http://www.lesswatts.org

Open Source Technology Center, Intel Corporation

Introduction

Saving power is a hot topic all over the Internet. For some it's longer battery life, for others it's the heat problem in their data center. For all of us it's about reducing the impact of computing on our planet.

In the data center, power consumption has been rising rapidly for the past several years. It is a growing problem that is reaching a boiling point. Based on predictions made by the US Environmental Protection Agency[1], without improvements, usage will continue to grow to the point where, in 2011, 10 additional power plants will be required to handle the requirements. Most of this increase in power consumption will come from volume servers, and a significant percentage will be running the Linux* operating system and other open source software.

While the computer hardware industry has been making strides in providing capabilities to reduce power consumption, there has not been sufficient focus on enabling these advances in software, especially in the world of open source software.

LessWatts.org is an open source project and a community of end users, open source developers, and operating system vendors (OSVs), focused on delivering the components and tools needed to reduce the power used by systems running Linux.

This paper describes the technologies and tools embodied in the LessWatts.org open source project and gives insight into the key areas that must be addressed in order to slow the seemingly unending quest for more electricity.

The following topics will be discussed:

- What is the current situation of data center power consumption?
- What recent advancements, in microprocessor and platform architecture, can software developers exploit to reduce power consumption?
- What types of improvements can one expect by using these technologies?
- What can the community do to drive acceptance of these capabilities into the mainstream?

[1] "Report to Congress on Server and Data Center Efficiency, Public Law 109-431", U.S. Environmental Protection Agency ENERGY STAR Program, 8/2/2007

Power Consumption Rises to Unprecedented Levels

According to a recent report presented to the United States Congress by the US Environmental Protection Agency, data centers in the US consumed 61 billion kilowatt hours (kWh) of electricity, more than doubling from 28.2 billion kWh in 2000. This represents approximately 1.5% of the total power used in the US and is equivalent to the energy used by 5.8 million average households. Much of this increase is due to the rapid growth of volume servers (servers costing less that $25,000) that have increased from an installed base of approximately 4.9 million units in 2000 to a projected 15 million units in 2010, a compound annual growth rate (CAGR) of 12% per year.

In parallel to the tremendous growth experienced by the server market, the Linux operating system has also increased its footprint within the market. In fact, Linux usage continues to grow in both servers and mobile devices. According to IDC, the market segment share of Linux in the US volume server market will increase to 27%, resulting in shipments of 3.7 million new Linux-based volume servers between 2007 and 2010.[2] On the mobility front, Linux is projected to have 14% market share in 2007, while growing with a CAGR of 13% from 2006-2011

Amidst all this growth, there is an increasing awareness among end users of server power consumption and its associated costs. Ann Livermore, EVP of Technology Solutions at HP, recently said in a LinuxWorld keynote presentation, "The energy efficiency of servers has become the issue 'Every customer wants to talk about'". In a recent paper[3], Google noted that "...the cost of powering server systems has been steadily rising with higher performing systems, while the cost of hardware has remained relatively stable... if these trends were to continue, the cost of the energy consumed by a server during its lifetime could surpass the cost of the equipment itself."

At an average power consumption of 225 watts each, new Linux-based servers will use over 7.2 billion kilowatt hours of electricity between 2007 and 2010, costing $475 million dollars and generating 4.8 million tons of carbon emissions.[4]

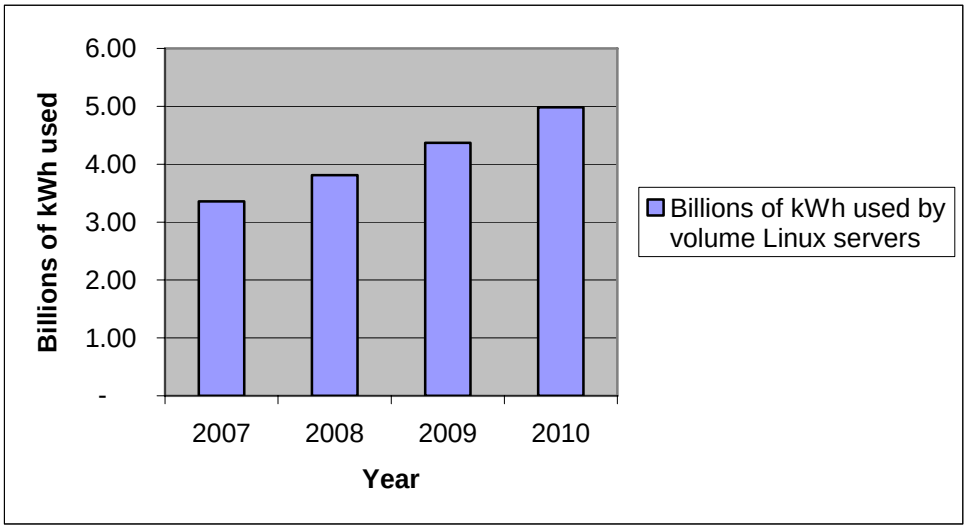

Figure 1- Projected Electricity Consumption of Linux-based Volume Servers, 2007-2010

[2] IDC US and Worldwide Server Installed Base 2006-2010 Forecast, 2/6/2007
[3] "Power Provisioning for a Warehouse-size Computer", Google Inc, 6/2/2007, http://labs.google.com/papers/power_provisioning.pdf
[4] IDC #206585

Improving Power Consumption through Hardware and Software

In years past, the race in the world of microprocessors and computer systems was focused on improving performance with little attention given to power consumption. But in recent years, vendors have realized that power consumption is a critical factor in end-users' purchasing decisions, often more important than raw speed. This is true not only of mobile platforms, but also in desktop and server platforms.

In response to the growing customer requirement for improved server power efficiency, Intel has introduced a number of new power management technologies in its server platforms. The new Intel® Server System S7000FC4UR[5] (formerly known as "Caneland"), which uses the Quad-Core Intel® Xeon® processor 7300 series, brings the first set of new technologies, and the scope will expand extensively with future microprocessors, such as Penryn[6] and Nehalem[7], and future platforms that will use these microprocessors. These new technologies bring key aspects of the power management features from Intel® Centrino® processor technology to Intel server platforms.

The technology transfer from mobile CPUs to Intel's next latest server platform is a natural progression for Intel. Power consumption in a constrained laptop environment has always been a key consideration for mobile users who demand maximum battery life and excellent performance, all in a continually lighter and thinner notebook platform. Historically, Intel addressed these customer needs through its Intel® Centrino® processor technology, which includes a number of power saving features enabled largely by internally developed software. By extending the reach of the mobile CPU feature set and software to the server market, Intel is now positioned to pass along similar power and energy efficiencies.

While a number of new technologies have been developed to reduce power consumption at the hardware level, software that exploits these new features often lags significantly behind the introduction of the hardware. One of the goals of the LessWatts project is to accelerate the pace at which these new technologies make their way into commercial products.

As an example, let's look at the impact that can be achieved by the integration of a single power optimization called "tickless idle" into the Linux kernel.

[5] http://www.intel.com/design/servers/platforms/S7000FC4UR/index.htm
[6] http://www.intel.com/technology/architecture-silicon/45nm-core2/index.htm?iid=technology_45nmcoreproducts+tabs_overview
[7] http://www.intel.com/technology/architecture-silicon/next-gen/description.htm?iid=technology_next-gen+body_description

Saving Power Through "Tickless Idle"

One of the keys to achieving significant power savings is to put the processor into a lower power state when it is idle and keep it there as long as possible.

The following table shows the various power states of an Intel® Core® 2 Duo processor. The first column in the table is the processor power state, or C-state. C-states are a term used in the Advanced Configuration and Power Interface (ACPI) specification[8] to describe the processor power state. The second column shows the maximum power used in each C-state.

C-state	Maximum CPU power consumption
C0	35 watts
C1	13.5 watts
C2	12.9 watts
C3	7.7 watts
C4	1.2 watts

Table 1 – Intel® Core™ 2 Duo (Standard Voltage) Maximum Power Consumption at Various C States[9]

As can be seen from the table, significant amounts of power can be saved by placing the processor into one of the lower power states. However, deciding which power state to use is complicated by the fact that using these lower power states comes at a price in terms of latency and energy consumption. The deeper the C-state, the longer it takes to leave it and the more energy is consumed in the process.

Once the operating system has put the processor into a lower power state, it would be best to stay there as long possible to maximize energy savings. Unfortunately, earlier versions of the Linux kernel made this impossible because it relied on a regularly occurring interrupt called a "timer tick".

The timer tick is used by the Linux kernel to notify it that it is time to perform a number of housekeeping tasks. In earlier version of the kernel, the frequency of this tick was either 100 Hz (1 tick every 10 milliseconds), 250 Hz (1 tick every 4 milliseconds) or 1 KHz (1 tick every millisecond). Based on this, one can see that the maximum amount of time the CPU could be idle was 1, 4 or 10 milliseconds before the timer interrupt would need to be processed. This significantly limited the useful of the deeper C-states on the processor.

With the Linux kernel 2.6.21 for i386 architecture, a new feature called "tickless idle" was introduced. In essence, tickless idle removes the periodic timer interrupt and allows the CPU to remain idle until the next timer event is scheduled to occur.

In theory, tickless idle allows the kernel to set the processor into a lower power state when it is idle and keep it there for a much longer time than previously allowed. However, in practice, on a typical running system, it is not uncommon to see hundreds of timer events per second which keep the CPU from entering a lower power state because much of the code in the Linux kernel (such as device drivers) and user space code was not written with power awareness as a primary concern.

[8] ACPI specification Version 3.0b, 10/6/2006, ACPI Working Group, http://www.acpi.info/spec.htm
[9] "Intel® Core™ 2 Duo Processor for Intel® Centrino® Duo Processor Technology Datasheet", April 2007, http://www.intel.com/design/mobile/datashts/314078.htm

To provide developers with tools they can use to help them optimize their code for power efficiency, Intel developed the Linux PowerTOP tool, one of the projects that is part of the LessWatts.org project. PowerTOP allows developers and end users to observe which processes on their system are taking the processor out of its lower power sleep states and even provides recommendations on techniques that can be used to reduce power consumption.

To learn more about the technical details of tickless idle in the Linux kernel, see the article "Absolute Power" by Arjan van de Ven at http://www.linux-mag.com/id/4037

Now, let's look at the impact of integrating tickless idle onto a current generation Intel server running Linux.

Impact of LessWatts.org Technology

The table below shows the impact of applying only one of the LessWatts.org technologies, tickless idle, on the power consumption of a server based on the Intel Xeon processor 5300 series (formerly known as "Clovertown") processor. The Intel Xeon processor 5300 series combines 4 processor cores onto a single multi-chip processor module and is optimized for use in single or dual processor configurations.

	Savings at 100% idle	Savings at 85% idle
Server with Intel Xeon processor 5300 series (G stepping)	4.5%	3.65%

Table 2 Server Power Savings Resulting From Applying Tickless Idle Patch[10]

As can be seen, switching tickless idle on reduces system power consumption by 4.5% when the system is idle and by 3.65% if the system is running at 15% load (which is still the average for systems in today's data center). Higher load will lead to smaller savings. When applied over large numbers of servers in a data center, these savings can be significant.

It is important to note that these results reflect the application of only one technology on a current generation platform. It is expected that the impact will grow significantly with additional technologies and next generation server platforms.

[10] Based on Intel internal measurements, September 2007. Actual performance may vary.

Significant power savings can be seen when applying LessWatts technologies to notebook computers as well. The figure below shows the impact to power consumption of a notebook computer based on the Intel® Core™ 2 Duo processor when several techniques documented on the LessWatts.org web site are applied.

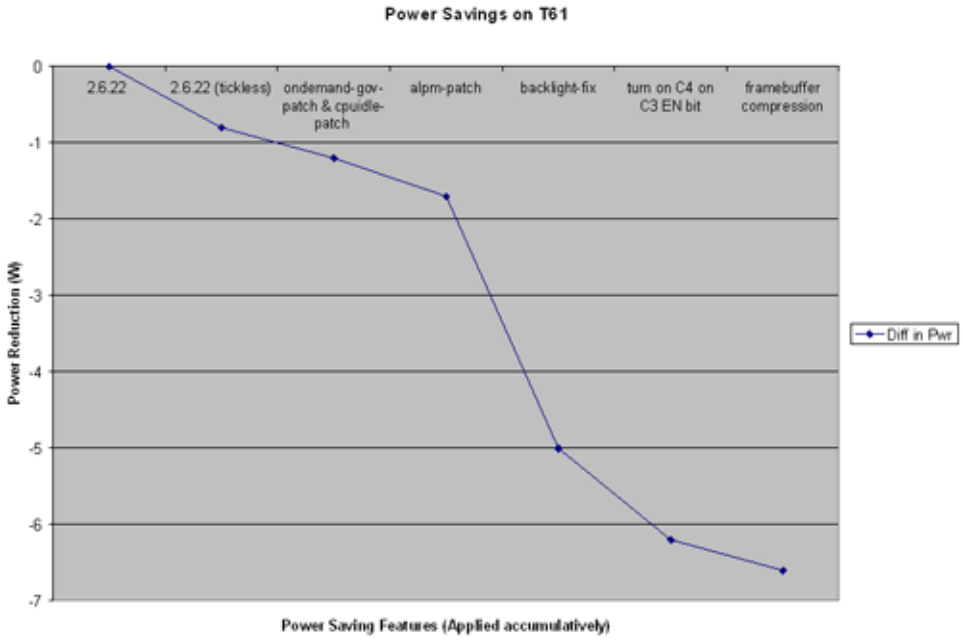

Figure 2 – Notebook Power Consumption with Various Patches Applied[11]

For a detailed description of the various patches that enable these power savings, see http://www.lesswatts.org/projects/power-performance

LessWatts.org Projects

LessWatts.org hosts several open source projects focused on reducing the power consumption of computers running Linux. These projects include:

- **PowerTOP.** PowerTOP is a Linux tool that helps you find programs that are consuming extra power when your computer is idle. You can see the power savings immediately within the tool. A lot of the issues have already been found. The "Tips & Tricks" page offers fixes for known issues.
- **Tickless Idle.** "Tickless idle" eliminates the periodic timer tick in the Linux kernel when the CPU is idle. However, the benefits of tickless idle are lost if the CPU is frequently awakened by unnecessary timer event. This project is a collection of various kernel enhancements and patches based on the tickless idle feature.
- **Applications Power Management.** PowerTOP helps you identify misbehaving applications on a kernel enabled with tickless idle. This project is a collection of the known application issues and fixes for those problems. Collectively, they enable the kernel to remain in deep idle state for longer periods, saving power for the end user.
- **Processor Power Management .** Intel processors support many power management features. This project will be a one stop shop to find out everything you wanted to know about Intel Processor Power Management-related features, solutions, and enhancements that are being integrated into the Linux kernel.

[11] Based on Intel internal measurements, September 2007. Actual performance may vary.

- **Power and Performance Measurement.** The Linux kernel community developers implemented tickless idle and other features to take advantage of the potential of the hardware power savings infrastructure. This project presents the results of implementing specific power saving features on mobile, desktop, and server platforms.
- **Linux Battery Life Toolkit (BLTK).** The Linux BLTK consists of a test framework and six example workloads. The toolkit framework launches the workload, collects statistics during the run, and summarizes test results. The framework could launch any workload but currently has six example workloads: Idle, Reader, Office, DVD Player, Software Developer, and 3D Gamer.
- **Power Policy Manager (PPM).** The PPM is a system-wide power policy framework that allows multiple power policies to be active at once. You can use flexible plug-ins to extend subsystem power management and you can create custom power policies. This architecture makes PPM a flexible and comprehensive power management framework.
- **Power Quality of Service (QoS).** QoS power management enables aggressive power management by subsystems while honoring the QoS needs of applications and other subsystems. This project will provide a central place for developing QoS power management infrastructure and applications to do aggressive, but not too aggressive, power management.
- **Advanced Configuration & Power Interface Component Architecture (ACPICA).** ACPICA provides an OS-independent reference implementation of the ACPI specification. It can be easily adapted to any host OS.
- **Linux Advanced Configuration & Power Interface (ACPI).** The Linux ACPI project is focused on making Linux run well on all ACPI-enabled platforms. Linux/ACPI is based on Intel's ACPICA reference implementation.
- **Device and Bus Power Management.** Devices and buses are responsible for a significant amount of overall system power consumption so it is important to maximize their power savings while the devices are running and while they are idle. This project seeks to lower overall system power consumption by making sure that devices and buses are using power appropriately. Initial focus in this area will be on Serial ATA (SATA), USB, and PCI Express (PCIe).
- **Display and Graphics Power Management.** Intel's graphics devices, such as the Intel® G965 family of chipsets, support many advanced graphics features, and—due to their flexible design—they also support many power saving features. This project aims to minimize graphics power consumption in general and exploit those features where possible—without sacrificing performance.

For detailed project descriptions or to get involved, visit: www.lesswatts.org/projects

Intel is working with its Linux operating system vendor co-travelers to ensure that these technologies get integrated into commercial Linux distributions as soon as possible. As an example, Intel would like to see OSVs include a power aware installation option that sets several tunings for power and selects the power optimized kernel as the default.

Harnessing the Power of the Community

When using the word "community", as it relates to open source software, one often thinks of the thousands of software developers who design, code, debug and test the software that powers the open source movement. However, in terms of power consumption, this view is limiting and excludes a huge number of individuals and organizations who can influence positive change. When it comes to reducing power consumption, the community includes:

- Information technology executives who specify and approve the purchase of computing hardware and software.
- IT staff who plan, deploy, and maintain data centers
- End users of computing equipment
- Computer system manufacturers and resellers

- Government agencies such as the US ENERGY STAR program
- Press and analysts

Active participation by all members of the community is key to driving the market to adopt solutions that are more energy efficient.

LessWatts.org is one of several organizations focused on reducing energy consumption. Some other organizations, to which Intel belongs, include:

The Climate Savers Computing Initiative (http://climatesaverscomputing.org, http://www.intel.com/technology/eco-technology/climatesavers.htm) – Initially developed by Intel and Google, the Climate Savers Computing Initiative is working to deliver significantly increased PC and server energy efficiency by uniting industry, consumers, government, and conservation organizations.

The Green Grid – (http://thegreengrid.org) - The Green Grid is a consortium of information technology companies and professionals seeking to lower the overall consumption of power in data centers around the globe. The organization is chartered to develop meaningful, platform-neutral standards, measurement methods, processes, and new technologies to improve energy efficient performance of global data centers.

The LessWatts.org project complements these other organizations by focusing on improving the power efficiency of Linux and other open source software. The LessWatts.org project gives the broader community insight into the advancements taking place in reducing power to drive the adoption of these technologies into products they can purchase and deploy.

Intel is also working closely with OEMs to drive the adoption of technologies that will comply with the new US ENERGY STAR program guidelines through a series of implementation guides and other collateral. More information can be found at http://www.intel.com/go/energystar/

Conclusion

Intel believes firmly that technology can be a catalyst for change and can improve the lives of those who use it. The LessWatts.org project is one of the ways that Intel is working with the community to drive improvements in power consumption that will lead to a cleaner environment and allow companies to spend less money powering their IT infrastructure. We look forward to you helping us drive this vision forward by contributing at http://ww.lesswatts.org