Low Power Server, Server Systems and Design Methodology Thereof

Inventor: Andrew Sharp

Background

Technical Field

This invention relates to the design methods and the design of various components in the creation of ultra-efficient computer servers, storage systems, and other Information Technology devices such as network or storage switches.

Background Art

Currently, the design of high-volume standard computational servers tends to follow one path – high performance. This invention concerns the creation of a new path: energy efficiency.

Designing a server for energy efficiency generally relates to the process and methods of reducing the average power consumed by a server while still providing requisite performance and functionality as similar servers which are not energy-efficient. Such power efficienct servers are typically designed by selecting non-standard off-the-shelf components from vendors offering lower-power-components. Additionally, major components such as printed circuit boards, fans and power supplies can be specifically designed with the intention to maximize efficiencies in: electrical power distribution and conduction; cooling; and emissions in auditory and electromagnetic categories, without sacrificing needed levels of system performance.

[Can you fill in some more background art about what is currently done for low power today that will contrast the benefits of your invention? Note to Andy: Is there currently software that controls power consumption in servers? ]

There really isn't anything done today by conventional manufacturers, unless you count minor attempts such as “shared fans”, which, while slightly (5%?) better than conventional designs, is actually the wrong way to go. Which means the following text doesn't make much sense, but maybe I'll go back and add some stuff about that.

The above methodology, while providing small reductions in power consumption, does not provide sufficient gains in efficiency to really be considered as an efficient design or as a valuable proposition in the market place. Rather the servers produced under this standard methodology are generally specified by the size of the power supply provided to power the server and do not yield the significant power savings possible through the use of this art.

In order to design for maximum energy efficiency, all components of the server must be considered and measured for their contribution to power consumption and power dissipation (in the case of fans and air flow controllers) and the server design must be must have a repeatable, deterministic methodology for measuring maximum power consumed. For this embodiment, both Idle Power Consumption IPC) and Total Design Power (TDP), which represents the maximum power a server will consume under full computational load, will be considered as the measure of energy efficiency. IPC represents the power consumed by the device when it is powered up and ready, but has negligible computational load (idle). TDP is important to server customers because it allows a well-engineered design of the power distribution infrastructure and the HVAC provisioning in order to have a smooth running data center and high reliability and availability of the services the server is intended to provide. IPC is important because most servers spend 80-90% of their existence in the idle state. Accurate measurements of the TDP of a server are currently difficult to find specified on standard servers. All servers will require varying amounts of power depending on the instantaneous load on the computational components as well as on the power supply and cooling system, and their power consumption follows a curve such as shown in Figure 1/coventional. As can be seen as load increases, so does the power consumed by the server. Once the components are fully loaded, all electrical components in the server will be theoretically operating at their maximum and TDP can then be measured. Figure 2 shows the TDP curves for servers designed under a standard methodology exhibit TDP curves similar to curve TDP1 whereas servers designed under the art of this invention will exhibit TDP curves similar to TDP2 where the maximum power is substantially lower and reached earlier in the load cycle. [Andy- did I understand this correctly?] Nope

Designing a highly efficient server requires power consumption to be minimized while maintaining requisite performance throughput. While obtaining a reasonably accurate IPC figure for a server is relatively straightforward, doing so for TDP is quite difficult. An embodiment of this invention utilizes a testing process, methodology and algorithm to most accurately determine the TDP rating of a computer server. This design methodology requires Special Copyrighted Software Programs (SWTDP’s) to fully exercise all computational components to create as close to 100% loaded conditions as practicable at which point power consumption is measured, thus yielding TDP for a given design. Designs are iterated using the flowchart shown in Fig 3 to select the most energy efficient design components and configurations. Power reductions in the computational components and motherboards have a ripple effect on secondary components such as fans and power supplies. As the server becomes more and more power efficient, less power is dissipated requiring less powerful cooling components and lower total-power-output power supplies which further reduce the TDP of the server. By using SWTDP’s to fully exercise the system components, a minimized TDP is specified as the maximum power the server will consume under virtually any condition. This max TDP can be used as a design parameter of systems consisting of 100’s or 1000’s of servers configured in racks in commercial applications such as businesses and data centers. By utilizing a significantly lower max TDP than standard designed servers, server racks and data centers can be designed with lower energy consumption requirement and lower cooling requirements. Since powering and cooling of systems of servers also consumes its own power, these subsystems also become more energy efficient. Thus, the ability to design a server to a significantly lower, highly efficient power consumption causes a ripple effect and virtuous circle of power reduction and energy efficiency gain.

[Andy How do we weave in the idea of equivalent performance? – “…. while maintaining overall server performance design goals” “…without sacrificing performance” – is this the overall thinking? Where to capture? I think I did by use of the words `requisite' and `required'. We could also say `required application performance'.

Designing the non-computational components for higher efficiency becomes possible as the power of computational components is reduced. The focus of the designs are to improve cooling efficiency by more effectively controlling air flow within the server and having the cooling system dissipate less power. These designs fall into 8 categories: A) efficient motherboard design, B) greater fan efficiency C) utilization of airflow controlling devices to smooth and direct cooling airflow inside a server, D) use of 42mm system fans, E) fans with variable speed that can also be shut off, F) BIOS/boot up/ACPI code, G) Server case design for efficient airflow, H) Air intake and exhaust grill design

A) Computer server motherboard design elements for increasing overall power efficiency

(a) Use of 5 volt fans instead of 12 volt fans

5-Volt fan headers, allowing the use of 5V fans which typically can use much less power than traditional 12 volts fans, in efficient servers that don't need high speed, and/or high pressure, cooling fans.

(b) Onboard components and component connector placement

Memory, cable connectors and other components that are wider in one aspect than another, are mounted with the thinner aspect in line with the direction of cooling air flowing though the computer, thus increasing cooling efficiency. This is shown in low detail in Fig. 4a.

(c) I/O device connector port placement

Conventional designs located these headers on the left side of the board, which is usually in the direct path of the air flow cooling the CPU. This new design locates the rear I/O ports, and any other vertical components, on the opposite side of the motherboard from the CPU or CPUs to prevent them from obstructing the cooling air flow. This substantially improves air flow over the motherboard (and therefore through the computer case) of cooling air propelled by cooling fans. With the rear I/O headers, ports and other large components moved out of the way, the cooling air can flow across the CPU and other hotter parts, and exit the case substantially easier, allowing the use of less powerful fans to cool the components of the computer server, which in turn reduces the amount of energy consumed overall by the computer server. This applies to the I/O ports, but also to other tall components that could hinder air flow. Also shown in Fig. 4 A.

(d) Use of high efficiency components

Use of components, especially I/O controllers, that are designed for the high efficiency requirements of mobile computing. Including, but not limited to: network controllers and the associated PHY; CPUs; fan controllers; disk controllers and associated integrated circuit devices; video controllers (if any); USB controllers; and keyboard/mouse controllers.

[Andy – we think we have the case where you can describe a redesign of motherboards for low power and high efficiency that may not have been capture above. If so, please add.] Motherboards with network controllers designed for mobile (laptop) applications; low power PHYs as used in mobile applications; use of mobile (notebook/laptop) processors with greater levels of frequency scaling to further reduce IDC.

B) Fans with disappearing frame segments

(a) Fans have “disappearing” frame segments on the top and/or bottom of the fan, so that the blades come as close as possible to the top and/or bottom of the enclosure they are located in, maximizing fan blade size to the available space in the enclosure. See figure 5. This means that the external dimensions of the entire fan assembly will not be square, but in fact wider in the horizontal than in the vertical. This allows for a larger fan blade to be used in a fan that is located in an enclosure and therefore doesn't need an encompassing frame to protect the fan from foreign objects or to protect foreign objects from the spinning fan blades.

(b) Fans are often made of plastic: plastic frame and plastic blades. These fans will have soft touch plastic at least on the bottom and top of the frame, so as to damp vibrations coming from the fan, minimizing the amount of vibrations transmitted to the rest of the computer server and its components. Also, this soft touch plastic, a common material often found in car dashboards and other consumer products, will aid in holding the fans in place. For fans made of metal or other material, a thin coating of soft touch plastic, or other vibration damping material such as rubber or silicone, can be applied to the frame material to achieve the same result. See Fig 6.

C) Device(s) to smooth and direct cooling airflow inside a computer

(a) Small devices that can be attached to the computer motherboard to aerodynamically smooth the flow of cooling air around and/or above protuberances and blockages (I/O ports; I/O headers; electrical components such as capacitors, transistors, coils, etc.; cable attachment headers; the motherboard itself) that normally have a rectangular or other non-aerodynamic shape which blocks efficient air flow. Optimizing the air flow around such blockages allows for more efficient cooling of the computer, which allows for a reduction in the amount of fan power required.

(b) Batteries, speakers, electrical coils, capacitors and other components on the motherboard. The shape of air flow blockers inside a computer case are usually rectangular or of a rectangular cross section. This is because these blockages are most commonly the rear I/O ports attached at the back of the motherboard. But other devices or objects can also block the flow of cooling air. The motherboard itself is not aerodynamically optimized, hence one of the devices is designed specifically to direct all air flow coming from the fans above the motherboard where it can do the most good.

(c) In the case of wider blockages such as multiple I/O headers mounted side by side on the motherboard, the devices are intended to be used in groups. The size of the devices would be chosen to match the height and width of the blockage(s).

(d) The device(s) will have various shapes, including:

1. Basic vertical wedge shape, as seen in Fig. 7 (aka 1c) (side view sketch). This shape is designed to transition air flow over the top of the blockage.

1. Basic blended vertical and horizontal wedge, used to transition air flow over and around a single blockage. This shape would more or less come to a point at the leading edge, resembling a cone cut in half from top to bottom and then laid flat, with the pointed end pointed towards the air source, most likely a fan or fans. The tail end of the device would not necessarily be rounded like a simple cone, but be a match for the vertical cross section of the blockage itself, and that shape gradually shrinking towards the leading end of the device.

[drawings and illustrations] – Andy do we want to add more?

1. Devices that are essentially device shape #2 cut into right and left halves. These are to be used in conjunction with device shape #1 to create a complete air flow transition around larger blockages possibly consisting of multiple I/O headers mounted side by side or nearly so.
2. Rounded versions of the above described shapes.
3. The devices are made of a very smooth, hard material, most likely plastic for cost and light weight, facilitating good air flow.
4. The devices may have a thin, soft layer on the motherboard side that is coated with reusable adhesive. The soft layer is malleable and does not rebound, allowing the device to be pressed down over small surface mount components that may be on the motherboard in front of the blockage(s). Not applicable when device shapes are built in to components or headers.
5. Designs or shapes described above, but integrated into the external case of the I/O header or component itself.

D) 42mm system fans for use inside a computer server case

(a) Computer server cases that are 1U tall (1.7 inches or 43mm) suffer from lack of efficient cooling fan solutions, because the internal case volume is relatively small. Fans that are situated inside such cases are referred to as case fans. This is a design for 42mm case fans specifically for optimizing efficiency of cooling.

(b) Current 1U case fans are 40mm square, and anywhere from 5mm to 25mm deep, or deeper. This fan design specifies 42mm fans, thereby leaving no amount of internal space empty.

(c) Fans have “disappearing” frame segments on the top of the fan, so that the blades come as close as possible to the top (or bottom) of the case, maximizing available space. This means, among other things, that the fans can reasonably only be installed in the case one of two ways: that is, with the open frame side facing the top or bottom of the case. It also means the external dimensions of the fans may not be square, but in fact wider in the horizontal than in the vertical.

(d) Plastic frame and plastic blades. Fans are generally made of plastic. These fans will have soft touch plastic at least on the bottom and top of the frame, so as to damp vibrations coming from the fan, minimizing the amount of vibrations transmitted to the rest of the computer server and its components. Also, this soft touch plastic, a common material often found in car dashboards and other consumer products, will aid in holding the fans in place.

E) Variable speed fans which can be completely powered down

(a) Computer system fans that can, besides varying their RPM and thus varying the amount of power they draw and the amount of air they move, can completely stop the fan motor, essentially drawing no power. Either through a communication signal via an interface, or merely by setting the fan supply voltage to zero, the fan will stop. The fan can start again when signaled to do so or when supply voltage is raised to nominal levels. This will significantly reduce the amount of power consumed by a computer when it is in idle or low power mode, and passive convection cooling is sufficient to cool the system or parts of the system.

F) BIOS/Boot up code for ultra-efficient server

(a) Boot up and self-test ROM or NVRAM code (herein called BIOS or POST code) that greatly decreases the amount of power consumed during the boot and self test phase.

Current BIOS code in PC derived computer systems including servers utilizes a continuous loop design meant to refresh the video display. The use of directly connected video displays is a long since deprecated practice for servers. Our boot up code would dispense with the continuous loop video refresh code altogether and utilize an interrupt driven design, thereby drastically reducing power consumption during boot up, system test, and interactive operations.

(b) Fan management code capable of turning one or more of the cooling fans off, or put another way, RPM of zero, when fan operation isn't required to cool the system.

(c) Significantly increase the effectiveness of thermal management code, especially in the particular area of fan control, in the BIOS to help reduce power consumption of the system during operation.

(G) Computer server case and motherboard design elements for more efficient air flow though the case

(a) This description covers new computer server motherboard design elements, and matching case design elements, to facilitate less restricted air flow through the case, thus increasing cooling efficiency of the running computer server.

(b) Fig. 8 shows the motherboard design with the rear I/O ports located to the right side of the rear edge of the motherboard. This allows substantially less restricted airflow through the case of cooling air moved by internal cooling fans. With the rear I/O headers and ports moved out of the way, the cooling air can exit the case substantially easier, allowing the use of less powerful fans to cool the internal components of the computer server, which in turn reduces the amount of energy consumed overall by the computer server.

(H) Air intake and exhaust grill design

(a) Currently, all computer server cases consist of a grill pattern stamped or cut out of sheet metal, usually aluminum. The thickness of the grill hatching is usually 2 — 2.5mm. Sometimes it is as thin as 1mm, but usually in a higher density, thus negating the thinness of the hatching in terms of easing air flow. The grill element of these cases is designed this way because these server cases use the grill metal as structural elements of the case. However, the thickness of the hatching in these grill designs severely retards air flow, badly affecting cooling efficiency.

(b) Our design is to have separate structural elements akin to pillars in a building, that are outside of the critical cooling air flow, and to utilize non-structural, non-load bearing grill material that has very thin hatching, in the 0.1mm or less thickness range, in the areas of the case that in the critical path for intake and exhaust of cooling air, thereby substantially reducing the power needed to move cooling air through the case. This will decrease the amount of air that cooling fans have to push or pull through the system, thus reducing the overall amount of power consumed by the system.

The design of highly efficient servers facilitates the design of future highly-efficient data centers. Server design methodology utilizing the described TDP testing the procedure can be used to create servers and/or other computer related appliances (such as: computer storage machines; communication or storage networking machinery) with accurately specifiable power consumption profiles. These can then be used to design the next generation power-optimized data centers using application centric data center modules. These modules can include server designs as well as highly efficient network routers and other devices which were are designed utilizing a similar TDP methodology. See Figure 4b.

The ultra-efficient design methodology also facilitates designing servers for special application computing needs. Some examples: big data processing; image/video/audio search; speech recognition. Servers maybe optimized for use of general purpose processors, or GPUs, FPGAs, and ASICs. The ability to deploy allows for the Design of Next Generation All Purpose Data Center(NGAPD) utilizing application-centric data center sub-modules comprised of ultra-efficient servers designed for that application category along with a method for analyzing existing/running applications and migrating them to the appropriate submodule of the NGAPD intended to most efficiently run them.

The ultra-efficient design methodology also facilitates a blade server design implementing variable load power supplies, preserving high efficiency levels regardless of load (population level of blade chassis).