


Is Cloud Computing Green Computing?

By Jack Newton





loud computing and environmentalism are two of the most significant movements of the past decade. Although the two topics are often

discussed separately, the ever-increasing impact of information technology (IT) on the environment can't be ignored. A recent McKinsey report estimates that IT, taken as a whole, produces nearly 1 gigaton of emissions a year, accounting for about 2 percent of total global emissions. The world's rapidly increasing demand for computation and data storage will see such emissions increase by the year 2020 to 1.54 gigatons, or 3 percent of global emissions, twice the total output of the United Kingdom today. By some estimates the emissions associated with IT will, by 2020, be greater than those associated with the airline industry. Although the ever-increasing demand for computing cannot be controlled, we must ask if there are ways we can deliver computing in a more efficient, environmentally friendly way. Cloud computing promises such efficiencies, giving rise to the natural question: Can moving more of our computational demands to the cloud help save the environment?

Understanding the Cloud

Examining the underlying concepts and technologies that represent cloud computing will help clarify the broader effect cloud computing is likely to have on the environment.

The term "cloud computing" has been used broadly enough that its definition is accordingly nebulous, but it could be

Jack Newton is cofounder and president of Clio (www.goclio.com), a leading provider of cloud-based practice management software. He holds an M.Sc. in computer science and holds three software-related patents in the United States and the European Union. He has also spoken at CLE seminars across the United States about how cloud computing can help law practices run more effectively and efficiently. He may be reached at jack@goclio.com.

considered simply as a metaphor for the Internet (which is often, appropriately enough, depicted as a cloud in network diagrams). More specifically, cloud computing can be regarded as both an infrastructure and business model, where software and data, rather than being stored locally on your own servers and computers, are delivered to you in real time via the Internet.

Microsoft and Google, two pillars of the computing business, serve as a useful study in the transformational effect cloud computing is having on the industry. Microsoft has, traditionally, been an advocate of the "client-server" model of computing, where each business buys its own servers and workstations, purchases expensive software licenses for everything from file sharing to e-mail services to word processing, and hires IT staff to keep everything running. For the past 20 years this has been a wildly successful—and profitable—enterprise for Microsoft.

Conversely, Google, one of the pioneers of modern-day cloud computing, espouses a completely different model of computing. Rather than hosting e-mail and file servers on-premise, running database servers, and purchasing myriad software licenses, businesses simply use Google's products—such as Gmail and Google Docs—through a web browser. Google claims more than 2 million businesses have signed up for its cloud-based suite of products, and its adoption rate is only increasing. The swift consumption of Google's cloud-based services is now presenting a real threat to the Microsoft Office productivity suite, the company's most important revenue source after Windows. Microsoft's recently released Office 2010 is making an effort to combat Google by offering free cloud-based editions of Word, PowerPoint, and Excel.

Cloud Computing: A Second Industrial Revolution?

The underlying technological shift driving cloud computing in many ways parallels the forces at work during the industrial revolution. Prior to Henry Ford's pioneering work to centralize and streamline the assembly of vehicles, cars were assembled

Cloud computing is the most energy-efficient method of addressing our demand for computation and data storage.

one-by-one by craftsmen in specialized shops. The advent of the modern assembly line ushered in an era of mass-produced, well-built, and affordable vehicles. The continued iteration and refinement of the assembly line process led Ford to build the world's first factories.

Similarly, the software and servers of the "client-server" computing model are set up one-by-one by the craftsman of a new generation: "the IT guy." The inefficiencies of each business setting up an IT system that is, largely, identical to the IT systems of its peers echo the inefficiencies Ford saw in his predecessors' methods for assembling cars.

Like the industrial revolution before it, the cloud computing revolution centers around the concept of a factory. The factories are being built not by Ford, but by Google, Amazon, Apple, and Microsoft, and they produce one thing: computational power.

The innovation that underlies these cloud computing factories is a technology called "virtualization." Prior to the development of virtualization, each server tended to be dedicated to a given task. An e-mail server, for example, couldn't also act as a file server and as a database server. The reasons for this were often software- and performance-related, as the various server software products tended to interfere with each other and compete for system resources. As a result, system administrators tended to err on the side of conservatism and dedicate a server to a particular role, choosing performance and reliability over efficiency.

Virtualization allows multiple "virtual machines" to be run on a single physical server. The individual virtual machines are completely isolated from one another, providing the performance and reliability benefits of a dedicated server while taking fuller advantage of the computation resources of the host physical server. If one virtual machine crashes, the other virtual machines will remain completely unaffected.

Virtualization is nothing new—IBM pioneered the concept in the 1960s—but its potential has only been fully realized with the development of massive new data centers. These new data centers,

coupled with virtualization technology, allow for processing power to be purchased piecemeal, just as power can be purchased from the electricity grid. The transformation of computing into a utility that can respond with increased (or decreased) power to ever-changing, elastic demand is one of the defining traits of cloud computing.

The Dirty Secret Hiding in the (Server) Closet

For the typical on-premises IT system, there are more dirty secrets hiding in the server closet than you might guess. A McKinsey study found that, on average, server utilization is only 6 percent. This means that 94 percent of the time servers are sitting idle—all while they continue to draw significant amounts of power. The factors contributing to this underutilization include the "one task, one server" mentality that sees a powerful server dedicated to a single function. Additionally, servers are often more powerful than they need to be because they are purchased by businesses planning for tomorrow's needs. Lastly, and perhaps most obviously, most servers sit idle outside of the typical eight-hour workday.

Worse yet, the same study found that nearly 30 percent of servers worldwide are not used at all—skeletons of a disused IT infrastructure. System administrators, afraid of accidentally shutting down an important business function, simply leave servers running to be on the safe side.

The traditional client-server computing model and ever-decreasing costs of servers has resulted in the proliferation of servers in businesses of all sizes. Market research firm IDC predicts nearly 40 million servers will be in operation by 2011, up from 19 million in 2001.

Cloud to the Rescue

The gross inefficiency of the traditional client-server model suggests the computational needs of these 40 million servers could be met with a mere 2.4 million servers operating at 100 percent capacity. Cloud computing, thanks to virtualization, operates servers at levels closer to their theoretical maximum. More importantly, however, cloud computing providers automatically power down servers and

resources that aren't needed to meet current demand levels. A broad shift to cloud computing could, theoretically, result in a nearly 17-fold reduction in the number of servers required to meet 2011's computation demands.

Although cloud computing continues to be embraced by businesses large and small, universal adoption of cloud computing is unlikely ever to occur. Some businesses, owing either to regulatory requirements or internal policies, are unable or unwilling to leverage the public cloud computing resources offered by Amazon, Google, and others. However, these companies are taking advantage of virtualization to create their own data centers. These "private clouds" offer many of the same economic and environmental benefits of the public cloud while allowing companies to retain absolute control over their IT environment. HP is one company to have made the shift to private cloud computing; the company consolidated what used to be 85 data centers staffed by 19,000 IT workers to six cloud data centers with half the IT staff.

Cloud with a Sooty Lining?

Cloud computing promises to dramatically reduce IT emissions by way of increasing server utilization and overall efficiency, but still more can be done to decrease the ecological impact of cloud computing infrastructure.

Although the term "cloud computing" evokes an almost ethereal image of an industry with little or no environmental impact, the physical infrastructure running the cloud has a very real, and growing, environmental footprint. The scale of cloud computing data centers is almost hard to imagine: Google is rumored to operate more than 1 million servers across three dozen data centers worldwide. Microsoft recently opened a new data center near Chicago that spans more than 500,000 square feet and holds 400,000 servers; the company is adding servers at a rate of 40,000 per month in an effort to catch up with Google. The Smart 2020 report by the Climate Group estimates energy consumption of these and other cloud data centers to be 330 billion kilowatt hours per year.

Because power is the primary cost associated with operating such data centers, companies choose to locate cloud data centers near cheap, plentiful, and reliable power. One of Google's largest data centers is located near Portland, Oregon, where inexpensive hydroelectric power is drawn from the Columbia River. As an added attraction, the area has much underutilized fiber optic capacity, a hangover of the dot-com era. Nearby Quincy, Washington, has attracted nearly a half-dozen data centers for the same reasons.

While hydroelectric power provides green energy to these particular data centers, most data centers derive their power from coal and other sources of dirty energy. Microsoft's massive data center near Chicago, for example, draws a mere 1.1 percent of its energy from renewable sources, according to a recent Greenpeace report. Similar data centers run by Google, Apple, and Yahoo! meet only 1 percent to 10 percent of their power needs with renewable resources.

Although cloud computing promised to dramatically reduce the number of servers required to meet the world's computing demands, the true challenge is to shift as much of the energy consumption of the cloud to green, zero-emission sources as possible. To accomplish this, the cloud computing providers' desire for cheap energy must be aligned with environmentalists' desire for reduced emissions. Taxes on emission-producing energy sources,

whether by a cap-and-trade system or carbon tax, will accomplish this goal.

One of the primary drawbacks of green energy sources such as wind and solar farms is their requisite distance from major population centers. Transmission lines to major population centers are tremendously expensive to build, and up to 30 percent of the energy generated can be lost in transmission. But cloud data centers need not be near population centers. Locating cloud data center hubs near remote, zero-emission energy sources eliminates power line transmission losses and construction costs while providing cloud providers with several perks they factor into siting decisions: remoteness (for security reasons) and affordable land.

Toward a Green Cloud

Cloud computing is the most energy-efficient method we have to address the ever-accelerating demand for computation and data storage. Although the architecture of cloud computing is an order of magnitude more efficient than traditional on-premises server solutions, the promise of truly green cloud computing lies with locating cloud data centers near clean, renewable sources of energy. Policy decisions that encourage the consumption of green energy sources will balance cloud computing providers' need for affordable energy with the need to reduce the overall environmental impact of cloud computing and will ensure that cloud computing is, in fact, green computing. GPSOLO

Do I Need Permission from *GPSolo* to Reprint an Article?



Everything in *GPSolo* is copyrighted and may not be reproduced without permission. Requests for permission to reprint articles should be made to: www.abanet.org/reprint or made in writing to: Permissions, ABA Publishing, 321 North Clark Street, Chicago, IL 60654-7598, fax 312/988-6030