

Evaluation of IEEE802.3az on Linux Ethernet Driver

2011.12.9

富士通コンピュータテクノロジーズ

松本博郎

- 背景
- IEEE802.3azとは
- Linuxカーネルの対応状況
- 測定方法
- 測定結果
- まとめ

- 組込み機器では消費電力の低減が要件として取り上げられる。バッテリー駆動時間の延長であったり、24時間電源ONのまま待機する必要があったり、用途は様々ある

- Linuxカーネルのデバイスドライバでは、消費電力を低減させる機能のひとつとして、suspend / resumeベースの仕組みを利用できる
 - 「echo mem > /sys/power/state」⇒ ドライバのsuspend処理
 - 電源押下等のイベント ⇒ 各々のドライバのresume処理

- 本資料では、suspend / resumeベースでない、消費電力を低減する仕組みであるIEEE802.3azについて、Linuxカーネルのデバイスドライバ対応状況を調査し、実際に動作させて消費電力を測定する

■ 有線LANの帯域が使用されていない時に、有線LANコントローラのPHY / MAC層で消費電力を抑える規格

- EEE (Energy Efficient Ethernet) と記載されることが多い
- Task Forceのページ

- <http://www.ieee802.org/3/az/index.html>

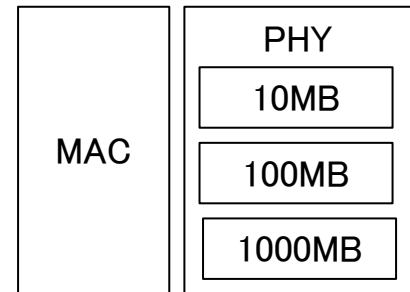
- PHY / MAC層より上位層はIEEE802.3azを意識しないで良い

- Linuxの場合、ドライバのみで対応

- **LPI (Low Power Idle) と RPS (Rapid PHY Selection)**

- LPIはMACへ電源供給を一時的に止める手法
- RPSはLANの帯域によってPHYを動的に切り替える手法
- ITproのEEEの記事

- <http://itpro.nikkeibp.co.jp/article/Keyword/20090331/327539/>



LANコントローラの概観
ITproのEEEの記事より引用

■ 24時間電源ONで待機するような機器向け

■ IEEE802.3azに対応した機器同士でないとは動作しないらしい

■ IEEE802.3az対応デバイスとデバイスドライバ

- linux-3.2-rc1にて、LPI、EEE、802.3azというキーワードが使用されているコミットログを検索

デバイス	デバイスドライバ
82575	igbドライバ
82579	e1000eドライバ
BCM5718, BCM5719 ...	tg3ドライバ
BCM54618E ...	bnx2xドライバ
RTL8111E	r8169

82575 ... Intel® 82575EB Gigabit Ethernet Controller

82579 ... Intel® 82579 Gigabit Ethernet Controller

BCM5718 ... Broadcom® BCM5718 - x2 PCI Express® Dual-Port Gigabit Ethernet Controller

BCM5719 ... Broadcom® BCM5719 - X4 PCI Express® Quad-Port Gigabit Ethernet Controller

BCM54618E ... Broadcom® BCM54618E - Single-Port 10/100/1000BASE-T Energy Efficient Ethernet GbE Transceiver

RTL8111E ... Realtek® RT8111E Integrated Gigabit Ethernet Controller for PCI Express Applications

■ メインラインカーネルのe1000eドライバ

■ デフォルトでは無効なので、ソースコードを変更して有効にする

- eee_disableをfalseにしておけば、LPIのレジスタにアクセスできるようになる

```
diff --git a/drivers/net/ethernet/intel/e1000e/ich8lan.c b/drivers/net/ethernet/intel/e1000e/ich8lan.c
index e2a80a2..88f7086 100644
--- a/drivers/net/ethernet/intel/e1000e/ich8lan.c
+++ b/drivers/net/ethernet/intel/e1000e/ich8lan.c
@@ -821,7 +821,7 @@ static s32 e1000_get_variants_ich8lan(struct e1000_adapter *adapter)

    /* Disable EEE by default until IEEE802.3az spec is finalized */
    if (adapter->flags2 & FLAG2_HAS_EEE)
-       adapter->hw.dev_spec.ich8lan.eee_disable = true;
+       adapter->hw.dev_spec.ich8lan.eee_disable = false;

    return 0;
}
```

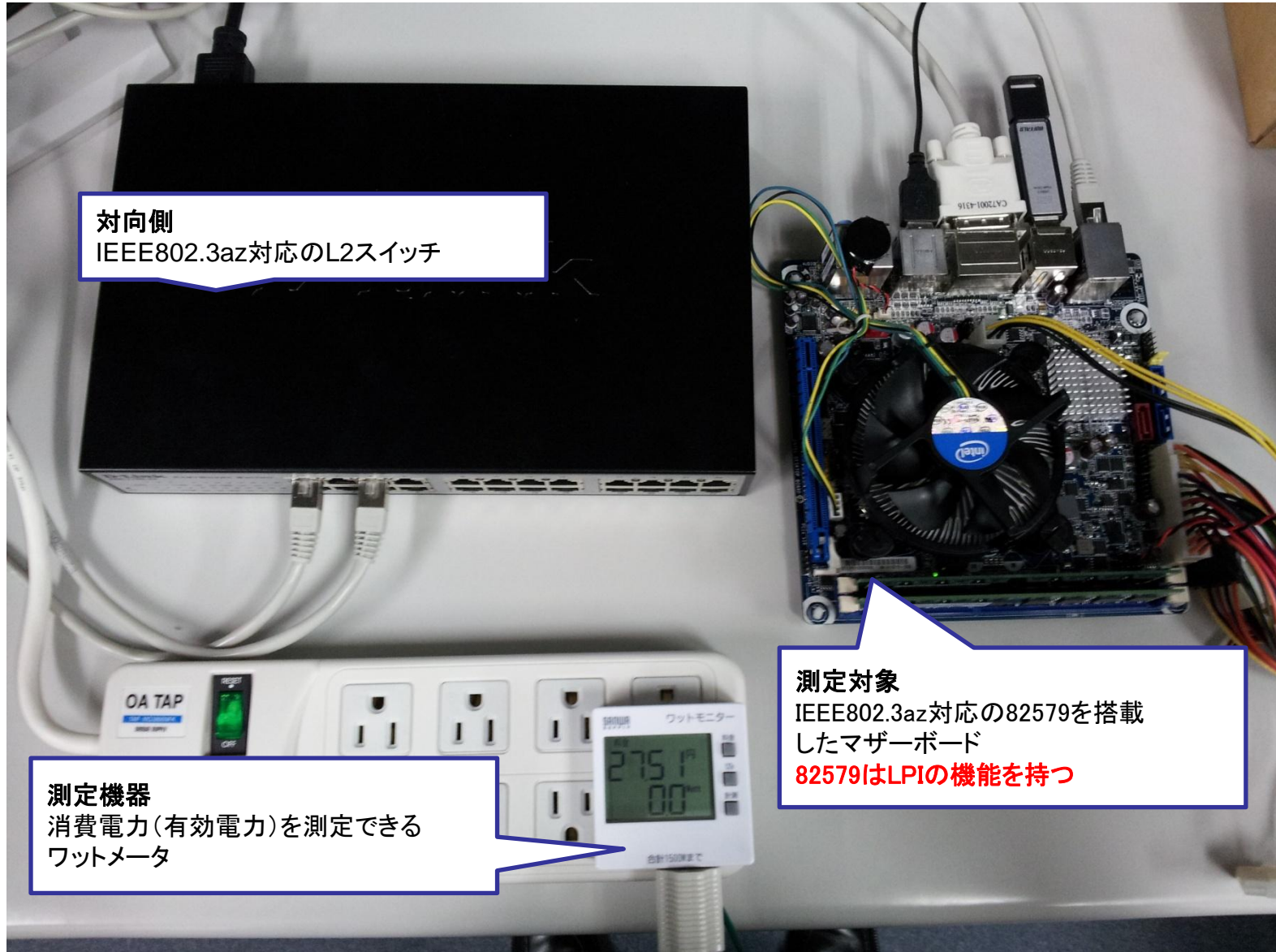
■ Intel Wired Ethernetプロジェクトのe1000eドライバ

■ Source Forgeのページ

- http://sourceforge.jp/projects/sfnet_e1000/

■ デフォルトで有効 (“insmod e1000e.ko EEE=<0 | 1>”で切り替え)

測定環境(ハードウェア)



■ カーネル

- linux-3.0.8を使用

- Intelのイーサネットデバイスのデバイスドライバを無効にしてビルド

■ デバイスドライバ

- Intel Wired Ethernetのe1000e-1.6.3を使用

■ ルートファイルシステム

- initrdを使用

■ ネットワーク帯域

- 1000MB / Full Duplex

```
# insmod e1000e.ko EEE=1
[ 231.470192] e1000e: Intel(R) PRO/1000 Network Driver - 1.6.3-NAPI
[ 231.470301] e1000e: Copyright(c) 1999 - 2011 Intel Corporation.
[ 231.470423] e1000e 0000:00:19.0: PCI INT A -> GSI 20 (level, low) -> IRQ 20
[ 231.470532] e1000e 0000:00:19.0: setting latency timer to 64
[ 231.470706] e1000e 0000:00:19.0: (unregistered net_device): EEE Support Enabled
[ 231.553158] e1000e 0000:00:19.0: irq 42 for MSI/MSI-X
[ 231.786000] e1000e 0000:00:19.0: eth0: (PCI Express:2.5GT/s:Width x1) e0:69:95:35:10:c8
[ 231.786159] e1000e 0000:00:19.0: eth0: Intel(R) PRO/1000 Network Connection
[ 231.786296] e1000e 0000:00:19.0: eth0: MAC: 11, PHY: 11, PBA No: FFFFFFFF-OFF
```


■ IEEE802.3azが有効だとLANケーブルと同じ消費電力になる

LANケーブル	ドライバ	IEEE802.3az	消費電力
抜いた状態	ロードしていない	無効	28.0W
挿した状態	ロードしていない	無効	28.6W
挿した状態	ロードしている	無効	28.6W
挿した状態	ロードしている	有効	28.0W

■ IEEE802.3azを有効にすることで0.6Wの消費電力を低減できる

ワットメータの精度が±1Wである為、厳密な値ではない
表示上は確かに0.6W程度異なる

■ ネットワークアイドル時のリンクステータスについて

- イーサネットデバイスのLEDは1000MB
- ethtoolでも1000MBの表示のまま

■ IEEE802.3az動作時のパケットについて

- L2スイッチのミラーポート機能とWiresharkやtcpdumpで見れるレベルでは、IEEE802.3az有効 / 無効で差異なし

■ 対向側がIEEE802.3az対応ではない場合について

- LANの疎通はできる
- 消費電力はIEEE802.3azが無効の場合と同様である


■ 今回の測定で分かったことについて

- IEEE802.3azの手法であるLPIの消費電力を測定できた
- リンクステータスが変化しないので動作しているかどうか分かりづらかった
- デバイスドライバはレジスタを叩くだけでよく、ソフトウェア開発は難しくない
- LPIはネットワーク対向側がIEEE802.3az非対応でも動作するらしい
 - それとも82579は特別？（消費電力の低減は機能しないようだが）

■ 今後の課題について

- もうひとつの手法であるRPSの消費電力も測定したい
 - こちらは動的にリンクスピードが変更されるはずなので、動作していることは分かりやすいはず
 - LPIはネットワークアイドル時に効果を発揮するが、RPSはネットワークが使用されている時にネットワークの帯域に合わせて効果を発揮できそう（LANの帯域に合わせて動的にPHYを切り替える）
 - RPSはネットワーク対向側がIEEE802.3az非対応の場合にLANの疎通がどうなるかが気になるところ

※文中に記載されている他社の製品名、サービス名等はそれぞれ各社の商標です。



FUJITSU

shaping tomorrow with you