



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

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

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

