ACADEMIA
Accelerating the world s research.

Energy Management for Commercial Servers

W s F lt r

IEEE Computer

Cite this paper

Downloaded from Academia edu

Get the citation in MLA, APA, or Chicago styles

Related papers

Download a PDF Pack of the best related papers

Introducing the Adaptive Energy Management Features of the Power7 Chip
Alan Drake, Malcolm Ware

Conserving disk energy in network servers
Eduardo Pinheiro

Data Center Energy Consumption Modeling: A Survey
JORGE ANDREE VELASQUEZ RAMOS

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/2956025

Energy Management for Commercial Servers

Article *in* Computer · January 2004

DO : 0. 09/MC.2003. 250880 Sou ce: EEE Xp o e

C A ONS
290

READS
56

6 authors, including:

Charles Lefurgy
IBM
69 PUBL CA ONS **2,370** C A ONS
SEE PROF LE

Michael Kistler
IBM
26 PUBL CA ONS **1,701** C A ONS
SEE PROF LE

Tom W. Keller
IBM
29 PUBL CA ONS **1,357** C A ONS
SEE PROF LE

All content following this page was uploaded by Charles Lefurgy on 20 March 2013.

he user has requested enhancement of the down oaded f e A n text references under ned n b ue are added to the or g na document and are nked to pub cat ons on ResearchGate, ett ng you access and read them mmed ate y

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

Energy Management for Commercial Servers

As power increasingly shapes commercial systems design, commercial servers can conserve energy by leveraging their unique architecture and workload characteristics.

Charles Lefurgy
Karthick Rajamani
Freeman Rawson
Wes Felter
Michael Kistler
Tom W. Keller
IBM Austin Research Lab

In the past, energy-aware computing was primarily associated with mobile and embedded computing platforms. *Servers*—high-end, multiprocessor systems running commercial workloads—typically included extensive cooling systems and resided in custom-built rooms for high-power delivery. In recent years, however, as transistor density and demand for computing resources have rapidly increased, even high-end systems face energy-use constraints. Moreover, conventional computers are currently air cooled, and systems are approaching the limits of what manufacturers can build without introducing additional techniques such as liquid cooling. Clearly, good energy management is becoming important for all servers.

Power management challenges for commercial servers differ from those for mobile systems. Techniques for saving power and energy at the circuit and microarchitecture levels are well known,[1] and other low-power options are specialized to a server's particular structure and the nature of its workload. Although there has been some progress, a gap still exists between the known solutions and the energy-management needs of servers.

In light of the trend toward isolating disk resources in separate cabinets and accessing them through some form of storage networking, the main focus of energy management for commercial servers is conserving power in the memory and microprocessor subsystems. Because their workloads are typically structured as multiple-application programs, system-wide approaches are more applicable to multiprocessor environments in commercial servers than techniques that are primarily applicable to single-application environments, such as those based on compiler optimizations.

COMMERCIAL SERVERS

Commercial servers comprise one or more high-performance processors and their associated caches; large amounts of dynamic random-access memory (DRAM) with multiple memory controllers; and high-speed interface chips for high-memory bandwidth, I/O controllers, and high-speed network interfaces.

Servers with multiple processors typically are designed as symmetric multiprocessors (SMPs), which means that the processors share the main memory and any processor can access any memory location. This organization has several advantages:

- Multiprocessor systems can scale to much larger workloads than single-processor systems.
- Shared memory simplifies workload balancing across servers.
- The machine naturally supports the shared-memory programming paradigm that most developers prefer.
- Because it has a large capacity and high-bandwidth memory, a multiprocessor system can efficiently execute memory-intensive workloads.

In commercial servers, memory is hierarchical. These servers usually have two or three levels of

0018-9162/03/$17.00 © 2003 IEEE Published by the IEEE Computer Society

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

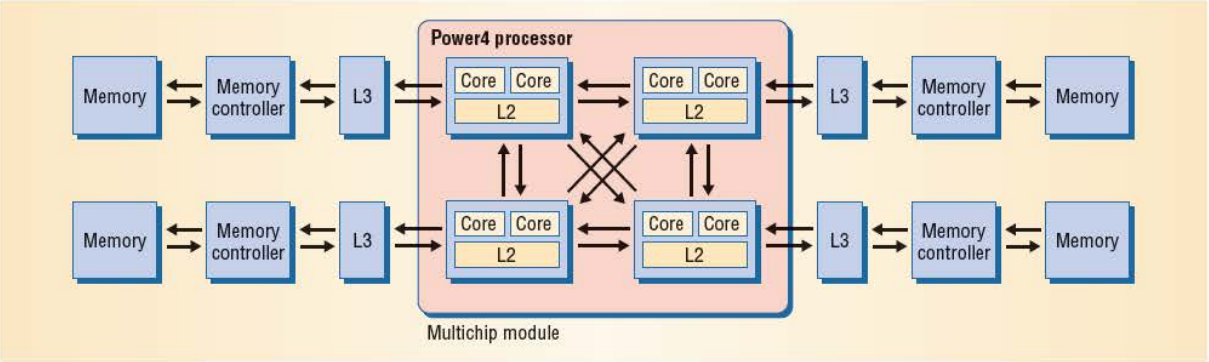

Figure 1. A single multichip module in a Power4 system. Each processor has two processor cores, each executing a single program context. Each core contains L1 caches, shares an L2 cache, and connects to an off-chip L3 cache, a memory controller, and main memory.

Table 1. Power consumption breakdown for an IBM p670.

IBM p670 server	Processors	Memory	I/O and other	Processor and memory fans	I/O component fans	Total watts
Small configuration (watts)	384	318	90	676	144	1,614
Large configuration (watts)	840	1,223	90	676	144	2,972

cache between the processor and the main memory. Typical high-end commercial servers include IBM's p690, HP's 9000 Superdome, and Sun Microsystems' Sun Fire 15K.

Figure 1 shows a high-level organization of processors and memory in a single multichip module (MCM) in an IBM Power4 system. Each Power4 processor contains two processor cores; each core executes a single program context.

The processor's two cores contain L1 caches (not shown) and share an L2 cache, which is the coherence point for the memory hierarchy. Each processor connects to an L3 cache (off-chip), a memory controller, and main memory. In some configurations, processors share the L3 caches. The four processors reside on an MCM and communicate through dedicated point-to-point links. Larger systems such as the IBM p690 consist of multiple connected MCMs.

Table 1 shows the power consumption of two configurations of an IBM p670 server, which is a midrange version of the p690. The top row gives the power breakdown for a small four-way server (single MCM with four single-core chips) with a 128-Mbyte L3 cache and a 16-Gbyte memory. The bottom row gives the breakdown for a larger 16-way server (dual MCM with four dual-core chips) with a 256-Mbyte L3 cache and a 128-Gbyte memory.

The power consumption breakdowns include

- the processors, including the MCMs with processor cores and L1 and L2 caches, cache controllers, and directories;
- the memory, consisting of the off-chip L3 caches, DRAM, memory controllers, and high-bandwidth interface chips between the controllers and DRAM;
- I/O and other nonfan components;
- fans for cooling processors and memory; and
- fans for cooling the I/O components.

We measured the power consumption at idle. A high-end commercial server typically focuses primarily on performance, and the designs incorporate few system-level power-management techniques. Consequently, idle and active power consumption are similar.

We estimated fan power consumption from product specifications. For the other components of the small configuration, we measured DC power. We estimated the power in the larger configuration by scaling the measurements of the smaller configuration based on relative increases in the component quantities. Separate measurements were made to obtain dual-core processor power consumption.

In the small configuration, processor power is greater than memory power: Processor power accounts for 24 percent of system power, memory power for 19 percent. In the larger configuration, the processors use 28 percent of the power, and memory uses 41 percent. This suggests the need to supplement the conventional, processor-centric approach to energy management with techniques for managing memory energy.

The high power consumption of the computing components generates large amounts of heat, requiring significant cooling capabilities. The fans driving the cooling system consume additional power. Fan power consumption, which is relatively fixed for the system cabinet, dominates the small configuration at 51 percent, and it is a big component of the large configuration at 28 percent. Reducing the power of computing components

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

would allow a commensurate reduction in cooling capacity, therefore reducing fan power consumption.

We did not separate disk power because the measured system chiefly used remote storage and because the number of disks in any configuration varies dramatically. Current high-performance SCSI disks typically consume 11 to 18 watts each when active.

Fortunately, this machine organization suggests several natural options for power management. For example, using multiple, discrete processors allows for mechanisms to turn a subset of the processors off and on as needed. Similarly, multiple cache banks, memory controllers, and DRAM modules provide natural demarcations of power-manageable entities in the memory subsystem. In addition, the processing capabilities of memory controllers, although limited, can accommodate new power-management mechanisms.

Energy-management goals

Energy management primarily aims to limit maximum power consumption and improve energy efficiency. Although generally consistent with each other, the two goals are not identical. Some energy-management techniques address both goals, but most implementations focus on only one or the other.

Addressing the power consumption problem is critical to maintaining reliability and reducing cooling requirements. Traditionally, server designs countered increased power consumption by improving the cooling and packaging technology. More recently, designers have used circuit and microarchitectural approaches to reduce thermal stress.

Improving energy efficiency requires either increasing the number of operations per unit of energy consumed or decreasing the amount of energy consumed per operation. Increased energy efficiency reduces the operational costs for the system's power and cooling needs. Energy efficiency is particularly important in large installations such as data centers, where power and cooling costs can be sizable.

A recent energy management challenge is leakage current in semiconductor circuits, which causes transistors designed for high frequencies to consume power even when they don't switch. Although we discuss some technologies that turn off idle components, and reduce leakage, the primary approaches to tackling this problem center on improvements to circuit technology and microarchitecture design.

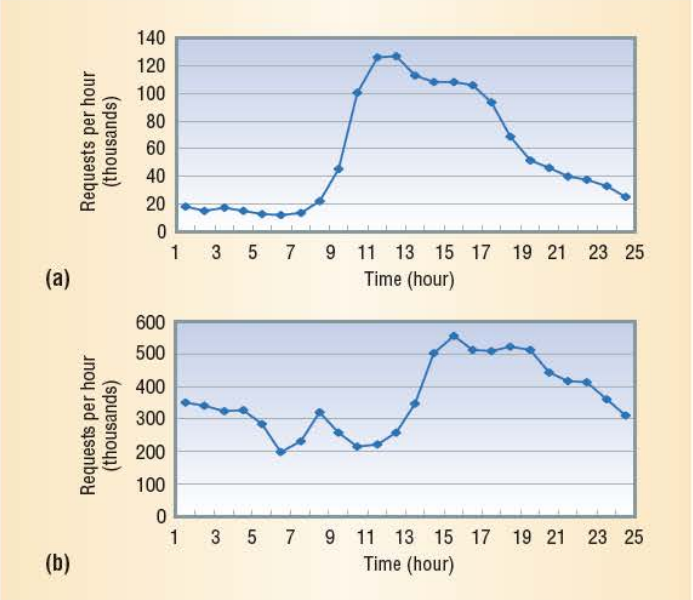

Figure 2. Load variation by hour at (a) a financial Web site and (b) the 1998 Olympics Web site in a one-day period. The number of requests received varies widely depending on time of day and other factors.

Server workloads

Commercial server workloads include transaction processing—for Web servers or databases, for example—and batch processing—for noninteractive, long-running programs. Transaction and batch processing offer somewhat different power management opportunities.

As Figure 2 illustrates, transaction-oriented servers do not always run at peak capacity because their workloads often vary significantly depending on the time of day, day of the week, or other external factors. Such servers have significant buffer capacity to maintain performance goals in the event of unexpected workload increases. Thus, much of the server capacity remains unutilized during normal operation.

Transaction servers that run at peak throughput can impact the latency of individual requests and, consequently, fail to meet response time goals. Batch servers, on the other hand, often have less stringent latency requirements, and they might run at peak throughput in bursts. However, their more relaxed latency requirements mean that sometimes running a large job overnight is sufficient. As a result, both transaction and batch servers have idleness—or *slack*—that designers can exploit to reduce the energy used.

Server workloads can comprise multiple applications with varying computational and performance requirements. The server systems' organization often matches this variety. For example, a typical e-commerce Web site consists of a first tier of simple page servers organized in a cluster, a second tier of higher performance servers running Web applications, and a third tier of high-performance database servers.

Although such heterogeneous configurations pri-

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.