How to Reduce Energy Consumption, Waste, and Costs while Increasing Efficiency and ROI

Microsoft recognizes the tough challenges that data center managers, industry operators, and IT businesses face today as they struggle to support their businesses in the face of budget cuts and uncertainty about the future. It's natural that environmental sustainability is taking a back seat in many companies at this time. But the fact is, being "lean and green" is good for both the business and the environment, and organizations that focus their attentions accordingly will see clear benefits. Reducing energy use and waste improves a company's bottom line, and increasing the use of recycled materials is a proven way to demonstrate good corporate citizenship to your customers, employees, and the communities you do business in.

That said, it isn't always easy to know where to begin in moving to greener and more efficient operations. With that in mind—along with Microsoft's commitment to share best practices with the rest of the data center industry—this paper presents the top ten best business practices for environmentally sustainable data centers. The items in this list were submitted by senior members of Microsoft's Global Foundation Services (GFS) Infrastructure Services team. Their backgrounds include expertise in server and chip development, data center electrical and mechanical engineering, power and cooling architecture and design, research and development, and business operations and administration.

Microsoft has followed the practices below for several years now and found that in addition to helping protect the environment, they lead to optimal use of resources and help teams stay aligned with core strategies and goals:

1. Provide incentives that support your primary goals: Incentives can help you achieve remarkable results in a relatively short period of time if you apply them properly. Take energy efficiency as an example. A broad range of technology improvements and best practices are already available that companies can use to improve efficiency in the data center. However, industry adoption for these advances has been relatively low. The main reason is that the wrong incentives are in place. For instance, data center managers are typically compensated based on uptime and not efficiency. Microsoft now provides specific incentives to reward managers for improving the efficiency of their operations, using metrics such as Power Usage Effectiveness

(PUE), which determines the energy efficiency of a data center by dividing the amount of power entering a data center by the power used to run the computer infrastructure within it.

The current global PUE average for the data centers that Microsoft owns is 1.53 and we are working aggressively to drop this average yearly PUE below 1.2 for all new data center designs by 2010. Uptime is still an important metric, but it is now being appropriately balanced against the need to improve energy efficiency.

Another outmoded incentive in the industry involves how data center hosting costs are allocated back to internal organizations. Most often these costs are allocated based on the proportion of floor space used. These incentives drive space efficiency and ultra-robust data centers, but they come at a high cost and typically are not energy efficient. Space-based allocation does not reflect the true cost of building and maintaining a data center.

Microsoft has achieved substantial efficiency gains by moving to a model that allocates costs to internal customers based on the proportion of energy their services consume. Not long after Microsoft's GFS organization implemented this change, engineers from internal groups began contacting GFS to ask how they could architect upcoming releases so they didn't use as much energy. And product groups began evaluating their server utilization data to make sure they didn't already have unused capacity before ordering more servers.

It's important to note that GFS didn't simply change its billing practices and then leave it up to the product teams to figure out how to reduce their energy use. The migration to this new process was thoughtfully rolled out with supporting data, tools, and guidance to our internal teams so they could integrate these improvements into their practices. However, without financial incentives, it is doubtful these tools would have been used to the extent and success that they were.

2. Focus on effective resource utilization: Energy efficiency is an important element in Microsoft business practices, but equally important is the effective use of resources deployed. For example, if only 50 percent of a data center's power capacity is used, then highly expensive capacity is stranded in the uninterruptible power supplies (UPSs), generators, chillers, and so on. In a typical 12 Megawatt data center this could equate to $4-8 million annually in unused capital expenditure. In addition, there is embedded energy in the unused capacity since it takes energy to manufacture the UPSs, generators, chillers, and so on. Stranding capacity will also force organizations to build additional data centers sooner than necessary. This wouldn't

happen had they fully utilized existing data center infrastructure first.

3. Use virtualization to improve server utilization and increase operational efficiency: As noted in the point above, underutilized servers are a major problem facing many data center operators. In today's budgetary climate, IT departments are being asked to improve efficiency, not only from a capital perspective, but also with regard to operational overhead. By migrating applications from physical to virtual machines and consolidating these applications onto shared physical hardware, Microsoft data centers are increasing utilization of server resources such as central processing unit (CPU), memory, and disk input/output. It is quite common to see several instances in data centers where server resources are under-utilized. Industry analysts have reported that utilization levels are often well below 20 percent. Microsoft is using technologies such as Hyper-V to increase virtualization and thus utilization year over year, which in turn helps increase the productivity per watt of our operations. GFS is also actively working on broad-based adoption of Microsoft's upcoming Windows Azure cloud operating system, which uses virtualization in its core. On Windows Azure, an application typically has multiple instances, each running a copy of all or part of the application's code. Each of these instances runs in its own virtual machine (VM). These VMs run 64-bit Windows Server 2008, and they're provided by a hypervisor that's specifically designed for use in the cloud.

One immediate benefit of virtual environments is improved operational efficiency. Microsoft operations teams can deploy and manage servers in a fraction of the time it would take to deploy the equivalent physical hardware or perform a physical configuration change. In a virtual environment, managing hardware failures without disrupting service is as simple as a click of a button or automated trigger, which rolls virtual machines from the affected physical host to a healthy host.

A server running virtualization will often need more memory to support multiple virtual machines, and there is small software overhead for virtualization. However, the overall value proposition measured in terms of work done per cost and per watt is much better than the dedicated underutilized physical server case. Key benefits of virtualization include:

• Reduction in capital expenditures
• Decrease in real estate, power, and cooling costs
• Faster time to market for new products and services
• Reduction in outage and maintenance windows

4. Drive quality up through compliance: Many data center processes are influenced by the need to meet regulatory and security requirements for availability, data integrity, and consistency. Quality and consistency are tightly linked, and can be managed through a common set of processes. Popular approaches to increasing quality are almost without exception tied to observing standards and reducing variability.

Compliance boils down to developing a policy and then operating consistently as measured against that policy. The extended value that can be offered by standardized, consistent processes that address compliance will also help you achieve higher quality benefits. Microsoft has seen many such examples as we achieved certification to the international information security standard, ISO/IEC 27001:2005. For instance through monitoring its data center systems for policy compliance, the company has exposed processes that were causing problems, and found opportunities for improvements that benefitted multiple projects.

5. Embrace change management: Poorly-planned changes to the production environment can have unexpected and sometimes disastrous results, which can spill over into the planet's environment when the impacts involve lower energy utilization and other inefficient use of resources. Changes may involve hardware, software, configuration, or process. Standardized procedures for the request, approval, coordination, and execution of changes can greatly reduce the number and severity of unplanned outages. Data center organizations should adopt and maintain repeatable, well-documented processes, where the communication of planned changes enables teams to identify risks to dependent systems and develop appropriate workarounds in advance.

Microsoft manages changes to its data center software infrastructure through a review and planning process that is based on the Information Technology Infrastructure Library (ITIL) framework. Proposed changes are reviewed prior to approval to ensure that sufficient diligence has been applied. Additionally, planning for recovery in the case of unexpected results is crucial. Rollback plans must be scrutinized to ensure that all known contingencies have been considered. When developing a change management program, it is important to consider the influences of people, processes, and technology. By employing the correct level of change management, Microsoft has increased customer satisfaction and improved service level performance without placing undue burden on its operations staff.

Other features that your change management process should include:
• Documented policies around communication and timeline requirements
• Standard templates for requesting, communicating, and reviewing changes
• Post-implementation review, including cases where things went well

6. Invest in understanding your application workload and behavior: The applications in your environment and the particulars of the traffic on your network are unique, and the better you understand them, the better positioned you'll be to make improvements. Moving forward in this regard requires hardware engineering and performance analysis expertise within your organization, so you should consider staffing up accordingly. Credible and competent in-house expertise is needed to properly evaluate new hardware, optimize your request for proposal (RFP) process for servers, experiment with new technologies, and provide meaningful feedback to your vendors. Once you start building this expertise, the first goal is to focus your team on understanding your environment, and then working with the vendor community. Make your needs known to them as early as possible. It's an approach that makes sense for any company in the data center industry that's working to increase efficiency. If you don't start with efficient servers, you're just going to pass inefficiencies down the line.

7. Right-size your server platforms to meet your application requirements: A major initiative in Microsoft data centers involves "right-sizing the platform." This can take two forms. One is where you work closely with server manufacturers to optimize their designs and remove items you don't use, such as more memory slots and input/output (I/O) slots than you need, and focus on high efficiency power supplies and advanced power management features. With the volume of servers that Microsoft purchases, most manufacturers are open to meeting these requests as well as partner with us to drive innovation into the server space to reduce resource consumption even further.

Of course not all companies purchase servers on a scale where it makes sense for manufacturers to offer customized stock-keeping units (SKUs). That's where the second kind of right-sizing comes in. It involves being disciplined about developing the exact specifications that you need servers to meet for your needs, and then not buying machines that exceed your specifications. It's often tempting to buy the latest and greatest technology, but you should only do so after you have evaluated and quantified whether the promised gains provide an acceptable return on investment (ROI).

Remember that you may not need the latest features server vendors are selling. Understand your workload and then pick the right platform. For example, Microsoft recently replaced a high-end SQL four-socket SKU with a well engineered two-socket SKU based on the latest microprocessor technology that provided higher capacity, similar performance, and much lower power. Conventional wisdom has been to buy something bigger than your current needs so you can protect your investment. But with today's rapid advances in technology, this can lead to rapid obsolescence. You may find that a better alternative is to buy for today's needs and then add more capacity as and when you need it. Also, look for opportunities to use a newer

two-socket quad-core platform to replace an older four-socket dual-core, instead of overreaching with newer, more capable four-socket platforms with four or six cores per socket. Of course, there is no single answer. Again, analyze your needs and evaluate your alternatives.

8. Evaluate and test servers for performance, power, and total cost of ownership: Microsoft's procurement philosophy is built around testing. Our hardware teams run power and performance tests on all "short list" candidate servers, and then calculate the total cost of ownership, including power usage effectiveness (PUE) for energy costs. The key is to bring the testing in-house so you can evaluate performance and other criteria in your specific environment and on your workload. It's important to not rely on benchmark data, which may not be applicable to your needs and environment.

For smaller organizations that don't have resources to do their own evaluation and testing, SPECpower_ssj2008 (the industry-standard SPEC benchmark that evaluates the power and performance characteristics of volume server class computers) can be used in the absence of anything else to estimate workload power. In addition to doing its own tests, Microsoft requests this data from vendors in all of its RFPs. For more information visit the Standard Performance Evaluation Corp. web site at www.spec.org/specpower.

9. Converge on as small a number of stock-keeping units (SKUs) as you can: One of Microsoft's leading data center initiatives is the move to a server standards program where internal customers choose from a consolidated catalogue of servers. Narrowing the number of SKUs allows Microsoft to make larger volume buys, thereby cutting capital costs. But perhaps equally important, it helps reduce operational expenditures and complexities around installing and supporting a variety of models. Complementing this approach, Microsoft's server selection process is built around a 12- to 18-month cycle, so new models of servers aren't constantly being brought on board. This increases operational consistency and results in better pricing, as long-term orders are more attractive to vendors. Finally, it provides exchangeable or replaceable assets. For example, if the demand for one online application decreases while another increases, with fewer SKUs it is easier to reallocate servers as needed.

10. Take advantage of competitive bids from multiple manufacturers to foster innovation and reduce costs: Competition between manufacturers is a good thing, which Microsoft encourages through ongoing analysis of proposals from multiple companies that puts most of the weight on price, power, and performance. Microsoft develops hardware requirements, shares them with multiple manufacturers, and then works actively to develop optimized solutions. After a preliminary analysis, detailed development work continues with the company that has the best proposed design. Energy efficiency, power consumption, cost effectiveness

and application performance per watt each play key roles in hardware selection. The competition motivates manufacturers to be price competitive, drive innovation, and provide the most energy efficient, lowest total cost of ownership (TCO) solutions. In many cases, online services do not fully use the available performance. Hence, it makes sense to give more weight to price and power. Remember that power impacts not only energy consumption costs but also data center capital allocation costs.

Conclusion Beyond the business practices listed above, Microsoft's Global Foundation Services' team is taking significant steps in four areas important to environmental sustainability:

• **Using recycled resources whenever practical:** The Microsoft data center in San Antonio, Texas, for example, uses approximately eight million gallons of recycled water a month from the city's waste water system during peak cooling months.
• **Using renewable resources whenever available:** The Microsoft data center in Quincy, Washington, uses 100 percent renewable hydropower from the Columbia Basin River. The San Antonio facility obtains its electricity from a utility that derives more than 10 percent of its peak capacity from renewable energy-including wind, solar, and landfill gas. And the Dublin, Ireland data center will use outside air for cooling, thereby reducing the need for energy-intensive coolers.
• **Reducing waste in operations:** One example of Microsoft's focus on reducing waste is the company's transition to using standard shipping containers to house thousands of servers apiece. Ordering servers by the truckload eliminates the need for large amounts of packaging and other materials previously required when servers were delivered individually or in racks.
• **Taking part in industry environmental groups:** Microsoft is a co-founder and active participant in the Climate Savers Computing Initiative and The Green Grid - industry organizations focused on improving computer systems and data center energy efficiency and establishing a firm methodology for measuring Power Usage Effectiveness (PUE) accurately and consistently. In part through these groups Microsoft is advocating that the industry move to a broader range of operating environments that will enable data centers to run without chillers in many parts of the globe, thereby saving large amounts of power and capital expense.

Microsoft has also implemented a number of best practices and policy guidelines that drive its construction and facility operations worldwide. Examples include benchmarks for the design, construction, and operation of high performance green buildings, high efficiency electric motors for pumps and fans, electronic variable speed drives, electronic ballasts for fluorescent lamps, and occupancy dimmers. In short, GFS leaves no stone unturned in optimizing its use of power and natural resources.

Global Foundations Services' focus on the environment is consistent with Microsoft's commitments in this area. Most recently Microsoft announced in March 2009 that it is taking a proactive corporate approach to reduce its carbon emissions per unit of revenue by at least 30 percent compared with 2007 levels by 2012 to help reduce the company's carbon footprint. Because data centers are a significant component of Microsoft's carbon footprint, GFS will play a vital role in Microsoft's efforts to meet this corporate goal.

Contributors: *Arne Josefsberg, general manager of Infrastructure Services; Christian Belady, power and cooling architect; Dileep Bhandarkar, distinguished engineer; Daniel Costello, director of Data Center Software Services; and Jawaid Ekram, general manager of Data Center Software Services*