DIS project no.: 100375 DIS contract no.: 150466	Electricity programme	On behalf of the Swiss Federal Office of Energy

Final report, Juli 2004

Findings relating to the use of ACPI in Windows servers

prepared by
Alois Huser, Thomas Grieder
Encontrol GmbH
Bremgartenstrasse 2, 5443 Niederrohrdorf

energieschweiz

This study has been carried out on behalf of the Swiss Federal Office of Energy. All contents and conclusions are the sole responsibility of the authors.

Please refer to the following web site for additional information concerning the Swiss Federal Office of Energy "Electricity" programme: www.electricity-research.ch

Contents

Abstract

In a practical test, the possibilities and limits of the use of energy-saving sleep modes were compiled for servers of the lower price segment Many of these servers are used by small to medium-sized enterprises and are left running around the clock, although they are not used at night or during the weekend.

The implementation of the sleep modes is based on the so-called *ACPI* specification *(Advanced Configuration and Power Interface),* which was created by an industrial consortium. The most advanced of these is the implementation for the *Windows Server 2003* operating system. With *Linux,* the sleep modes are not yet fully available. The test was therefore carried out with the *Windows Server 2003* operating system.

In many applications, and in particular when the servers are used as file or web servers, no operational disadvantages result from the use of the rest states.

By making use of the lowest possible sleep mode during the night and at weekends, the energy consumption of the server can be halved.

For users, the results have been summarised in a two-page fact sheet, which is enclosed with this report as an appendix.

1 Summary

Power management has already been in use on desktop computers for a number of years. The power management of *Windows®* is based on the *Advanced Configuration and Power Interface (ACPI)* specification, which was created by a consortium headed by *Intel®* and *Microsoft®*. *ACPI* concerns the computer software and hardware to an equal extent. It incorporates all elements, from individual hardware devices through to device drivers, operating system and applications software.

The aim of this project was to gain practical experience with *ACPI* functionalities on low-end servers[1] and obtain findings regarding their limits. The findings are intended to encourage manufacturers to continue development work on the implementation of *ACPI* based power management and to motivate end-users to use this tool and better understand its functions, and thus to utilise pre-defined sleep modes. The results have been summarised for users in a separate two-page fact sheet.

No practical findings are available at the international level, nor are Swiss importers aware of any enquiries from customers. As before, the prevailing opinion is that servers always have to be switched on 24 hours a day, 7 days a week. Searches on the Internet nevertheless yielded a number of studies on power management in major computer centres, and it is conceivable that such studies may provide impulses in the future for improving power management in the area of low-end servers.

By contacting product managers in Switzerland it was possible to identify a low-end server that supports certain *ACPI* sleeping states. With this test system several operating modes were established and the behaviour of the server was recorded.

The server has two energy-saving states, namely *standby mode* and *hibernation mode*. In *standby mode* the reduction of power consumption is only marginally lower than when the server is operating without any active applications. It is only when it is in *hibernation mode* that consumption is reduced to below 10 W. If this mode is used at night and over the weekend it is possible to halve the server's energy consumption.

No problems arise when the system is used as a data and web server. However, there are certain limitations when it is used as a *domain controller* and *DCHP server*, though it is possible to avoid such difficulties if the server is reactivated via a scheduled task at fixed times of day. It is not possible to use sleep modes with *RAS servers* and for operation in *clusters*.

As a result of the findings obtained from this study some recommendations to manufacturers of hardware and software can be given:

Manufacturers of server hardware should consistently implement sleep states S3 and S4 (*Suspend to RAM* and *Suspend to Disk*). For this purpose the power supply of the server has to be optimised for the *sleep modes*.

Microsoft has drawn up a set of minimum requirements on systems and devices entitled *Microsoft® Windows® Logo Program*, with the aim of guaranteeing high-quality system performance under *Windows*. Within the scope of this program, *Microsoft* should insist on the implementation of states S3 and S4 as a compelling requirement. This would exert greater pressure on device manufacturers to fully implement power-saving modes.

[1] Low-end or entry-level server: processor and power supply normally not redundant, price below 30,000.00 Swiss francs.

The application software also has a decisive influence on the use of sleep modes. Program developers should set out to ensure that open network connections do not prevent the shutdown of clients and servers.

We recommend buyers and operators of low-end servers to already start demanding compatibility with *ACPI* when making their acquisitions. Suppliers of low-end servers should indicate the power consumption in normal and sleep modes. The enclosed fact sheet contains additional information concerning installation and operation.

The *Federal Office of Energy* should emphasise their actions on the following two items:

The findings of this study should be made available to a large group of system administrators by spreading the fact sheet. Thus a certain demand for sleep states in low end servers will be raised.

In addition, the *Office of Energy* can apply the findings described in the fact sheet for their own IT infrastructure. Acquisition and proper configuration of such servers can be coordinated with other federal offices, in order to obtain as broad an effect as possible.

2 Introduction

A number of studies on the topic of automatic shut down of servers and use of energy-saving operating modes have been carried out within the scope of the *Swiss Federal Office of Energy's* "Electricity" research programme. A representative survey was carried out among 400 small and medium-sized companies in order to determine the degree of interest in such solutions (Gubler M. & Peters M., 2000). 57% of the companies participating in the survey would welcome an automatic shut down system for servers. More than 90% of the servers in the companies concerned run 24 hours a day. The reasons for this are seldom of a practical nature, but rather are the result of uncertainties, habit and doubts concerning technical feasibility. This study aims to help eliminate such doubts and show that switching servers automatically is technically feasible.

There are no standards or internationally recognised guidelines in existence concerning the power consumption of servers. In view of this, the implementation of a power management system is especially important. On the manufacturer side there are various approaches to implementing power management in servers. The situation in the *Windows®* environment is the clearest example: here, the *ACPI* standard which was defined for personal computers, has been adopted. Although various manufacturers guarantee the system compatibility of certain low-end servers with this standard, it is still not known how *ACPI* sleep modes function in practice in a server environment, nor what their limits are.

The aim of this project was to gain practical experience with *ACPI* functionalities on low-end servers and obtain findings regarding their limits. The findings are intended to encourage manufacturers to continue development work on the implementation of *ACPI* based power management and to motivate end-users to use this tool and better understand its functions, and thus to utilise pre-defined sleep modes.

This project went in a different direction from others that are already well advanced and focused on the automatic_shut down of servers in order to save energy (Huser, A. 2001 and Sauter, B. 2003). Please refer to chapter 8 for a comparison between the two methods.

3 Principles of the *ACPI* standard

Power management[2] has already been in use on desktop computers for a number of years. The *Advanced Configuration and Power Interface (ACPI)* specification was created by an industry consortium with the aim of transferring power management from *BIOS*[3] to the operating system (Advanced Configuration and Power Interface Specification 2004). *ACPI* is oriented on both software and hardware. It incorporates all elements of a computer, from hardware to software device drivers through to applications software (Fig. 3-1).

[2] Definition of power management according to the ACPI specification: Mechanisms in software and hardware to minimize system power consumption, manage system thermal limits, and maximize system battery life. Power management involves trade-offs among system speed, noise, battery life, processing speed, and alternating current (AC) power consumption. Power management is required for some system functions, such as appliance (for example, answering machine, furnace control) operations.

[3] Basic Input Output System

However, it does not specify the criteria that have to be met in order for a server to qualify as compatible with *ACPI*. A manufacturer may claim compatibility with *ACPI* without having to include all the pre-defined modes.

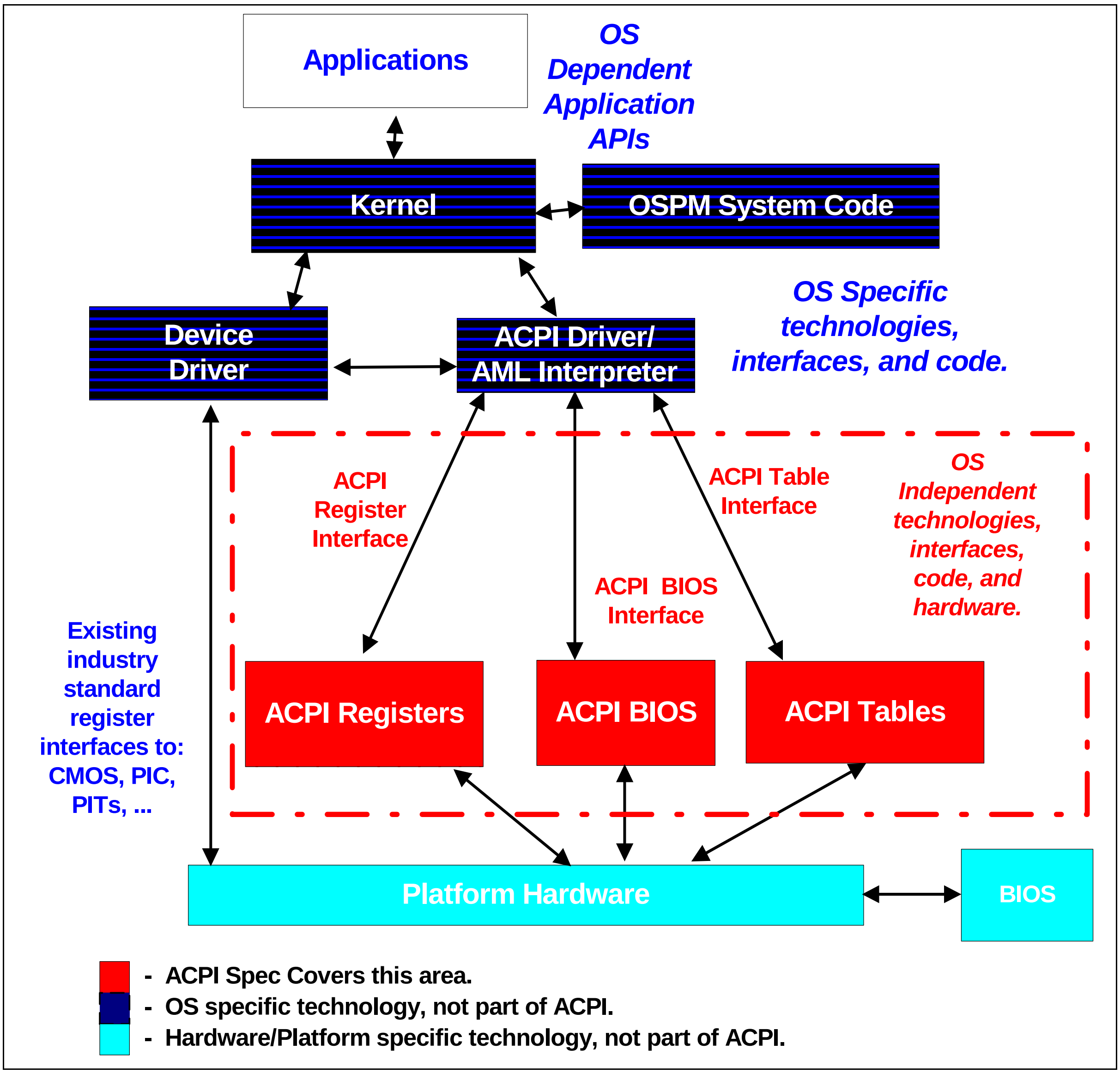

Fig. 3-1 ACPI elements (explanation of terms[4])

[4] *ACPI Machine Language (AML):* Pseudocode for a virtual machine supported by an ACPI-compatible OS and in which ACPI control methods and objects are written. The AML encoding definition is provided in section 17, “ACPI Machine Language (AML) Specification.”

OS: operating system (e.g. Windows®, Linux)

Operating System-directed Power Management (OSPM): A model of power (and system) management in which the OS plays a central role and uses global information to optimize system behavior for the task at hand.

3.1 System states

Basically a computer system may be equipped with four global states, referred to by the abbreviations G0 to G4 (Table 3-1).

State	Software running	Delay for switching to G0	Power consumption	OS restart required	Exit state electronically[5]
G0 Working	Yes	0	20 to 500 W	No	Yes
G1 Sleeping	No	< 1 sec to 1 min depending on standby state	1 to 100 W	No	Yes
G2/S5 Soft off	No	> 10 sec to 1 min	1 to 20 W	Yes	Yes
G3 Mechanical off	No	> 10 sec to 1 min	0 [6]	Yes	No

Table 3-1 ACPI global states

Four sub-states (S1 to S4) have also been defined for G1 (Table 3-2). In G2 (*soft off*), the system context is not preserved and the operating system has to be completely restarted.

Please refer to the appendix for a detailed description of the various states.

State	Software running	Delay for switching to G0	Power consumption	Deactivated system components
S1	No	< 1 sec	20 to 500 W	---
S2	No	< 1 sec	20 to 500 W	CPU, cache
S3 (Suspend to RAM)	No	Several seconds	1 to 50 W	CPU, cache, chip set
S4 Suspend to disk	No	Several seconds to 1 min	1 to 20 W	All devices, system context preserved on hard disk
S5 (Soft off)	No	> 10 sec to 1 min	1 to 20 W	All devices, system context not preserved

Table 3-2 ACPI sleeping states

[5] For example, by moving the mouse, pressing a key, incoming request from network, etc.

[6] Exceptions: support battery for system clock still active; with PCs and laptops, external power supply unit may still be active

Each system component may assume certain device states (D0 = fully operational to D3 = off), regardless of the global system state. These are normally not visible to the user.

For the central processing unit (CPU), further operating states have been defined as a subset of G0 (working), They are named C0 (full load) to Cn.

Normally the computer will automatically switch between G0 and G1 (*standby*) – see Fig. 3-2. Independently of this, some system components may be switched to a status with reduced power consumption (D1 to D3). In G0, the CPU will also be transferred to energy-saving states (C0 to Cn), depending on the load.

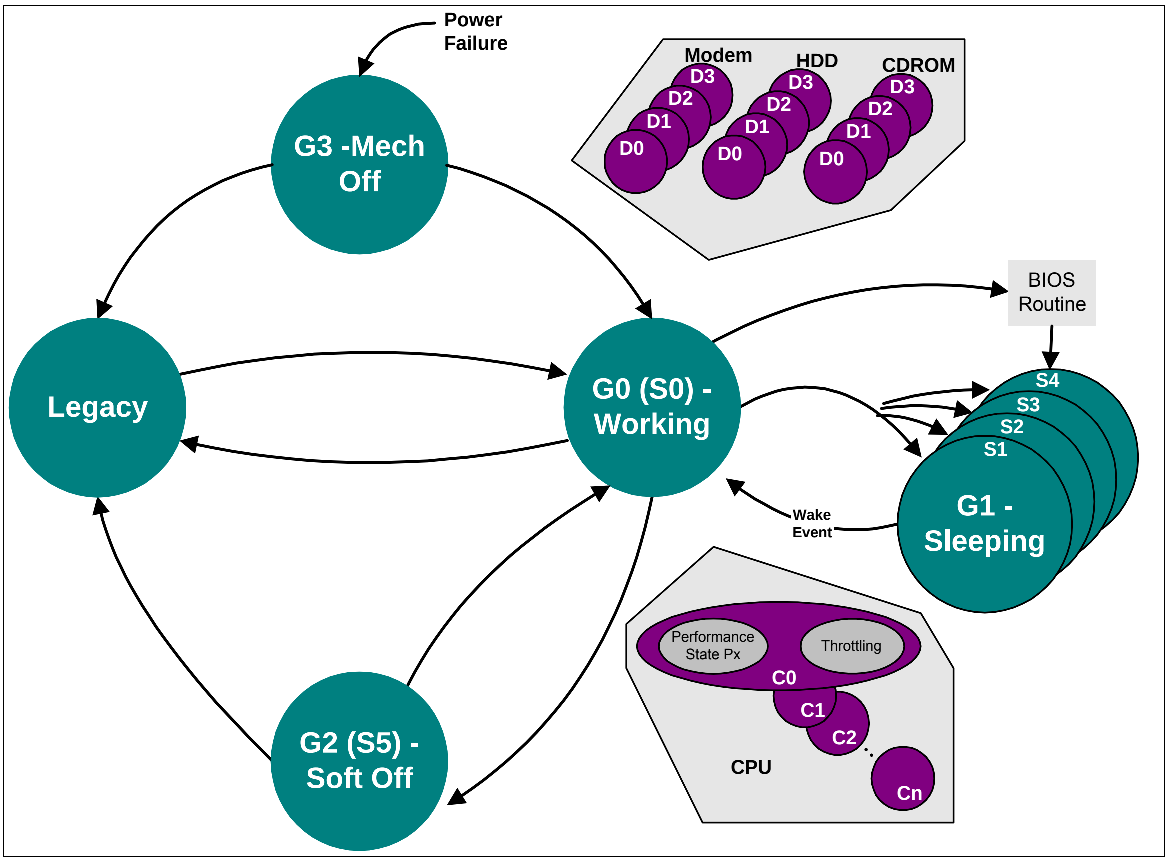

Fig. 3-2 Overview of ACPI operating states

3.2 *ACPI* and servers

The *ACPI* specification also applies to servers. In principle, they may be treated in the same way as workdesk computers. The following text has been taken from *ACPI* specification 2.0c and deals with the topic of servers:

"Perhaps surprisingly, server machines often get the largest absolute power savings. Why? Because they have the largest hardware configurations and because it's not practical for somebody to hit the off switch when they leave at night.

Day Mode. In day mode, servers are power-managed much like a corporate ordinary green PC, staying in the Working state all the time, but putting unused devices into low-power states whenever possible. Because servers can be very large and have, for example, many disk spindles, power management can result in large savings. OSPM allows careful tuning of when to do this, thus making it workable.

Night Mode. In night mode, servers look like home PCs. They sleep as deeply as they can and are still able to wake and answer service requests coming in over the network, phone links, and so on, within specified latencies. So, for example, a print server might go into deep sleep until it receives a print job at 3 A.M., at which point it wakes in perhaps less than 30 seconds, prints the job, and then goes back to sleep. If the print request comes over the LAN, then this scenario depends on an intelligent LAN adapter that can wake the system in response to an interesting received packet."

3.3 Compatibility with Windows

Microsoft has drawn up a set of minimum requirements on systems and devices entitled *Microsoft® Windows® Logo Program* (or *WLP*), with the aim of guaranteeing high-quality system performance under *Windows* (Microsoft Windows Logo Program System and Device Requirements, 2004).

For servers, this program demands compatibility with *ACPI Specifications Version 1.0b*, and at least *standby state* S5. Servers used in the home and by small businesses (*SOHO*[7] servers) also have to support S1, while other sleeping states (S1, S2 and S4) are not a requirement[8].

Fig. 3-3 Microsoft label for the distinction of compatible servers

Microsoft has released a special kit for testing compatibility of hardware with its operating systems[9]. It includes an *ACPI* stress test that can also be used on servers. One part of this test consists in determining whether network connections are restored within 20 seconds after the server wakes up from a sleeping state.

---

[7] SOHO = Small Office, Home Office

[8] Server Note – Clarification: At a minimum, SOHO servers must support S1. Other server systems are not required to support S1 or S3. Server systems may choose not to support S4. All server systems must be AC PI-compliant and should support S5, which enables operating system support of S4. Servers with more than eight physical processors are not required to be ACPI-compliant.

[9] Microsoft Windows Hardware Compatibility Test Kit: http://www.microsoft.com/whdc/hwtest

The *Windows Server 2003* operating system comprises the following products:

- *Standard Edition*
- *Enterprise Edition*
- *Datacenter Edition*
- *Web Edition*
- *Small Business Server*

In special versions, *Enterprise* and *Datacenter* support 64-bit architecture. The *ACPI* specification only supports this architecture as of version 2.0. Under *Windows Server 2003*, *Enterprise Edition 64-bit version* and *Windows Server 2003, Datacenter Edition 64-bit version*, some functions are missing that are available in the 32-bit versions of these operating systems. This includes most of the *ACPI* functionality. Only the *ACPI* 2.0 tables that were adapted for the 64-bit environment are supported. For *SOHO* servers however, which are the types normally in use in small and medium-sized businesses, these 64-bit versions do not yet come into the question.

3.4 Network cards

Network cards have to be capable of waking a server from the sleeping states. For this purpose it is sufficient that "meaningful" data are forwarded to the server via the network. Examples include:[10]

- Receipt of a *Magic Packet*™
- Remote access by an administrator
- Network data that are directly addressed to the system

[10] Extract from ACPI specification: Chapter A.9.3.2: Wake frame events are used to wake the system whenever *meaningful* data is presented to the system over the network. Examples of meaningful data include the reception of a Magic Packet™, a management request from a remote administrator, or simply network traffic directly targeted to the local system.

4 Processing international findings

4.1 Research on the Internet

Research on the Internet concerning the topic of power management in servers yielded very little relevant information.

4.1.1 Intel

Chip manufacturer *Intel* describes its efforts with respect to new technologies in an issue of *Technology@Intel Magazine* (Bodas D. 2004). Here the main goal is to keep the generation of heat in large servers within reasonable bounds. To accomplish this, *Intel* is pursuing a three-level strategy:

Demand-based switching (*DBS*) adapts the power consumption of processors to the load. This is accomplished by adjusting the operating voltage and clock rate. Similar technologies are being applied by the mobile technology Centrino™. The savings potential attributed to this technology is up to 30 percent.

At the data centre level, *Intel* is working on *ACPC* (*automatic control of power consumption*) technology, by means of which it is possible to monitor the power consumption of a server via a feedback mechanism, and to limit it as required. In this way, administrators will be able to limit the heat load in their data centre in crisis situations.

ACPC is supplemented by an overlapping technology, *EPTM* (*Enterprise Power and Thermal Manager*). This is a software tool that combines the options provided by *DBS* and *ACPC*. It enables the dynamic allocation of power consumption at various organisational levels, from rack to complete data centre. In this way it is possible to optimise the performance and power consumption at a higher level.

Although the incentive to develop this technology came from the world of large-scale data centres, *demand-based switching* would also be able to have a positive impact in the area of SOHO servers.

4.1.2 American Megatrends

On its website, *BIOS* manufacturer *American Megatrends* presents a product called *AMIBIOS® Server* with the following features[11]:

- Fully ACPI ready
- ACPI system states S1, S2, S3, S4 and S5
- ACPI device states D1, D2 and D3
- Remote wake up via modem and network

To accomplish full implementation of *ACPI* power-saving states it is essential that all components in the chain (operating system, *BIOS*, drivers and hardware) work optimally together. *Windows Server 2003* is able to meet this requirement as an operating system. With *AMIBIOS*, manufacturers also have a *BIOS* at their disposal that functions as the next link in the chain and also meets all requirements. The fact that there are no existing low-end servers with full implementation of power-saving states indicates that either there are still compatibility problems concerning drivers and hardware devices, or manufacturers are not interested in full implementation.

---

[11] Source: http://www.ami.com/support/doc/amibiosserver2000.pdf

4.1.3 Duke University

Some interesting comments are to be found in a “position summary” by *Duke University*, North Carolina, USA (Chase J. & Doyle R.). This report indicates that *ACPI* sleeping states could also be of importance for large-scale server clusters in the future. This would trigger major development activity from which *SOHO* servers would also stand to benefit.

The authors state that “under light load [of a server cluster[12]] it is most efficient to use power management (e.g. *ACPI*) to step some servers to low-power states ... our premise is that servers are an appropriate granularity for power management in clusters ... simply ‘hibernating’ idle servers provides adequate control over on-power capacity in large clusters, and it is a simple alternative to techniques (e.g. voltage scaling) that reduce server power demand under light load ... our system is based on reconfigurable switches that route request traffic toward the [already] active servers and away from inactive serves ... it enables the system to concentrate request traffic on a subset of servers running at higher utilization.”

4.1.4 HP Business Support Center

From two documents in *HP’s Business Support Center* it may be deduced that certain products within its range of servers offer *ACPI* functions. However, *HP’s* product management in Switzerland was unable to provide any more detailed information:

In one *support document* it is noted that the products of the *ProLiant* family have been developed in accordance with the requirements of the *Microsoft Windows Logo Program*. This means that at least states S1 and S5 would have to be implemented. In the device documentation, compatibility with *ACPI* 1.0b is cited, but there is no information included concerning standby modes and associated consumption levels.

A second *support document* refers to problems experienced with *ProLiant servers* relating to waking from *hibernation mode* when connected to a number of *virtual LANs (VLANs)*. From this it may be deduced that state S5 has been implemented and has also been used on a number of occasions. Here, too, *HP* product management in Switzerland was unable to provide any more detailed information.

4.2 Manufacturers’ Internet forums

An enquiry concerning power management functions in servers was placed in the user forums of two manufacturers. In one case, no reply was given within a period of two weeks. In the other case, two replies were posted on the forum. The first of these simply referred to a general description of power management in desktop PCs, while the second may be regarded as representative of the still prevailing opinion concerning this topic. It stated that there is no power management in *ProLIant* servers … the hardware is not slowed down or switched into a power-saving mode … [*Proliant servers*] have powerful ventilators, extremely efficient power supply devices, an excellent thermal layout and thermal shut-down in the event of malfunction … in other words, the entire hardware is active at all times, and this is correct and absolutely essential for ensuring a fast response time.

[12] A computer cluster, which is usually called simply a cluster, refers to a number of networked computers that work together simultaneously to process segments of a shared major task.

5 Findings from enquiries to distributors operating in Switzerland

A survey carried out among the main distributors in Switzerland yielded remarkably uniform findings. On initial enquiry, most product managers for servers for use in small offices and homes (*SOHO* servers) did not know what the term *ACPI* referred to. They were not aware of any enquiries from clients and did not have any experience in the area of power management. It was only after further enquiries had been made at support centres of their parent companies that some of them were able to provide statements on this topic. The following manufacturers were involved in the survey:

- *Dell*
- *Fujitsu-Siemens*
- *IBM*
- *HP*

5.1 Dell

Statement by product manager H. P. Odermatt:

"Entry-level servers from many manufacturers are usually desktops dressed up as servers, and therefore include *ACPI* support. *Dell*'s product for this market – the *PowerEdge 400SC* – is a good example of this. Power management in servers is normally controlled by the operating system. Both *Windows XP Client* and *Windows 2003 Server* support *ACPI*. However, the question of the usefulness of standby functions on servers still needs to be addressed. The main features of a server are high availability and rapid access, not standby functions with correspondingly slow response times upon reactivation."

However, in the accompanying documentation to the *Poweredge 400SC* the term *ACPI* is not mentioned. There are no specifications concerning sleeping states or consumption levels for these states.

5.2 IBM Switzerland

IBM was the only company we addressed that had an employee in Switzerland responsible for product security and environmental aspects. We therefore had a highly qualified contact person at our disposal for this study.

According to B. Oldani, buyers of servers show very little interest in questions relating to power management. *IBM Germany* has occasionally been confronted with such questions, but in Switzerland, enquiries from development departments have only been triggered as a result of this project. In view of this, the topic of power management will be addressed at a forthcoming internal conference[13].

There is no information about *ACPI* and sleeping states in any *IBM* data sheets. Upon request, *IBM* will also provide environmental declarations which are based on the criteria of the *ECMA*[14].These documents, however, also lack information concerning power consumption in any given sleeping mode, since these specifications are not listed as a requirement by the *ECMA*. It is only by studying

[13] IBM: Environmental Leadership Conference, Washington DC, USA (to be attended by Silvio Wehren, IBM Germany)

[14] European Computer Manufacturing Association, http://www.ecma-international.org/

certain internal specifications that it becomes apparent that servers in the *x series* are in fact compliant with *ACPI 1.0b*.

5.3 HP

No enquiries have been received to date concerning standby states. However, according to feedback from Support there are clients (in the small and medium-sized companies category) who would prefer to switch off their servers at weekends, and have in fact carried out according trials. But in each case, difficulties in reactivating the server prompted the operator to abandon the attempt and leave the server running permanently.

6 Test system

6.1 Hardware

A number of product managers were contacted in order to find a standard commercial low-end server that, according to its product description, supports *ACPI* specifications. *IBM* has such a server in its product range that is *ACPI* compatible, including the *RAID*[15] controller. Its technical specifications are as follows:

Product description:	IBM eServer xSeries 225 (model 8647)
Processor:	Xeon 2.67 GHz
Mainboard:	Dual processor
RAM:	512 MB
HDD:	3 x 36.4 GB Ultra 320 SCSI Hot-Swap SC
Network card:	Broadcom NetXtreme Gigabit Ethernet
Power supply unit:	Max. capacity 425 W
Ventilation:	Three casing ventilators (no processor ventilators)
BIOS:	Phoenix V. 6.0

The network card is integrated on the mainboard and is based on a *Single Chip Gigabit Ethernet solution.* A control for power management and a *Wake on LAN* are integrated into this switching unit[16].

6.2 Software

Windows 2003 Server Standard Edition was the operating system installed on the server. In order to ensure that *ACPI* functions faultlessly, it is necessary to use the latest drivers, and in some cases an upgrade needed to be installed.

6.3 Declaration of product-related specifications in accordance with ECMA

The manufacturer's declared product-related specifications in accordance with *ECMA*[17] were as follows (extract):

Weight:	20.7 kg (maximum configuration)
Dimensions:	46.6 x 16.5 x 66 cm (height, width, depth)
Power consumption:	
Working:	425 W (max.)
Standby:	110 W
Off:	10 W

[15] RAID (= Redundant Array of Inexpensive / Independent Disks) systems are used for organising multiple hard disks on a computer. In this way it is possible to enhance the level of operating security, performance and/or capacity. A variety of options have been defined as RAID levels.

[16] Broadcom: On-chip power circuit controller and Wake on LAN power switching circuit

[17] European Computer Manufacturing Association, http://www.ecma-international.org/

Acoustic emission (sound power / sound pressure levels)
Working: 6.1 Bel / 46 dB(A)
Standby: 6.0 Bel / 45 dB(A)

Heat emission: 425 W

The system is compatible with *ACPI*, Version 1.0a.

6.4 Measurements

6.4.1 Results

The power consumption was measured in three configurations and different operating modes using an *EMU 124.K* measuring device:

Configuration 1: 3 x SCSI HDD with RAID controller

Configuration 2: 1 x IDE HDD

Configuration 3: 1 x IDE HDD, 3 x SCSI HDD with RAID controller

State	Config. 1	Config. 2	Config. 3
G0 Working, CPU at full load	153 W	98 W	165 W
G0 Working, no active applications	118 W 132 VA power factor 0.89	88 W 102 VA power factor 0.86	126 W
G1 Standby (ACPI S1 or S2, mainboard does not support S3)	116 W	82 W	121 W
S4 Suspend to disk	8 W 39 VA power factor 0.21	8 W	8 W

Table 6-1 Measurement of power consumption in different operating modes

The sound pressure level was measured at a distance of 1 metre using a *TES1350 digital sound level meter*. The reading was 57 dB(A) in both normal and *standby* mode.

6.4.2 Evaluation

The difference in power consumption between operation without any active applications and *standby* mode was only 2%. This very low reduction may be attributed to the fact that the hardware does not support state S3 (*Suspend to RAM*). According to the manufacturer, S3 is only implemented on main-boards with a single processor. For two processors, additional development would have to be initi-ated.

The power requirement of the *RAID* controller and the three SCS hard disks is 40 W. In operating mode without active applications, this is equivalent to a third of the total server capacity.

7 Findings from the use of ACPI in the test system

7.1 BIOS settings

Power management settings are configured in *BIOS.* In order to use the *standby* and *hybernate states* it is necessary to set the respective *Wake Up* functions to *Enabled* (Fig. 7-1).

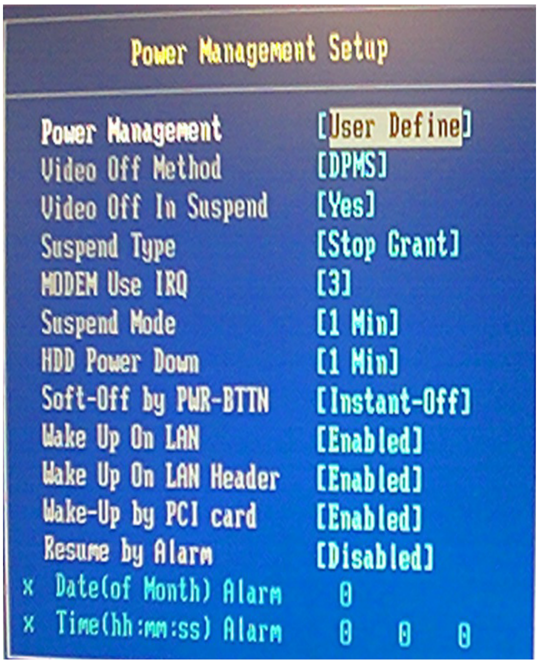

Fig. 7-1 Power management setting in BIOS

7.2 Operation system settings

Windows power management settings are configured in the same way as screensaver settings (= via *Control panel* → *Display* → *Screensaver* → *Power management*). Here the delay times may be configured for all available power-saving modes into which the server is to switch automatically after a specified period of inactivity.

Our measurements revealed that wake-up times from these states are very slow: *standby mode*, 15 to 25 seconds, *hybernate mode* approximately 75 seconds. For this reason, we configured the delay times so that the server only switched into *hybernate mode* outside of normal working hours (settings: *standby* 15 minutes, *hybernate mode*, 1 hour). In this way it would still be possible to use *standby* mode during working hours.

In order for the server to be reactivated via the network, the option "*Device can activate computer from standby mode"* has to be selected from the *Power management* menu (= via *Control panel* → *Hardware* → select network card → *Power management).*

7.3 Use as data server

The test system was used as a data server for 1 week by a workgroup comprising several users. Each evening it automatically switched into *hybernate mode* one hour after it had last been used. The following morning, the system was reactivated as soon as the first user logged on. The power consumption for the 7-day period was 9.4 kilowatt hours. This is equivalent to an average capacity of 56 watts, which represents a reduction by 52 percent versus consumption in permanent operation (118 W).

The users integrated the data server into the local system as a network drive. When the first user logged on each morning, a warning appeared stating that this network path could not be found and connection had therefore not been restored. The reason for this was that the server had not yet been fully reactivated (waiting time, around 75 seconds). Once the server had been fully reactivated, the network connection was restored without the need for further intervention by the user. This behaviour therefore does not have any negative consequences for working with the server after reactivation.

If this behaviour is regarded as unacceptable, it is possible to avoid it by configuring the data server to start up automatically at a fixed time each day prior to the first user log-in (cf. chapter 7.9).

Another restriction that was identified was that it proved impossible to switch either the local client device or the server into *standby mode* if an *Office* application (*Word*, *Excel*, *Access*, *Powerpoint*) was still accessing documents or application components on the network.

7.4 Use as domain controller

When the server is configured as a domain controller[18], *Windows* automatically sets the power management options to *permanent operation*, and thus deactivates switching into *standby* mode. It is therefore necessary to reconfigure the power management settings in the same way as described in chapter 7.2 above.

When a user logs in to the system, the client computer requests the stored profile from the *domain controller*. All modifications to the profile are automatically saved on the server when the client device is shut down. If the *domain controller* is not available, the client computer uses a locally stored copy of the profile. In this case, modifications can only be saved locally and will be overwritten with the profile stored on the server next time the device is powered up.

This behaviour gives rise to a restriction relating to the use of sleep states. If the *domain controller* should be in *standby* or *hybernate mode* after a given period of inactivity, it will be woken up following the next log-in by a client computer, and first has to be reactivated. During this period, the user profile stored on the server is not available and the client computer therefore accesses the local one. This behaviour does not give rise to any problems as long as no modifications have been made to the profile. However, if any modifications have been made and should be saved on the server, the user will have to restart his/her client computer before and after the modifications with the server already activated. This is necessary in order to ensure that the modifications are not lost.

7.5 Use as DHCP server

DHCP[19] enables the integration of a new computer into an existing network without the need for additional configuration. When starting up a client computer, the server first has to issue the client

---

[18] The term *domain* refers to a group of networked computers. The domain controller manages access to the computers in a given domain.

[19] DHCP (*Dynamic Host Configuration Protocol*) enables the dynamic allocation of IP addresses and other configuration parameters to computers in a network.

with an IP address[20] before the user logs in. This has to take place without any delays, otherwise the log-in attempt will fail.

When sleep modes are used, a problem arises here that prevents the proper functioning of DHCP. The first client that logs on to the network is unable to return the server to normal operation sufficiently quickly due to the DHCP search. As a result, the client issues itself with an own IP address (auto-configuration) and will therefore not be identified by the domain controller. This means the computer concerned will not have access to the network.

There are two possible solutions to this problem:

- The computers in the network can be configured with a fixed IP address. This is feasible in a small network with a low number of connected computers.
- It is also possible for the DHCP functions to be taken over by a router or firewall. In this case, the component concerned must be kept permanently in operation. The power consumption of these devices is significantly lower (generally speaking, at 10 to 20 watts) than that of a server.

7.6 Use as web server

Using the system as a web server did not give rise to any problems relating to *ACPI*.

7.7 Use as RAS server

When used as an LAN / WAN router or RAS[21] server, RAS prevents automatic switching into sleep states. After the configured waiting time has expired, an error message will appear.

However, it is possible for a user with the necessary access rights (e.g. a system administrator) to switch the server to *standby mode*. Here, the user concerned does not receive any information regarding existing connections to client computers, and all active connections are interrupted.

7.8 Use of server in clusters

According to *Windows Help*, *ACPI* power management functions must not be used on members of server clusters[22]. A cluster member that deactivates its disk drives or switches into *system standby* or *system hybernate mode* can trigger a cluster failure. If a power-saving state has been activated on several cluster members, this can lead to the failure of the entire cluster. Cluster members have to use a power management scheme in which the option *Deactivate hard disks* is set to *Never* – e.g. the scheme *permanent operation*.

Small and medium-sized companies are unlikely to use several servers in a cluster, however, so this problem is of no relevance to the present study.

---

[20] IP addresses permit the logical addressing of computers in networks. Version 4 IP addresses normally appear as series of four groups of digits separated by dots (e.g. 192.168.0.34, 127.0.0.1).

[21] RAS (*Remote Access Service*) is a Windows service for enabling remote access

[22] A computer cluster, which is usually called simply a cluster, refers to a number of networked computers that work together simultaneously to process segments of a shared major task.

7.9 Automatic start-up based on predefined schedule

Restrictions arose for some uses of the server since wake up from sleep states was relatively slow. It is possible to avoid these drawbacks if the server is automatically powered up on working days before the first client logs in. This can be accomplished under *Windows* by defining a *scheduled task*.

Procedure: *Start* → *Settings* → *Control panel* → *Scheduled tasks*. In this context the type of task is irrelevant, for example an input request can be selected. However, it is essential to select *Reactivate computer for performing the task*, since this will ensure that the computer is reactivated by the task. To access this option, it is necessary to activate the *Open additional settings for this task after clicking on 'Finish'* box before closing the Assistant. When the corresponding pop-up window appears on screen, the option *Reactivate computer for performing the task* has to be selected from the menu.

8 Comparison between off and sleep modes

Many servers are not used at night and over the weekend, yet they are left on in full operating mode and consume a great deal of electricity (cf. Table 1, "G0, no active applications"). This study focused on automatically switching a server into the *standby* or *hybernate modes* with reduced power consumption via an *ACPI*-defined function.

As an alternative, the server can be completely switched off during these periods, and this may also be automated, for example with the aid of an intelligent UPS[23] device or via an electronically controlled multiple socket outlet [24]. The feasibility and effectiveness of such solutions have been studied and documented in several pilot projects during the past few years (Held M. 2003, Huser A. 2001, Sauter B. 2003), and the necessary devices are now available for commercial use.

In *hybernate mode*, the server is not completely separated from the power supply – its consumption is still around 10 watts. Switching off the server completely means that this power consumption is eliminated altogether. There are other advantages associated with separation from the power supply that cannot be realised with *ACPI* sleep states:

- Separation reduces the risk of fire during periods when the office is not occupied.
- It reduces the risk of harm to the system due to power supply disturbances (lightning, power surges, etc.).
- Completely switching off the server prevents unauthorised access to data, both internally and externally.
- Daily rebooting means that the server has to be faultlessly installed and configured, and thus increases the reliability of the system.

Drawbacks associated with this more radical solution include:

- External devices are required (UPS, controlled outlet) for an automated solution.
- The server has to be actively switched, either via a programmed schedule or manually by a user on site (e.g. by pressing a corresponding button on a UPS device, or activating the server via a controlled outlet). Logging in to a client computer does not trigger an automatic start-up.

---

[23] UPS = Uninterruptible Power Supply device that, with the aid of a battery, continues to supply electricity to the connected equipment for a certain length of time following a power failure.

[24] e.g. Memo Switch Netcontrol, www.emt.ch

9 Conclusions

In a variety of studies the Swiss Federal Office of Energy's "Electricity" research programme has demonstrated that both switching and using energy-saving modes are advantageous in low-end servers, from the point of view of energy efficiency as well as operational aspects. This study set out to gain practical findings using a standard commercial low-end server.

For buyers and operators of servers, the findings of this study have been summarised in a separate two-page fact sheet (cf. Appendices), which contains practical hints for using power-saving states (*standby* and *hybernate mode*), from the purchase stage through to the installation and operation of a server.

9.1 Scope of action for the Federal Office of Energy

The findings of this study point out clearly that the ACPI sleep states can be used on a large number of servers without any operational disadvantages. This information should be made available to the system administrators of small and medium-sized companies. The attached fact sheet was designed for this purpose. Making it available to a large group of system administrators shall result in an encouragement for suppliers to improve the energy efficiency of their products.

In addition, the *Office of Energy* can apply the findings described in the fact sheet for their own IT infrastructure. Acquisition and proper configuration of such servers can be coordinated with other federal offices, in order to obtain as broad an effect as possible.

9.2 Recommendations for hard- and software manufacturers

The following conclusions may provide manufacturers of server hardware and software with encouragement to further improve the energy-efficiency of their products.

9.2.1 Optimisation of power supply

In studies carried out by Aebischer B. & Huser A. (2002) it was ascertained that the efficiency of a power supply unit declines sharply when the load is below 20% of the nominal level. Developers of such devices should therefore set out to ensure that the output capacity is adapted as closely as possible to actual requirements. Unnecessarily generous reserve capacities lead to higher losses. According to the data sheet, the maximum power consumption of the server measured in this study was 425 watts, but at full load and in a configuration with 2 processors and 4 hard disk drives, the effective power requirement was only 165 watts. This means that the power supply was overdimensioned by more than 100%.

Furthermore, developers of power supply units should also consider the question of consumption in the sleep modes. With the server measured in this study, the power consumption in the hybernate mode was still in the region of 10 watts, but with certain technical measures it would be possible to reduce this to below 1 watt. For this purpose it is necessary to install a separate 5-volt path in the power supply unit that has been optimised to ensure the lowest possible power consumption. Such electronic circuits are widely used today in PC power supply devices.

9.2.2 Implementation of ACPI states S3 and S4

Server mainboards which are able to accommodate several processors only have the ACPI state S1 implemented (if there are any *ACPI* functions implemented at all). The resulting reduction in power consumption is very limited. The availability of states S3 and S4 (*Suspend to RAM* and *Suspend to disk*) is a prerequisite for attaining significant reductions, so it is important for these to also be implemented.

9.2.3 Operating systems

In its logo program *Microsoft* only cites minimal demands with respect to the implementation of *ACPI* states (Microsoft Windows Logo Program System and Device Requirements, 2004). If it were to include S1 to S4 in its list of requirements in the same way as for desktop computers, this would exert greater pressure on manufacturers to fully implement *ACPI*.

9.2.4 Applications

In the course of this study it was found that, when the test system was used as a data server, neither the local client computer nor the server could be switched into *standby mode* as long as an *Office* application was still accessing documents or application components across the network. This means that one single user of a client computer who does not close an *Office* program may prevent the server from switching into *standby* or *hybernate* mode. Applications should therefore be developed so that they automatically close any files that were accessed across the network after a specified period of inactivity.

10 References

Aebischer B. & Huser A. (2002): *Energieeffizienz von Computer Netzgeräten*, CEPE & Encontrol, on behalf of the Swiss Federal Office of Energy, Berne, 2002), http://www.electricity-research.ch

Advanced Configuration and Power Interface Specification (2004): Compaq Computer Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd., Toshiba Corporation, Revision 2.0c, August 2003

Bodas D. (2004): *New Server Power-Management Technologies Address Power and Cooling Challenges,* Technology@Intel Magazine,2004, http://www.intel.com/update/contents/sv09031.htm

Chase J. & Doyle R.: *Position Summary: Energy Management for Server Clusters,* Department for Computer Science, Duke University, USA, www.cs.duke.edu/ari/publications/bop-statement.pdf

Gubler M. & Peters M. (2000): *Servernutzung in Klein – und Mittelbetrieben*, IPSO, on behalf of the Swiss Federal Office of Energy, Berne, http://www.electricity-research.ch

Held M. (2003): *Auswirkungen von periodischem Ein- und Ausschalten auf die Server-Hardware-Zuverlässigkeit*, EMPA, on behalf of the Swiss Federal Office of Energy, Berne, 2003), http://www.electricity-research.ch,

Huser A. (2001): *Schalten von Servern in KMU's, Marktbedarf und Pilotanlagen*, Encontrol, on behalf of the Swiss Federal Office of Energy, Berne, January 2001), http://www.electricity-research.ch,

Microsoft Windows Logo Program System and Device Requirements (2004): *Design and testing requirements for server, desktop, and mobile systems and devices that run the Microsoft Windows family of operating systems*. Microsoft, Version 2.2, January 2004

Sauter B. (2003): *Energieeffizientes Servermanagement*, Energy Management Team AG, on behalf of the Swiss Federal Office of Energy, Berne, September 2003), http://www.electricity-research.ch,

Appendices

Appendix 1: Descriptions of system states (extracts from ACPI specifications)

Global states

G3 Mechanical Off

A computer state that is entered and left by a mechanical means (for example, turning off the system's power through the movement of a large red switch). Various government agencies and countries require this operating mode. It is implied by the entry of this off state through a mechanical means that no electrical current is running through the circuitry and that it can be worked on without damaging the hardware or endangering service personnel. The OS must be restarted to return to the Working state. No hardware context is retained. Except for the real-time clock, power consumption is zero.

G2/S5 Soft Off

A computer state where the computer consumes a minimal amount of power. No user mode or system mode code is run. This state requires a large latency in order to return to the Working state. The system's context will not be preserved by the hardware. The system must be restarted to return to the Working state. It is not safe to disassemble the machine in this state.

G1 Sleeping

A computer state where the computer consumes a small amount of power, user mode threads are not being executed, and the system "appears" to be off (from an end user's perspective, the display is off, and so on). Latency for returning to the Working state varies on the wake environment selected prior to entry of this state (for example, whether the system should answer phone calls). Work can be resumed without rebooting the OS because large elements of system context are saved by the hardware and the rest by system software. It is not safe to disassemble the machine in this state.

G0 Working

A computer state where the system dispatches user mode (application) threads and they execute. In this state, peripheral devices (peripherals) are having their power state changed dynamically. The user can select, through some UI, various performance/power characteristics of the system to have the software optimize for performance or battery life. The system responds to external events in real time. It is not safe to disassemble the machine in this state.

S4 Non-Volatile Sleep

A special global system state that allows system context to be saved and restored (relatively slowly) when power is lost to the motherboard. If the system has been commanded to enter S4, the OS will write all system context to a file on non-volatile storage media and leave appropriate context markers. The machine will then enter the S4 state. When the system leaves the Soft Off or Mechanical Off state, transitioning to Working (G0) and restarting the OS, a restore from a NVS file can occur. This will only happen if a valid non-volatile sleep data set is found, certain aspects of the configuration of the machine have not changed, and the user has not manually aborted the restore. If all these conditions are met, as part of the OS restarting, it will reload the system context and activate it. The net effect for the user is what looks like a resume from a Sleeping (G1) state (albeit slower). The aspects of the machine configuration that must not change include, but are not limited to, disk layout and memory size. It might be possible for the user to swap a PC Card or a Device Bay device, however.

Notice that for the machine to transition directly from the Soft Off or Sleeping states to S4, the system context must be written to non-volatile storage by the hardware; entering the Working state first so that the OS or BIOS can save the system context takes too long from the user's point of view. The transition from Mechanical Off to S4 is likely to be done when the user is not there to see it.

Because the S4 state relies only on non-volatile storage, a machine can save its system context for an arbitrary period of time (on the order of many years).

Global system state	Software runs	Latency	Power consumptio n	OS restart required	Safe to dis-assemble computer	Exit state electroni cally
G0 Working	Yes	0	Large	No	No	Yes
G1 Sleeping	No	>0, varies with sleep state	Smaller	No	No	Yes
G2/S5 Soft Off	No	Long	Very near 0	Yes	No	Yes
G3 Mechanical Off	No	Long	RTC battery	Yes	Yes	No

Notice that the entries for G2/S5 and G3 in the Latency column of the above table are “Long.” This implies that a platform designed to give the user the appearance of “instant-on,” similar to a home appliance device, will use the G0 and G1 states almost exclusively (the G3 state may be used for moving the machine or repairing it).

Standby states

Sleeping states (S*x* states) are types of sleeping states within the global sleeping state, G1. The S*x* states are briefly defined below. For a detailed definition of the system behavior within each S*x* state, see section 7.3.4, “System _S*x* States.” For a detailed definition of the transitions between each of the S*x* states, see section 9.1, “Sleeping States.”

S1 Sleeping State

> The S1 sleeping state is a low wake latency sleeping state. In this state, no system context is lost (CPU or chip set) and hardware maintains all system context.

S2 Sleeping State

> The S2 sleeping state is a low wake latency sleeping state. This state is similar to the S1 sleeping state except that the CPU and system cache context is lost (the OS is responsible for maintaining the caches and CPU context). Control starts from the processor’s reset vector after the wake event.

S3 Sleeping State

> The S3 sleeping state is a low wake latency sleeping state where all system context is lost except system memory. CPU, cache, and chip set context are lost in this state. Hardware maintains memory context and restores some CPU and L2 configuration context. Control starts from the processor’s reset vector after the wake event.

S4 Sleeping State

> The S4 sleeping state is the lowest power, longest wake latency sleeping state supported by *ACPI*. In order to reduce power to a minimum, it is assumed that the hardware platform has powered off all devices. Platform context is maintained.

S5 Soft Off State

> The S5 state is similar to the S4 state except that the OS does not save any context. The system is in the “soft” off state and requires a complete boot when it wakes. Software uses a different state value to distinguish between the S5 state and the S4 state to allow for initial boot operations within the BIOS to distinguish whether or not the boot is going to wake from a saved memory image.

Appendix 2: Microsoft Windows Logo Program System and Device Requirements

Design and testing requirements for server, desktop, and mobile systems and devices that run the Microsoft Windows family of operating system. Version 2.2, January 2004

A1.1.2 ACPI system board and ACPI firmware

ACPI BIOS must be Windows-ready:

x86-based client and server systems: ACPI-compliant as defined in ACPI 1.0b, Section 1.6 The system board and BIOS must support Advanced Configuration and Power Interface Specification, Revision 1.0b.
System board devices not power managed or configured via standard bus specifications must comply with ACPI 1.0b.

Itanium-based systems: ACPI 2.0 plus 64-bit fixed tables; see A5.4.1.

A1.1.2.2 Power management and Plug and Play capabilities are ACPI compliant.

A1.1.2.2.1 All system-board power management or Plug and Play features comply with the ACPI 1.0b.

This requirement applies even if a particular feature is not a specific requirement or recommendation in ACPI 1.0b.

A1.1.2.2.2 System supports S3, S4, and S5 states.

Desktop and single-processor workstation systems must support the S5 (soft-off) state as required in the ACPI 1.0b specification, plus the S3 and S4 states. Every system must support wake from all implemented sleep states, except S4 and S5, for all wake-capable devices and buses.

Mobile PC Note
Mobile PCs are required to support S1 or S3, in addition to S4 and S5. Mobile PCs are not required to wake from S3 or S4.

Server Note - **Clarification**
At a minimum, SOHO servers must support S1. Other server systems are not required to support S1 or S3. Server systems may choose not to support S4. All server systems must be ACPI-compliant and should support S5, which enables operating system support of S4. Servers with more than eight physical processors are not required to be ACPI-compliant.

A1.1.2.2.3 If software fan control is implemented, thermal design and fan control comply with ACPI 1.0b.

A thermal model and fan control must be implemented as defined in Section 12 of the ACPI 1.0b specification as a means of running the PC quietly while it is working and of turning the fan off while it is sleeping. Notice that hardware override is permitted only to handle thermal conditions when the operating system is not running, the operating system has put the system in a sleep state, or safe operating parameters have been exceeded.

Appendix 3: Power management with Windows Server 2003 (fact sheet)

(next two pages)

Power Management with Windows® Server 2003

Leaflet for System Administrators

Electricity Program

The utilisation of Power Management in servers offers you three major advantages:
① Less heat will be produced in the server area; it may be possible to do without air conditioning.
② The server is immune to power losses at night and during the weekend.
③ You save 50% of your electricity costs.

Energy consumption of a server

Depending on its loading, a server of the entry level class, with one processor and three fixed disks, requires 120 to 160 W of electrical power (without a screen) (Fig. 1).

A server with ACPI[1] functionality automatically goes into a state with reduced power consumption after a longer period of inactivity (Fig. 1):

Standby: Parts of the server run at reduced levels or are switched off (e.g. reduced clock frequency, fixed disks switched off), thereby resulting in lower power consumption.

Hibernate: The data that is currently being used is written to the fixed disk, and all components that are not needed to reawaken the system are switched off. The server is then in the operating state with the lowest possible power consumption. Switching back to the normal mode takes longer than for the *Standby* mode.

Advantages of active power management

The *Standby* and *Hibernate modes* are managed by the power management of the *Windows®* operating system.

In principle, power management should be activated for all servers that do not provide a permanent service and that, for example, are inactive for some hours during the night. This would include, as examples, locally used *File* and *Database servers* that are only used during certain times.

Thanks to switching to the *Hibernate mode?*

- It may be possible to do without air conditioning
- The server will be immune to power losses at night and during the weekend
- Up to 50% of the current consumption will be saved.

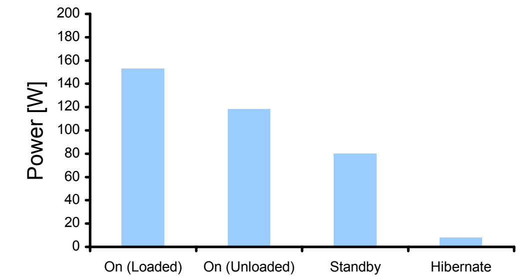

Fig. 1 Power requirement of a server

Valid for new operating systems

The procedure described in the following has been tested with *Windows® Server 2003, Standard Edition*.

Power management is not available for *RAS*[2] *servers* and only with limited performance for *DHCP*[3] *servers* (the server must be in the normal operating mode when logging in a client). Use as a *Domain-Controller* is possible.

Procedure for acquisition, installation and operation

Proceed according to the following steps. The recommended set-values can be taken directly from the pictures (Figs. 2, 3, 4).

1. Purchase of ACPI-compatible hardware

Have your supplier confirm that the server in your final configuration will support at least one of the two *ACPI*-modes S3 (*Suspend to RAM*) or S4 (*Suspend to Disk*). Request the electrical power figures in operation and in the *Standby* and/or *Hibernate modes*. Compare the data from different suppliers.

Ensure that the network card makes it possible to activate the server over the network.

2. Installation

Activate all the functions of the *Power Management* in the *BIOS Set-up*. Ensure that the *ACPI*

[1] ACPI: Advanced Configuration and Power Interface

[2] RAS: Remote Access Service
[3] DHCP: Dynamic Host Configuration Protocol

functionality will be integrated during the installation of the operating system.

Install the latest driver versions of all hardware components. In doing this, you increase the chance that the power-saving modes will be supported.

Activate the following two options in the device properties of the network card: *Computer can switch off devices in order to save energy* and *Device can activate the computer from the Standby-Mode* (Fig. 2)

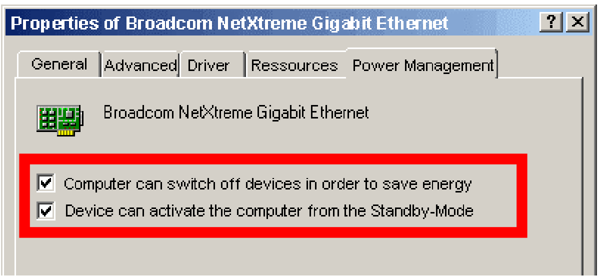

Fig. 2 Network card settings

3. Setting up the Power Scheme

With the right mouse button, click on the free area of the desktop and select the item *Properties*. Open the *Power Schemes* tab and select the *Minmal Power Management* scheme. Now enter the times according to Figure 3. With these settings, the server should only go into the *Hibernate mode* after longer periods of inactivity, i.e., in the evening.

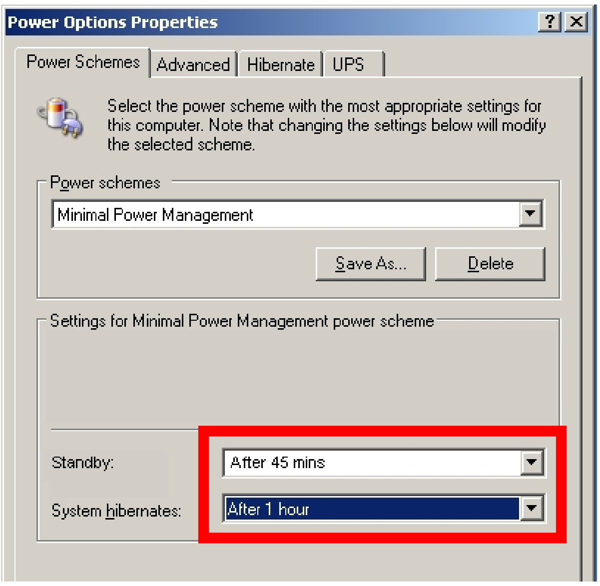

Fig. 3 Power Scheme settings

4. Setting up the Hibernate Mode

For the *Hibernate Mode*, space must be reserved on the fixed disk. To do this, activate the *Enable Hibernation* field in the *Hibernate* tab (Fig. 4). The required memory space depends on the size of the RAM memory.

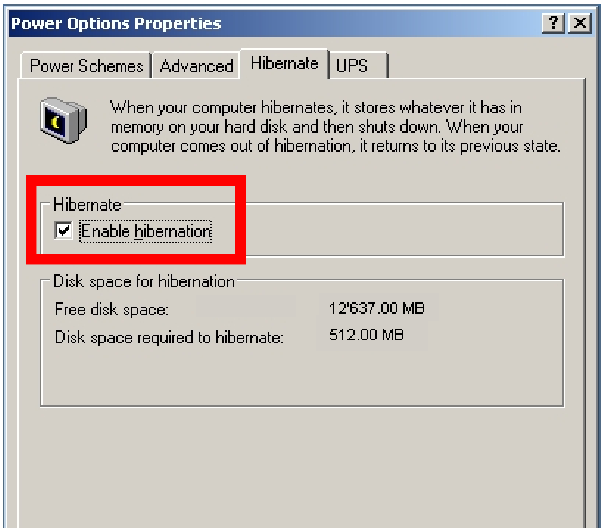

Fig. 4 Hibernate Mode settings

The server is not completely separated from the mains supply in the *Hibernate mode*. A complete switch-off via an external device (for example a UPS system or an intelligent multiple socket outlet, www.emt.ch) offers additional advantages:

- Increase of reliability through the daily re-booting of the system
- Prevention of access by third persons
- Protection against mains failures

For detailed information see the leaflet “Switching of IT Servers” mentioned below.

5. Operation

The server will only switch into the *Hibernate Mode* if no client is still using files on the server. You should therefore close all programs on the connected client computers in the evening, and then switch the computers off.

To find out more ...

- The corresponding Help themes in Windows: *Overview of Energy Options, Overview of ACPI*
- Findings relating to the use of ACPI in Windows servers, 2004
  Download: www.electricity-research.ch
- Leaflet *Energy Efficiency in the Invitations to Tender for IT units*, 2004
  Download: www.electricity-research.ch
- Leaflet *Switching of IT Servers*, 2004
  Download: www.electricity-research.ch

Federal Office for Energy, April 2004
Download: www.electricity-research.ch

Prepared by Alois Huser (Encontrol GmbH, Niederrohrdorf)
Order: ENET, No. 240 051, www.energieforschung.ch

energieschweiz