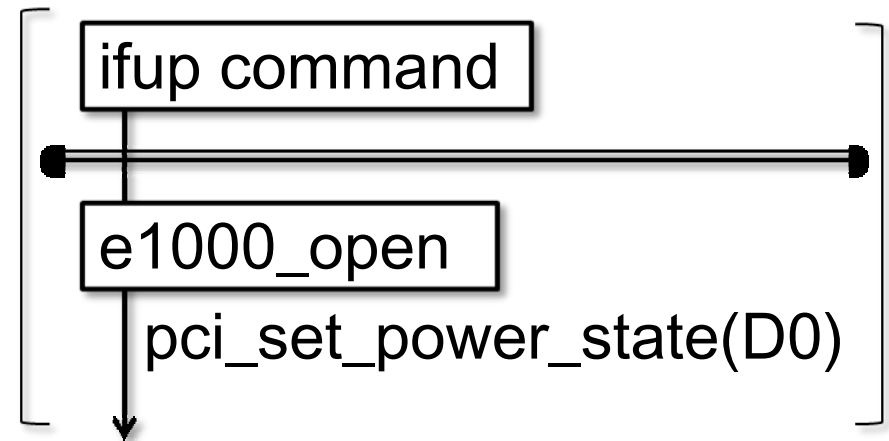
[problem]

- doesn't let the power down when running ifdown command.

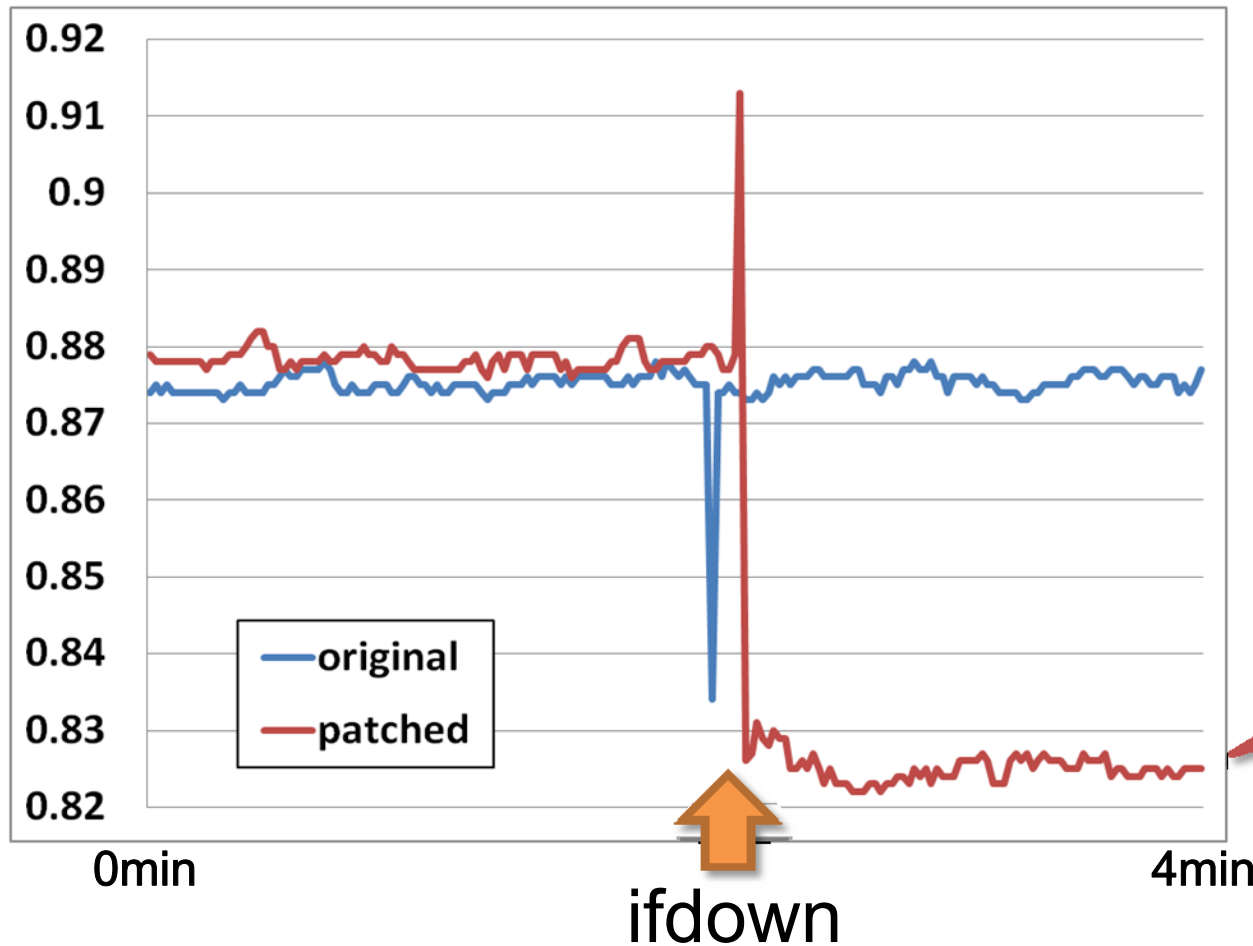


[solution]

- set to the D3hot state at the end of e1000_close().



Modified e1000 Driver



- Modified driver worked expectedly.

NIC Driver Modification



- Link-down/ifdown behaviors are different among drivers
- e100/e1000 hacks
 - e100(link-down): depending on polling period
 - e1000(ifdown): effective
- Same approach can be also applied to embedded area
- Future plan
 - Other enhancements
 - Other drivers

Proposal



- Total control tools for users / admins
- Benchmark software
- Integrated cluster monitoring / management tool
- Aggressive device shut-down
- **In-direct indicators other than powertop**

Summary



- Surveyed about standards / organizations
- Linux implementation status
- Measured actual power reduction effects
- Proposed development ideas
- Started Linux improvements (e1000)

Questions



- What is standard / good way to measure power consumption?
- How do you think about,
 - Benchmark suite
 - GUI configurator
 - Integration in cluster management
- Any other idea?



Do the Next, Open your Window

MIRACLE