Facebook's Open Compute Project: an Industry Wide Collaborative Initiative to Accelerate Enterprise Datacenter Energy Efficiency

Glenn Hansen, PECI

ABSTRACT

To date, datacenters have largely been designed based on "catalog engineering" where products are selected from established vendors. What if you could custom design and build systems and products that meet your needs in a cost-effective manner? That is what Facebook's Datacenter Development Team has done. Initial results indicate Facebook's new Prineville, Oregon datacenter is 38% more energy efficient than their leased facilities and less expensive to build than other enterprise datacenters (OCP 2012a).

After building its Prineville datacenter, inspired by the model of open source software, Facebook formed the Open Compute Project: an industry-wide initiative to reshape the infrastructure hardware industry to be more open and more focused on energy efficiency. By sharing its design and soliciting ideas from others, Facebook hopes to accelerate the rate of innovation in web-scale datacenter design.

This paper describes Facebook's design approach, the energy measures implemented at the Prineville datacenter and the Open Compute Project.

Project Background

Facebook is an Internet-based social utility serving more than 900 million active users worldwide. This scale requires a substantial amount of computer processing power, electronic data storage, and internet connection bandwidth. Even though the power efficiency of server chips has improved by a factor of 16 over the past decade, datacenters are consuming increasingly more power because products and users are growing at an even faster rate (Belady 2007, Koomey 2011). This is certainly the case for Facebook, where an expanding customer base and new features are driving the need for additional datacenters.

Facebook historically leased space in datacenters, which offered the flexibility it required as a small company. But rapid growth caused Facebook to reexamine the economics of leased space. At the same time, Facebook was working to minimize environmental impacts in its business operations. Building Facebook's own highly efficient datacenter represented an opportunity to align the company's growing infrastructure needs with sustainability goals.

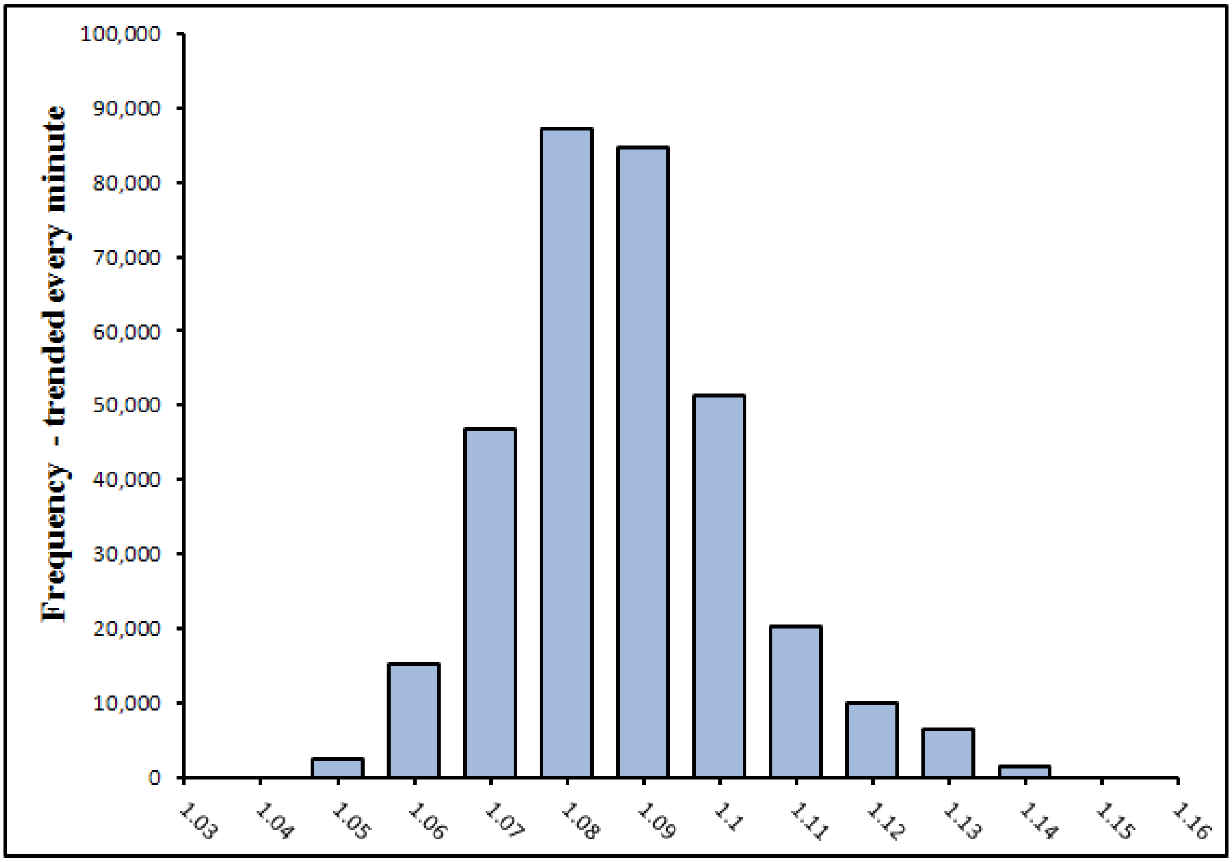
Efficiency Results

Facebook's Prineville, Oregon datacenter opened in 2011 with a 38% energy use reduction compared to other datacenters the company leases. The Prineville datacenter's power usage effectiveness (PUE),[1] an industry benchmark of datacenter efficiency, averaged 1.08 during the first five months of operation, as shown in Figure 1 below. This is a significant

[1] Power Usage Effectiveness (PUE) is defined as the Total Facility power divided by the IT Power. An ideal data center would have a 1.0 PUE where all power is used for IT Power.

improvement over Facebook's leased facilities that average 1.5 PUE. It is also far better than the industry average of 1.9 PUE, and beats the EPA criteria for "best practice" and "state of the art" datacenters (1.3 PUE and 1.2 PUE, respectively) (EPA 2007).

Figure 1. Prineville Datacenter Initial Power Usage Effectiveness (PUE) Period April 14, 2011 – September 30, 2011



Source: Facebook

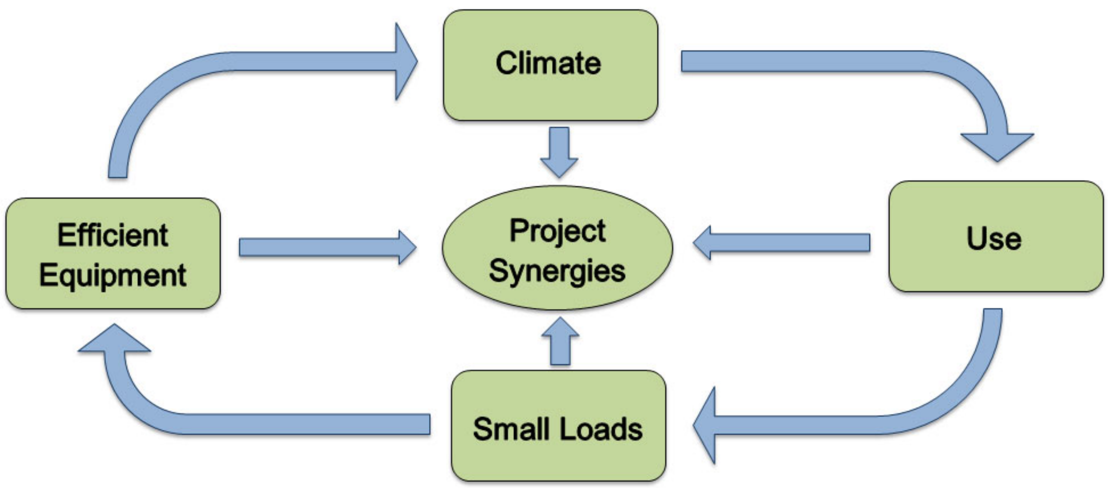
Design Approach

In 2009 Facebook directed its datacenter development team to design an energy-efficient datacenter that advanced beyond industry best practice. The team was also challenged to deliver the project below the cost of a typical large enterprise datacenter. Instead of building the datacenter using an established design and selecting from vendor pre-packaged hardware offerings, Facebook allowed the team, comprised of in-house engineers and Alfa Tech, to custom design and build its software, servers, and datacenters from the ground up. Thus, the team had considerable latitude in pursuit of Facebook's goal: "build one of the most efficient computing infrastructures at the lowest possible cost."

Integrated Design

The datacenter development team decided that the best way to achieve Facebook's goals was through an integrated design approach. In integrated design, individual design decisions are made with a whole building perspective and the design process is carefully sequenced to optimize the whole building based on: site conditions, facility use, design criteria, and equipment selection. This approach creates synergistic effects that reduce total energy use, minimize or reduces upfront construction costs, and achieve greater overall building and system performance [EDR 2002, NEEA 2012]. Integrated design is most effective when there is opportunity to reduce loads. Figure 2 below shows the relationships between key design decisions in an integrated design approach.

Figure 2: Integrated Design Elements



Source: NEEA

Development of the Prineville datacenter benefited greatly from the integrated design approach. Some of the most significant design decisions for the Prineville datacenter are summarized below.

Site Conditions. Web-scale datacenters enjoy more flexibility in selecting a site than many other types of facilities. Recognizing this, the development team conducted a careful analysis of possible sites' ambient temperature and humidity profiles. Central Oregon's high desert climate, which is dry and relatively cool, provided opportunity for maximizing economizer and evaporative "free" cooling potential.

Facility Use. Standard datacenter designs are cautious, seeking to minimize equipment downtime by tightly controlling environmental conditions. Facebook performed environmental chamber testing for their server and found it was acceptable to widen the range of environmental conditions. Temperature and humidity operating parameters were set to ASHRAE upper limits, which reduced cooling loads.

Design Criteria. The electrical power and HVAC system capacities are determined by selecting design criteria and calculating corresponding loads. The critical assumption that impacts these datacenter Facility Support (FS) systems is the estimation of the IT equipment load. Instead of sizing FS systems based on peak IT nameplate values, Facebook applied utilization rates which lowered FS capacities. Also, the extra electrical power and HVAC system capacities found in many mission critical datacenters were also reduced since the Facebook datacenter could tolerate facility shut down. This is permissible since Facebook can distribute computing functions to other datacenters that they operate.

Equipment Selection. Once loads are minimized and equipment capacities are appropriately determined, the last step is selecting the most efficient components, equipment and systems available. In Facebook's case they took this a step further and developed entirely new products that were not available from standard vendor product lines. Facebook's efficient equipment selection was comprehensive, which affected both IT equipment and FS systems.

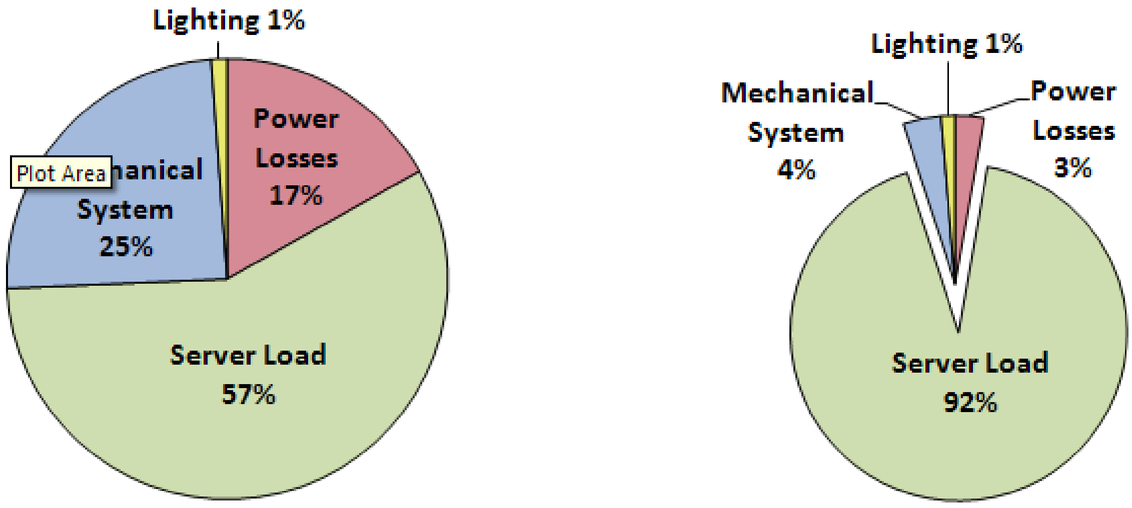
Following a successful integrated design process that minimized loads, cooling loads could now be met with an air economizer and evaporative "free" cooling system. This eliminated the significant capital cost of a chilled water plant. The server's tolerance to high temperatures also reduced the HVAC system's design airflow capacity, which reduced energy use and lowered capital cost even further. Eliminating the chilled water plant and reducing the HVAC airflow capacity in turn allowed the team to use a smaller electrical power distribution system.

Innovative Energy Efficiency Measures

This section describes the energy efficiency measures incorporated into the Prineville datacenter that were supported by the Energy Trust of Oregon's (ETO) New Buildings Program: IT - Servers, Facility - Power, Facility - HVAC Airflow and Cooling. As ETO's Program Management Contractor, PECI engineering staff worked with Facebook's team to determine energy model baseline assumptions and assess final kWh savings.

An estimate of incremental energy savings is provided for each efficiency measure. This percent savings value is based on equipment or system energy use, not on a total facility energy use basis. The greatest energy savings are from the IT servers since they represent the largest load and efficiency gain combination. The other Facility Support efficiency measures experienced greater individual percent savings reductions, but are applied to a smaller loads. Figure 3, as shown on the following page, shows the allocation of energy use for Facebook's existing typical datacenter compared to the Prineville datacenter. This can be thought as a graphical representation of PUE and indicates the Prineville datacenter uses minimal facility energy to support the IT server load. As the Prineville datacenter efficiency improves, the PUE lowers and the server load as a percentage of total facility loads increases. The size reduction of the Prineville datacenter energy use allocation figure represents the smaller facility loads achieved with the energy efficient design.

Figure 3. Datacenter Energy Use Allocation



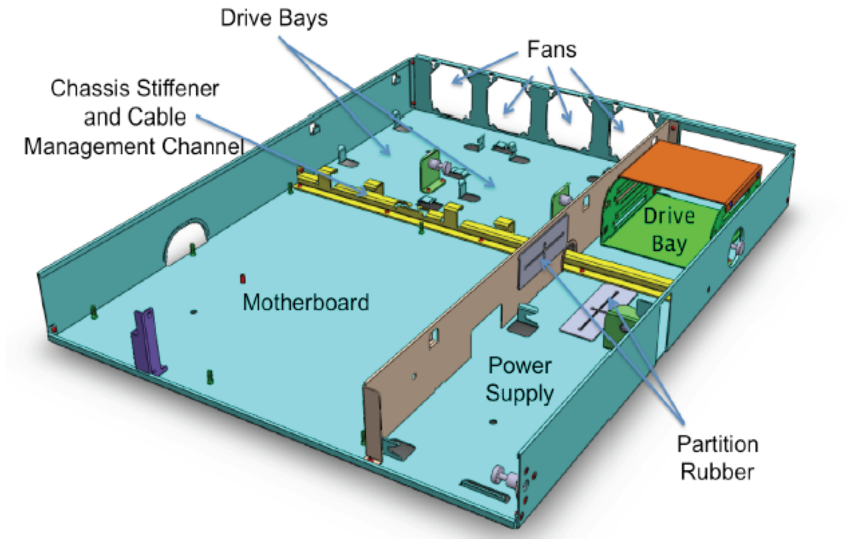
Existing Datacenter **Prineville Datacenter**

Source: Facebook

IT - Server (Estimated Savings - 38%)

A server design that tolerates high temperatures is a key project success factor. The Facebook Open Compute server had an initial design inlet temperature of 80°F and a 20°F temperature raise (or Delta-T). High temperatures were allowed by designing: 1) larger heat sinks and selection of electronic components that tolerate higher design temperatures, 2) larger chassis (1.5 U) that accommodated larger fans and no front panel to create more free area for improved: airflow, enhanced heat transfer and reduced air resistance. The server 20° F Delta-T is maintained at lower computing loads by using variable speed server fans that are adjusted based on server outlet temperature. Next, the highest efficient power electronics were selected for the central processing unit (CPU), power supply (PS), and voltage regulating modules (VRM). The dual-feed (277V and 48VDC) power supply enabled the design of the facility power distribution system. Lastly, Facebook's custom server lacks unnecessary features or components (e.g. video card, metal top to server chassis), resulting in each server weighing six pounds less than servers used at other Facebook datacenters. Figure 4, as shown on the following page, shows a 3D schematic drawing of the Open Compute server. Note this drawing accurately portrays the absence of a front panel and top cover. The server bottom sitting above is used as the lower server's top cover.

Figure 4. Open Compute Vanity-Free Server Design Elements

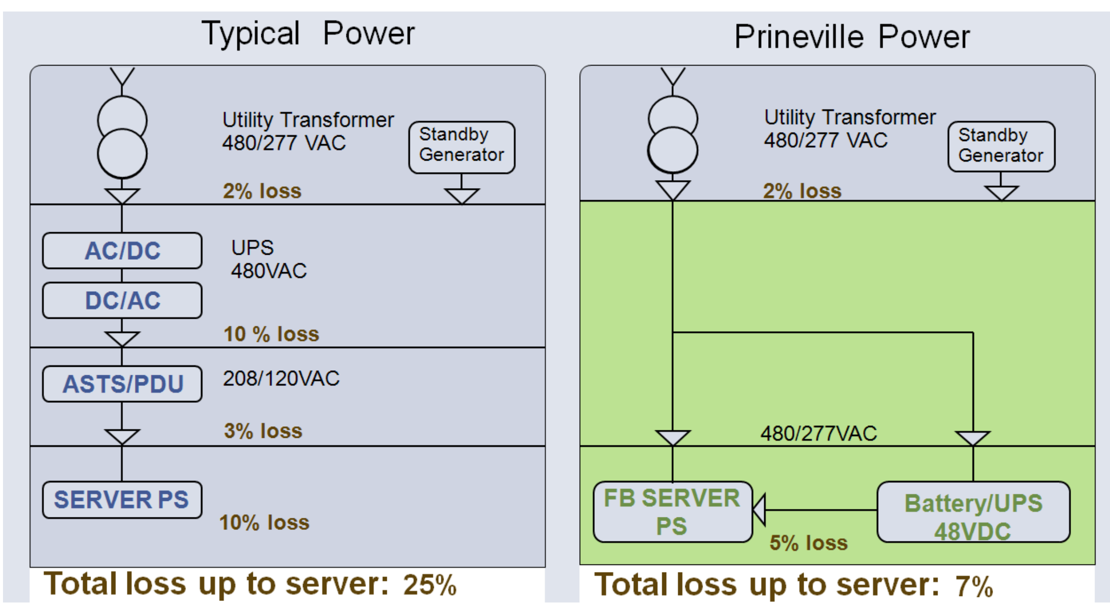


Source: Facebook

Facility - Power Distribution (Estimated Savings - 18%)

The power distribution design, as shown in Figure 5 below, eliminates the central uninterruptable power supply (UPS) and the power distribution units (PDU's) used in most datacenters. The centralized UPS has been replaced with distributed batteries. The distributed batteries are charged by a 480/277 VAC to 48VDC battery charger located in each battery cabinet. By supplying 48VDC power directly to the servers, the PDU's commonly placed between the central UPS and server rack, were eliminated. This design is efficient because it minimizes the number of power conversions and distributes higher voltage closer to the load.

Figure 5. Open Compute Power Distribution Diagram

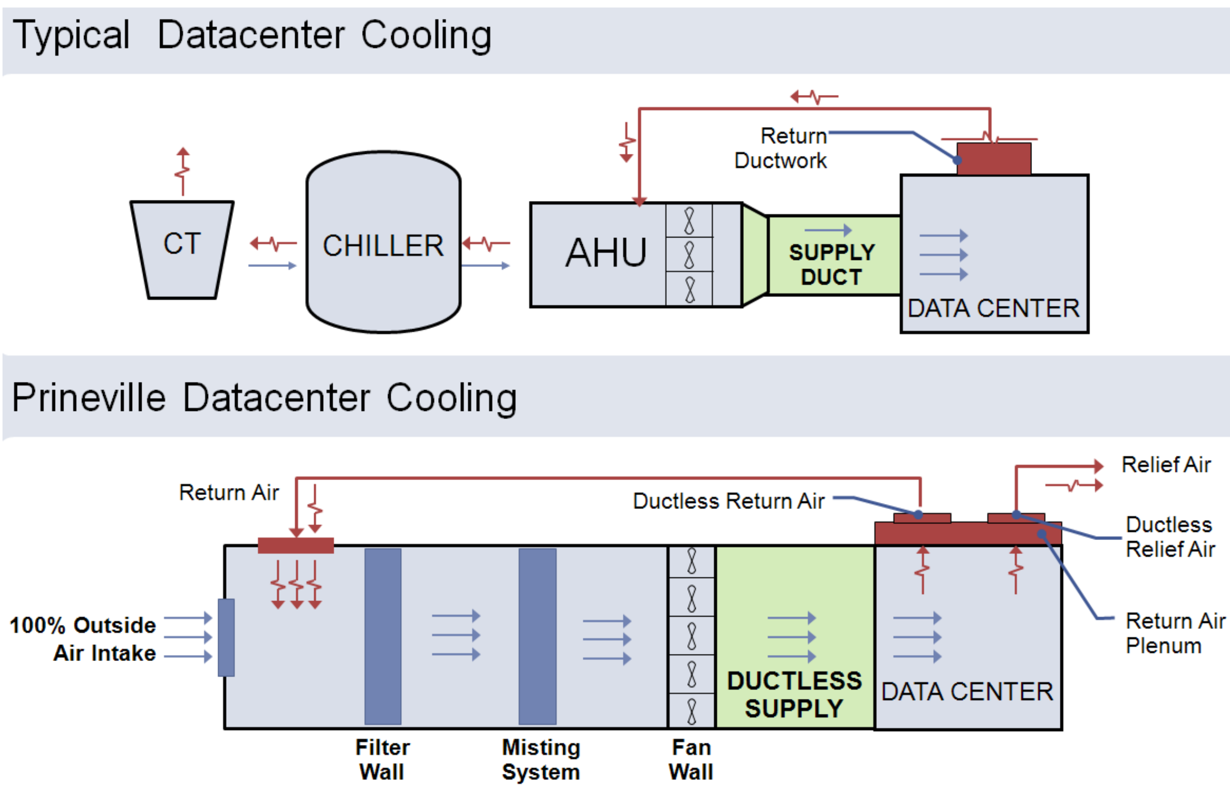


Source: Facebook

Facility - HVAC Airflow (Estimated Savings – 60%)

The HVAC airflow system, as show in Figure 6 below, saves fan energy by reducing airflow rates and lowering the air delivery pressure. This design consists of four energy saving features not commonly found in datacenters: 1) the efficient circulation of air through large built up fan rooms (penthouse) vs. smaller air handling units (AHUs) which lowers component (e.g. air filters) pressure losses, 2) formation of cool and hot aisles with full containment to keep supply air and return air separate as it flows through the server racks, 3) elimination of metal duct work and use of large air plenums that deliver air in and out of cold and hot aisles, and 4) variable airflow fans with VFDs that power direct drive supply fan/motor sets which are speed-controlled using pressure differential sensors measured across the server racks. Datacenters built without cold and hot aisle containment experience large (30-40%) bypass rates, where the supply air and return air mix together. This requires more energy since colder supply air and higher airflow rates are required to achieve the same equivalent cooling.

Figure 6. Open Compute HVAC Air Distribution Diagram



Source: Facebook

Facility - HVAC Cooling (Estimated Savings - 50%)

The HVAC cooling system, as shown in Figure 6 above, eliminates the chilled water plant or DX cooling units commonly used in datacenters and use an air economizer and evaporative "free" cooling system. This design consists of: 1) an air economizer mixing section that contains the outside air and return air dampers and 2) a direct evaporative cooling misting /

fogging system.[2] The air economizer acts as the 1st stage of cooling and uses cool outside air to flush heat out of the datacenter. The evaporative cooling system acts as the 2nd stage of cooling and runs when the outside temperature exceeds approximately 80°F (server design inlet temperature). The direct evaporative cooling system is also used during the winter economizer operation to aid moisture to the dry cool air in order to maintain minimum humidity levels.

The typical datacenter cooling schematic in Figure 6 above indicates the AHU is a return air only system and excludes an air economizer. This is representative of past design practice (Integral Group 2011). However, there has been an increased acceptance of air economizers in datacenters and many newer enterprise datacenters are being designed with air economizers. ASHRAE's updated Energy Standard 90.1-2010, which is the basis of many state energy codes, defines prescriptive requirements that datacenters incorporate either an air or a water economizer.

The Open Compute Project

Upon completing construction of its Prineville datacenter, Facebook took the unprecedented step of launching the Open Compute Project to share its design innovations with others and harness the creative talent of a community of problem-solvers. The Open Compute Project's mission is to design the most efficient datacenter equipment and facility for scalable computing - infrastructure that can be readily expanded as needed. By openly sharing ideas, specifications, and other intellectual property, the Open Compute Project aims to maximize innovation and reduce operational complexity (OCP 2012b).

In late 2011, Facebook established the Open Compute Project Foundation with board members representing a wide array of hardware industry stakeholders, including vendors and customers (datacenter users). This industry collaboration will test the value proposition that open sharing of best practices will drive continued innovation and improvement. Increased participation of datacenter users in the design process is expected to promote flexible standards that can be tailored to meet users' needs. Open Compute Project Foundation board member Jason Waxman says the Open Compute Project will "democratize and bring together much more choice in the industry for people to get efficient platforms. If you present people with an open spec, everyone can innovate."

Engaging with the Open Compute Project

There are many opportunities for individuals and companies to get involved with the Open Compute Project community. After becoming a member by signing the membership agreement, one can: join a current project and offer comments and suggestions, suggest a new project to the incubation committee, test out a prototype design, or contribute a hardware or software prototype to a project. The following groups have been established and communicate through mailing lists: general information, datacenter design, hardware management, Open Rack, storage, and virtual I/O. Interested participants can also attend the Open Compute Project summits which currently occur twice a year. Summits offer a chance for the Open Compute Project community to meet and report out on latest developments, such as, "Enabling Innovation where it Matters" from the May, 2012 San Antonio summit (OCP 2012c).

[2] Future Facebook datacenters are evaluating using a wetted fiber pad as the evaporative media.

Open Compute Project Ongoing Efforts

The Open Compute Project reports on operational findings through their web site. As real operational experience is gained, elements of the Open Compute design will be refined. Below are some findings, challenges and new opportunities that Open Compute has identified so far.

Environmental design conditions. A failed HVAC control sequence resulted in the build-up of high humidity levels in the datacenter, resulting in water condensing on computing equipment. The HVAC control sequence was fixed, but the high moisture event led Facebook to do additional work in their test lab. This additional testing led to greater knowledge on server component performance and allowed for even wider environmental design parameters. These wider environmental parameters will reduce HVAC loads further and save additional energy. For example, this experience allowed Facebook to downsize the DX mechanical cooling capacity in its North Carolina datacenter. Table-1 below indicates how the datacenter environmental design conditions have changed. The listed parameters from top to bottom are: Server Inlet Temperature Range / Upper Humidity Range / Server Delta – T.

Table - 1: Facebook Datacenter Environmental Design Conditions

Initial	Interim	Latest
65 - 80ºF inlet	65 - 85ºF inlet	65 - 85ºF inlet
65% humidity	80% humidity	90% humidity
20ºF ΔT	22ºF ΔT	22ºF ΔT

Source: Facebook

Storage efficiency. In February 2012, Facebook announced an open-source storage hardware solution was under development. A preliminary specification has been released through Open Compute, which is referenced as the Open Vault (Storage) (OCP 2012d).

HVAC controls – optimize control sequences. There is opportunity to integrate the computing management system with the HVAC automation system to map out computing load as it occurs on the datacenter floor. Mapping energy flow directly versus the current design of a disturbed array of temperature sensors will allow for a more responsive HVAC system.

Conclusion

Facebook has demonstrated that a deep examination of design requirements and exercising an integrated design process can result in significant efficiency gains at reduced capital project cost. Facebook's design embraces many new innovations which involved a high degree of critical thinking to solve technological challenges. Project success has placed Facebook as an enterprise datacenter design leader.

With an innovative design in hand, Facebook took an unprecedented step by releasing their design through the formation of the Open Compute Project. This collaborative initiative is an attempt to accelerate and expand the adoption of energy efficient innovations in datacenters. Demonstrating these innovative designs can work will increase demand for new energy efficient products, expand into the mid-tier datacenter computing market and drive down costs.

The Facebook open source hardware design approach serves as a case study for other energy intensive industries to examine and potentially replicate. Where a group of users define their requirements and suppliers innovate and bring highly efficient produces to the market. Policy groups, think tanks, and utilities are also encouraged to examine this project and think about a potential role of initiating, facilitating or funding this kind of market transformation.

References

ASHRAE. 2010. *Standard 90.1: Energy Standards for Buildings except Low-Rise Residential Buildings.* Atlanta, GA: ASHRAE.

Belady, C. L. 2007. *In the Data Center, Power and Cooling Costs More than the IT Equipment it Supports*. February 1, 2007. Electronics Cooling. http://www.electronics-cooling.com/2007/02/in-the-data-center-power-and-cooling-costs-more-than-the-it-equipment-it-supports/.

[EDR] Energy Design Resources. 2002. *Design Brief: Integrated Building Design.* San Francisco, CA: California Public Utilities Commission. http://www.energydesignresources.com/media/1711/EDR_DesignBriefs_design.pdf.

[EPA] US Environmental Protection Agency, Energy Star Program. 2007. *Report to Congress on Server and Data Center Energy Efficiency, Public Law 109-431*. Washington, DC: US Environmental Protection Agency. http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf

Facebook. 2012. Operating data and design graphics provided by Facebook.

Integral Group. November 30, 2011. Energy Efficiency Baselines for Data Centers: Statewide Customized New Construction and Customized Retrofit Incentive Programs. Oakland CA: http://www.pge.com/includes/docs/pdfs/mybusiness/energysavingsrebates/incentivesbyindustry/hightech/data_center_baseline.pdf

Koomey, Jonathan. 2011. Growth in Data center electricity use 2005 to 2010. Oakland, CA: Analytics Press. August 1, 2011. http://www.analyticspress.com/datacenters.html.

[NEEA] Northwest Energy Efficiency Alliance. 2012. *Integrated Energy Engineering and Performance Modeling into the Design Process.* Portland, OR: Northwest Energy Efficiency Alliance. http://www.betterbricks.com/graphics/assets/documents/Performance_Modeling_FINAL-WEB.pdf.

[OCP] Open Compute Project. 2012a. *Hacking Conventional Computing Infrastructure.* Palo Alto, CA: Facebook. http://opencompute.org/

[OCP] Open Compute Project. 2012b. *Mission and Principles.* Palo Alto, CA: Facebook. http://opencompute.org/about/mission-and-principles/.

[OCP] Open Compute Project. 2012c. Enabling Innovation where it Matters. http://opencompute.org/2012/05/02/enabling-innovation-where-it-matters

[OCP] Open Compute Project. 2012d. Open Vault (Storage). http://opencompute.org/projects/open-vault-storage/