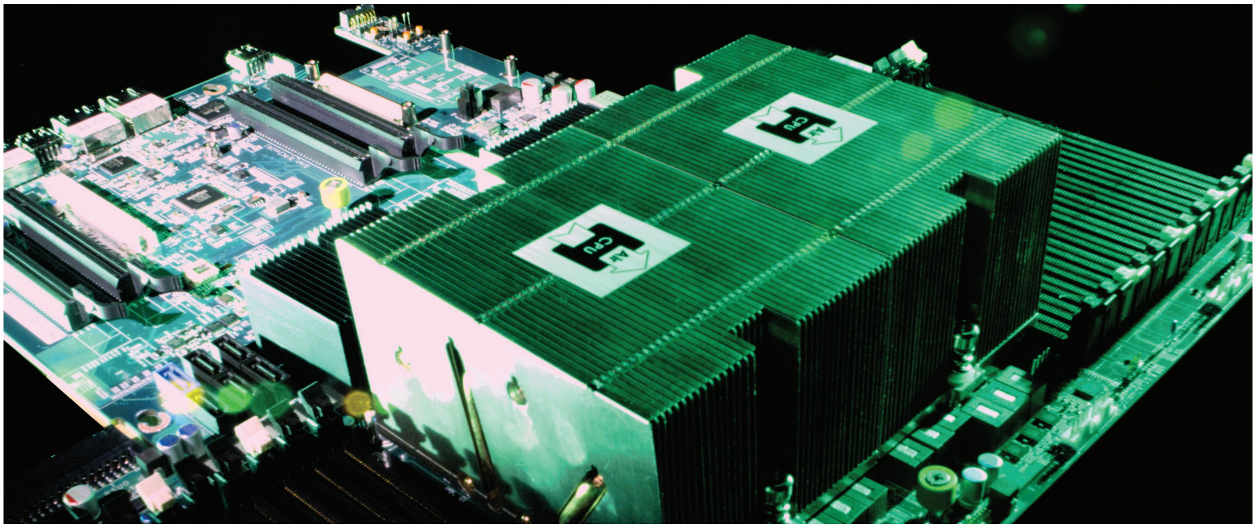

Open Compute Project

Overview

ABSTRACT

Organizations today are facing compute and storage challenges that existing infrastructure technologies cannot handle efficiently. A radical rethinking of server and datacenter design was required to address the challenges flowing from the recent surge of data, analytics, and compute power requirements. The answer is the Open Compute movement, which was formed to create open source hardware that rewrites the economic rules of the large data center. Beginning with Facebook's self-described initiative to "hack the data center", Open Compute is gaining steam, creating standards for a new class of affordable, open, and energy efficient servers. This thriving community of innovators creates specifications in the open to ensure Open Compute technologies are as scalable, efficient, and innovative as possible. This paper looks at the origins of Open Compute, the drivers of change that made it a reality and its likely impact on the data center. Using the new AMD Open 3.0 motherboard as a point of reference, the paper explores the technical and economic potential of Open Compute.

The Revolution Will Not Be Over-Provisioned

Open Compute and the Evolution of the Data Center

The world of IT hardware and infrastructure never stands still. Every year brings newer, more efficient and more powerful hardware. Nevertheless, the technology industry is at a major turning point today as people and businesses become more connected. Devices and applications are generating more compute load and data than existing infrastructure can handle efficiently and cost-effectively. IT departments are having difficulty handling this surge of demand, in some cases running up against hard limits on space, power and cooling availability. The Open Compute Project (opencompute.org) was founded in 2011 to address these challenges. Styled on the open source software community model, the Open Compute Project brings people and organizations together in a transparent effort to improve the economics of computing infrastructure without compromising on performance.

THE ORIGINS OF THE OPEN COMPUTE COMMUNITY AND STANDARDS

To understand why the Open Compute Project came into existence, and why it is getting traction in the enterprise, consider the growth in demand that's currently being placed on IT infrastructure. Storage, for example, exploded to 2.7 zettabytes (2,700 exabytes) in 2012, a 48% gain from 2011 according to eWeek.[1] Cloud computing, "Big Data" and other trends are pushing increasing volumes of data onto existing server instances. The Cisco Global Cloud Index 2011–2016 predicts that annual global data center IP traffic will reach 6.6(zb) by the end of 2016, up from 1.8(zb) in 2011, as shown in Figure 1.[2] The Cisco report also projects that the number of workloads per installed traditional server will increase from 1.5 in 2011 to 2.0 by 2016 while workloads per cloud server will double within that timeframe.[3]

Figure 1 - Global Data Center IP Traffic Growth (Source: Cisco Cloud Index: 2011-2016)

IT managers are under pressure to handle this increase in demand without overspending. And, the larger the enterprise, the more acute the challenge will be. Managers of massively scaled data centers, such as those found in financial services firms and cloud hosters, are tasked with delivering significant compute capabilities while holding down costs. At the same time, workloads continually shift in unpredictable ways. Put another way, the manager of a large-scale infrastructure must reduce hardware acquisition costs (CapEx) and operating expense (OpEx) while retaining maximum flexibility:

Lower CapEx

- Lowering capital costs while updating aging infrastructure
- Avoiding paying for overprovisioned server hardware
- Building infrastructure for evolving, sometimes unclear future workloads

1 http://www.eweek.com/c/a/Data-Storage/IBM-and-Storage-Top-Five-Trends-for-2012-568230/
2 http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html
3 ibid

Reducing OpEx

- Maximizing computational power at a low power envelope
- Simplifying server management
- Optimizing IT staff utilization
- Achieving maximum infrastructure flexibility

Facebook was facing these types of challenges when it decided to revamp its own data center. Their efforts later emerged as the core of the Open Compute Project in 2011. Following in the footsteps of Google, Amazon and others, Facebook's goal was to build the most efficient computing infrastructure at the lowest possible cost as it scaled up to handle almost one billion users worldwide. With their level of scale, Facebook could easily have had custom designed hardware manufactured for them. However, after studying the economics of their IT needs, Facebook realized that they could realize considerable cost savings by building their own solution and sharing their efforts with the industry as a whole. Facebook took the open hardware path for two basic reasons: First, it could save them a great deal of money. In addition, if their open designs gained widespread adoption, Facebook's costs would decline further as Open Compute gained economies of scale in the industry.

The Open Compute Project has followed the lead of Facebook's data center redesign team, who "honored their hacker roots[4]" by founding a "Rapidly growing community of engineers around the world whose mission is to design and enable the delivery of the most efficient server, storage and data center hardware designs for scalable computing."[5] As a community of engineers, Open Compute works on the principle that infrastructure managers know what they need and want. They can collaborate in the open to achieve the most innovative designs. The group is committed to transparency, sharing knowledge and allowing for free implementation of specifications so that everyone involved can improve on each other's work.

THE SUCCESS OF THE OPEN COMPUTE APPROACH

Facebook designed and built their software, servers and data centers from scratch. They created a datacenter in Prineville, Oregon that is able to conduct computing operations using 38% less power than earlier generations of Facebook infrastructure[6]. The facility has a Power Usage Effectiveness (PUE) of 1.09, which is extremely low when compared to the EPA best practice recommendation of 1.5 and the industry norm of 1.9[7]. PUE is a measure of how much power is used for compute versus non-compute needs. For example, a data center with a PUE of 2.0 uses one watt for non-compute needs, such as air conditioning and lights, for every watt used for actual compute. Facebook's 1.09 PUE, meaning that 91% of the energy consumed is used for compute, makes it one of the most power efficient data centers in the world, surpassing the 1.12 PUE benchmark set by Google.[8]

Overall, the Prineville facility was 24% less expensive to build and run than other state-of-the-art data centers. The success of the Prineville facility speaks to Facebook's holistic approach to creating new thresholds of computing efficiency. How did they do it? The savings are the result of several bold design decisions. Facebook got rid of the traditional inline UPS system, instead using a standby 48VDC UPS system at the server cabinet level.[9] Facebook also eliminated power distribution Units (PDUs) and now powers equipment using a 277VAC distribution panel. Facebook changed its approach to power transformation as well, removing the 480V to 208V transformation. In contrast to a traditional data center, which loses 11-17% of its energy in transformation, Facebook only loses 2% in this process in the Prineville facility.[10] LED lights powered by Ethernet keep lighting costs down. Outside air evaporative cooling replaced cooling towers and centralized chillers. Racks designed for improved airflow and air distribution ductwork further reduced energy requirements for cooling.

Given that major economic advances in technology flow from thorough cycles of change, Open Compute embraces the totality of computing. The Open Compute movement is about far more than just hardware. Open Compute specifications address a range of enterprise computing subjects that align to the goal of improving costs without reducing performance:

[4] http://www.opencompute.org/
[5] http://www.opencompute.org/about/mission-and-principles/
[6] http://www.opencompute.org/about/energy-efficiency/
[7] http://www.vertatique.com/no-one-can-agree-typical-pue
[8] http://www.google.com/about/datacenters/efficiency/internal/
[9] https://www.facebook.com/note.php?note_id=10150148003778920
[10] http://www.opencompute.org/about/energy-efficiency/

- **Motherboard design** – Open Compute's specifications for motherboards, of which AMD's Open 3.0 is an example, are designed to be minimalist, with few components that are not essential for the defined workload. Open Compute motherboards are projected to be less expensive to acquire and run than comparable hardware from established original equipment manufacturers (OEMs).
- **Open Rack** – Open Compute has published the Open Rack standard, which calls for a 21" equipment space, rather than the standard 19" currently in use. The rack column itself, however, is still 24" wide in total, keeping it consistent with existing floor deployment patterns in the industry.[11] Open Rack has a cable-free power distribution system. Servers in Open Racks don't have their own power supplies. They can plug into bus bars at the back of the rack. Open Rack also introduces a slightly larger OpenU, which is 4.5 mm higher than the standard U. The OpenU enables easier cabling and better thermal management. Open Rack is intended to facilitate use of new configurations of disk drives and other peripherals. Potentially, Open Rack can drive the disaggregation of compute components, such as CPUs, NICs, hard drives, and so forth. The result of disaggregation will be reductions in waste and a cost.
- **Storage** – The Open Compute storage project is developing a number of innovations to help data centers save money and operate storage more efficiently. For example, Cold Storage is an emerging specification to enable data that is not likely to be accessed often (or ever) to be stored on disks that can spin down when not in use. This can drive significant power savings and increase the hardware's longevity. Fusion IO, another Open Compute initiative, is a 3.2 Terabyte I/O PCI Express adapter card designed speed up storage I/O. Finally, Hyve is a new concept for a storage server that can accommodate 15 3.5" drives in a 2U form factor.
- **Data center design** – As the Facebook Prineville example shows, Open Compute is leading the way with new approaches to the overall data center design that can save on energy and operating costs.
- **Virtual IO** – Open Compute is advancing the specifications for Virtual IO, which virtualizes, or disengages, the data IO required for storage, compute and network, from physical devices. Virtual IO streamlines the provisioning of compute environments for tenants in a large data center.
- **Hardware Management** – Open Compute is developing the Open Machine Hardware Management Specification to improve the efficiency of remote machine management in large data centers. Tasks such as rebooting of remote servers, temperature monitoring, lights-out management, and so forth, become more efficient with a standardized, non-proprietary

Evolving OCP Specifications

The Open Compute (OCP) specifications that began with the Facebook project continue to evolve and mature. For instance, the OCP 1.0 servers in the Prineville facility were built for a custom rack that was designed specifically for Facebook's needs. Servers conforming to the subsequent OCP 2.0 specification were designed for the new Open Rack Specification. Now, with OCP 3.0, Open Compute-based servers, such as AMD Open 3.0, can fit on both open racks and standard data center racks. Other improvements in OCP 3.0 include expanded processor chip choices, new options for network interface cards (NICs), increased memory capacity per core, and an increased number of PCI card slots and ports. The table below summarizes the differences between the OCP 2.0 specification and OCP 3.0, as implemented by AMD:

[11] http://www.opencompute.org/2012/05/02/introducing-the-open-rack/

OCP 2.0	AMD Open 3.0	AMD Open 3.0 Advantage
Designed for Facebook	Designed for Financial Service Companies with designed in compatibility for Facebook's infrastructure.	▪ More flexibility ▪ Proven for use in financial services – extensible to comparable large workloads in other verticals.
ODM: Quanta	ODM/SI: Tyan, Quanta/Penguin, Avnet, Appro / ZT Systems	More choice and variety of vendors
Supports Opteron 6200 series up to 2 sockets, up to 115W Thermal Design Power (TDP).	Supports all Opteron 6300OPNs. Lowest TDP of 85.	Reduced power use
6.6x20" motherboard	16" x 16.5" motherboard	First and only Open Compute motherboard to meet Open Rack standard.
16 DIMMs memory support (1.25/1.35v/1.5V)	24 DIMM Memory Support (1.25/1.35v/1.5V)	More memory per core
Intel 82576 NIC	Customer choice of NIC	Greater choice and cost flexibility for using NIC of choice.
1 x16 PCI-E, 2 USB, 6 SATA II, Aspeed 1100	▪ 6 SATA, 2 USB, 2 Serial ▪ HPC = 2 LP PCI-e slots, ▪ General Purpose= 1 PCI-E and 1 LP PCI-E, ▪ Storage= 4 full height, short PCI-E ▪ 1 GbE BCM5725, and 10GbE via add-on mezzanine card	Flexibility to support different usage scenarios, e.g. PCI slots allowing for different controllers to be plugged in depending on whether it is for cloud or storage workload.
Open Hardware Management not built in.	Open hardware management	Ability to manage servers Out of Band (OOB) – powering up servers remotely, rebooting, re-imaging the OS from remote sources, monitoring temperate, fan speeds, etc.
Twin size form-factor (half size)	Adheres to standard 19" rack size. (But, available in 1, 1.5, 2, 3U versions)	Variety of form factors but also able to fit into a standard rack enclosure.
Non-recurring engineering (NRE) required for standard rack support and validation.	Standard validation (Adheres to Openrack and other Open Compute standard form factors.)	Lower cost to acquire and deploy.

A New, Economical Approach to Infrastructure

Data centers have traditionally run on a set of well understood economic rules and tradeoffs. Power and space are usually finite resources. It's costly to add either one. Different workloads, such as storage or email server operation, have usually required heterogeneous hardware with associated acquisition and operational costs. These costs include paying for multiple OEM platforms and a variety of built-in components that may or may not be necessary. Multiple server platforms translate into multiple machine images, patching programs, maintenance skill sets, spare parts inventories, and so forth. A given number of servers required a given ratio of support staff. Open Compute is starting to change this reality. Tradeoffs between expense and performance are becoming less relevant with Open Compute. It is now possible to reduce CapEx and OpEx costs, gain efficiency but not compromise on performance.

Open Compute potentially changes the essential role of the infrastructure manager. With more direct control over the design and customization of hardware, the infrastructure manager can become a smarter hardware buyer or even a hardware builder. Rather than selecting from a menu of pre-built options, the infrastructure manager can tailor order his or her own custom hardware with Open Compute.

These potentialities align with the goals of managers who want to move their facilities up the scale on the IT Capabilities Maturity Model. The model provides a framework for mangers to evaluate and implement process improvements that lead to optimal performance results. It is a way that world-class organizations can measure the effectiveness of their IT infrastructure. Open Compute helps create the economies of scale and control over infrastructure that leads to being self-supporting – the highest and most desirable level of the maturity model.

Open Compute enables IT managers to rethink their infrastructure altogether. In contrast to the traditional process of acquiring different servers for different tasks, Open Compute allows for one base machine that can be adapted across multiple workloads. For example, Open Compute, as realized by AMD Open 3.0, offers a single base motherboard design that can be used for general purpose and cloud, big data, and storage. The parts are interchangeable. The consistency in motherboard design and components is the same even if the boards are produced by multiple ODMs. It is possible to swap components from one board to another, if that is required. A common base motherboard confers numerous practical and financial advantages that extend to the entire datacenter:

- **Sourcing choices expand** with the open specifications inherent in Open Compute. In addition to the main original equipment manufacturers (OEMs), original design manufacturers (ODMs) are now producing hardware based on OCP specifications. ODMs involved in Open Compute hardware production manufacture products according to designs provided by partners such as AMD. The ODM is responsible for the low level mother board design and based on drawings and specifications from the client, the ODM lays the transistors and other elements onto the motherboard. The resulting servers are generally less expensive to acquire and operate, as the Facebook experience demonstrates.

- **Feature correct designs** that do not over-provision application-specific integrated circuits (ASICs). For example, most servers in the enterprise use shared storage. As a result, most servers do not have, nor do they need, a storage array inside the box. However, many conventional servers ship with a storage controller. Having an over-provisioned component like this adds to both acquisition and operating costs. The component has to be purchased by the OEM, installed, and supported. Markups and warranty allowances must be added to the base component cost. In production, the component draws power that is not needed for the server's specific job.

- **A vanity-free approach** avoids the costs associated with corporate branding. Buying a major OEM brand means paying for corporate overhead and marketing expenses that must get allocated to each machine that the OEM sells. In some cases, the ODM can reduce the hardware price by not adding a logo bezel for the machine. Eliminating the bezel also cuts down on fan use, as air can circulate through the machine more naturally. Using less plastic is also a nod to "green" computing. A 1U bezel requires approximately 1 pound of plastic.[12] In a 20,000 server data center, getting rid of bezels saves about 10 tons of plastic!

- **Volume purchasing discounts** become possible when ordering thousands of the same base server, regardless of workload.

- **Lower acquisition costs result** from the lack of corporate branding overheads, non-essential components. AMD projects that its cost differential compared to OEMs can be greater than 50%.

- **Interchangeable parts** simplify maintenance and cut costs. With AMD Open 3.0, the components are the same even if the boards are produced by multiple ODMs. It is possible to swap components from one board to another, if that is required. As Frank Frankovsky, VP of hardware design and supply chain at Facebook and chairman of OCP's Open Compute Foundation noted, monolithic designs don't let IT managers replace components such as obsolete CPUs. "You might be five generations behind on the processor while 80% of the other components of the server are still good," Frankovsky said. "Smarter technology refreshes not only make sense from a financial perspective; think about how much less IT equipment is going to go into the waste stream."[13] Alternatively, when server replacement is the standard procedure for a hardware problem, the reduction in SKUs based on the Open Compute lowers the investment in carrying a large inventory of replacement machines.

- **IT staff utilization typically improves** with less heterogeneous infrastructure. For example, an article about Unix environments in the online journal Softpanorama notes that 30 servers per admin is standard when running complex instances[14], while another publication cites a Microsoft datacenter with 300,000 machines that is run by 30 admins.[15] The discrepancy between the two numbers can be explained by system diversity. 300,000 uniform Microsoft Windows Servers can be run by 30 people at a FTE to

[12] http://www.amazon.com/HEWLETT-PACKARD-OPTIONS-Hewlett-Packard-SECURITY/dp/B00B2T0Z9Y/ref=sr_1_2?ie=UTF8&qid=1367160596&sr=8-2&keywords=server+bezel
[13] http://www.informationweek.com/software/informaion-management/facebook-open-compute-project-shapes-big/240146481
[14] http://www.softpanorama.org/Admin/number_of_servers_per_sysadmin.shtml
[15] http://www.datacenterknowledge.com/archives/2009/12/30/how-many-servers-can-one-admin-manage/

machine ratio of 1:10,000 while a diverse Unix environment might be at 1:30. The Softpanorama article makes a similar point, saying, “If all servers are from the same manufacturer (for example, Dell) and use single flavor of Unix and that runs standard set of applications like LAMP stack (this can be for example Web hosting providers environment) that’s one game. If the company has all flavors of Unix that exist and several types of hardware that run different application such as ERP, Oracle database, monitoring system like Openview, etc. that’s another game. Generally completely homogeneous environment permit servicing twice as much server as high diversity environment (even three or more flavors of Unix and three of more hardware vendors adds a lot to complexity).”[16]

Image management can also improve with Open Compute. Traditionally, each server SKU (e.g. storage from OEM X, general purpose server from OEM Y) had to have its own distinct server image containing operating system, drivers, and so forth. With a common baseboard built on OCP specifications, there can be a reduction in SKUs and related images to support.

- **Hardware management efficiencies** flow from the emerging Open Machine Management Specifications. Out-of-band management tasks such as remote rebooting and machine crash analysis have historically meant relying on a proprietary management platform from each OEM. Instead, hardware built to Open Compute specifications can feature a common baseboard management controller (BMC) that enables out-of-band management using Open Machine Management Specifications. This reduces or eliminates the licensing cost required for proprietary out-of-band management platforms and mitigates against the risk of vendor lock-in. AMD Open 3.0 offers support for several open source platform management tools, including Intelligent Platform Management Interface (IPMI), Data Center Manageability Interface (DMCI) and the Systems Management Architecture for Server Hardware (SMASH) standard.

New Potential for Infrastructure Flexibility

Open Compute hardware is flexible in terms of physical deployment. For example, there is now the option of placing a server’s power supply on the left or right side of the board. This makes it possible to install OCP servers on racks with either left or right side power. Or, it’s possible to have redundant power on both sides. New form factors make possible better use of space, packing more compute power into the same racks. In addition to the conventional 1, 2 and 4 U heights available for servers, Open Compute supports such alternative sizes as 1.5U, and 3U. The result can be a more efficient use of vertical server rack space.

Flexible configuration options enable infrastructure managers to optimize Open Compute hardware for specific needs without overspending. The network interface card (NIC) on the AMD Open 3.0 motherboard offers an example. The Open 3.0 motherboard comes with a Broadcom NIC. However, the board is designed to take a PCI or mezzanine card that can expand the network IO capacity from between 1 and 10 Gigabytes. The infrastructure owner has the ability to configure servers with just the right amount of network IO capacity.

THE AMD REALIZATION OF THE OPEN COMPUTE 3.0 SPECIFICATION

AMD got involved with Open Compute early in the life cycle of the community, engaging with the specifications development process and contributing its expertise. The move was the latest in a long history of embracing technological change. AMD was the first to break the 1GHz barrier in 2000 with the Athlon processor, for instance. It was also the first to offer an x86 processor that performed simultaneous 32 and 64 bit computing. AMD produced the first quad core x86 processor and the first sixteen-core processor. Additionally, AMD was the first to break the one petaflop barrier.[17] Open Compute is the next frontier in this pioneering history.

In keeping with the Open Compute principles, the AMD Open 3.0 motherboard, AMD’s realization of the Open Compute Project’s Motherboard and Server Design 3.0 specification, takes a minimalist design approach. The 16” x 16.5” board is universal, serving as the base for general, HPC and storage editions of the product. It is the first Open Compute-based product to be able to fit into a standard rack. It also fits a variety of non-proprietary 1U, 1.5U, 2U, 3U and 4U mechanical stamped sheet metal enclosures that support 2.5″ and 3.5″ disc drives. Open 3.0 supports the full AMD Opteron 6000 Series processors along with AMD’s highest memory capacity and bandwidth operation.

[16] http://www.softpanorama.org/Admin/number_of_servers_per_sysadmin.shtml
[17] http://sites.amd.com/us/Documents/Final_WhyAMDPublic.pdf

AMD has created three versions of Open 3.0, all of which share the same base motherboard:

- 1U high performance compute (HPC) server
- 2U general purpose/Cloud server
- 3U storage server

Financial Industry Origins

AMD Open 3.0 was designed in collaboration with the financial services community, defined and reviewed by over 15 leading financial services institutions. However, AMD Open 3.0 is intended for use in the enterprise datacenter regardless of industry. It was designed to meet the unusually stringent requirements that the industry has for TCO, performance and power use. Financial services infrastructure must deliver robust compute performance. Outages can be prohibitively expensive when billions of dollars are at stake at any given moment. AMD Open 3.0's design goal was to cover up to 80% of the workloads and usage models in the financial services industry, defined in the following way:

Workload	Recommended AMD Open 3.0 hardware
HPC grid computing	1U HPC server
Virtual Desktop Infrastructure	1U
Web application utility infrastructure server for Linux distributions	2U general purpose/cloud server
Cloud server IaaS & PaaS nodes	2U general purpose/cloud server
Open source SQL servers	2U general purpose/cloud server
Multi-hypervisor next generation VDI	2U general purpose/cloud server
Web application server developer platforms	1U and 2U general purpose/cloud server
Developer tools utility infrastructure servers	2U general purpose/cloud server
Big Data scale-out servers	1U HPC server
Cloud scale-out distributed object store nodes	3U storage server
Data caching	3U storage server
Messaging	2U general purpose/cloud server
Content management & search	3U storage server

For detailed information on the Open 3.0 offerings, please see the AMD motherboard specification on the Open Compute site.

The end product has been well received. George Brady, Executive Vice President, Technology Infrastructure, Fidelity Investments, commented on Open 3.0, saying, "Like the success and dynamics of open source software, open source hardware enables open access to a broad range of engineering and technology talent, and a new rate of innovation and standards creation."

A Price/Performance Breakthrough

The AMD Open3.0 servers function on an equal footing with any comparable OEM product equipped with the same AMD processor. There is no denigration of compute performance, memory, energy management or heat characteristics in the utilization of the AMD Opteron 6300 series processor with the Open Compute implementation. What there is, however, is a striking set of gains in overall acquisition cost savings and operational efficiency.

- **Price -** The acquisition price of the AMD Open 3.0- based servers is projected to be significantly lower than that of comparable OEM products. Comparing the 1U Open 3.0 edition with an HP Proliant 360 server with similar memory and related components, the AMD-based product would be priced at $4,589 compared to $10,800 for the Proliant.[18] The compute performance of the two machines would be essentially the same, however.
- **Performance –** Independent testing from Principled Technologies shows that Open 3.0-based machines

[18] Based on comparison between HP Proliant 360 and 1U DP, both with Dual Opteron G34, 24 DIMM slots, 2x GigE, 1GigE dedicated to IPMI, OCP network card expansion slot (optional), 1x PCIe2.0 x8, 1x PCIe2.0 x8 for Quanti mezzanine card, 1x PCIe2.0 x8 for OCP network mezz card.;16-Core 2.4GHz, 115W; AMD G34 Passive Heat Sink; 16GB, DDR3-1600MHz, Reg.ECC, 2-Rank, Chipkill; 8GB, DDR3-1600MHz, Reg. ECC, 2-Rank, Chipkill; 24 SATA, 2.5in Constellation, 1TB, 7200RPM, 64MB Cache, 6Gb/s; 1 ST91000640NS ConnectX®-3 EN network interface card, 10GigE, dual-port SFP+, PCIe3.0 x8 8GT/s, tall bracket, RoHS R6; 1 MCX312A-XCBT Std Power Cord, 1.2m, C14 to C13

(OCP Appro) server matches the performance characteristics in a virtual desktop infrastructure (VDI) when compared with an identically provisioned HP Proliant 360 server. As part of its independent review of the Appro and HP Proliant servers, Principled Technologies performed an analysis that compared the two servers in the use case of VDI. The VDI study compared the equipment, power and infrastructure required to run 10,000 virtual desktops. Running a test program that simulated virtual desktop usage, the study found that the OCP hardware could run 120 VMs per machine vs. 117 on the Proliant. For 10,000 desktops the OCP approach was $528,301 lower than that required for OEM in acquisition costs. That means an organization could enjoy comparable performance while spending 57% less.

- **Power use –** In addition to the existing power advantages of the AMD Opteron 6300 processor series, the Open 3.0 products offer a low overall power budget. Not being over-provisioned with non-essential components, each of which draws power, the Open 3.0 server requires less power to operate than comparable OEM products. The table below shows power budgets for the three AMD Open 3.0 products in anticipated configurations. Though estimating power use by servers is far from simple and subject to variations given different load levels, and so forth, the power levels are less than what are understood to be industry averages of approximately 1000 Watts in an established range of 500-1200 watts.[19]

		pwr	qty	SE Power	pwr	qty	Standard Power	pwr	qty	Standard Power
Processors		140	2	280.0	115	2	230.0	115	2	230
Ajax + CPU + Memory + IO				376.8			422.8			445.9
2.5” SSD		6	1	6.0						
SATA HDD	2.5” SATA				5.0	20	100.0	5.0	0	0
SAS HDD	2.5” SAS 15k RPM	16.3	0	0.0	16.3	0	0.0	16.3	35	570.5
Fans		13.2	12	158.4	4.6	8	36.5	2.76	6	148.8
System power (Twin + Ajax) - HPC Node				**541.2**	**General Purpose - 2U**		**559.2**	**Storage Node - 4U**		1,165.2

Source: *http://www.opencompute.org/wp/wp-content/uploads/2012/05/Open_Compute_Project_AMD_Motherboard_Roadrunner.pdf*

REIMAGINING IT FOR MULTIPLE INDUSTRIES

Open 3.0 is well suited to the financial sector, where self-building and custom designed hardware are common. However, the design achievements that benefit financial services can also work in numerous other industry use cases:

- **Telecom Organizations –** Telcos face a challenge in meeting variable levels of compute demand while keeping infrastructure costs down. Open Compute gives telecom infrastructure managers the ability to provision and manage massive server deployments cost-effectively.
- **Large Cloud Providers/Hosters –** The cloud computing infrastructure industry is one where units of computational power have become a commodity. The business value adds in the cloud occur from service, management, software, and so forth. The basic compute power at the core is a given – and it’s priced at commodity levels. In this economic environment, Open Compute gives cloud providers low cost, custom-designed hardware. The combination of low acquisition costs and efficient management makes it possible to be profitable in the cloud even when pricing compute power competitively.
- **Pharmaceutical/Biotech/Life Sciences Organizations –** The pharmaceutical industry imposes significant demands on its infrastructure managers, even at smaller enterprises. Pharma is data intensive, with most companies working with the expectation that years of research study data and other repositories of information will be highly available. Managing demand for giant computer workloads and storage requirements pushes infrastructure managers to find more and more cost effective compute resources. Open Compute enables pharma infrastructure mangers to squeeze more compute, data and storage capacity into existing facilities while holding steady on OpEx.

19 Tech Republic - Cost comparison toolkit: Calculating server power usage, 2013

AMD AND OPEN COMPUTE: AN ONGOING COMMITMENT

With Open 3.0, AMD is taking the first step toward the data center of the future. Going forward, many innovative products are envisioned as AMD partners with Open Compute on the realization of new, maturing standards and specifications. In this spirit, the AMD roadmap includes numerous potential breakthroughs in cost and power efficiency based on Open Compute. For example, the anticipated introduction of ARM into the Open Compute offerings promises new, far lower thresholds of power use. Or, as the Cold Storage standard solidifies, AMD expects to make hardware offerings available to support it. Going further, AMD's upcoming datacenter "fabric" based on the SeaMicro technologies (acquired by AMD in 2012) portends yet another significant transformation of infrastructure. In this case, it will be the disaggregation of the entire stack, with order-of-magnitude improvements in component flexibility. Out of this, AMD will be bringing IT managers a 6U "data center in a box". It's a never-ending journey.

RESOURCES

Visit www.amd.com/opencompute for more information.

Sponsored by AMD