COMPUTE MORE, CONSUME LESS: SMART POLICIES UNLEASH DATA CENTER PRODUCTIVITY

Dr. Albert Esser, vice president of power and data center infrastructure solutions at Dell, discusses how changing your mind-set can help meet compute demand for years to come. By setting operational policies around virtualization, regular hardware refreshes, and best-practices data center design, IT leaders can help improve productivity and lower power consumption—enabling much more useful work to be performed within facilities that are already in place.

Standard measures of data center efficiency focus exclusively on how a computing infrastructure uses the power flowing into it. Given that many data centers are reaching the limits of their power and cooling capabilities, these are important metrics. However, a second and equally important consideration can also affect the balance sheets: server utilization. To unlock the true potential of the data center, enterprises must shift their focus from power consumption patterns to the overall productivity of their IT environments.

Operational policies designed to increase server utilization and advance overall performance and efficiency can lead to dramatic improvements in data center productivity without increasing power consumption. Adopting these policies enables administrators to support compute demands for growing business requirements within their existing data centers—instead of having to build a new capital-intensive facility to process the required workload.

Dr. Albert Esser, vice president of power and data center infrastructure solutions at Dell, spoke with *Dell Power Solutions* recently about a new metric to measure overall data center effectiveness, why utilization is so important to data center productivity, and how organizations can dramatically improve their data center productivity while still staying within the boundaries of limited power supplies.

WHAT IS AN ACCURATE GAUGE OF DATA CENTER EFFECTIVENESS?

One way is to look at how efficiently power is consumed. The industry-standard metrics commonly used to measure data center effectiveness are Power Usage Effectiveness (PUE) and Data Center Infrastructure Efficiency (DCiE). Both are useful metrics, but they can be misleading because they do not tell the whole story. Although PUE and DCiE effectively measure the efficiency of power utilization within a data center, they are not designed to capture the amount of actual productive work being completed. For example, a data center may be very power efficient—that is, have an excellent PUE rating—yet still not be operating anywhere near its full compute potential.

At Dell, we believe a useful measure of data center effectiveness must include actual work completed per watt, not just power efficiency. We have proposed a new metric, *data center performance per watt*, which directly measures actual work completed relative to power consumption (see Figure 1). This type of measurement captures not only power efficiency, but also the effectiveness of computing resources in doing actual work.

A useful analogy is car mileage—no one is too concerned about how many gallons a gas tank holds; what people really care about is how many miles a car gets per gallon of gas. Same with a data center: for administrators who have hit the wall in terms of

power availability, the pivotal question is, how far can their data center go on the same amount of power? Another useful metric for measuring data center productivity is *data center IT utilization*, which measures actual work completed relative to data center compute potential (see Figure 1). This measurement captures how effectively a data center is taking advantage of the compute power that is already in place.

In practice, actual processor and network utilization are very good indicators for IT gear in production; exact measures for what constitutes useful work are still being developed, but it is important to start thinking about data center productivity in this way to help determine suitable operational policies. The proposed productivity measures scale from server to facility, and can be used to guide purchase decisions. Today, industry-standard load simulators are good estimators of actual performance in production. At Dell, we believe a holistic metric will help drive efficiency and productivity and enable data centers to compute more while consuming less—leading to a reduction in servers, space requirements, and power consumption.

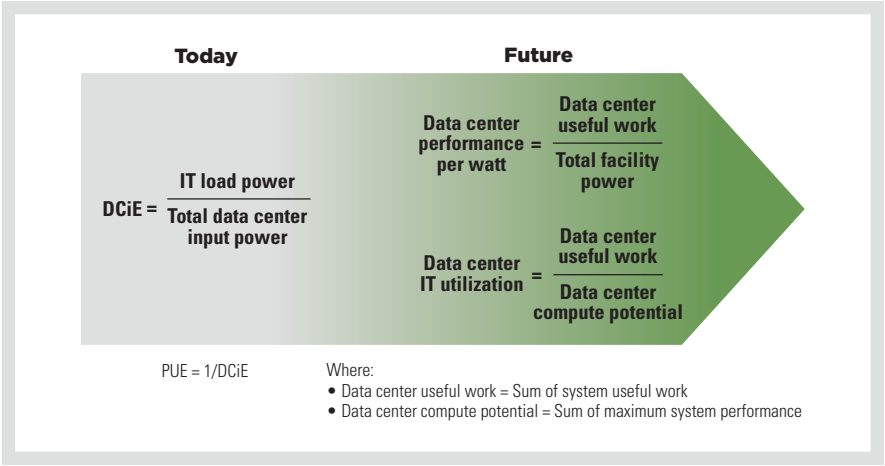


Figure 1. *Evolving data center metrics to focus on overall productivity*

HOW CAN ADMINISTRATORS ENHANCE DATA CENTER PERFORMANCE PER WATT?

There are really two fundamental ways: improving infrastructure efficiency and increasing IT productivity. Infrastructure efficiency is important, but many data centers have already achieved most of the available gains in this area. That said, there are still important infrastructure efficiency considerations and strategies that can have a significant effect on overall data center performance per watt, such as using energy-efficient equipment, optimizing data center temperature, and utilizing best-practices data center design.

The most promising advances, however, lie in improving IT productivity; that is, increasing the actual amount of computing work completed in the data center relative to the amount of power used.

CAN YOU TELL US MORE ABOUT THE IMPACT OF DATA CENTER TEMPERATURE AND BEST DESIGN PRACTICES?

Data center temperature and design can have a tremendous impact on energy efficiency and consumption. For example, data center managers have historically assumed that colder is better. Actually, the opposite is true: a data center with a higher set point for air and water temperatures can actually improve energy efficiency. In particular, our research has shown that in a typical data center, a temperature of roughly 77 degrees Fahrenheit is optimal for Dell hardware. Temperatures lower than 77 degrees typically increase the burden on the air-conditioning system, and temperatures higher than 77 degrees typically trigger an increase in server fan activity (the latter being OEM specific). At Dell, we have focused on industry-leading, low-flow/high-efficiency fan technology. Seventy-seven degrees is, in most cases, the inflection point at which the combined air-conditioning and server fan activity consume the least power.

Data center design can also have a significant impact on overall power efficiency. Fortunately, simple modifications such as moving cooling capabilities closer to equipment, using data center aisles to augment the plenum raised floors, and spacing servers optimally throughout racks to eliminate hot spots can have a major impact on cooling efficiency and power consumption.

HOW DOES HARDWARE UTILIZATION RELATE TO IT PRODUCTIVITY?

Utilization is really the key to accessing the full potential of a data center. Although the compute power of every generation of servers continues to increase dramatically—think of Moore's Law—the actual *utilization* of each subsequent generation of servers has decreased, according to a recent production server utilization study.[1] What this means is that data centers are not only *not* taking advantage of huge gains in compute power, they are actually taking less advantage of each new generation than the generation before it. This is a huge opportunity lost.

The same study found that approximately one-quarter of the servers handle the lion's share of actual work. This scenario leaves an extraordinary amount of server capacity virtually unused, and therefore wasted. These underutilized servers deliver very little work for the amount of power they draw.

Such inefficiencies are routinely incurred by legacy software applications and operating systems that simply were not designed to take advantage of dramatic increases in hardware performance. Also,

[1] Production Server Utilization Study, Dell Labs, November 2008.

the traditional one-application-per-server computing model typically leaves many servers massively underutilized. So we need to think of other ways to take full advantage of the incredible computational power available in today's generation of servers.

HOW CAN VIRTUALIZATION HELP IMPROVE DATA CENTER PRODUCTIVITY?

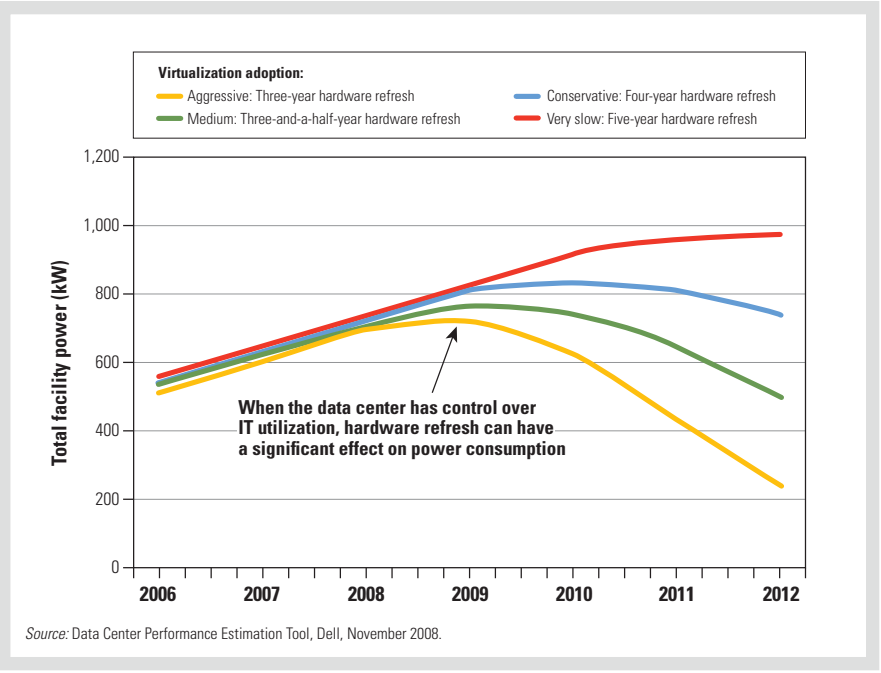
Virtualization is one of the most effective policies an organization can adopt to improve utilization and enhance overall data center productivity. In addition, virtualization can deliver dramatic increases in productivity while reducing overall power consumption and space requirements in data centers.

For example, although a fully utilized server characteristically draws more power than an underutilized one, virtualization enables organizations to reduce the overall number of physical servers required. Virtualization can offer a huge net reduction in power consumption, which enables IT organizations to increase the overall amount of work completed without increasing power consumption—leading to increased data center performance per watt.

There are other ways to improve data center productivity as well. In particular, adopting a policy of regular hardware refreshes can be critical to meeting requirements for year-over-year increases in compute demand without adding to the power requirements or footprint of the data center. A refresh cycle of three years is recommended; however, even organizations that employ four- or five-year refresh cycles can benefit from performance improvements in updated hardware. However, it is critical that old equipment is decommissioned and removed to help reduce operating and capital expenses.

HOW CAN VIRTUALIZATION AND REGULAR HARDWARE REFRESHES HELP LOWER POWER CONSUMPTION?

Policies such as virtualization and regular hardware refreshes not only help improve productivity, they can also help reduce power consumption dramatically. As Figure 2 shows, organizations that pursue aggressive virtualization strategies combined with three-year hardware refresh cycles can meet year-over-year increases in compute demand and reduce power consumption at the same time. Even organizations that adopt relatively conservative virtualization strategies and set less frequent hardware refresh cycles can significantly slow down the increase in power consumption. By adopting a few relatively simple practices, organizations can meet future compute demand within their existing data centers.



Source: Data Center Performance Estimation Tool, Dell, November 2008.

Figure 2. *Applying operational policies to help reduce power consumption in data centers*

WHAT SPECIFIC STEPS CAN ADMINISTRATORS TAKE TO IMPROVE DATA CENTER PRODUCTIVITY?

First, decommission unused equipment. All too often, equipment that has not been used in years is still drawing power and taking up space. In many cases, servers are still in place because no one knows what they do and they are afraid to turn them off.

The second step is to increase server utilization through virtualization and consolidation and to improve energy efficiency by optimizing data center temperature and best-practices data center design. To help with these steps, Dell offers a range of consulting services including virtualization services that help identify opportunities such as data center design improvements.

Once the data center has been purged of unused equipment, virtualized to maximize utilization, and optimized for energy efficiency, the next step is to refresh hardware regularly—ideally every three years. Also, as compute demand increases, organizations can take advantage of the open space created by virtualization and consolidation to add server power to the data center.

Fortunately, these gains are within immediate reach. By following a few simple steps, organizations can help increase data center productivity and help reduce—or at least dramatically decelerate the increase in—power consumption to meet compute demand for years to come.

Albert Esser, Ph.D., serves as vice president of power and data center infrastructure solutions at Dell, where he is responsible for enhancing Dell's enterprise-class IT solutions. Albert has an M.S. and a Ph.D. in Electrical Engineering from the University of Aachen.