

Zastosowanie ACPI w Linuksie 2.6

Oszczędzanie energii

W nowym jądrze wykorzystano zaawansowane funkcje ACPI. Standardowa funkcja konfigurowania sprzętu powinna być znana większości użytkowników systemów z jądrem 2.4, ale funkcja zarządzania energią została poszerzona o kilka trybów oszczędzania energii. **TIMO HÖNIG**

Oto historia prawdziwego sukcesu. Praktycznie każdy laptop, sprzedawany w zestawie komputer, urządzenie i system operacyjny obsługują ACPI. Od czasu wprowadzenia pierwszej wersji ACPI (Advanced Configuration and Power Interface) w 1996 roku, specyfikacja ta wyparła z rynku standard APM (Advanced Power Management) oraz mało elastyczne rozwiązanie plug & play dla BIOSu.

ACPI jest często traktowane tylko jako inna wersja APM. Pogląd ten mija się jednak z prawdą: litera C w skrócie oznacza słowo konfiguracja (configuration). Dzięki temu, że konfiguracja jest integralną częścią rozwiązania ACPI, użytkownik otrzymuje jednolity, działający niezależny od platformy sprzętowej interfejs konfiguracyjny. Linux 2.6 używa ACPI do przekierowywania przerwań PCI. Jeżeli chcesz na komputerze jednoprotocowym korzystać z kontrolera APIC (Advanced Programmable Interrupt Controller) zamiast kontrolera PIC, aby uniknąć konfliktów zasobów lub współdzielenie przerwań, nie masz innego wyjścia, jak tylko zdecydować się na ACPI. Podsystem ACPI przekazuje jądru informacje, które urządzenia PCI są typu hot-plug oraz jak je konfigurować. Odpowiednie procedury można znaleźć w `drivers/pci/hotplug/acpihp_g glue.c`.

Interfejs ACPI korzysta z opracowanego na jego potrzeby języka ASL (ACPI Source Language), stanowiącego osobną część oraz języka AML (ACPI Machine Language)



Esther Keller, vishnix.com

w celu stworzenia wirtualnej warstwy dla funkcji zależnych od określonej platformy sprzętowej. AML opisuje sprzęt i sposób uzyskania dostępu do niego. Każdy system operacyjny obsługujący ACPI ma interpreter AML dla kodu AML.

Dwuwarstwowa specyfikacja

Specyfikacja ACPI [1] dzieli architekturę ACPI na dwie warstwy. Pierwsza z nich, niska, obejmuje następujące elementy architektury:

- Tabele ACPI
- BIOS ACPI
- Rejestry ACPI

Tabele ACPI opisują sprzęt obsługujący ACPI oraz jego konfigurację w AML. Centralna struktura danych każdego systemu ACPI korzysta z bloków definicji do określania, w jaki sposób można uzyskać dostęp do sprzętu. Podczas uruchamiania systemu BIOS ACPI przechowuje tabele ACPI w pamięci. BIOS

jest także potrzebny w operacjach wstrzymania i przywracania pracy komputera (Suspend-to-RAM, Suspend-to-Disk).

Druga warstwa, wysoka, jest częścią systemu operacyjnego, korzysta z jądra oraz sterowników ACPI i udostępnia interfejs API, którego mogą używać komponenty ACPI należące do niższej warstwy. Interpreter AML jest częścią tej warstwy.

ACPI w systemie Linux 2.6

Konfiguracja jądra obejmuje funkcje zarządzania energią ACPI [2] w menu rozwijanym opcji *Power management options (ACPI, APM)* (patrz Rysunek 1). Jeżeli chcesz korzystać z funkcji Suspend-to-RAM, musisz najpierw wybrać opcję *Prompt for development and/or incomplete code/drivers* poniżej pozycji *Code maturity level options* w celu wyświetlenia opcji *Sleep States* wraz z innymi opcjami ACPI.

Timo Hönig studiuje informatykę w Augsburgu, w Niemczech. Od ponad dziesięciu lat ma do czynienia z sieciami, terminalami, Open Source i Unixem. Jego obecnym hobby jest projekt Open Source o nazwie FnFX pracujący nad demonem ACPI pod Linuksa dla laptopów Toshiba (<http://fnfx.sourceforge.net>).



Rysunek 1. Opcje jądra 2.6 do zarządzania energią oparte na mechanizmach ACPI.

Podczas uruchamiania systemu jądro wywołuje funkcję `acpi_boot_init()` w celu uzyskania dostępu do tabel ACPI i analizuje je przy użyciu ACPI BIOS. Szczególnie ważna jest tabela DSDT (Differentiated System Description Table). Zawiera ona specyfikacje urządzeń kompatybilnych z ACPI oraz ich konfigurację.

Powszechnie wiadomo, że niektórzy producenci używają w swoich systemach błędnych tabel DSDT. Nawet jeśli producent nie udostępni uaktualnionego BIOSu z poprawną tabelą DSDT, a użycie opcji (`CONFIG_ACPI_RELAXED_AML`) jądra *Relaxed AML* nie rozwiązuje problemu, nie ma powodu do zmartwień. Niezawierające błędów tabele DSDT dla wielu systemów dostępne są pod adresem [3]. Kolejne kroki procedury można znaleźć na stronie [4].

Jeżeli nie możesz znaleźć poprawnych tabel DSDT dla swojego systemu, możesz spróbować naprawić tabelę ręcznie. Intel udostępnił kompilator ASL oraz dezasembler AML `iasl` [5]. Umożliwiają one kopiowanie błędnych tabel z lokalizacji (`cat /proc/acpi/dsdt > dsdt.aml`) oraz dezasembrowanie ich za pomocą kompilatora AML. Wykonanie polecenia `iasl -d`

`dsdt.aml` powoduje zdezasembrowanie tabeli DSDT i przekazanie kodu źródłowego w ASL do pliku `dt ds.dsl` (patrz Listing 1).

Następny krok polega na zlokalizowaniu błędów w kodzie źródłowym ASL, usunięciu ich i ponownej kompilacji kodu (`iasl -tc dt ds.dsl`). Spróbuj wyeliminować wszystkie ostrzeżenia i błędy, wprowadzając stosowne poprawki w kodzie ASL (`dt ds.dsl`), przed użyciem ponownie skompilowanej tabeli DSDT. Porady dotyczące naprawy błędnych tabel DSDT znajdziesz pod adresem [6].

Po uruchomieniu systemu Linux 2.6 obsługującego ACPI, możesz przejrzeć wpisy ACPI w lokalizacjach `proc` oraz `sys`. Wpisów w `/proc/acpi` oraz `/sys` można używać do uzyskania dostępu do interfejsu ACPI. Do zapisu w powyższych lokalizacjach wymagane są uprawnienia root. Jeżeli `sysfs` nie został jeszcze zamontowany, użyj w tym celu poniższych poleceń:

```
mkdir /sys
mount -t sysfs sysfs /sys
```

Wpis `sysfs /sys sysfs defaults 0 0` w pliku `/etc/fstab` powoduje automatyczne montowanie `sysfs` podczas uruchamiania systemu.

Tryby oszczędzania energii

Poszczególnymi trybami oszczędzania energii (patrz Tabela 1) można zarządzać przy użyciu `/sys/power/state`. W `cat /sys/power/state` można sprawdzić, jakie tryby są obsługiwane w systemie, a polecenie `echo -n „Sleep_State” /sys/power/state` przelacza system w określony tryb.

Tryb Suspend-to-RAM (S3) wyłącza cały system, z wyjątkiem pamięci głównej, czyli CPU, pamięci cache, chipsetu i urządzeń pe-

ryferyjnych (dysków twardych, USB itd.). Aby zapobiec utracie danych przechowywanych w pamięci, zasilana jest także pamięć RAM. Teoretycznie tryb Suspend-to-RAM umożliwia przejście do trybu oszczędzania energii i wznowienie pracy systemu w ciągu kilku sekund. Z moich doświadczeń wynika jednak, że tryb ten rzadko działa w taki sposób. Bardzo słusznie został on oznaczony w jądrze jako eksperymentalny (*experimental*).

Tryb Suspend-to-Disk (S4) pozwala zaoszczędzić najwięcej energii spośród wszystkich trybów, ponieważ powoduje on zapisanie bieżącej zawartości pamięci, rejestrów i stanu urządzeń peryferyjnych na dysku twardym (Linux używa do tego celu dużej partycji swap) przed całkowitym wstrzymaniem systemu. Podczas ponownego uruchamiania systemu z dysku odczytywana jest zawartość pamięci głównej oraz poprzedni stan z systemu, które są następnie przywracane. Linux 2.6 obsługuje tryb Suspend-to-Disk, niezależnie od obsługi standardów ACPI i APM. Dla porównania: Linux 2.4 nie obsługuje wcale trybu Suspend-to-RAM, ale po zaimplementowaniu patchy dla jądra opracowanych w ramach projektu Swsusp [7], użytkownik może korzystać z trybu Suspend-to-Disk.

Tryby Software Suspend i Suspend-to-Disk

Linux 2.6 oferuje dwa alternatywne warianty trybu Suspend-to-Disk (S4): Software Suspend (`CONFIG_SOFTWARE_SUSPEND`) oraz Suspend-to-Disk (`CONFIG_PM_DISK`). Pierwszy z nich jest oznaczony jako eksperymentalny (*experimental*), a drugi jest obsługiwany bez większych problemów.

Tryb Software Suspend wymaga podania parametru `resume` podczas uruchamiania sys-

Listing 1. Dezasembrowanie tabeli DSDT za pomocą `iasl`

```
01 [root@sunshine:~]$ cp /proc/ac-
02 pi/dsdt ~/dsdt
03 Intel ACPI Component Architecture
04 ASL Optimizing Compiler / AML Di-
05 sassembler version 20030918 [Sep
06 18 2003]
07 Copyright (C) 2000 - 2003 Intel
08 Corporation
09 Supports ACPI Specification Revi-
10 sion 2.0b
11 * Intel ACPI Component Architec-
12 ture
13 * AML Disassembler version
14 installed and loaded
15 20030918
16 Pass 1 parse of [DSDT]
17 *
18 Pass 2 parse of [DSDT]
19 * Disassembly of dsdt, Thu Dec 11
20 17:28:16 2003
21 */
22 DefinitionBlock ('DSDT.aml',
23 'DSDT', 1, 'TOSHIB', '2000 ',
24 537003284)
25 {
26 Name (\_S0, Package (0x04)
27 [root@sunshine:~]$
```

temu – niestety nie można skompilować go z jądrem. Opcja określa partycję swap, na której mają być przechowywane zawartość pamięci i informacje o stanie. Dobrym rozwiązaniem jest wpisanie w pliku konfiguracyjnym stałego parametru, na przykład `resume=/dev/hda2`.

Alternatywny tryb Suspend-to-Disk umożliwia określenie partycji swap w parametrze jądra `pmdisk`. W odróżnieniu do parametru `resume`, parametr `pmdisk` można powiązać z jądrem.

Kiedy system czuwa

Dla obu trybów, Suspend-to-RAM oraz Suspend-to-Disk, przyjęto założenie, że zainstalowane sterowniki obsługują wstrzymywanie i wznowianie pracy systemu. W przypadku Linuksa 2.6 spełnienie tego założenia wymagało zmodyfikowania modelu sterownika. Sterowniki urządzeń muszą udostępniać funkcje `suspend` i `resume`, które jądro wywołuje odpowiednio przy przejściu w tryb oszczędzania energii i wznowieniu pracy systemu.

Jeżeli system ciągle się zawiesza lub nie wznowia pracy, spróbuj rozwiązać ten problem w następujący sposób. Z uprawnieniami root przejrzyj pliki rejestru jądra, następnie skompiluj ponownie sterownik powodujący problem, o ile nie zostało to już wcześniej zrobione. Przed wywołaniem `echo Sleep_State > /sys/power/state` usuń moduł sterownika, a następnie przywróć go po wznowieniu pracy systemu. Cały proces przechodzenia w tryb oszczędzania energii można łatwo wykonać za pomocą skryptu. Listing 2 przedstawia przykład skryptu umożliwiającego przejście w tryb Suspend-to-RAM, mimo iż kontroler hosta USB (moduł sterownika `ohci-hcd.o`) stwarzał duże trudności przy wznowianiu pracy systemu.

Jeżeli masz problemy zarówno z przejściem do trybu Suspend-to-Disk (Software

Suspend lub Suspend-to-Disk), jak i wznowieniem pracy, możesz pokusić się o ręczne edytowanie parametrów jądra. Parametry `resume=noresume` i `pmdisk=off` spowodują, że jądro nie będzie podejmować próby przywrócenia poprzedniego stanu podczas uruchamiania systemu. Wariant Software Suspend wymaga dodatkowo ponownej inicjalizacji partycji swap, w naszym przypadku `mkswap /dev/hda2`, oraz ponownego przypisania jej do systemu: `swapon /dev/hda2`.

Dynamiczna zmiana szybkości zegara CPU w urządzeniach przenośnych

Nowe CPU, stosowane w urządzeniach przenośnych, mogą zmniejszać i zwiększać szybkość zegara w zależności od bieżącego obciążenia CPU (Intel Speedstep oraz AMD Power Now!). Aby wykorzystać tę możliwość w Linuksie 2.6, włącz funkcję `cpufreq`, tak jak to przedstawiono na Rysunku 2, a następnie wybierz odpowiedni sterownik `cpufreq`.

Interfejs `cpufreq` zmienił się podczas tworzenia Linuksa 2.5. Jego wersję ostateczną można znaleźć w `sysfs`. Jeżeli chcesz używać jądra skompilowanego z tymi funkcjami, użyj `dmesg` | `grep cpufreq` w celu sprawdzenia, czy jądro może inicjalizować tę funkcję (patrz Listing 3).

Jeżeli funkcja `cpufreq` wywołuje komunikat o błędzie, sprawdź najpierw, czy masz odpowiedni sterownik `cpufreq`. W naszym laboratorium musieliśmy włączyć obsługę Plug & Play BIOS w jądrze (`CONFIG_PNP` i `CONFIG_PNPBIOS`), aby uruchomić sterownik ACPI funkcji `cpufreq`.

Wpisy

Zakładając, że funkcja `cpufreq` została zainicjalizowana poprawnie, pod pozycją `/sys/devices/system/cpu/cpu0/cpufreq` w `sysfs` powinny być dostępne następujące wpisy.

`cpufreq_max_freq` i `cpufreq_min_freq` zawierają

wartości maksymalnej i minimalnej szybkości procesora bieżącego CPU. `scaling_max_freq` i `scaling_min_freq` definiują maksymalną i minimalną szybkość zegara tworzącego zakres, w którym będzie pracować CPU. Wpisy `echo speed > scaling_max_freq` i `echo speed > scaling_min_freq` umożliwiają użytkownikom z uprawnieniami root zmianę wartości. `speed` musi mieć wartość z zakresu wyznaczonego przez wartości wpisów `cpufreq_max_freq` i `cpufreq_min_freq`.

Wpis `scaling_available_governors` określa, jakie strategie dostosowywania prędkości zegara system udostępnia: `powersave`, `userspace` lub `performance`. Aktywna w danej chwili strategia jest przechowywana w `scaling_governor`. Użytkownik z uprawnieniami root może użyć polecenia `echo Governor > scaling_governor` w celu uaktywnienia innej strategii (`Governor` musi być jednym z trzech ciągów zdefiniowanych w `scaling_available_governors`).

Aby wyeliminować potrzebę stosowania poleceń `echo`, w ramach projektu `Cpudyn` [8] opracowano demona używającego interfejsu `cpufreq`. `Cpudyn` ułatwia modyfikowanie strategii dostosowywania szybkości zegara CPU do własnych potrzeb.

Sterowanie szybkością w ACPI

Wiele CPU, które nie obsługują dynamicznej zmiany szybkości zegara, może korzystać z funkcji sterowania szybkością, opartej na mechanizmach ACPI. Zastosowanie takiego rozwiązania pozwala zwykle zmniejszyć ilość wydzielanego ciepła oraz generowanego hałasu. Warto rozważyć ten pomysł, zwłaszcza w przypadku serwerów wykorzystywanych jedynie w ciągu standardowych godzin pracy.

CPU musi oczywiście obsługiwać funkcję sterowania szybkością. Jeżeli CPU ma taką funkcję, w `/proc/acpi/processor/CPUx/throttling` znajdziesz listę zawierającą jej dostępne stany. Stan `T0` wskazuje, że CPU pracuje z maksymalną szybkością. W przypadku każdego innego stanu, CPU pracuje z szybkością mniejszą o wartość procentową zdefiniowaną w `/proc/acpi/processor/CPUx/throttling`.

ACPI na laptopach

Opisane w niniejszym artykule funkcje ACPI są szczególnie interesujące dla użytkowników laptopów. Informacje o stanie baterii komputera można znaleźć w `/proc/acpi/battery/BATx/info`. Dane o poziomie naładowania, pojemności i napięciu są przechowywane w `/proc/acpi/battery/BATx/state`. Jeżeli dana

Listing 2. Skrypt procesu przechodzenia do trybu oszczędzania energii

```
01 #!/bin/sh
02 rmmmod usbblp usb-storage hid; # Remove USB Module
03 rmmmod ohci-hcd; # Remove Host Controller Module
04 rmmmod usbcore; # Remove USB Core Module
05
06 echo -n „mem” > /sys/power/state; # Suspend to RAM
07
08 modprobe usbcore; # Load USB Core Module
09 modprobe ohci-hcd; # Load Host Controller Module
10 modprobe usbblp usb-storage hid; # Load USB Module
11
12 exit 0;
```

Listing 3. Test funkcji cpufreq

```
01 [root@sunshine:~]$ dmesg | grep
cpufreq
02 cpufreq: CPU0 - ACPI performan-
ce management activated.
03 cpufreq: *P0: 750 MHz, 22000
mW, 250 uS
04 cpufreq: P1: 350 MHz, 9800 mW,
250 uS
```

wartość spadnie poniżej określonego progu, który można zdefiniować poleceniem `echo threshold /proc/acpi/battery/BATx/ alarm`, ACPI może zaalarmować użytkownika za pomocą zdarzeń zdefiniowanych w `/proc/acpi/events`.

Wpis `/proc/acpi/ac_adapter/ADPx/state` informuje użytkownika, czy zasilacz sieciowy jest podłączony. Zdarzenia podłączania i odłączania zasilacza są rejestrowane w `/proc/acpi/events`. Sterownik przycisku (opcja jądra `CONFIG_ACPI_BUTTON`) oraz sterownik temperatury procesora (`CONFIG_ACPI_THERMAL`) działają na tych samych zasadach. Sterownik przycisku sygnalizuje każde naciśnięcie przycisku zasilania, wstrzymania oraz przycisku służącego do zamykania obudowy. Sterownik temperatury monitoruje temperaturę CPU (bieżące wskazania są dostępne w `/proc/acpi/thermal_zone/THRM/temperature`) i sygnalizuje osiągnięcie wartości zdefiniowanej w `/proc/acpi/thermal_zone/THRM/trip_points`.

Właściciele laptopów mogą korzystać z demona ACPI (`acpid` [9]) do odpowiadania na komunikaty `/proc/acpi/events`. Kiedy pokrywa komputera jest zamknięta, `acpid` automatycznie powoduje przejście komputera w tryb `suspend to disk`. Mimo wielu dostępnych funkcji ACPI, większość użytkowników nowoczesnych laptopów szybko zdaje sobie sprawę, że nie mogą korzystać ze skrótów klawiaturowych w celu zmiany jasności wyświetlacza pod Linuxem oraz że klawisze funkcyjne [Fn] nie działają prawidłowo.

Dobre wieści mamy za to dla właścicieli laptopów Asus i Medion: opcja sterownika

ASUS/Medion Laptop Extras (`CONFIG_ACPI_ASUS`) tworzy wpisy poniżej `/proc/acpi/asus` umożliwiające zmianę jasności wyświetlacza oraz kontrolowanie diod LED. Sterownik konwertuje kombinacje klawiszy funkcyjnych [Fn] na wydarzenia ACPI. Dla wszystkich innych funkcji dostępny jest demon [10].

Użytkownicy komputerów Toshiba mogą wykorzystać specjalny sterownik ACPI, znany jako *Toshiba Laptop Extras* (`CONFIG_ACPI_TOSHIBA`). Niestety moduł używa `/proc/acpi/toshiba/keys` do przekazywania naciśnięć klawiszy funkcyjnych zamiast korzystać ze standardowych zdarzeń ACPI. Inne wpisy w `/proc/acpi/toshiba` służą do ustawiania jasności i sterowania pracą wiatraka. Dla urządzeń Toshiba są również dostępne demony [11].

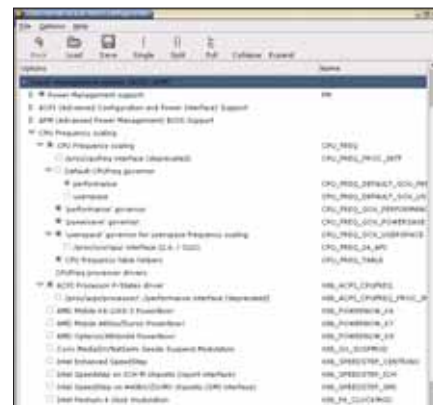
A co z dokumentacją?

Użytkownicy borykający się z problemami związanymi z Linuxem 2.6 i ACPI mogą zdać się na własną intuicję lub próbować wyszukać w sieci dokumentację. Niestety, materiałów na ten temat praktycznie nie ma. Autorzy publikacji starannie unikają tego zagadnienia. Wyjątek stanowi tutaj dokumentacja jądra, jednak nie zawiera ona wielu informacji na temat funkcji zarządzania energią, opartych na ACPI. Jedynie, co udało mi się wygooglać, to lista mailingowa `Acpi4linux` [12] oraz kilka grup dyskusyjnych, dzięki którym można wymieniać opinie z użytkownikami próbującymi rozwiązać podobne problemy.

Wnioski

Coraz więcej sprzętu pojawiającego się na rynku obsługuje ACPI, co wymusza na deweloperach podjęcie działań z tym związanych. Wielu użytkowników pamięta z pewnością wielką klępkę ACPI towarzyszącą Windowsowi 2000.A przecież cele specyfikacji ACPI są ważne – systematyczna konfiguracja sprzętu i jednolite zarządzanie energią.

Warto zaznaczyć przy tym, że przekierowywanie przerwań PCI przy użyciu ACPI pod Linuxem jest stabilne. Jednakże inna doskonała funkcja zarządzania energią w Linuxie 2.6, `Suspend-to-RAM`, jest



Rysunek 2. Opcje ACPI `cpufreq` dla CPU w urządzeniach przenośnych.

oznaczona jako eksperymentalna (*experimental*), ponieważ działa poprawnie na licznych komputerach.

Co najgorsze, w wielu systemach dostępnych na rynku znajdują się błędne tabele DSDT. Co ciekawe, część z nich jest licencjonowana przez producentów należących do konsorcjum ACPI. Użytkownikom pozostaje jedynie oczekiwać na poprawione tabele DSDT. Niestety, ACPI wymaga jeszcze wiele żmudnej pracy, aby mógł przejąć rolę centralnego interfejsu konfiguracyjnego do wolny sprzęt. ■

INFO

- [1] ACPI: <http://www.acpi.info>
- [2] ACPI4Linux: <http://acpi.sf.net>
- [3] Omówienie tabel DSDT ACPI4Linux: <http://acpi.sf.net/dsdt/index.php>
- [4] DSDT initrd patch: <http://gaugusch.at/kernel.shtml>
- [5] Intel IASL: <http://www.intel.com/technology/iapc/acpi/downloads.htm>
- [6] Naprawianie błędnych tabel DSDT: http://www.cpqlinux.com/acpi-howto.html#fix_broken_dsdt
- [7] Tryb Software Suspend dla Linuxa: <http://swsusp.sf.net>
- [8] CPU Dynamic Frequency Control: <http://mnm.uib.es/~gallir/cpudyn/>
- [9] Demon zdarzeń ACPI: <http://acpid.sf.net>, http://sf.net/project/showfiles.php?group_id=33140
- [10] ASUS/Medion Laptop Extras: <http://julien.lerouge.free.fr>
- [11] Narzędzia ACPI dla Toshiby: <http://fnfx.sf.net>, <http://sourceforge.net/projects/tclkeymon/>
- [12] Lista mailingowa ACPI4Linux: <http://acpi.sf.net/maillinglists.html>

Tabela 1. Tryby oszczędzania energii ACPI w Linuksie 2.6

Tryb oszczędzania energii	ACPI	Sleep_State off Opis
Standby (S1)	<i>standby</i>	System jest zasilany, CPU nie przetwarza danych.
Suspend-to-RAM (S3)	<i>mem</i>	System jest wyłączony z wyjątkiem pamięci głównej.
Suspend-to-Disk (S4)	<i>disk</i>	System jest wyłączony, pamięć główna i stan są zapisywane na dysku.

Uwaga: Stabt S0 (Running) oraz S5 (Soft Off) nie są trybami oszczędzania energii ACPI.