



Gas Mileage and Server Sprawl

How server consolidation
can help your business
run on all cylinders

By Randy Watson

Back in the '80s, an American luxury car manufacturer introduced a new fuel-saving technology called “cylinder deactivation.” Conceptually, it was pretty simple: Under conditions of lower vehicle loads, the engine-management computer deactivated up to half of the car’s eight cylinders by cutting off the fuel supply and ignition spark to two or four of the cylinders, turning the V-8 into a 6- or 4-cylinder car. Gas mileage improved—a little.

New technologies are traditionally introduced on high-end vehicles, which are better able to carry the increased R&D costs in their high-margin pricing. Problem was, the geniuses at the luxury car manufacturer never bothered to think through the market for fuel-efficient, high-end, large, high-powered cars. As it turned out, this was an oxymoron, as the typical luxury car buyer didn’t care about gas mileage and certainly didn’t need to worry about the money. Having a large, powerful, high-end car is a statement, and part of that statement was never going to be about saving a little gas, let alone the

planet. The noble experiment flunked in the marketplace.

What’s that got to do with server sprawl and server consolidation? Server sprawl is kind of the opposite effect of cylinder deactivation. Servers have become so inexpensive (due to “learning-curve economics” prevalent in electronics manufacturing—think flat-screen TVs) that, for perhaps 15 years, the IT community has accepted that the easiest and fastest way to add new applications or new capacity was simply to install an additional server. So IT departments have been adding “cylinders,” most of which are only 10-

to 15-percent utilized. Conventional wisdom says adding cheap boxes is easier than trying to extract more utilization out of existing inventory.

But with server sprawl came incredibly low average server utilization, high and growing electricity consumption (terrible “gas mileage”), and galloping administrative costs.

Enter Server Consolidation

Think of server consolidation as cylinder reactivation. Like cylinder deactivation, server consolidation is also a simple concept: Find underutilized resources and put their unused capacity to work. The benefits are immediate: Do the same work, use fewer servers, use less electricity, use less floor space, use less air conditioning—all by using stuff you already own. But you can’t just smash workloads from multiple servers together onto one any more than you could open the hood of your car and implement cylinder deactivation with a monkey wrench.

Server consolidation requires several key elements. First, you must know the characteristics of each workload on the existing server farm. When do loads peak? Do they coincide? Are individual workloads likely to grow or shrink? What resources will those projected workloads require in a consolidated environment? And don’t forget to project three years into the future. Are the servers to be consolidated (and the workloads they support) all in the same time zone, or are they spread around the world? The batch window in New York is going to run right into the start-of-day in London—does server consolidation make that better or worse?

Next, as you consider combining workloads, think about combining hardware. Typically, some of the hardware is for keeps and some is at least technologically obsolete, if not physically tired or incompatible. A detailed inventory of equipment and systems software resources is mandatory. In addition to squeezing more work from fewer resources, server consolidation usually provides significant real-dollar savings in hardware maintenance and leasing costs as well as software licensing and maintenance fees for operating

systems, related systems software and applications software. Make an accurate, detailed inventory of hardware, systems software and other licensed product programs to document savings.

No consolidation assessment should even be considered without looking at high availability (HA) and disaster recovery (DR). Server consolidation makes these technologies absolutely mandatory, but at the same time helps facilitate them. Savings from consolidation often justifies HA and DR solutions, where in a server sprawl, HA was a nonstarter on economics alone.

You can model the hardware resources required for consolidated workloads using tools ranging from simple spreadsheets to highly functional software designed for the purpose. The savings will be significant; don’t short-change the process at this point.

And any discussion of server consolidation must address the human factor. Server sprawl has engendered administrator sprawl, and in these economic times, this is fundamentally a jobs issue. But a large measure of the total economic benefit to be derived from a server-consolidation project can be in the human cost replacements.

Administering a server farm of like systems is like managing a classroom of adolescents—turn your back and somebody is immediately causing trouble. A consolidated group of servers can be more like a class of one genius student. Any teacher would choose the genius, if only because there’s just one student to manage. You can maintain eye contact at all times.

Consolidated servers are easier to administer

because so much of the admin work is automated. The basic technologies underpinning server consolidation provide the technologies to simplify administration.

Slick LPARs

In the IBM Power Systems* market, the first, most important technology that facilitates server consolidation is LPAR. Originally conceived and implemented in the IBM mainframe world more than 40 years ago, partitioning allows a large server to be sliced up into multiple pseudo-computers, called partitions, each with its own share of processor power, memory and disk space. An operating-system image governs each partition. With Power Systems, these images can be a flexible mix of IBM i, AIX* and Linux*. So far, so good.

But in a real-world consolidation project, the workloads assigned to a given LPAR won’t rest at their initial settings. An LPAR must grow or shrink according to peaks, valleys and end-of-month periods. A hypervisor helps

Benefits of Consolidation

Server consolidation involves taking many, individually simple servers and putting them into a large, more complex environment that’s easier and less expensive to manage and operate, while saving electricity, floor space, air conditioning, administrator labor, licensing costs and hardware-maintenance fees. In the process, you gain reliability, redundancy, economics of scale and greater peak resources should your volumes spike or your business grow.

—R.W.



the various operating-system images constantly retune themselves to the right size for the work required of a given LPAR. And wouldn't it be great if the granularity of CPUs were pretty small, like one-tenth of a processor? With IBM Power Systems servers, you've got it.

And in a really slick implementation of LPARs, the system would tune itself, taking resources from underutilized LPARs and temporarily assigning them to overworked ones to accommodate a spike. This is, of course, exactly how IBM Power Systems solutions work. And to make sure the response time on, for example, the self-service application

Fewer separate servers means a smaller footprint. This, in turn, means less electricity in and heat out.

for ordering at the corporate online store is always split-second, you'd need a way to control the LPAR that services that application so nobody else could rob resources from it. That's standard equipment with Power Systems servers.

Benefits in Spades

So at the end of the day, with LPARs implemented, controlled by a very smart hypervisor, in a virtualized environment spanning many images of different operating systems, what benefits from all of this technology will the server sprawl customer expect?

Believe it or not, one of the bigger (albeit intangible) benefits is **consistency and simplicity**. Rather than scurrying around a data center on roller skates to service operating-system requests from scores of separate server boxes, you control everything from one console, using a single, consistent interface, the Hardware Management Console.

Second is **security and redundancy**. Server consolidation pretty much means you'll be in an HA environment. In many cases, this will take the form of a mirrored data center providing full, near-real-time replication and failover capabilities.

Third, let's consider the **environmental benefits**. Fewer separate servers

means a smaller footprint. This, in turn, means less electricity in and heat out—a double benefit from the reduction in power and cooling. Our cylinder reactivation using LPAR technology can literally save fuel.

Next, let's count the ways you can **squeeze money out of the cost of a data-center infrastructure** by consolidating. Start with reducing the physical quantity of server boxes, and for every one of those you can eliminate, you're sending that many fewer dollars to a manufacturer for hardware maintenance, operating-system licensing, licensed product programs and lease payments.

Since many ISVs base software maintenance fees on individual server serial numbers, consolidation will potentially **reduce this meaningful component of data-center cost**. Some ISVs charge by the number of processor cores installed in a server, regardless of how heavily those processors are utilized. In a consolidated environment, servers with fewer, very fast cores can be configured with exactly the amount of processor resource needed, and no more, minimizing the maintenance paid to ISVs that happen to charge on a per-core basis. Sorry, Oracle.

What's the Catch?


You must know exactly what you're doing now, and what you expect to accomplish with server consolidation. The first step to see if you're a candidate for server consolidation is a thorough assessment of your current operations, hardware installed, software deployed, workload peaks and valleys, and the growth profiles of your business and the applications you deploy. It's vital that these preliminary steps be done carefully, ideally by a team that has specific experience in measuring workloads, hardware performance, utilizations and growth trends with an eye toward consolidation. Once these basic foundations of your server

consolidation are set, tools are available to make it simple to explore scenarios to find costs savings, efficiencies and opportunities before spending any money on equipment, software or implementation.

Dumbing Down or Smartening Up?

Let's go back to the cylinder-deactivation analogy. Deactivate some cylinders of a car you've already paid for to get a little better gas mileage and worse performance. You're still dragging around the entire car, whether or not it's filled with people and stuff. You've increased the utilization of the engine by making it appear half the size it was before. In effect, you're wasting half of the engine you paid for. We could call this dumbing down the engine to meet available workload. This is kind of a "feel-good" green. Not saving much gas, but that special logo on the car announces to everyone that you're trying.

With the server-consolidation model, you take an engine you've paid for, that's working at, say, 20-percent utilization, and by consolidating workloads, you can make that engine work at 90-percent utilization or more. Same electricity, same investment, same cooling, same operator, yet you're getting four times the work out of the server for the same inputs of energy and cost. We could call this smartening up the server to take on more workload. It's as if you could transport several people in the same car, more conveniently and with the same gas mileage. And oh, by the way, this kind of capacity planning is actually green.

Do you want to sport the green logo on your car or a truly green IT infrastructure with huge bang for the buck? You pick. 



Randy Watson is president of Midrange Performance Group (MPG), an IBM business partner and ISV providing performance management and capacity planning software and services for the IBM Power Systems platforms. Randy is aptly assisted at MPG and on this article by Jim Young, VP of sales.