



March 21, 2008

Ethernet's Role in Greening the Datacenter

by Dan Maltbie, Founder and Chief Product Officer at Woven Systems

According to a recent report by the U.S. Environmental Protection Agency (EPA), datacenters across the country consumed 61 billion kilowatt-hours of electricity in 2006 at a cost of \$4.5 billion -- twice as much as in 2000. That's more power than is required to operate the nation's 250 million television sets. The EPA predicts that by 2012, U.S. datacenters will consume 100 billion kilowatt-hours of electricity at a cost of \$7.4 billion.

IT managers have long known they pay "twice" for electricity in a datacenter: first to power up the many racks crammed full of IT equipment, and then to cool off all those power-hungry, heat-generating systems. Power is getting expensive, too -- up to half the cost of ongoing operations, according to some reports. And being wasteful is no longer politically correct given the tons of carbon dioxide being dumped into a warming atmosphere.

Some organizations face an even more pressing problem: they simply do not have the power or cooling capacity to grow. Gartner estimates that up to half of all datacenters will encounter this problem in this year. And datacenters are scaling ever larger to meet the growing demand for new services and to satisfy the need for more online storage, particularly given the increasing use of video content. Consolidation to fewer and larger datacenters is also driving growth as operators seek greater economies of scale. Many of these new "mega" datacenters are being located, quite consciously and prudently, near sources of readily-available and low-cost power.

The IT equipment in a typical datacenter consumes about half of the total power needed, with the balance going to transforming and distributing electricity, backup power sources, air conditioning, lighting and other requirements. Of the IT equipment, servers and storage are the major consumers of power at 40 percent and 37 percent, respectively, according to Gartner. The remaining 20 percent or so can be attributed to networking, and more efficient networking helps reduce total power demand both directly and indirectly. Direct network power consumption can be reduced substantially from current levels; the power required to network a single server, for example, can be cut from more than 50 Watts in some cases to less than 10 Watts, resulting in considerable savings for a datacenter with tens of thousands of servers. Indirect reductions in power consumption result from unifying the datacenter network to facilitate more efficient server virtualization and storage consolidation.

Mixed Motives

For some, the motivation for greening-up the datacenter is a noble effort to mitigate the effects of global climate change by reducing the facility's carbon footprint. For others, the motivation has a different green incentive: capital and operational expenditures for the power and cooling that currently represent nearly half of a datacenter's total cost of ownership, according to the Uptime Institute. Over a two year period, in fact, the total cost to power and cool servers can exceed the original cost to purchase those servers.

Reducing power consumption in datacenters is all about efficiency. And the best way to improve efficiency is to eliminate waste. For example, a dedicated server with direct attached storage (DAS) has a utilization of only 10-40 percent. Virtualized servers and consolidated storage can increase utilization rates to 80 percent or more.

Virtualization and consolidation, however, create new challenges satisfying the datacenter's networking needs. At a minimum a modern, virtualized datacenter requires both a local area network and a storage area network (SAN). Many datacenters also require high-performance computing (HPC) clusters. Operating two or three networks -- typically Ethernet, Fibre Channel and InfiniBand -- is inefficient.

Perhaps the most inefficient way to accommodate these separate networking needs is the all-in-one "Swiss Army Knife"

switch. These large, expensive and power-hungry general-purpose Layer 2/Layer 3 switch/router behemoths need to be full-featured to be able to support a full spectrum of applications, including many not applicable to the datacenter. The net effect of engineering for maximum application flexibility is power consumption growing to a factor of five or more over solutions engineered specifically -- and exclusively -- for the needs of the datacenter.

Unification via Ethernet Fabrics

A much more efficient way to unify datacenter networking can be found in a new class of Ethernet switches purpose-built for the datacenter -- with nothing more, but certainly nothing less than what is required. These switches strip out the unnecessary and power-hungry functions to reduce physical footprint to one-third the rack space and the carbon footprint to one-fifth the power consumption of general-purpose switches or routers.

Ethernet fabric switches are designed for maximum data throughput, minimum latency, and minimum power consumption. To achieve these objectives, its design emphasizes sophisticated congestion management for lossless, non-blocking throughput, extremely low latency with cut-through switching, high scalability, flexible bandwidth provisioning, and, last but certainly not least, minimal power and space consumption.

Such a unified Ethernet fabric begins with top-of-rack switches that take advantage of Ethernet interfaces available for virtualized server and consolidated storage bays. For example, most servers today have a Gigabit-per-second Ethernet (GigE) interface built into the motherboard or the blade chassis, so the power consumed is trivial -- and already accounted for in the server's power supply. A top-of-rack switch with 48 GigE ports and four 10 Gbps Ethernet (10 GigE) ports, for example, can consume as little as 160 Watts, or 3 Watts per port. This single, power-efficient network can be used for both storage access and inter-process communications among servers.

The top-of-rack Ethernet switches feed a core 10 Gbps Ethernet fabric capable of unifying the LAN, SAN and HPC onto a single switched network. These core switches vary considerably in their capabilities and power consumption. Traditional Ethernet switches are rarely suitable for the core of the datacenter for a variety of reasons. The primary being the lack of scalability beyond the port density of the largest switch at the very core of the fabric. Even in smaller datacenters where 128 or 256 ports might be sufficient, the lack of effective congestion management, along with excessive latency from store-and-forward switching, normally renders traditional Ethernet switches unsuitable for SAN and HPC applications.

Ethernet switches purpose-built for the datacenter overcome these and other limitations to deliver non-blocking, lossless throughput with sub-10 microsecond latencies across a resilient, multipath Ethernet fabric supporting upwards of 4,000 10 GigE edge ports. Just as significantly, the better implementations of these Ethernet fabric switches consume as little as 17 Watts per 10 GigE edge port.

The power required to deliver non-blocking throughput on an Ethernet fabric amounts to only about 8 Watts per GigE-connected server for both storage access and inter-process communications. Impressive indeed. But also readily achieved. Here's how. To realize fully non-blocking throughput, no server or storage array can be oversubscribed. Using the top-of-rack switch described above, 40 of the GigE ports would therefore be aggregated in the four 10 GigE uplinks, leaving eight of the GigE ports unused. In this configuration, the switch's 160 total Watts of power yields a per-server consumption of just 4 Watts (160 Watts ÷ 40 servers). At the Ethernet fabric core, each ingress/egress port also requires a corresponding port on both the leaf and spine switches, resulting in a total power consumption of 51 Watts (17 Watts/port x 3 ports). A single 10 GigE ingress/egress port serving 10 servers, therefore, yields a power consumption of about 5 Watts per server (or 51 Watts ÷ 10 servers). In practice, of course, a small amount of oversubscription may be tolerated at the access layer, further increasing the power efficiency of the network.

Other Shades of Green

Not only is a unified Ethernet datacenter fabric green environmentally owing to its efficiency, it's green economically for other reasons. GigE Ethernet switching is notoriously inexpensive on a per-port basis, and could be considered "free" at the server end where the interface is already built-in. The cost of 10 GigE switches has recently dropped dramatically, as well, to less than \$2000 per-port for a datacenter-class switch. Next-generation 10 GigE adapters or NICs are also becoming increasingly more affordable.

The power savings associate with unifying datacenter networking on a purpose-built Ethernet fabric switch even helps defray the cost of the switch itself. Calculations comparing an Ethernet fabric switch with another popular datacenter switch revealed that, at 10¢ per kilowatt-hour (an average rate in the U.S.), the power savings alone over five years pays just over half the cost of the Ethernet fabric switch.

The real cost savings of Ethernet, however, derives from its remarkably low ongoing operational expenditures. The familiarity, simplicity and flexibility of Ethernet combine to make it the ideal choice for unifying datacenter networking

onto a single fabric with a single network interface per server and storage array, a single set of network cabling and patching, and a single technology for the IT staff to master and manage. When all of these cost savings are combined, the conclusion is obvious: unified networking on an Ethernet fabric is the logical next step in the greening of the modern datacenter.

Of course, not all datacenters can go all-Ethernet at this time. Some have highly specialized HPC applications requiring continued use of InfiniBand. And many operators will want to preserve their significant investments in Fibre Channel -- at least for a while longer. But for datacenters undergoing server virtualization and storage consolidation initiatives, network unification is now an integral part of going green with the advent of Ethernet fabric switches.



About the Author

Dan is a 27-year veteran of the networking industry focused on the design and development of new technology and products ranging from storage switches and IP routers to video encoding systems and Ethernet fabric switches. Before co-founding Woven Systems, Dan served as vice president of engineering at Sanera Systems, a provider of director-class Fibre Channel switches, which was later acquired by McDATA. Before Sanera, he was vice president of engineering for Caspian Networks where his team developed a telecom-class IP router. Previously, he was vice president of engineering at DiviCom, a video encoding system provider which is now part of Harmonic Lightwave. Earlier Dan held engineering and management roles of increasing importance at Hewlett Packard, where he represented the company in Ethernet IEEE 802 standards activities. Dan studied computer science and astrophysics at Indiana University.

Find this article at:

<http://www.hpcwire.com/features/17910474.html>

Check the box to include the list of links referenced in the article.

Copyright © 1994-2008 Tabor Communications, Inc. All Rights Reserved.