building
performance
congress

High Performance Data Centers

Huston Eubank, Rocky Mountain Institute (RMI), Snowmass, CO, USA
Bernard Aebischer, Centre for Energy Policy and Economics (CEPE), Swiss Federal Institute of Technology (ETHZ), Zurich, Switzerland
Malcolm Lewis, Constructive Technologies Group, Irvine, CA
Jon Koomey, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, USA; and Senior Fellow, RMI
William Tschudi, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, USA
Peter Rumsey, Rumsey Engineers, Oakland, CA
Chris Hipp, Founder and former CTO of RLX Technologies; founding member, Server Blade Trade Association

Abstract

The internet is an increasingly important factor in our economy. However, rapid growth of "mission critical" server-farm and fiber-optic-node data centers is presenting developers and energy service providers with urgent issues. Resulting costs have broad financial and societal implications.

A recent RMI workshop resulted in recommendations that can reduce energy demand by an order of magnitude (89 percent) compared to today’s standard designs, while providing equivalent computing power, dramatically increasing reliability, and substantially lowering system capital cost. At this workshop, a broad range of high-level, technically deep industry experts dug deeply into questions of technology choice, whole-system design, system integration and business strategy, and discovered numerous significant benefits for developers and designers of data centers, manufacturers of computing equipment and components, utilities, their customers, and related industries.

The recommendations include strategies for reducing native loads, reducing loads from computer power supplies, increasing cooling efficiency, next generation cooling, efficient facility power supply, and improving operations. Charrette participants calculated that the full cost of each watt of power delivered to the server is at least US$4/W and can be as high as US$20/W. Looking at design decisions through this lens quickly focuses attention on major opportunities.

Comprehensively integrating the recommendations is critical to achieving the best results. Current thinking seldom distinguishes between component and system cost, and between first and lifecycle cost. A whole-systems approach recognizes that even though certain parts of the design may be more expensive, offsetting savings can make the whole project cost less and be of greater value.

High Performance Data Centers

While the internet continues to become an increasingly important factor in our economy, the resulting rapid growth of “mission critical” server-farm and fiber-optic-node data centers has presented developers and energy service providers with urgent issues. Resulting costs have broad financial and societal implications. Even in a sluggish economy where existing data centers can be bought for pennies on the dollar, there are tremendous opportunities to significantly improve the performance of new and existing centers.

Today, for every watt going into an internet server in a large datacenter, roughly two watts are being drawn to cool the computer and provide it with protected power. The use of massive quantities of energy to force functionality in data centers is rarely questioned. In addition to wasting energy, however, increasing power density increases the risks of system failure and highly expensive downtime. “The more power a CPU draws, the hotter it gets. The hotter a CPU gets, the more likely it will fail…unpublished empirical data from two leading vendors indicates that the failure rate of a compute node does indeed *double* with every 10° C increase.”[1]

Greater performance at the cost of energy efficiency and system reliability makes no sense; there are better solutions. As reliability is the most critical element in data center facilities, however, efficiency cannot compromise—and, indeed, must be shown to increase—reliability.

At a recent Rocky Mountain Institute (RMI) workshop, a broad range of high-level, technically deep industry experts dug deeply into questions of technology choice, whole-system design, system integration and business strategy, and discovered numerous significant benefits for developers and designers of data centers, manufacturers of computing equipment and components, utilities, their customers, and related industries.

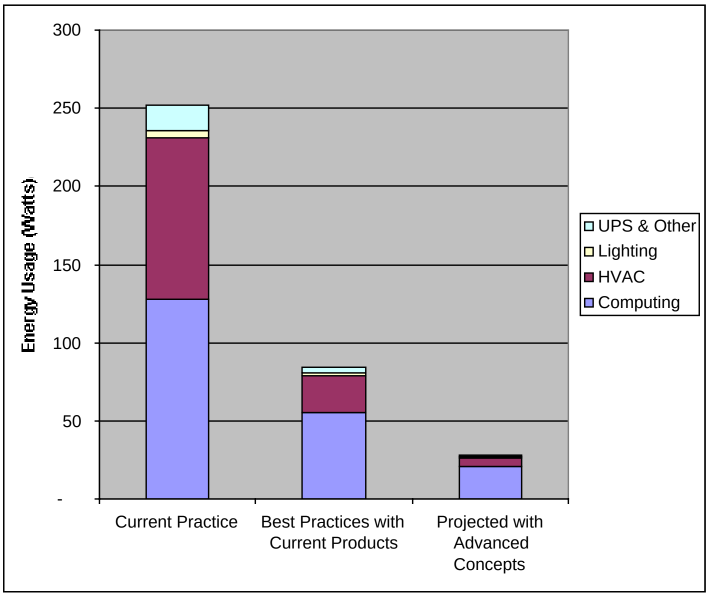

Figure 1: Overall power consumption in a data center

The RMI Low-Power Data Center Charrette produced design concepts that can reduce data center energy demand by an order of magnitude (89 percent) compared with today’s standard practice, while providing equivalent computing power, lower system capital cost, faster construction, and greater reliability. Using today’s existing technology a 66 percent reduction of power demand is feasible. While this estimate applies primarily to new sites, many of the charrette concepts are also applicable to retrofits of existing facilities.

Figure 1 shows projected overall energy consumption when the recommendations generated at the charrette are implemented. It identifies scenarios for energy consumption reduction in existing data centers, data centers using best practices with current products (currently available technology combined in smarter ways), and a projection for advanced concepts not yet “on the shelf.” Each of these scenarios provides equivalent computing power.

---

[1] “The Bladed Beowulf: A Cost-Effective Alternative to Traditional Beowulfs” by W. Feng, M. Warren, and E. Weigle (feng, msw, ehw_@lanl.gov), Advanced Computing Laboratory and the Theoretical Astrophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, undated, page 3.

Table 1: Absolute Units (Watts) Based on a Single 1U Box for Computing [2]

End Use	Current Practice	Best Practices with Current Products	Projected with Advanced Concepts	Assumptions
Computing	128	55	21	See Computing Worksheet
HVAC	103	24	5	See HVAC Worksheet; HVAC energy is computed as % of Computing Energy. But then where does HVAC efficiency show up?
Lighting	4	2	1	
UPS & Other	17	4	1	See UPS & Other Worksheet
Total:	**252**[3]	**85**	**28**	
% Energy Compared to Base Case:	**100%**	**34%**	**11%**	

The tables included in this report tabulate the various components that make up these results, and integrate the potential energy savings identified by the various working groups.

One metric for comparing the efficiency of data centers proposed at the charrette is total power to the facility divided by net power that goes into computing. Using this metric for each scenario yields the results shown in Table 2.

Table 2: Total power to the facility divided by net power that goes into computing

Total Power / Computing Power =	**With Concurrent Improvements in Computing**	**Holding Computing Power Constant**
Current Practice:	1.97	1.97
Best Practice Current Products:	1.54	0.38
Projected with Advanced Concepts:	1.36	0.13

Table 2 shows an obvious and expected “double-whammy” effect that results from best practices and advanced concepts. Because the energy required for data processing drops significantly as the efficiency of the computing devices themselves improve, the heat generated and the need to cool them decreases, often exponentially. Another factor is the issue of “oversizing.” Currently, oversizing is standard practice. It can cause the cooling-energy requirement to be as much as three times greater than is what *actually* required by empirical analysis. Thus, right-sizing represents a huge opportunity. “Best practices” assumes that variable cooling infrastructure is in place—systems and controls that adjust equipment use according to a user’s needs, as well as that equipment’s supporting infrastructure (chillers, fans, etc.).

The capital cost (new and retrofit) of these efficient systems was not estimated; however, the cooling team calculated that an annual return on investment (ROI) of 35–400 percent is achievable through improvements to HVAC systems alone.

---

[2] To make all commensurable and normalize to W/s.f. (watts per square foot) units this chart uses end-use percentages to correlate the different end-uses and ignores W/s.f.

[3] Rumsey Engineers, Inc., “Data Center Energy Benchmarking Case Study,” December 2002, Lawrence Berkeley National Laboratory. This study benchmarks the use of energy by data centers.

Integrated planning and design require that performance goals be identified *at the beginning of the process*. This allows the team to capture multiple benefits from individual features and optimize overall data center performance. It is important to recognize this unique integration process and *whole-systems* way of thinking when considering the use of the recommendations in this report. Many of them *cannot be considered in isolation* because their success and cost implications rely on the successful implementation of other recommendations.

Whole-systems thinking is a process in which the interconnections between systems are actively considered, and solutions are sought that address multiple problems at the same time. Some refer to this process as the search for “solution multipliers.”

Comprehensively integrating the recommendations is critical to achieving the best results. Design teams frequently use a value engineering approach, instead of a whole-systems approach, for system or product selection. Value engineering is a piecemeal approach that prices design elements one-by-one to find the cheapest available product. It fails to capture the benefits that can be achieved by recognizing that even though certain parts of the design may be more expensive, offsetting savings can make the whole project cost less and be of greater value.

The following paragraphs summarize the major findings of each working group at the charrette.

Native Loads

To achieve the greatest possible energy savings in a data center, it is best to begin with an examination of the native loads, and then follow the compounding savings from these native loads “upstream” toward the power source. As *Natural Capitalism* states:

> *Saving one unit of energy furthest downstream…avoids enough compounding losses... to save about ten units of fuel, cost, and pollution back at the power plant. Those compounding savings represent significant economic and environmental leverage…[enabling] each successive component, as you go back upstream, to become smaller, simpler, and cheaper. This…means that* downstream savings merit the greatest emphasis*.*[4]

The first step in increasing efficiency is to recognize and account for the *full cost of each watt of power delivered* to the server. For data centers this value is *at least* $4/W (average U.S. commercial electricity rates are $0.07/kWh) [5], while in places like Silicon Valley, New York City, etc., where electricity typically costs $0.14/kWh, this value is *at least* $8/W. In particularly inefficient data centers the value of each watt delivered to the servers is as high as $20/W. Note that power *always* saved (continuously) is worth several times as much as power saved intermittently.

Today’s typical 1-U[6] server uses approximately 128 watts. The “Hyperserver™” concept developed at the charrette would be much smaller than current servers and would run on 21 watts to match the computing power of current practice 128-watt servers. Its design is not radically different from the typical 1U server; only its packaging differs. It achieves high levels of efficiency because designers have reexamined the equipment necessary for a server and have removed as much energy intensive equipment as possible, notably fans and power supplies. Serendipitously, much current research is centered on creating low-power processors.

---

[4] Loosely quoted from *Natural Capitalism*, by Paul Hawkin and Amory and Hunter Lovins, pg 121-122

[5] These estimates are based on 1 W x 8766 hours/year x US$0.07/kWh x 0.001 kW/W x 3 year lifetime x 2 W-input/W-delivered = US$4/W. The last factor (2 W-input/W-delivered) is the “delivery factor,” which is the ratio of total data center demand to the demand of the servers. We have used the conservative value of 2 for these calculations, but it can easily be more than 10. In data centers with a delivery factor of 10, the value for each watt of power delivered to the servers is US$20/W or more.

[6] 1.75 inches, the height of a “pizza box.”

Dynamic resource allocation can solve the problem of unused system capacity by throttling resources up and down as demand dictates based on the true costs of those services. This can involve sharing resources across computers and software systems, across organizations, and across the globe.

Table 3: Computing End-Use Energy

End Use	Current Practice	Assumptions	Best Practices with Current Products	Assumptions	Projected with Advanced Products	Assumptions
CPU	70		20	Mobile CPUs	6	VLIW, low power, optimized CPU workload.
HDD	10		10	Optimized for energy	3	Remote high-efficiency HDD.
NIC/Misc	10		10	Assume no change	5	Optimized for energy.
Power Supply	33		10	Optimized for energy	6	One 3W rack-based power supply, high-efficiency; on-board converter (3 W).
Fan	5		5	Optimized for energy	1	Fluid cooling.
Total:	**128**		**55**		**21**	x2 to achieve comparable performance, x0.5 for savings for efficient resource allocation.

Computer Power Supplies

One of the most important missing ingredients in efficient data centers today is efficient, small-power supplies. As much as half of the energy that enters a computer is wasted in the power supply. Supplying and removing this wasted energy requires significant amounts of energy and capital.

While there is no shortage of ways to cost-effectively increase their efficiency, power supplies are being optimized—to the extent that they *are* being optimized—using the wrong numbers. Most power supply and equipment designers are wrong by two or three orders of magnitude in their basis of design and the actual cost of operation. Current thinking does not distinguish between component and system cost, nor between first and lifecycle cost. At a cost of at least $4–8/W for each additional watt of server power, the data center is paying dearly for the inefficiency of the power supplies used in typical servers.

If server purchasers were charged directly for the power and cooling loads they create, they would demand more efficient units from manufacturers. If designers and manufacturers understood that every watt saved is worth dollars per watt instead of cents per watt they would build significantly more efficient devices.

Next Generation Cooling

Water can conduct at least 3,500 times as much heat as the same volume of air. As temperatures on chips continue to rise and equipment loads continue to increase in density, liquid cooling becomes increasingly necessary. The first companies to move to liquid cooling will realize huge cooling-energy savings.

Resistance to liquid cooling may be greatly reduced if the liquid is kept well away from the chips by using techniques to move heat from electronic components to liquid located off the board or even outside the rack. Changing the thermal path from convective (air) to conductive (liquid) eliminates the need for fans and minimizes the number of heat transfer steps in the thermal path. Successful implementation of liquid cooling systems requires standardization of plug and play cooling connections, locations, and methods for heat removal.

One alternative to liquid cooling is to use more real estate to reduce the *compaction*, or spatial density of heat sources, without necessarily changing HVAC technologies in a radical way.

Cooling

A 35–95 percent efficiency improvement in cooling can be achieved with a 40–400 percent *annual* return on investment (ROI)—with no decrease in reliability. Onsite cogeneration can improve reliability and increase chiller efficiency (by using waste heat) for larger data centers.

Higher levels of efficiency are achieved by more elegant and lower cost solutions, such as air-side or water-side economizers and dry cooling. These solutions rely on the cooling potential of outside air whenever possible with minimum use of vapor-compression equipment. Other high-efficiency solutions include evaporative cooling in dry climates (where data centers typically need humidification) and thermal-based cooling systems that use waste heat from onsite cogeneration to drive the heat removal process.

Equipment failure rates are three times higher at top of rack than at the bottom because that’s where the heat collects.

Instead of operating data centers in the historically mandated 55–75-°F range, improving the management of airflow and using new technology make it possible to raise the supply air temperature—as high as 70–90 °F—while increasing reliability and cooling system efficiency.

In large, densely-populated data centers, the return air may embody larger total cooling loads (sensible + latent) than the outside air. In these cases, using outside air economizers will lower peak and average cooling loads. Data centers located in cool- and dry-climate regions can use natural cooling—which is free—by employing various techniques much of the year.

Typically, data center ventilation systems are designed, installed and operated at a constant rate for 8,766 hours per year. As a result, these systems frequently introduce far more outside air—that has to be conditioned—than is required. Except for command centers, few people continuously occupy data center critical space. Evaluating and minimizing ventilation rates can return big dividends in efficiency.

Chilled water systems with a capacity greater than 200 tons can operate at a total of 0.62 kW per ton. Systems greater than 60 and less than 200 tons can operate at a total of 0.83 kW per ton. These levels of performance have been achieved on real world facilities. However, the commitment of all members of the design, construction, and development team is required to realize them.

Optimizing existing control systems can provide a 20 percent reduction in total energy use on a typical HVAC system using only near term, no cost/low cost solutions. A 30 percent reduction in total energy use is possible using VFD (capital cost improvement) plus low cost or no cost measures. One of the simplest ideas—yet a concept with multiple benefits—is to network CRAC unit controls in order to optimize and economize cooling efforts, and to allow the CRAC units to cool selected zones independently of other areas.

In the future, self-correcting, truly fault-tolerant control algorithms with automated adjustments based on measured data could remove human error and lack of human responses to data. Building automation systems (BAS) could monitor rack/chip temperatures and return air temperatures to optimize operating conditions and energy use. And dynamic management tools could deliver cooling where the data-processing load is, and/or move the data processing load where the cooling is optimal.

Table 4: HVAC energy

End Use e	Computing Energy)	Assumptions	with Current Products	Assumptions	Advanced Products	Assumptions
Heat transfer out of rack	0	included in computing data for now.	0	included in computing data for now.	0	included in computing data for now.
Heat transfer out of room	0.23	Air-Based CRACs, constant volume (2 w/cfm).	0.11	Ducted from racks to plenum, VAV, auto-re-balancing (1 w/cfm).	0.11	Fluid cooling or heat pipe to central fluid cooling system (assume 50% efficiency improvement over air).
Heat rejection	0.58	Air-cooled, DX, poor part-load performance (2 KW/Ton).	0.32	Water-cooled, chilled water, high delta T, optimized part load performance, water-side economizer (0.62 KW/Ton).	0.15	Water-fluid cooling, absorption chiller, high delta T, optimized part load performance, water-side economizer (0.4 KW/Ton).
Utilization of waste Heat	0	None.		BCHP with absorption cooling.		Use waste heat to drive absorption cooling, plus BCHP.
Total:	0.81		0.43		0.26	

Question: How to handle recursion? Best practice computing will have less heat load *and* have higher efficiency HVAC.

Table 5: Ratio of HVAC energy to computing energy (ref: Koomey et al., LBNL):	**Total HVAC**	**Air handling as % total HVAC Energy**
	0.59	0.29
	0.29	0.24
Lightly loaded	1.52	0.32
Lightly loaded	0.84	
Average:	**0.81**	**0.28**

Table 6: Lighting Energy

End Use	Practice (as % Computing Energy)	Assumptions	Best Practices with Current Products	Assumptions	Projected with Advanced Products	Assumptions
Lighting	4.0%	Over-lit, uncontrolled, in lightly-loaded data center.	1.0%	Reduced lighting levels, occupancy sensor controls; zone to only illuminate areas of data center being used; in fully-loaded data center.	0.5%	Assumed further improvements in lighting efficiency.

Facility Power Supply

The facility electrical supply system is a critical part of data center design, as it drives capital cost, operating cost, and the essential criterion of system availability.

The standard industry measure of reliability—five to six “nines”—is an incomplete measure. In data centers, even short interruptions can result in long computer downtime, data loss, and significant revenue penalties. Thus the rate of failure or mean time between failure (MTBF) could be far more important than the power supply availability or duration of outages.

It is important to note that the charrette results indicate that a data center could operate at 600V or less.

The Power Supply Team recommended an onsite AC power distribution system. The choice of AC versus DC appears to be as much a cultural as a technical partiality, however, and the group analyzed both AC and DC options.

The primary power supply should be an on-site generation system with minimum double redundancy, using the grid as backup. The recommended design eliminates 50 percent of the losses of today’s systems. More efficient than the grid, this system uses its waste heat to power a thermal-based cooling system, further reducing overall electrical demand. The synergy between the data center’s requirement for reliable, on-site power and the ability of on-site generation to simultaneously satisfy the data center’s tremendous cooling requirement is a key strategy for reducing overall power consumption.

To add capacity as the size of the data center increases (modularly), single modules can be added as necessary.

At least at present, the recommended system should be connected to the grid to ensure reliability. Ideally, unused capacity could be sold back onto the grid to keep generators running at full load, thus making them optimally efficient and shortening the payback period of the total investment. Unfortunately, the combination of power export and high-reliability operation is problematic.

An optimally cost-effective system requires *both* the reliability benefits of standby operation and the energy savings of parallel operation. Although technically possible, it is difficult under present conditions to design *both* for power export to the grid and for premium reliability by island-mode operation during grid outages. Most distribution utilities will strongly discourage such a configuration. Thus, it is more practical today to design for premium reliability by island-mode operation during grid outages, and for parallel operation under normal conditions without the capacity to export to the grid.

Table 7: UPS & Other End-Uses

End Use	Computing Energy)	Assumptions	with Current Products	Assumptions	Advanced Products	Assumptions
UPS conversion losses	13%		7%	Reduce over-sizing inefficiency.	5%	Go to different technology for conversion and storage.

Operations

There are as many opportunities to improve the performance of data centers by correcting the perverse systems governing space, power, and cost relationships as there are by improving equipment and systems. The overarching principle is to make true performance and costs transparent, and get the incentives right. Incentives must be powerful and relevant, education must be a part of all data center considerations, and disconnected sectors need to work in unison.

Agents all along the organizational chain need to measure and to pay for the costs of the computing resources that they demand. The current system of charging users only on the basis of square feet encourages higher density of use and hence energy consumption well beyond the optimum. Current real estate models (design + construction relationships, lease + incentives) generate perverse signals because they do not reflect the true cost of the capital and operating expenses necessary to deliver electricity of the requisite reliability to the server. Aligning market incentives with *desired performance* should eliminate today's perverse incentive structures. Instead of charging on a per-square-foot basis, data center developers, designers, and managers need to select from a diverse menu of interrelated incentives: per watt, per power density, per teraflop, etc.—whatever metrics are practical and efficient.

A major misconception in space-to-power density ratios is that cost per unit of computation comes down as power density increases. If properly calculated, as briefly discussed above, the cost of supplying energy can be as high as $20,000 per kilowatt. The major cost is in the infrastructure to supply the cooling and power. This leads to radically different conclusions about the economics of further technology compaction. This is mainly a cost of power *density*, so pricing per square foot and per watt can help spread the costs and power density optimally.

There are myriad disconnects between the narrow foci and missions of the individual sector specialists—real estate, facilities, finance, vendors, IT, and end users—and the data center as a whole. All individuals involved in the planning, designing, siting, construction, operation, and maintenance of data centers need to share *goals* and *information* and any "*pain"* throughout all stages of the process. One sector should not be penalized so that other sectors might be rewarded; all should share in successes and failures in terms of energy consumption.

Reliability is the most critical element in data center facilities and is the easiest to sell. Therefore, efficiency cannot compromise reliability, and success will be facilitated if efficiency is shown to increase reliability.

If people don't know what something costs and do not have to pay for it, they cannot be expected to optimize it. Thus, it is important that we develop full and disaggregated cost assessments for equipment and electricity, and give them to agents/users/customers all along the supply chain. It is also important that we develop methods to calculate life cycle cost/total cost of ownership. Using this information, private and public entities can make good decisions about computing, electricity, and other resources.

Performance-based fees provide incentives that encourage design teams to create buildings and equipment that are optimally efficient by rewarding the team for the savings they generate *from the savings they generate.* Creating standards to measure efficiency provides incentives to improve efficiency.

Gathering and benchmarking operating data is another key recommendation. Feedback on costs is essential both for operations (short run) and planning (long run). Comprehensible and useful metrics must be developed and benchmarked. A list of recommended metrics was developed at the charrette, and is developed further at this conference in a separate paper: “Energy Efficiency Indicators for Data Centers” (Aebischer et al., 2004).

Measurement and verification capabilities continue to improve rapidly while costs decline, allowing more cost effective real time monitoring and management of energy and buildings systems that increase systems performance (including energy savings) improve system reliability, and reduce mean time to failure.

Creating an independent organization to provide testing, experimentation, education, and demonstrations could produce significant improvements in cost-effective data center efficiencies. Many functions that such an organization could provide are discussed in this report. If necessary, it should be jump-started by state energy agencies that manage public-goods fees.

Conclusion

The charrette results clearly point out how quickly the value of saving one watt compounds throughout the entire computing center. We detailed a reduction of 83.5 percent in the computing equipment itself. This translated into a 94 percent reduction in all the other building system loads that support the equipment loads. This amplification illustrates how the savings in one system cascade into numerous related systems, not only saving energy but also reducing equipment size, complexity, capital cost, and causes of downtime. Additionally, simply looking at energy consumption does not measure other operational costs, such as the human costs, lost revenue from downtime and unreliable performance—not to mention the simple cost of maintaining the systems. Finally in the case of data centers, efficient design massively reduces the quantity of material resources needed to provide computing services.

References

Aebischer, B. et al., 2004: Energy Efficiency Indicators for Data Centers. International Conference on Improving Energy Efficiency in Commercial Buildings, 19 - 20 April 2004, Frankfurt, Germany

Hawkin, Lovins & Lovins, 1999: *Natural Capitalism*

Rocky Mountain Institute., 2003: Design Recommendations for High-Performance Data Centers. Report of the Integrated Design Charrette, 2-5 February 2003. Rocky Mountain Institute, Snowmass, CO http://www.rmi.org