Provisional Application for Patent Cover Sheet

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)

Inventor(s)

Inventor 1 Remove

Given Name	Middle Name	Family Name	City	State	Country i
Andrew		Sharp	Mountain View	CA	US

All Inventors Must Be Listed – Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Title of Invention	LOPOCO NEXT GENERATION ULTRA-EFFICIENT DATACENTER
Attorney Docket Number (if applicable)	2018004 / LP02prov

Correspondence Address

Direct all correspondence to (select one):

◉ The address corresponding to Customer Number	○ Firm or Individual Name
Customer Number	73091

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

◉ No.

○ Yes, the name of the U.S. Government agency and the Government contract number are:

Doc Code: **TR.PROV**
Document Description: Provisional Cover Sheet (SB16)

Entity Status

Applicant claims small entity status under 37 CFR 1.27

◉ Yes, applicant qualifies for small entity status under 37 CFR 1.27

○ No

Warning

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

Signature

Please see 37 CFR 1.4(d) for the form of the signature.

Signature	/Marc P. Schuyler/			Date (YYYY-MM-DD)	2018-01-26
First Name	Marc	Last Name	Schuyler	Registration Number (If appropriate)	35675

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or paten. Accordingly, pursuant to the requirements of the Act, please be advised that : (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, t o a n other federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

LOPOCO NEXT GENERATION ULTRA-EFFICIENT DATACENTER

1. Field of Design

1.1. The need to make data centers more energy efficient has been identified for several years. The focus of many innovations and advancements has been largely in the areas of cooling, backup power generation and non IT electrical equipment (i.e., lights, human inhabitant needs, etc.).

1.2. The focus of this design relates to the largely untouched area of efficiency, specifically the efficiency of the data center IT equipment, which includes primarily computer servers, network switching appliances, and computer storage appliances and related equipment. Use of such equipment in conjunction with new design thinking of the datacenter itself leads to the realization of a next generation ultra-efficient datacenter.

Background

1.3. The efficiency of data center IT equipment has been a particularly low priority, primarily because the provisioning of IT servers and equipment falls to individuals who are not aware of application selections that will be running on servers to be provisioned, and, even if application selection was taken into account using advanced information, the initial provisioning cannot take into account the future application mix that will be running on these servers. Therefore, coordination between the two disparate responsibilities of provisioning and application selection has been, and is being, attempted on some level by many companies, but it typically is a very slow process in which neither group is heavily invested, and therefore this coordination often fails. In compensation for lack of coordination of provisioning and applications selection, data centers are almost entirely provisioned with a one-size-fits-all (OSFA) server configuration, which can't be designed to any successful level of efficiency. Even if some provisioning is done of a server configuration specifically matched to an application category, when that application is retired which sometimes might be as little as six months later, there exists no criteria by which those servers can be re-purposed, and so are replaced with OSFA servers in a continuing cycle of inefficiency.

1.4. Currently, the design of high-volume standard computational servers and storage systems tends to follow one path: high performance. This design concerns the creation of a new path: energy efficiency. Almost nothing is done today by conventional server manufacturers to create energy efficient servers, except for minor attempts such as “shared fans”, which, while in some cases can result in slightly reduced power consumption versus conventional servers with non-shared fans, is less efficient when viewed at the data center level. If ultra-efficient servers are employed, shared fans are a markedly less efficient design. As such, this does not provide sufficient gains in efficiency to be considered as an efficient design or as a valuable proposition in the market place.

1.5. What is needed in the marketplace is the ability to know how much power a server consumes, and therefore the ability to compare that power consumption server to server, and then by extension the power consumption at the rack level, and the datacenter level. Conventional inability to compare and rate power consumption, both on a server-versus-server basis and on a server-versus-application-load basis typically means that conventional server manufacturers do not have the ability to know, and therefore to claim, that their products are efficient. If fact, the servers produced under this lack-of-methodology environment are generally specified by the capacity of the power supply included, and do not yield the significant power savings possible through the use of this art.

2. Current Approach

2.1. 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 compared to similar servers which are not energy efficient. Such energy efficient servers are currently designed by selecting non-standard off-the-shelf components from vendors offering lower power components. Cooling and power components are typically selected to as closely match the requirements of each specific model and configuration as possible, and no more, eliminating the waste that results from using overly powerful components. Additionally, major components such as printed circuit boards, fans and power supplies can be specifically selected 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.

3. A System Approach

3.1. In order to design for maximum server energy efficiency, all components of the server are to be considered, and as many as practicable measured for their contribution to power consumption and power dissipation (in the case of fans and air flow controllers), as well as the relationship between all those components in regard to the TDP (Total Design Power) of the server. As such, the design of an ultra-efficient server crucially relies on a repeatable, deterministic methodology for measuring maximum power consumed. For this embodiment, both Idle Power Consumption (IPC), which represents the power consumed by the device when it is powered up and ready but has negligible computational load (idle), and Total Design Power (TDP), which represents the maximum power a server will consume under full computational and I/O load, will be considered as the measure of energy consumption. TDP is typically 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 with the high reliability and availability of the services the IT equipment is intended to provide. IPC is typically important because most servers spend 80-90% of their operational existence in the idle state. Accurate measurements of the TDP of a server are currently unavailable on standard servers. Anecdotal TDP measurements for a server are sometimes available from third parties, however these are always very inaccurate (sometimes as much as 50% too low) and only applicable to the particular configuration tested by the third party. This last part is also typically important because conventional servers can have widely varying arrays of possible components, resulting in a configuration matrix for just one model of server that numbers in the tens of thousands.

3.2. All servers will consume varying amounts of power depending on the instantaneous load on the computational components as well as on the power supply and cooling system. As the computational load increases, so does the power consumed by the server. Once the components are operating at their maximum capacity, all electrical components in the server will be theoretically operating at their maximum and TDP should then be measured. Figure 1/conventional shows the power consumption curve for servers designed under a conventional methodology, and Figure 1/Lopoco under the art of this design. Comparing the two curves, it is evident that the power consumption is substantially lower for servers produced under the art of this design, yet the server throughput stays similar.

3.3. Designing a highly efficient server involves minimizing power consumption while maintaining requisite performance throughput. While obtaining a reasonably accurate IPC measurement for a server is relatively straightforward, doing so for TDP is quite difficult. An embodiment of this design utilizes a testing process, methodology and software algorithms to most accurately determine the TDP rating of a computer server. This design methodology utilizes the Special Copyrighted Software Program SWTDP 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 Figure 2 to help identify the most energy efficient components and configurations, as well as fully understand the effect on a server's power consumption of a particular component, compared to other components in the same category. These comparisons are components of the complex overall methodology of designing an ultra-efficient server. Further, power reductions in the computational components and motherboards have a virtuous feedback 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 to fully exercise the system components, a minimized TDP is determined as the maximum power the server will consume under virtually any condition or application load. This TDP can be used as a parameter of maximum power per server when designing collections of servers consisting of 100's or 1000's configured in racks for commercial applications such as business enterprises and data centers. By utilizing a significantly lower TDP than conventional servers, server racks and data centers can be designed with lower energy consumption and lower cooling requirements. Since powering and cooling collections of servers also consumes its own power, these datacenter subsystems, power distribution and HVAC, also become more energy efficient. Thus, the ability to design and specify a server to significantly lower, ultra-efficient power consumption causes a cascading ripple effect and a virtuous circle of power reduction and energy efficiency gain.

3.4. Techniques provided herein therefor can be variously applied or embodied as: (a) techniques and/or software to assist with dynamic application migration, as referenced below; (b) techniques and/or software to assist with predictive power need determination, based on (e.g., expected or actual application flows), and associated power and/or cooling management, supply and diversion (i.e., in a machine, or collection of machines, such as a data center); (c) improvements in computer, blade and/or network appliance design, or the design of a collection of machines (such as a data center); and (d) in various methods, apparatuses, software, machines or systems, as will be further apparent from the description below. Note that while several specific design techniques for computer/machine design are described, each of these should be viewed as optional, with any permutation of these or combination of these being potentially useful, depending on circumstance or design.

4. Brief Description of One Design

4.1. This design description will cover a system level procedure for implementing an ultra-efficient data center.

4.2. Design for attaining efficiency from IT equipment

4.2.1. In order to create a datacenter for high IT efficiency 3 essential operational methodologies can be used all or in part:

4.2.1.1. Servers and storage equipment that can operate efficiently, grouped into classes based on application workload footprint

4.2.1.2. The dynamic and/or manual ability to match applications with the server configuration/model that is best suited for it on an acceptable efficiency/performance scale, and the ability to seamlessly migrate or run those applications on the equipment selectively suited for it

4.2.1.3. Data center cooling zones based on the TDP (Total Design Power = maximum power the server dissipates when fully utilized) of the server classes housed within each zone – with right-sized HVAC (cooling) provisioned separately for each zone. Zones will be physically separated enough so that the HVAC resources provisioned for each zone area will not be burdened with cooling duties for other zones/areas. See Figure 11 as an example.

4.2.2. Therefore data centers should be divided into separate application class zones, each provisioned only with the servers appropriate for that application class, with the HVAC requirements accurately known for each zone, so cooling equipment requirements will be well defined for each zone and can be precisely provisioned, gaining great cost savings over traditional designs, where the cooling is provisioned based on the maximum power available to the entire data center, rather than the actual cooling requirements of the IT equipment. These zones may or may not be separate buildings or structures, so in fact a data center may be comprised of only one zone. See Figure 11 as an example.

4.2.3. Use of ultra-efficient servers with accurately specified TDP will substantially decrease the amount of cooling required vs. the throughput of the servers, again decreasing costs substantially. Most data centers are constructed with the HVAC capacity, and therefore the

HVAC equipment provisioned, according to some arbitrary rule of thumb regarding the ratio of total power provisioned for the whole data center. These rough estimates are intentionally very conservative, and the result is often a large disparity in the amount of HVAC capacity versus the needed HVAC capacity, wasting money at the time of purchase, and wasting power consumed in the operating phase. Redoing the HVAC later can reduce the amount of power that was being wasted, but is done at very great expense. Perhaps the worst is if the amount of cooling resources that is initially underestimated, resulting in repeated costly repairs and ultimately a redo of the HVAC, if even possible at all. All of this unnecessary expense is eliminated when the data center owner/designer is able to precisely know ahead of time the amount of cooling resources that will be needed.

4.2.4. In the same way as "rightsizing" the needed cooling resources in the ultra-efficient data center, the same may be done for battery backup power resources (UPS), generator backup power resources, and power distribution equipment such as internal data center power transmission wiring, and power cords, and power distribution units (PDU). We can aggregate all these into one group labeled "power infrastructure resources". Again allowing for accurate right-sizing of power infrastructure resources results in very large cost and power consumption reductions over a conventional data center. Use of ultra-efficient servers will allow for substantial increase in overall throughput of a data center v. the data center's total power budget, substantially increasing the value of a data center with a specific power budget.

4.3. Application Management and Control Software

4.3.1. Proprietary Application Management and control software and hardware will be utilized to identify and analyze applications with respect to their computing resource needs, and migrate or launch applications to the server category best suited to handle the application efficiently. This section covers the design and implementation specifics of that sub-system. Monitoring, migrating and controlling functionality is a combination of software and hardware. The hardware involved is typical power and environment monitoring hardware where data on power consumption and heat can be constantly monitored and gathered. This data, along with software monitoring metrics gathered from the servers themselves on the resource usage patterns of running applications, is also used as input to the application management and control software.

4.3.2. Application migration software

4.3.2.1. This part of the system utilizes the data from the sensors and application monitoring metrics to migrate applications via proprietary application management software from one data center class to another, or take no action if an application is allocated the intended amount of resources for the application class data center sub-module where it currently is running.

4.3.2.2. Processor resource usage

4.3.2.2.1. If an application is consistently using 80% or greater of the available CPU resources in its current sub-module class, and there is a sub-module class with a higher performance processor, then the application will be migrated to that class. If an application consistently is using no more than 10% of the available CPU resources in its current sub-module class, and there is a sub-module class with a lower performance processor, then it will be migrated to that class.

4.3.2.3. Network resource usage

4.3.2.3.1. If an application is consistently using 100% of the available networking resources in its sub-module class, and there is a sub-module class with higher throughput networking hardware, then the application will be moved to that sub-module class. If an application consistently is using less than 10% of the available network bandwidth in its sub-module class, and there is a sub-module class with lower speed networking, then it will be migrated to that sub-module class.

4.3.2.4. Storage resource usage

4.3.2.4.1. All storage in the data center needs to be accessible to all the servers in the data center, therefore the amount of storage capacity required by an application will have no bearing on which sub-module class it is executing in. However, the disk I/O throughput is a criteria. If an application consistently experiencing high I/O wait times for disk I/O operations, and there is a sub-module class with higher performance storage hardware, then the application will be migrated to that sub-module class. For this algorithm, high I/O wait times are when the CPU I/O wait time is 90% or greater.

4.3.2.5. The migration decision structure is weighted with CPU utilization the highest priority, then Network, followed by Storage.

4.3.2.6. There are two methods of application migration.

4.3.2.6.1. Start/Stop

4.3.2.6.1.1. The application is merely stopped in one place, and restarted in another, but not necessarily in that order. An example would be a web service application, which can be started in the destination sub-module class, and once started, stopped in the origin sub-module class.

4.3.2.6.2. Virtual Machine Migration

4.3.2.6.2.1. Most virtual machine frameworks these days include the ability to migrate an application from one server to another. Many large data centers now employ an application container framework. Such frameworks can be heavily employed to migrate services and application containers from one group of servers in a sub-module class to another.

5. DETAILED DESCRIPTION

5.1. Ultra-Efficient Server Design - Low Power Server, Server Systems and Design Methodology

5.1.1. This section 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.

5.1.2. 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 the following categories: 1)Efficient motherboard design; 2)Ultra-efficient power supply components; 3) Greater Fan Efficiency; 4) Utilization of airflow control devices; 5)Variable speed fans with full-stop capability; 6)BIOS and APC Code Optimization; 7)Server case design for efficient airflow

5.1.2.1. **Efficient motherboard design**

5.1.2.1.1. Efficient Computer server motherboard design elements are also employed for increasing overall power efficiency by the use of 5 volt fans instead of 12 volt fans: 5-volt fan headers, allow the use of 5V fans which typically can move equivalent amounts of air while using equivalent amperages in slower speed fans, but at 5 volts instead of 12 volts, they allow a significant reduction in wattage over traditional 12 volts fans. These can be incorporated into efficient server's designs that don't need high speed and/or high pressure cooling fans.

5.1.2.1.2. Onboard components and component connector placement/alignment: 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 through the computer, thus increasing cooling efficiency. This is shown in Figure 5A.

5.1.2.1.3. I/O device connector port placemen on Motherboard: 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 airflow 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 with substantially less restriction, 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 and to other tall components that could hinder airflow. Also shown in Figures 5A. and 5B.

5.1.2.1.4. Use of highly efficient 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 ethernet PHYs; CPUs; fan controllers; disk controllers and associated integrated circuit devices; video controllers (if any); USB controllers; and keyboard/mouse controllers. 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 IPC.

5.1.2.2. **Ultra-efficient power supply components**

5.1.2.2.1. Ultra-efficient power supply components: The importance of being able to use of ultra-efficient power supplies cannot be overstated. Conventional power supplies may claim to be 80% or 90% efficient, but in fact only achieve that level of efficiency when producing greater than 200 watts of power. Even when operating at those efficiency levels, they are very wasteful and such a deficit cannot be overcome. Power supplies such as those currently utilized by Lopoco are able to operate at 80-90% efficiency while outputting only 20 watts of power, and immediately jump to 97-99% efficiency at output levels above that. Therefore we can see that a conventional server that is consuming 200 watts is losing almost 40 watts just in the power supply area alone (80% efficiency @ 200W), which means the computational and cooling components of the server are only consuming 160 watts. The power supply consuming 40 watts itself means that it then also needs additional, and usually separate, cooling resources dedicated to it, which further contributes to its inefficiency. In contrast, the ultra-efficient Lopoco power supplies consume 1.2 watts for an equivalent system.

5.1.2.3. **Greater fan efficiency**

5.1.2.3.1. Fans with disappearing frame segments: 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 Figures 6 & 7. This means that the external dimensions of the entire fan assembly will not be square, but wider in the horizontal dimension 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 direct airflow or to protect the fan from foreign objects or to protect foreign objects from the spinning fan blades, because the case itself serves that purpose.

5.1.2.3.2. Fans 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. Reducing vibration increases reliability while reducing acoustic noise, which key part of efficient data center consideration. 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 Figure 7.

5.1.2.3.3. 42mm system fans for use inside a computer server case utilizing all the available vertical space: This section is a design for 42mm (1U) and 84mm (2U) case fans specifically for optimizing efficiency of cooling. 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.

5.1.2.3.4. Current 1U case fans are 40mm square and anywhere from 10mm to 150mm deep, or deeper. This fan design specifies 42mm or 84mm square fans, thereby leaving no amount of internal space empty.

5.1.2.4. **Utilization of airflow controlling devices to smooth and direct cooling airflow inside a server;**

5.1.2.4.1. Device(s) to smooth and direct cooling airflow inside a computer: 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 airflow around such blockages allows for more efficient cooling of the computer, which allows for a reduction in the amount of fan power required.

5.1.2.4.2. 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 is most needed. 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).

5.1.2.4.3. Device(s) will have various shapes, including: Basic vertical and horizontal wedge shapes, as well as the half cone shape, as seen in Figure 8 and Figure 10 (top view sketch). This shape is designed to transition airflow over the top of the blockage. A similar vertical wedge shaped device would be used to direct the airflow which would otherwise go underneath the motherboard or circuit board to the top of the board where most or all of the hot components are located.

5.1.2.4.4. Basic blended vertical and horizontal wedge, used to transition airflow 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 oriented 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.

5.1.2.4.5. Devices that are essentially the device shape shown in Figure 8 and Figure 10, but cut into right and left halves. These are to be used in conjunction with device shape of basic blended vertical and horizontal wedges to create a complete air flow transition around larger blockages possibly consisting of multiple I/O headers mounted side by side or nearly so.

5.1.2.4.6. Rounded versions of the above described shapes: The devices are made of a very smooth, hard material, most likely plastic for cost and lightweight, facilitating good air flow.

5.1.2.4.7. 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.

5.1.2.4.8. Designs or shapes described above, but integrated into the external case of the I/O header or component itself.

5.1.2.5. **Variable speed fans including a full stop capability**

5.1.2.5.1. Variable speed (RPM) fans including a full stop capability: Computer system cooling fans with a variable speed capability that have vastly greater RPM ranges, including the capability to completely stop the fan motor, effectively drawing no power. The cooling fans vary their RPM and thus varying the amount of power they draw and the amount of air they move, in response to heat sensors which indicate how much heat needs to be evacuated. For ultra-efficient servers the added abilities to vary the RPM across a wider range especially the low RPM range, as well as the ability to completely stop the fan motor. Either through a communication signal via an interface, or merely by setting the fan supply voltage to zero, the fan will stop, without raising any system alarms. 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. Full stop fans are commonly utilized in laptops and on some direct cooling applications, but have not been used in computer servers.

5.1.2.6. **BIOS/boot up/ACPI code**

5.1.2.6.1. BIOS/Boot up code for ultra-efficient Servers Offer High Efficiency: The use of Boot up and self-test ROM or NVRAM code (herein called BIOS or POST code) greatly decreases the amount of power consumed during the power on self-test phase.

5.1.2.6.2. Current BIOS code in PC derived computer systems including servers, among many other great inefficiencies, utilize a continuous loop design to refresh the video display, rather than letting the video hardware do that itself. This is due to the fact that in the 1980's IBM PC video controllers did not have the internal capability to refresh the display, and so the CPU had to execute code every refresh cycle (roughly 1/30 of a second) for that purpose. Even though there hasn't been a video controller produced in 20 years that needs that, conventional server BIOS software has never caught up. Many servers have no provision for video display at all, not to mention that the use of directly connected video displays is a deprecated practice in datacenters for years, further eliminating the need for such code. In one embodiment of the design, the boot up code dispenses with the continuous loop video refresh code altogether and instead uses an interrupt driven design, thereby drastically reducing power consumption during boot up, system test, and interactive operations. Similar power reductions can also be achieved using a light weight polling design as well.

5.1.2.6.3. Other Code Optimization Features in the current design include:

5.1.2.6.3.1. Fan management code capable of turning one or more of the cooling fans off, i.e. RPM of zero, when fan operation isn't required to cool the system.

5.1.2.6.3.2. Thermal management code redesigned to significantly increase the effectiveness and efficiency, especially in the particular area of fan control. This will allow the BIOS to help reduce power consumption of the system during operation as well as during POST code execution during boot up.

5.1.2.6.3.3. Significant decrease in the overall time of POST code execution before turning over control to the operating system.

5.1.2.7. **Server case design for efficient airflow, including air intake and exhaust grill areas.**

5.1.2.7.1. Utilization of fans which have "disappearing" frame segments on the top and bottom 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 sides facing the top and bottom of the case. It also means the external dimensions of the fans may not be square, but wider in the horizontal dimension than the vertical.

5.1.2.7.2. 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 case 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.

5.1.2.7.3. Computer server case and motherboard design elements for more efficient airflow though the case - This description covers new computer server motherboard design elements, and matching case design elements, to facilitate less restricted airflow through the case, thus increasing cooling efficiency of the running computer server.

5.1.2.7.4. Figure 5A shows the motherboard design with the rear I/O ports located to the left and right sides of the rear edge of the motherboard. Traditionally these have been placed according to some standards which often find them directly in the path of airflow over the server's hottest components, such as the processor and power transistors. Relocating them in this fashion 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 of the bulk of this air flow, 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.

5.1.2.7.5. Air intake and exhaust grill design

5.1.2.7.5.1. In the current server designs, 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 airflow, badly affecting cooling efficiency.

5.1.2.7.5.2. One embodiment of this design, in contrast, 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.

5.2. Ultra-Efficient Server Design Methodology Considerations

5.2.1. 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 and networking appliances) with accurately specified power consumption profiles. These profiles can then be used to design the next generation power-optimized ultra-efficient data centers using application centric data center modules. These modules can include server designs described here as well as ultra efficient network routers and other devices which were are designed utilizing a similar TDP methodology. See Figure 9.

5.2.2. 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, and so forth. Servers maybe optimized for use of general purpose processors, or GPUs, FPGAs, and ASICs. The ability to deploy ultra-efficient servers 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 sub module of the NGAPD intended to most efficiently run them.

5.2.3. 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).

5.3. Efficiency and Total Design Power (TDP) - measuring system and methodology

5.3.1. In order to be able to achieve the goal of an ultra-efficient data center and to use its IT equipment efficiently (see section 4) and Ultra Efficient Server Design (Section 5.1), an accurate and repeatable method of determining the power consumption footprint (e.g. IPC and TDP) of a server is necessary. For Section 5.1, it is utilized extensively in the research, engineering, design, and verification phases. For Section 4, the accurate and reliable power consumption specification of IT equipment, especially servers and compute appliances such as computer storage and networking equipment, is crucial to the process. In this section, we will describe the uniquely accurate test measuring software and methodology created by Lopoco to achieve accurate and reliable specifications for IPC and TDP for a server.

5.3.2. Server TDP Software Design (SWTDP): Summary - the software program attempts to utilize all the compute and I/O resources of a server at the same time, thereby causing all subsystems of the server to be using the peak maximum power that they could ever consume in operation.

5.3.2.1. *Overview* - modern computer hardware, both integrated and discrete circuits (often called chips and components, respectively), have some ability to curtail power consumption when some or all of their capabilities are not being used. For example, a processor may shut down a floating point computation unit when there are no floating point computations to be done, or a network controller might shut down parts of its logic during periods of little or no network traffic, etc. These periods of shutdown or power down may be extremely short, possibly measured in nanoseconds or microseconds. If there is no carrier signal detected on an Ethernet port, the Ethernet PHY may power down for 950 microseconds and power up for 50 microseconds every millisecond to check if a cable has been connected and a possible link can be established, thus reducing power consumption by 95% if a network port is unused. The SWTDP software is designed to defeat the various automatic power saving features of various server components and force the server to maximize its power consumption. The operating method is to run the SWP software on a target server, utilizing one or more helper servers to send and receive network traffic, and measure the power consumption of the target server during the time the software is running. Power consumption measurement tools have the ability to record the maximum instantaneous power consumption during an arbitrary time period, and these numbers (watts, PF, amps and volts) are stored and recorded in a database, so that the observed TDP for each and every server is known.

5.3.2.2. The detailed SWTDP algorithms for exercising each subsystem of a server are split out by each subsystem in the list below:

5.3.2.2.1. General Purpose Processor: Today's modern general purpose processors have multiple CPU cores, large memory caches, TLBs (virtual memory translation lookaside buffer), memory controllers, data buses and bus controllers, and I/O device controllers. Each CPU core has many multiple execution units and instruction pipelines. All of these are manipulated by the logic in the processor to save power when not being used. The challenge in determining TDP is to use software that keeps as much of the whole processor activated (i.e. not able to power save) as possible. Lopoco uses encryption key generation and computation software by the name of *openssl* that exercises all of the floating point and specialty instruction units as fully as possible, while simultaneously using sufficient amounts of memory. Z instances of this program are run, where Z = (the number of logical CPU cores available on the processor or processors in the system) + 2. This prevents the processor(s) from shutting down execution units, pipelines, memory controllers and other such parts to reduce power consumption. Running Z copies of the program means that a wide range of memory in the system is used. This maximizes memory system power consumption by greatly reducing the amount of power saving that the memory chips and memory controllers are capable of doing. Many software programs were tested for the purpose of maximizing processor and memory usage, and while many do well at this goal, the openssl key generation program produces results that cannot be equaled or exceeded by any other currently available program. Hence an ordinary user running any software on the system will not be able to cause the server to consume more energy via the processor and memory, which is the ultimate goal of TDP determination.

5.3.2.2.2. Disks and disk I/O controllers: Similar to processors, disks, SSDs and disk I/O controllers have electronics that are designed to dynamically save power. The TDP software discovers all the disks in the system, groups them by type and model, and creates a software RAID setup for each group. A separate program loop repeatedly unpacks a large file archive simultaneously onto each group of disks, causing a large number of disk writes to be performed on each disk in the group at the same time. For all types of disks (rotating media and SSD, to name two), the write cycle consumes the most power. Expanding these archive files has two additional effects: firstly a large amount of metadata reads are performed, and on rotating media substantial extra head seeks are incurred which increases power consumption; and secondly the archive contains many small and large files, which causes a large backlog of pending writes to be queued up by the disk controller devices, additionally maximizing power consumption in the system. A program loop runs in parallel for each group of disks, keeping all disks busy all the time.

5.3.2.2.3. Network devices: Utilizing at least one helper server, a network connection is set up from the target server to the helper server, and simultaneous bi-directional data transfers are performed on a continuous and repeating basis. Since most “normal” network traffic consists of one machine transmitting while the other receives, network controllers can save some power consumption by powering down the unused receiver or transmitter, respectively. Therefore the TDP software is specially designed to use bi-directional transfers to eliminate that power usage optimization, again maximizing power consumption. The software has a user selectable option to exercise an additional network port, and by extension an additional helper server on the other end of the network connection. This option is initially intended for servers with additional higher speed network devices, for instance a 10G, 40G or 100G network device, which currently can consume considerably more power when at idle as well as when fully utilized than a conventional 100BT or 1G Ethernet network interface.

5.4. TDP Testing Methodology

5.4.1. A brief description of the testing methodology that is employed to ascertain the idle and TDP power consumption measurements for a server – utilizing the SWTDP software is described in the following sections below. See Figure 4.

5.4.2. When a server is released from the production line and ready to be tested, its power cable is plugged into a lab equipment power meter. This test device must be both accurate and repeatable, and must have at least the following for power consumption readings: watts, power factor, amps and volts. The capability to capture and/or record the maximum numbers sensed during a certain time period is highly useful. A network cable is connected to at least one of the server's network ports. All disks and devices that the unit will ship with should be connected to the system and should power on at the same time as the system is powered on. If the system has a video connector and/or a keyboard connector, those should be attached to suitable devices so that the power the system would utilize to operate them will be consumed. Most servers do not use a mouse, so no such device need be connected.

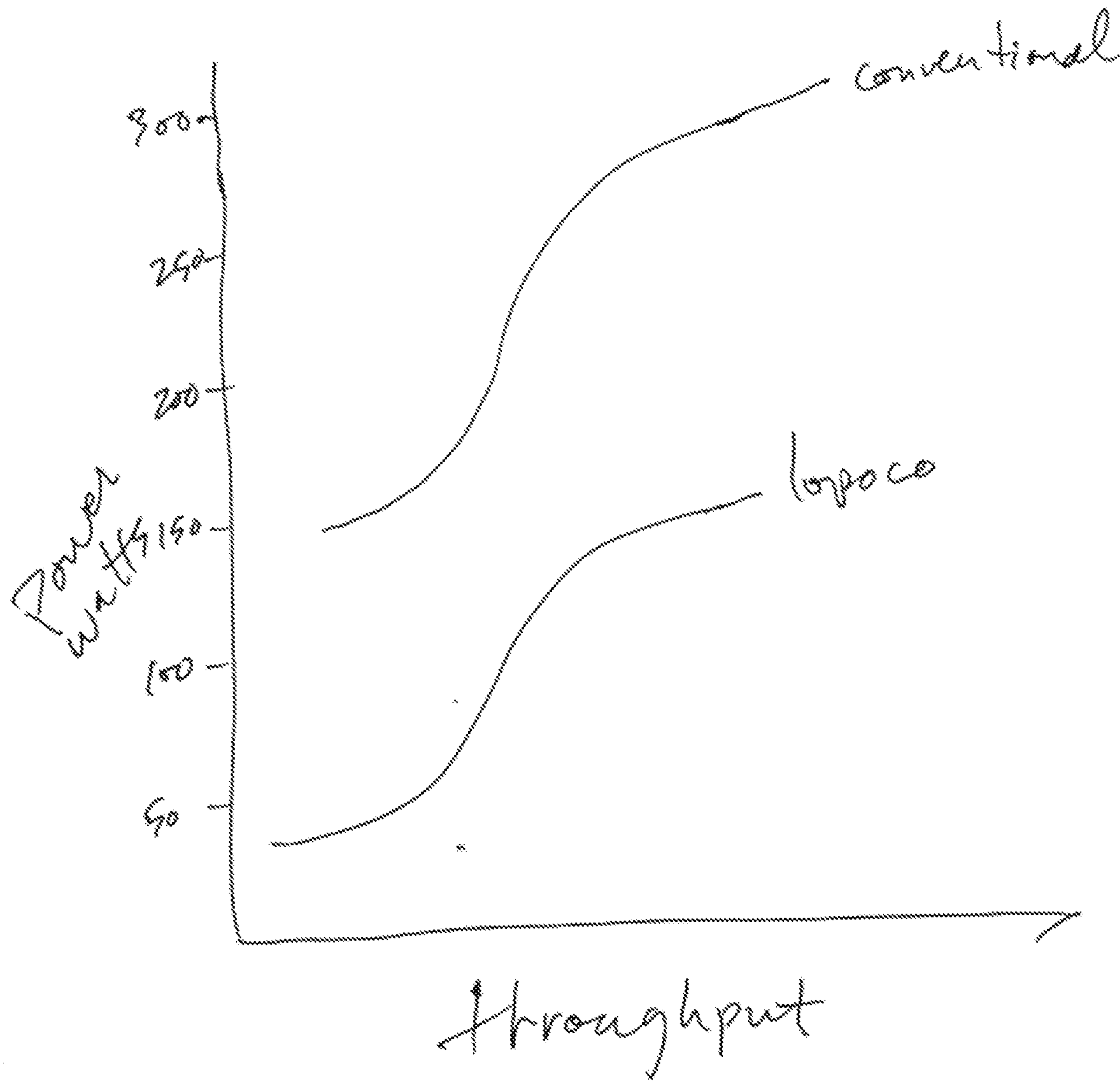
5.4.3. The server should then be booted, either from a non-disk medium such as a USB stick, or network booted, which means a portion of the system memory is set aside to load and run the operating system. Currently a version of the Debian Linux distribution is used, with a typical collection of services running but idle, such as the web server and a database server, as well as a few other sundry minor services. An IP address is obtained from a DHCP server via the network that is connected. After the system fully completes the boot up process, a suitable time is waited to allow the system to quiesce so that a reasonable idle power consumption measurement may be observed and recorded.

5.4.4. The idle power measurements are recorded (IPC).

5.4.5. The power meter should be set to record the maximums if such an additional setting is required on the meter to obtain the maximums.

5.4.6. The SWTDP software is run and after the software queries the hardware and determines the configuration, the main power exercising portion of the program begins. This portion of the program can take 15-30 minutes, roughly, depending on the performance capabilities of the server and the processor(s). When the power exercising portion of the program has completed, depending on the capabilities of the lab equipment, the maximum (TDP) power measurements may be read programmatically from the meter, or the software will ask the operator to enter the values, which will then be transmitted electronically to the operations database server for recordation. The configuration inventory of the server's memory, disks, network and disk controllers, and processor(s) which was collected earlier by the software is also electronically transmitted to the operations database server.

Figure 1



Notice how the curves don't align vertically. That's because we're willing to give up a small percentage of unneeded performance for large increases in efficiency. I tried to draw the curves to indicate less than 10% performance difference. For many applications that are not CPU bound, the actual throughput difference will be neglibilble.

Figure 2

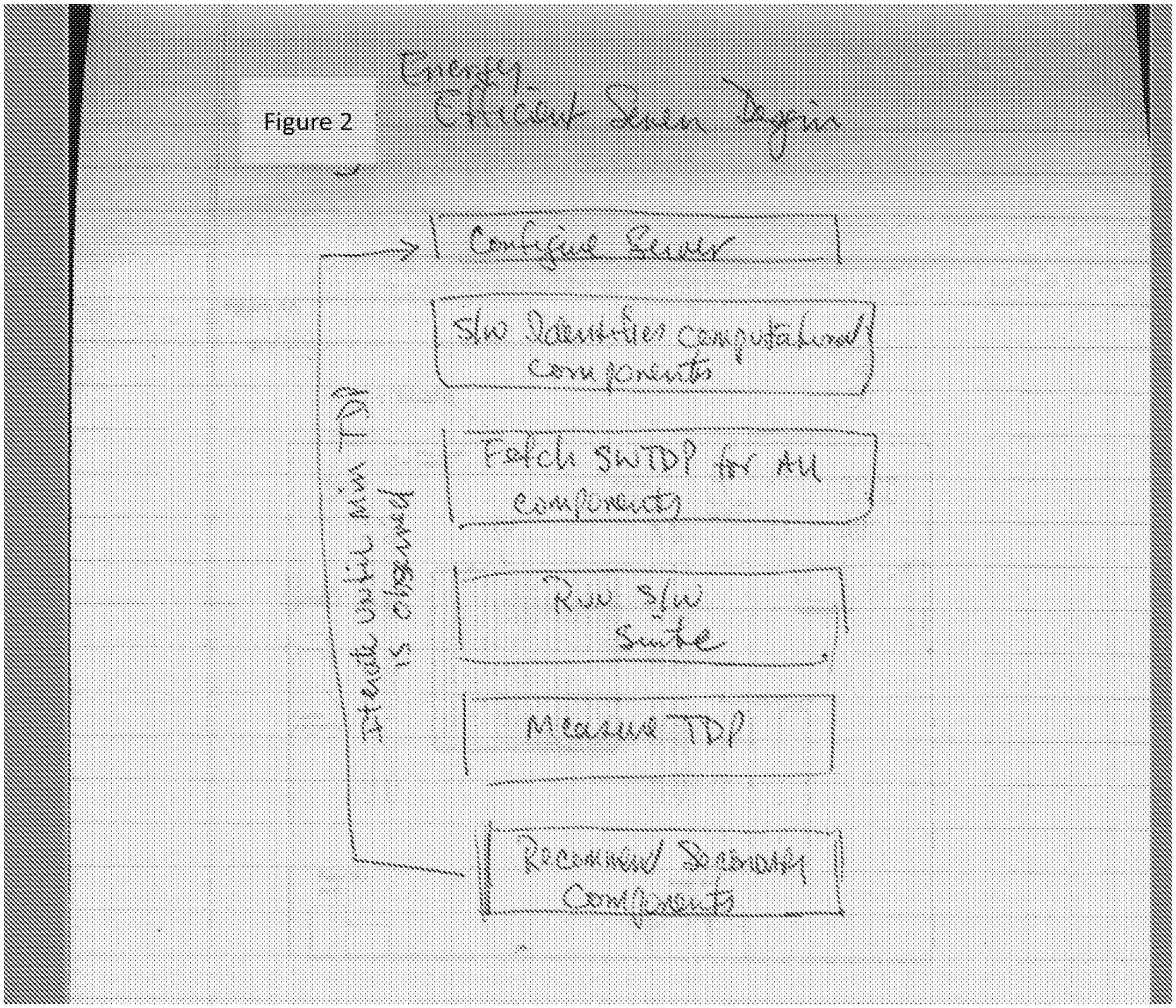


Figure 2
Energy Efficient Server Design
Configure Server
S/W Identifies computational components
Fetch SWTDP for All components
Run S/W Suite
Measure TDP
Iterate until min TDP is observed

Figure 4

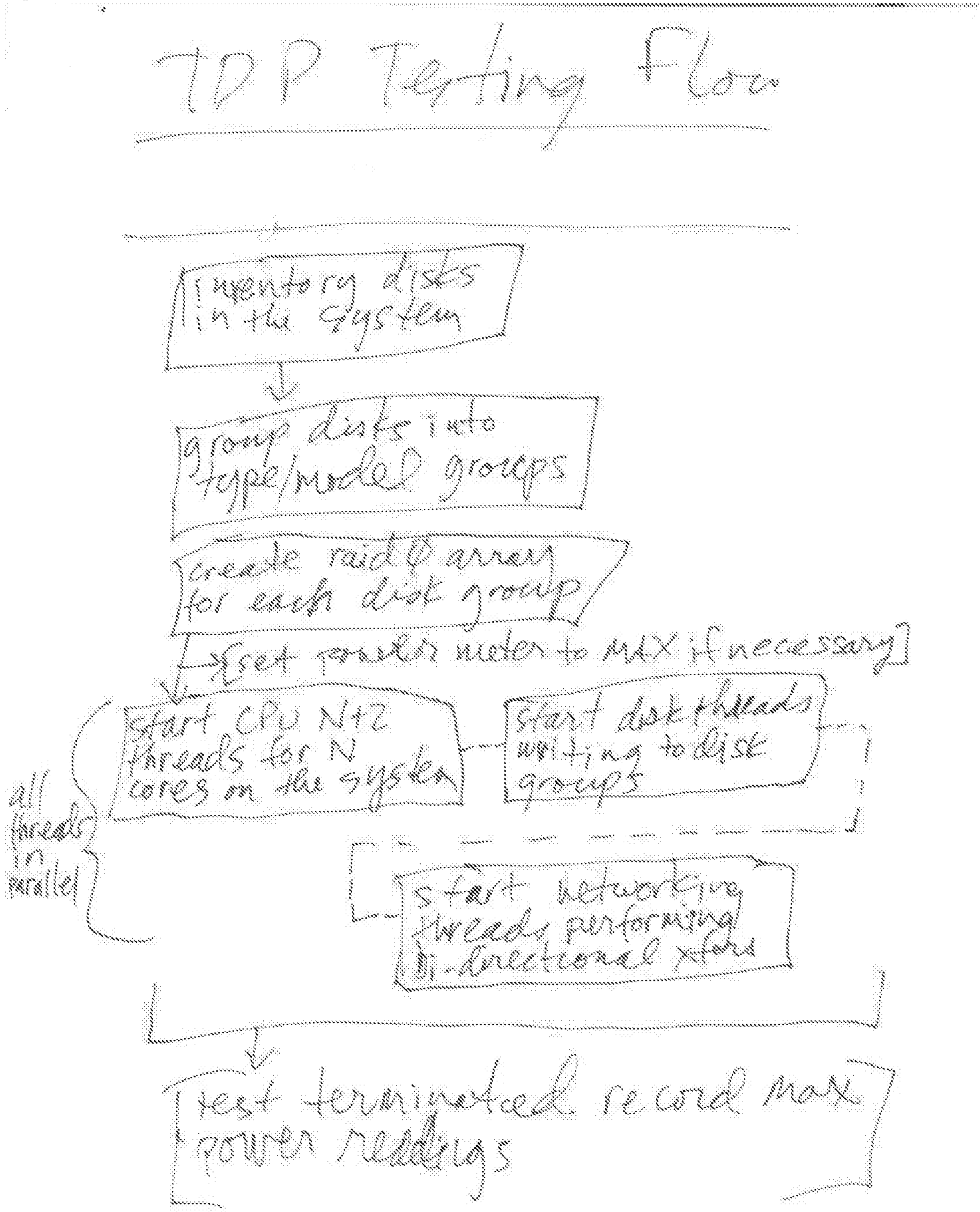

TDP Testing Flow
inventory disks in the system
group disks into type/model groups
create raid 0 array for each disk group
[set power mode to MAX if necessary]
start CPU N+2 threads for N cores on the system
start disk threads writing to disk groups
start networking threads performing bi-directional xfers
all threads in parallel
test terminated record max power readings

Figure 5A

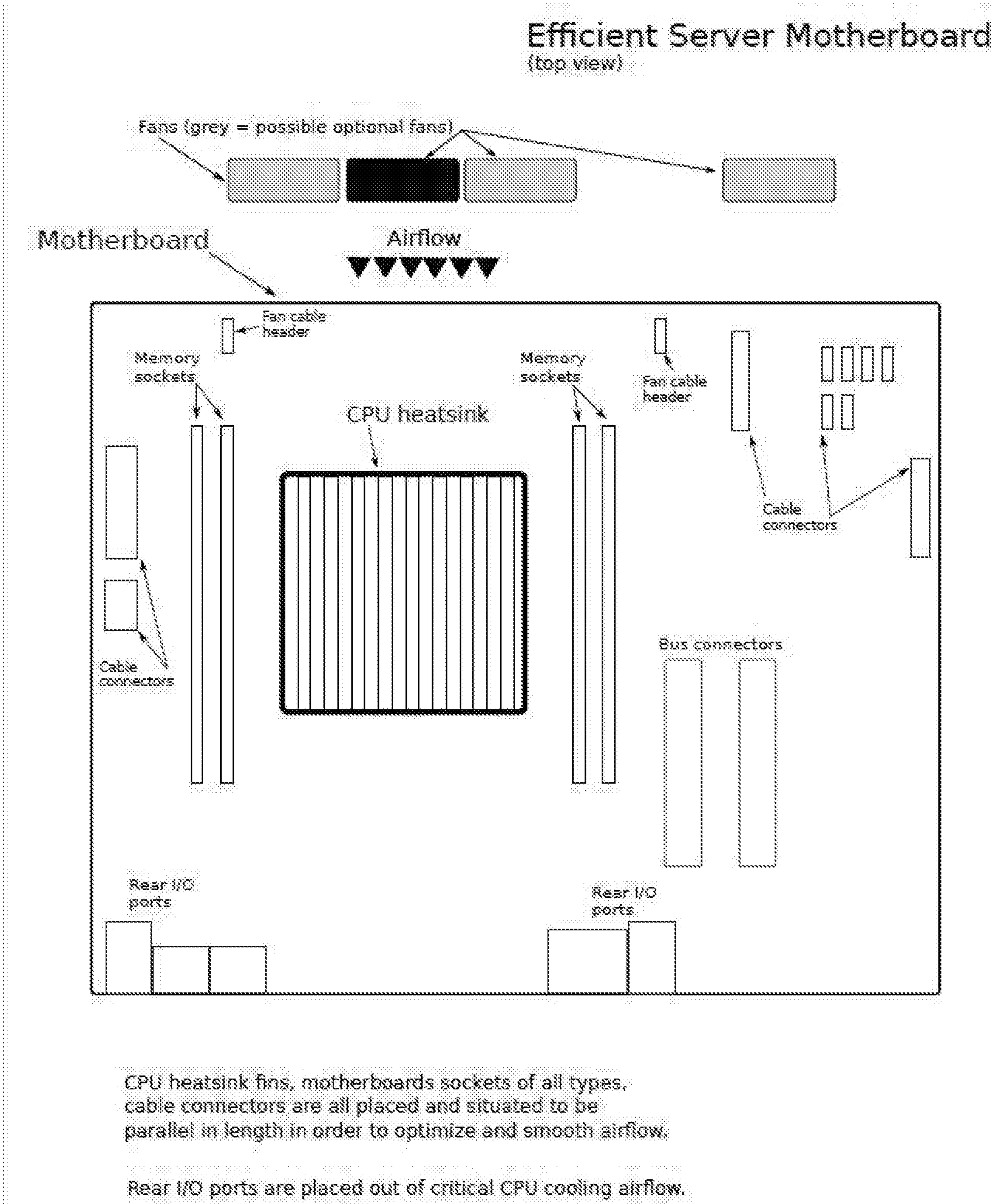

Efficient Server Motherboard
(top view)
Fans (grey = possible optional fans)
Motherboard
Airflow
Fan cable header
Memory sockets
Memory sockets
Fan cable header
CPU heatsink
Cable connectors
Cable connectors
Bus connectors
Rear I/O ports
Rear I/O ports
CPU heatsink fins, motherboards sockets of all types, cable connectors are all placed and situated to be parallel in length in order to optimize and smooth airflow.
Rear I/O ports are placed out of critical CPU cooling airflow.

Figure 5B Mother Board Side View Airflow

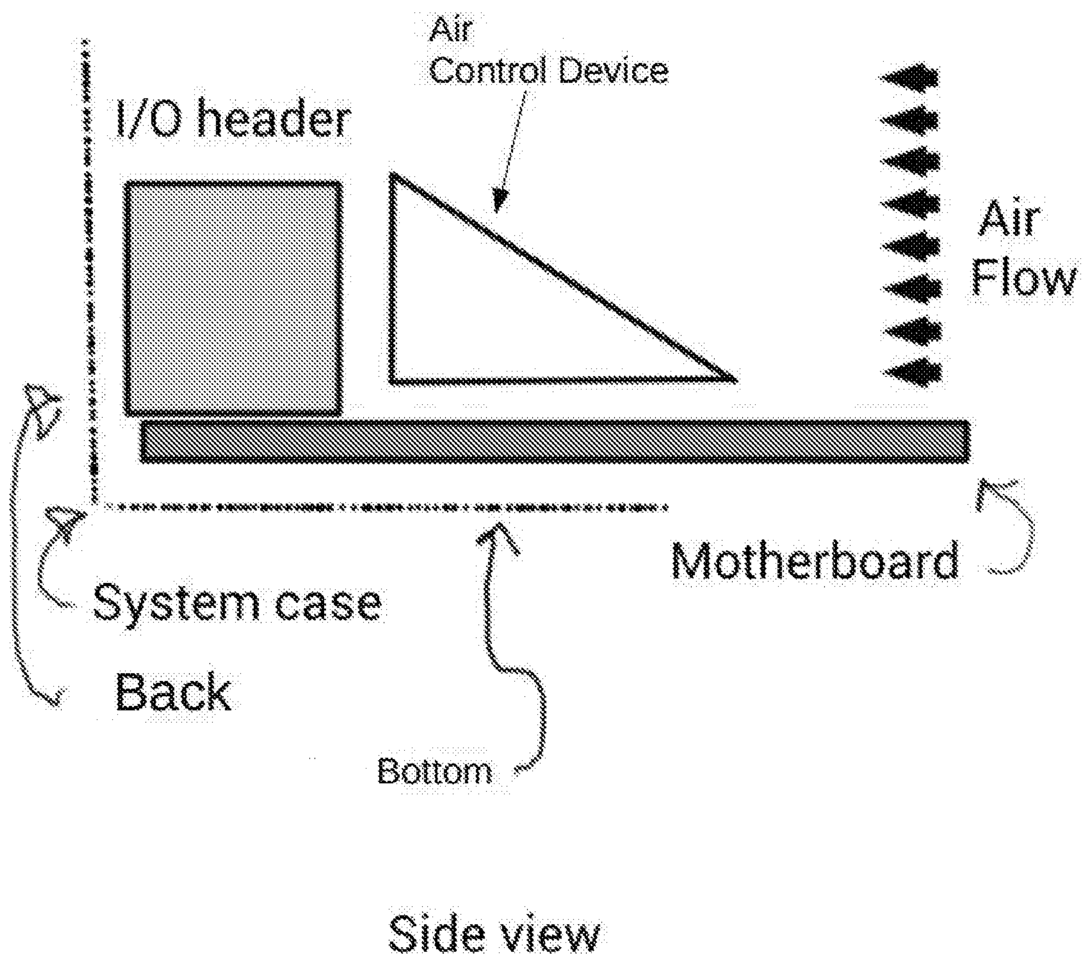


FIG 6

Example drawing of fan with disappearing top and bottom frame segments. Fan blades not shown (frame only).

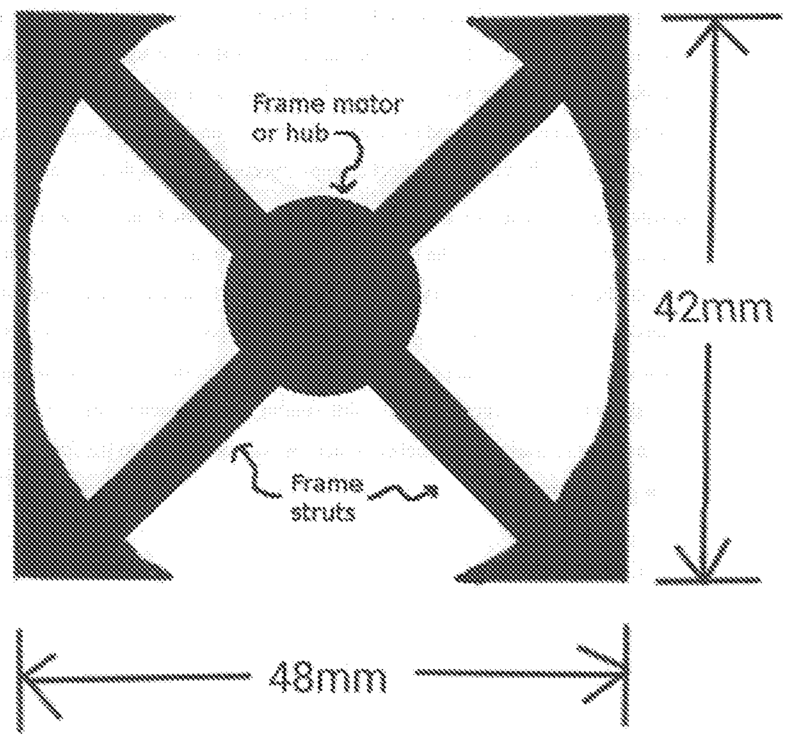


FIG 7

| Example drawing of fan with disappearing top and bottom frame segments.

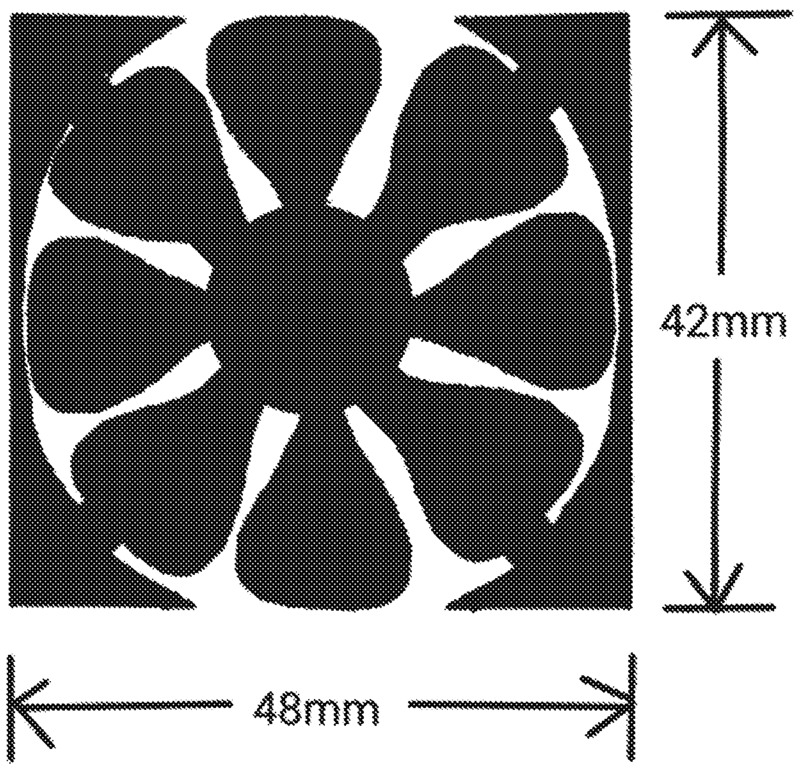
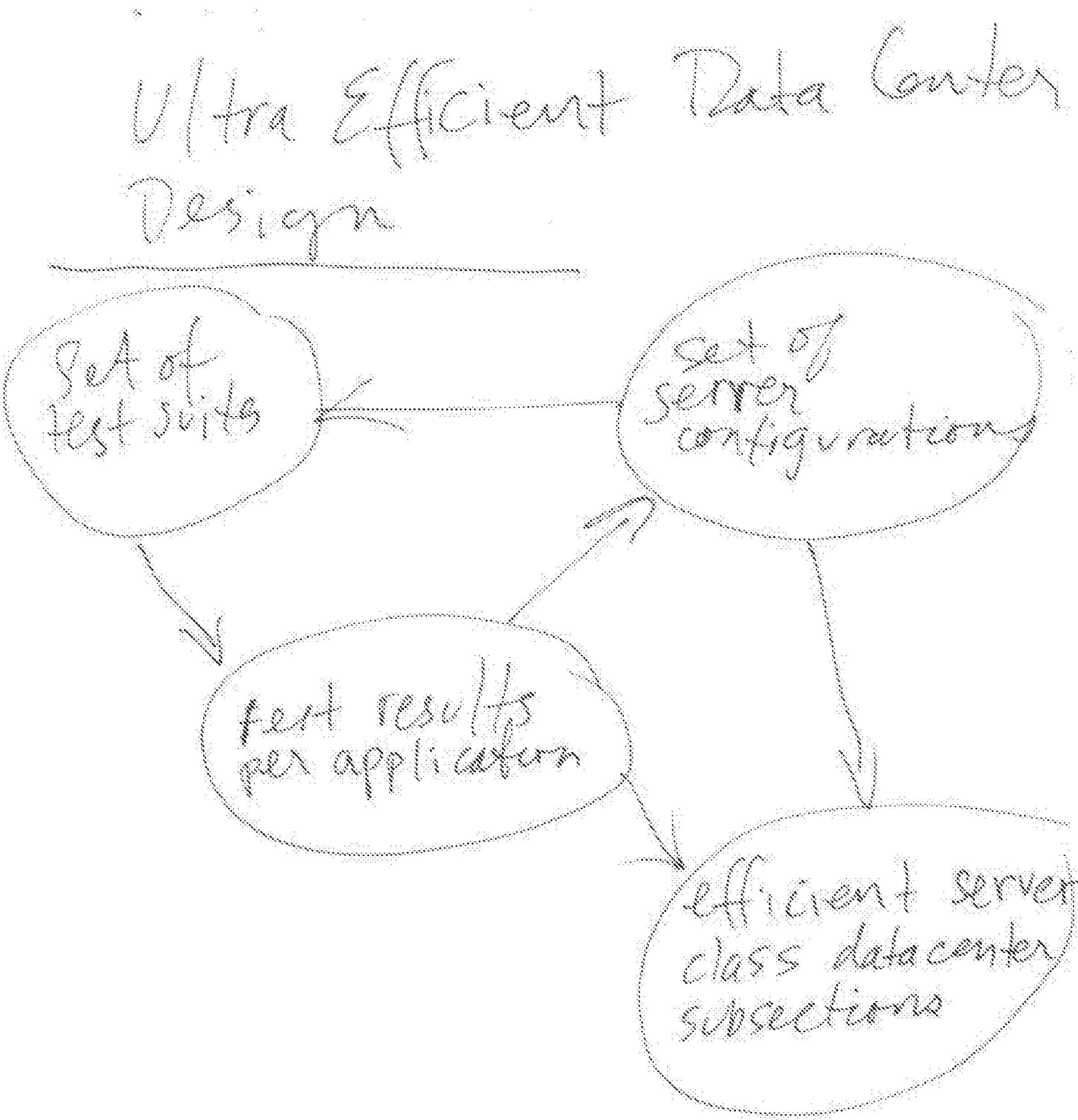


Figure 9



Ultra Efficient Data Center Design
Set of test suits
set of server configurations
test results per application
efficient server class datacenter subsections

Figure 8

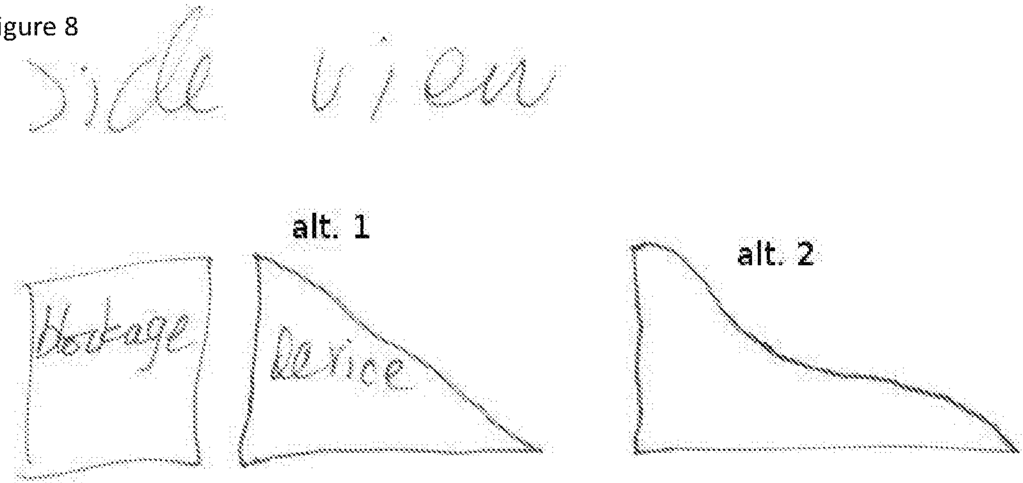


Figure 10

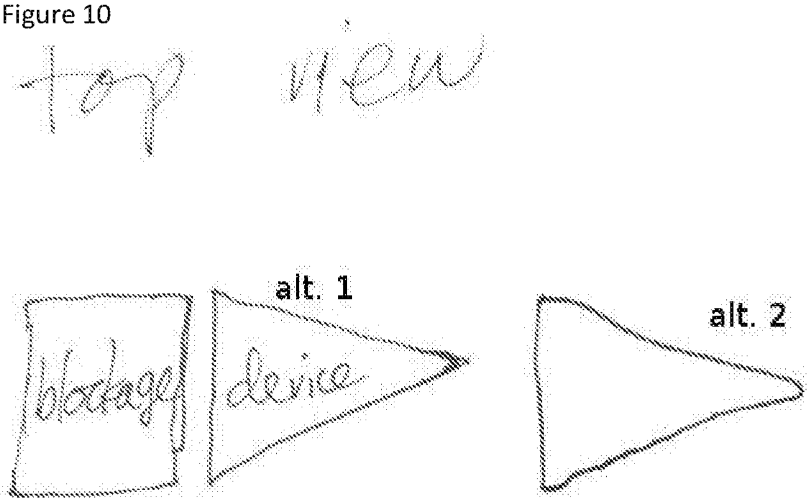


FIGURE 11

Ultra-Efficient Data Center

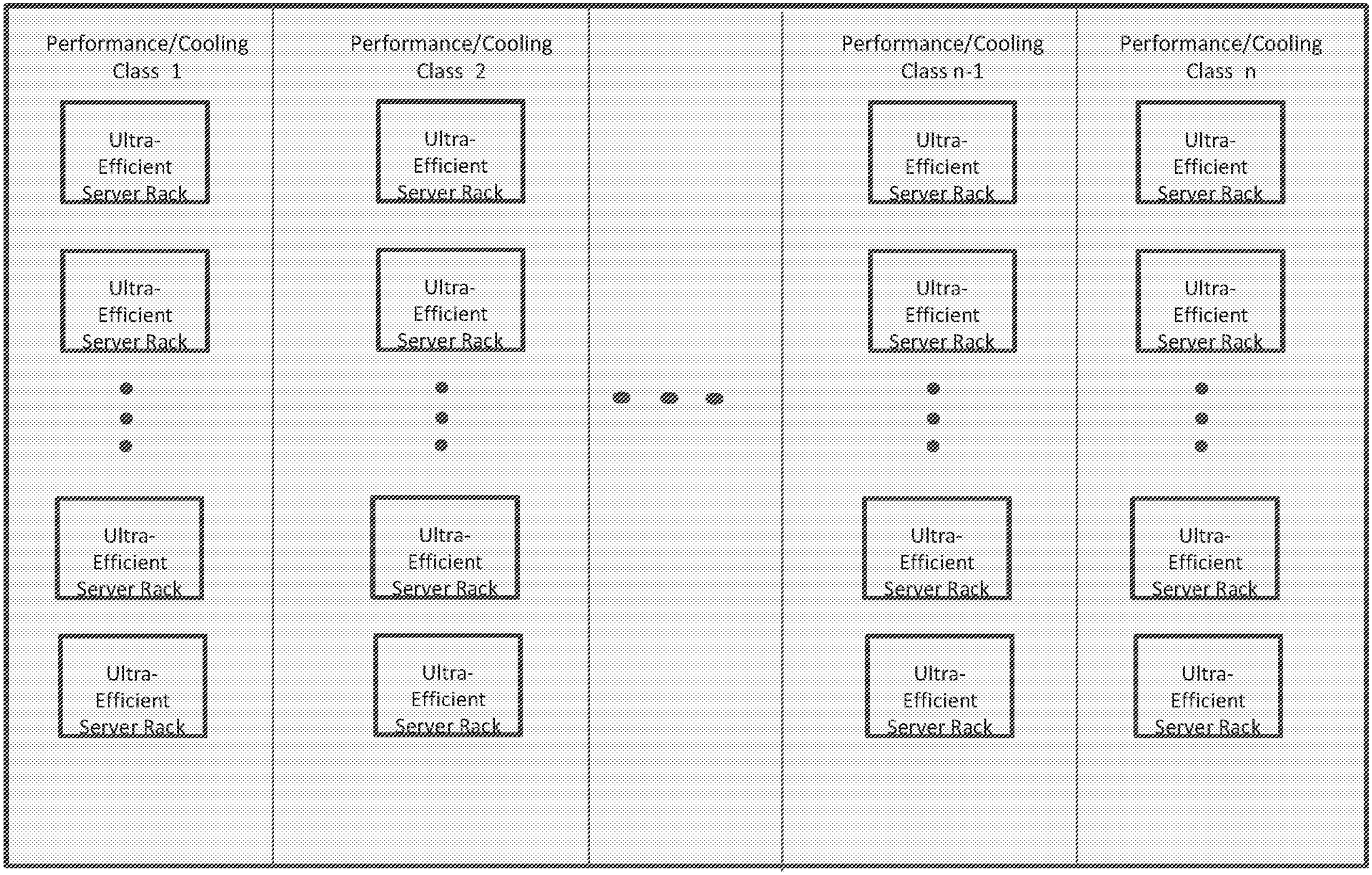


Diagram of Ultra-Efficient Data Center layout showing multiple racks of servers of different performance and cooling classes allowing highly efficient cooling provisioning and application migration based on application computational requirement. Highest performance is Class-1; lowest performance is Class-n.

FIGURE 11

Electronic Patent Application Fee Transmittal

Application Number:	
Filing Date:	
Title of Invention:	LOPOCO NEXT GENERATION ULTRA-EFFICIENT DATACENTER
First Named Inventor/Applicant Name:	Andrew Sharp
Filer:	Marc Philip Schuyler
Attorney Docket Number:	2018004 / LP02

Filed as Small Entity

Filing Fees for **Provisional**

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
PROVISIONAL APPLICATION FILING FEE	2005	1	140	140
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**140**

Electronic Acknowledgement Receipt

EFS ID:	31616016
Application Number:	62622281
International Application Number:	
Confirmation Number:	9169
Title of Invention:	LOPOCO NEXT GENERATION ULTRA-EFFICIENT DATACENTER
First Named Inventor/Applicant Name:	Andrew Sharp
Customer Number:	73091
Filer:	Marc Philip Schuyler
Filer Authorized By:	
Attorney Docket Number:	2018004 / LP02
Receipt Date:	26-JAN-2018
Filing Date:	
Time Stamp:	12:56:35
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$140
RAM confirmation Number	012618INTEFSW12573301
Deposit Account	
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Provisional Cover Sheet (SB16)	Part0-SB16-LP02.pdf	1536559 fe8cb0183cf28c5aae6271f481646828b7b41d97	no	3
Warnings:					
Information:					
2	Specification	Part1-Lopoco-Provisional-170131-FIN-AsFiled.pdf	111330 5fc5b250e0155aa84e380670e36ab51ac809b59a	no	12
Warnings:					
Information:					
3	Drawings-only black and white line drawings	Part2-DRAWINGS.pdf	1765280 f63268846ab8ceceb50e2b2ab22d854c56b4d292	no	10
Warnings:					
Information:					
4	Fee Worksheet (SB06)	fee-info.pdf	29586 625cab0cadd350db76716ee14f3259a0350f5a73	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	3442755		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.