Virtualization of server hardware for savings of energy and data center floor space

Matthew Buckley-Golder

University of Maryland University College

ITEC620 Section 9049

Monday, December 10, 2007

Abstract

Data centers are the homes for the many, many servers that corporations need these days in order to conduct their business efficiently. The environmental and cost pressures faced by these data centers are rapidly increasing in the face of natural resource constraints as a result of not only the exhaustion of domestic land in appealing areas, but also of rapid growth in developing countries such as India and China. Hardware virtualization contributes some solutions to this problem by offering the opportunity to consolidate multiple under-utilized servers into a single physical server, thereby reducing the physical and environmental footprint of data centers by reducing the physical space required to house them, and the energy inputs needed to power them.

The issues: data center floor space and energy consumption

Data center floor space

Data centers are increasingly becoming larger structures in order to house more and more servers so that they may process the rapidly increasing amounts of data that corporations have to deal with in the modern business world. With larger structures come larger heating and cooling costs, and since data centers are usually low-density structures, larger environmental impacts are present due to the roads and utilities required to support buildings with such a large perimeter that are usually located on cheap and distant land.

If the data center floor space requirement could be reduced, these environmental impacts would be lessened, as would the variable costs associated with operating a larger facility. The data center could also feasibly be located in more expensive spaces, potentially closer to environmentally-positive services such as mass transit, which would not be tenable if required to service a distant, low-density facility. These central locations would also be closer to large pools of high quality talent, making it easier to find staff to operate the facilities.

Reducing data center floor space would therefore have at least two benefits: reduced costs, and a smaller environmental footprint. The former has clear benefits, and the latter is increasingly becoming a corporate goodwill generator, as users of products and services develop an affinity for those that they see as being more environmentally-friendly and job candidates seek to work for environmentally-sensitive companies. It is also a shield against the inevitable cost increases associated with operating low-density spaces that will arise as the infrastructure required to support them becomes more expensive due to increased resource competition from developing countries like India and China.

Energy consumption

Energy consumption is an emerging challenge in the field of data center management. A recent study by Koomey (2007) estimated that data centers account for 1.2% of total US energy consumption. While this may not sound like much, total US summer electricity demand in 2006 was 789,745 MW (Energy Information Administration, 2007), meaning that data centers represent 9,477 MW of US electricity demand. This is the equivalent of three large coal power generation stations (Ontario Power Generation, 2007), or 3,159 large wind turbines (Creative Environmental Networks, n.d.). At a time when environmental concerns such as pollution and CO_2 emissions are becoming increasingly prominent social issues, reduction in power consumption can achieve significant corporate goodwill, be an attractive feature for environmentally-conscious employees, and help to provide brand distinction.

With a significant increase in the prices of the natural resources used for power generation due to high demand in developing countries with large populations, such as India and China, it is reasonable to expect that the price of energy will continue to increase. Further, power that is not used to perform processing work is usually dissipated as heat, meaning that more efficient use of power will reduce the amount of heat emitted from systems in the data center. Since data centers must reliably maintain a low temperature, cooling infrastructure is required to remove heat from the data center, and this cooling equipment uses significant amounts of power itself. Reduction in system power consumption therefore provides cost savings beyond the direct savings. For companies that depend on large data centers to power their operations – Google, for example – power is a very significant power expense (Harris, 2007).

There are additional challenges related to data center power consumption. A high power requirement limits the site choices available for a company that is trying to decide where to locate a data center because the data center must be located at a site that can be supplied with sufficient power. Harris (2007) notes that Google must do extensive evaluation of the intricate details of the power supply cost imposed by the local power utilities when deciding where to locate its data centers. It frequently chooses to locate its data centers near cheap sources of electricity such as hydroelectric power in order to get preferential rates (Mills, 2007). However, this may not be the optimal location for a data center when other factors such as environmental impact and the location of high quality job candidates are considered. If the power consumption of data centers can be reduced, optimization of these other factors may be more flexible.

A solution: hardware virtualization

What is hardware virtualization?

Hardware virtualization provides a means of running multiple, logically-isolated virtual systems within a single system. It allows the concurrent execution of more than one operating system, each inside its own virtual system such that it thinks it has exclusive access to the hardware, and is not aware of nor affected by any of the other virtual systems running on the same single system. With hardware virtualization you can, for example, run an off-the-shelf copy of Microsoft Windows Server 2003 concurrently with an off-the-shelf copy of Microsoft Windows Server 2000 concurrently with a downloaded, open source copy of Red Hat Linux. Each operating system will run inside its own isolated virtual machine with a malfunction in one virtual machine being unable to bring down any of the other virtual machines.

Hardware virtualization virtualizes the hardware resources of a “host” system within multiple virtual “guest” systems. Each of the guest systems functions like a genuine physical system – it goes through a bootstrap process, loads the operating system, loads device drivers for the devices on the system (which are virtualized copies of those on the host system, but the guest system is not aware of this), and then executes the operating system and applications as if they had dedicated access to the entire system. Virtual machine software provides these virtualized services to the virtual machines and allows tailoring of virtual system configuration to the needs of each guest operating system.

The hardware resources of a single physical system are therefore shared amongst many virtual systems. A physical system with 8GB of memory, for example, may allocate 1GB of that memory to the host system (the system that runs the virtualization software), and 1GB of memory to each of 7 guest virtual machines running within the physical system. These virtual systems would then share access to the physical system’s other resources: CPU time, hard disk storage, optical drives, network, and USB devices. The system administrator decides which virtual machines have access to which resources, and to what extent. In some cases, devices can be explicitly dedicated to a particular virtual machine. Hardware virtualization therefore allows the consolidation of many under-utilized servers onto a single server, whose capabilities would be more fully utilized, thereby making better use of an entity’s hardware investment.

VMware: a popular product for hardware virtualization

VMware is the most popular of the currently-available hardware virtualization software products and comes in two key forms: one product runs atop an industrial operating system such as Microsoft’s Windows Server products (VMware Inc, 2007b), and the other product hosts virtual machines without the requirement of an underlying host operating system (VMware Inc, 2007a). They also offer an embedded product that resides within the system hardware and streamlines the operation of virtual machines on the host.

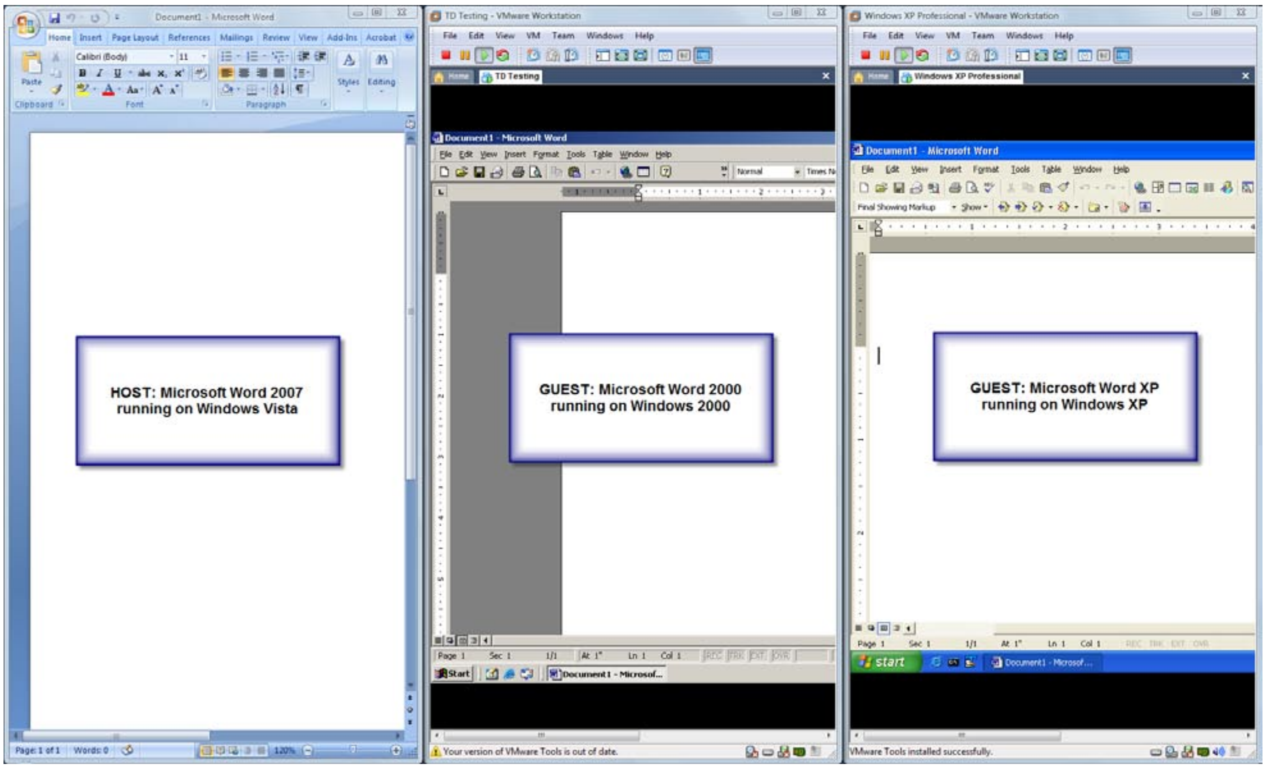

Figure 1 A virtual machine arrangement showing different applications and operating systems running on the Host and two Guests

To illustrate the concept, consider Figure 1. For demonstration purposes, this was created using VMware Workstation – a desktop product that uses the same virtualization concepts as exist in the server products. Figure 1 shows, on the left, a copy of Microsoft Word 2007 running on Microsoft Windows Vista (the physical Host system); in the middle, Microsoft Word 2000 running on Microsoft Windows 2000 Professional (a virtual Guest system); and, on the right, Microsoft Word XP running on Microsoft Windows XP Professional (another virtual Guest system). This demonstrates the concept of three different operating systems concurrently running three different versions of the same application running on a single physical system. The hard drives of the two Guest systems are contained within files on the hard drive of the Host system, and each virtual system has different specifications according to the needs of the Guest application.

While the above demonstration shows each Guest running Microsoft Word, it could just as easily be running a web server, DNS or WINS server, or a domain controller. Each Guest has its own networking resources and can be assigned its own IP address. In addition, if required, physical network controllers can be dedicated to a particular Guest system, as can hard disks.

My example uses a common desktop PC configuration: VMware claims that corporate customers are, on average, consolidating between 8 and 15 virtual machines onto one physical server; in some cases, as many as 20 or 30 virtual servers on one machine (Taylor, 2007).

Potential benefits

Cost savings: an example

The cost savings of hardware virtualization can be readily quantified.

For this example, I will focus on a small virtualization project and extrapolate its broader meanings. Based on the measurements taken by Ertl (n.d.), the following system configuration, representing that of a typical single purpose server configuration that might be used as a domain controller, departmental e-mail system, or DNS or WINS server, will be referenced:

- 2 x Intel Xeon 5160 CPU
- 24GB DDR2-667 RAM (12*2GB FB-DIMMs)
- 2 x 400GB SATA hard disks spinning
- DVD-RW drive
- Tagan 700W power supply

- multiple fans
- Linux 2.6.17

Ertl's measurements show the following power consumption measurements across various states of operation. The results are in Table 1.

CPU clock speed	idle	load1	load2	load3	load4
2000MHz	283W	290W	297W	305W	311W
2333MHz	284W	296W	309W	317W	326W
2666MHz	285W	305W	324W	335W	347W
3000MHz	286W	313W	340W	354W	368W

Table 1 Server power consumption under various loads

The "Load" results of this test are not significant to this discussion beyond what they contribute to a simple observation: this server uses between 283W and 286W of power simply to remain turned on and not performing any processing tasks. This is a baseline power consumption measurement that will be similar across all similar servers and, where different, will not vary significantly because it is simply a characteristic of modern servers (Bianchini, & Rajamony, 2004).

The implication is that significant power savings can be gained by consolidating underutilized servers: if four underutilized servers can be consolidated, using hardware virtualization, into one server, we will effectively be saving the sum of three idle power consumption figures, minus whatever increased power consumption is produced on the server from the additional load. As the table demonstrates, the effect of increasing the processing load of a server on power consumption is minimal compared to the assignment of a dedicated server to achieve pieces of that processing load. The worst-case is the spread between "idle" and "load4" states on the 3000MHz clock speed setting: an increase of 82W.

Ertl notes in his statistics that a maximum power consumption of 423W was achieved at the 3000MHz clock speed by maximizing use of all available system resources. Assuming 4 servers having the lowest clock speed of 2000MHz and running at “load1” processing conditions, this situation has the following power consumption:

290W * 4 = **1160W**

If we could consolidate these servers into a worst-case processing power situation of the 3000MHz CPU running under maximized conditions, and therefore using 423W of power, we have therefore saved:

1160W – 423W = **737W**

One electricity supplier quotes a rate of 5.10 cents per kWh for “Medium Commercial Customers” (BC Hydro, 2007). Based on this figure, for a data center operating 24 hours a day, 7 days a week, the annual cost savings are as follows:

0.737 kWh * $0.051 / kWh * 24 hours * 365 days = **$329.26**

This means that, across a data center that is able to identify a modest 100 such consolidation opportunities, an annual savings of $329,260 would be realized.

Further to this, there are savings beyond that of the raw electricity cost: system hardware costs are reduced; cooling costs are reduced (the wasted electricity is often dissipated as heat and must be removed using cooling infrastructure); and real estate and property tax costs resulting from the reduced system footprint are also diminished. In addition, there is likely an achievable gain in corporate goodwill from the improvement of environmental factors, and the potential to locate the data center in more central, smaller quarters, closer to mass transit, and closer to high quality employees.

Environment

Beside the cost savings achieved with hardware virtualization, there are also significant environmental benefits to society-at-large. Such environmental efforts are useful in brand promotion and goodwill generation; and their side-effects can be attractive to potential employees.

Firstly, the smaller system footprint enabled by virtualization allows smaller data centers to be built. Currently, data centers are most often built as large, single-storey structures in areas where land is relatively cheap (Dininny, 2006), and this requires a large investment by the hosting region in terms of infrastructure construction: each meter of building perimeter requires larger surrounding roads; and electricity, water, and sewer services must be stretched further in order to service them and the other businesses in the area. Secondly, the location of data centers in far-away places reduces the opportunity to service the employees of these centers with mass transit because the density of population is not high enough. By reducing the size of server hardware, it is possible that it may become much more affordable to locate data centers as in-fill structures in existing, central areas of higher density and, in doing so, would remove the requirement of building new roads and utilities while also providing the opportunity to service the facilities with mass transit. This will become especially important as the costs of servicing remote, low-density structures increases due to global resource constrictions brought on by the growth of countries like India and China, and as people begin to question the viability and affordability of using personal transportation to get to and from their place of work. The knowledge gained will be of paramount importance if a company wishes to locate their data centers in India and China in future, because these countries already recognize and discourage the use of low-density spaces (Pucher, Peng, Mittal, Zhu, & Korattyswaroopam, 2007).

Environmental benefits can also be derived from the savings from power consumption noted in the previous section. If the mentioned 100 consolidation opportunities were to be found, it would result in an ongoing power consumption reduction of:

737W * 100 = 73,700W ~= **74kW**

If 100 data centers could provide similar opportunities, this represents a savings of 7400kW, or 7.4MW. This is approximately the output of two large wind turbines (Creative Environmental Networks, n.d.), or 0.2% of the output of a large coal-based power plant (Ontario Power Generation, 2007). Considering that this not a large number of consolidation opportunities, and that there would be no reduction in capability, this is an impressive saving and does not include the derivative power savings mentioned in the previous section.

Conclusion

The direct and indirect cost savings and environmental benefits of hardware virtualization are clearly worth at least an investigation into the possibility of consolidating some of an organization's under-utilized servers into virtual servers running on a smaller number of physical servers. The result will be a reduced physical data center footprint, significant cost savings in electricity, and numerous indirect benefits resulting from these factors such as increased access to high quality talent; flexibility of location and increased potential to locate near mass transit; and the viability of locate facilities away from low density and other environmentally-sensitive land. Each of these factors is also powerful corporate goodwill generators. Considering that each of the cost factors are likely to continue to increase in future as developing countries with large populations continue to develop at a rapid pace and compete for natural resources, the savings will increase with each subsequent year.

References

BC Hydro. (2007, February 1). *Business rates*. Retrieved November 6, 2007 from http://www.bchydro.com/policies/rates/rates759.html

Bianchini, R., & Rajamony, R. (2004, November). Power and energy management for server systems. *IEEE Computer, 2004*(11), 68-74. Retrieved November 17, 2007 from IEEE Digital Library database.

Creative Environmental Networks. (n.d.). *Wind power: energy blowing in the wind*. Retrieved November 10, 2007 from http://cen.org.uk/developer/wind.asp

Dininny, S. (2006, March 28). *Server farms could be a cash crop*. Retrieved November 2, 2007 from http://seattlepi.nwsource.com/business/264525_ruraltech28.html

Energy Information Administration. (2007, October 22). *Electric power annual*. Retrieved November 10, 2007 from http://www.eia.doe.gov/cneaf/electricity/epa/epa_sum.html

Ertl, A. (n.d.). *How much electricity does a Computer consume?*. Retrieved November 5, 2007 from http://www.complang.tuwien.ac.at/anton/computer-power-consumption.html

Fox, S. (2007, February 26). The greening of IT. *InfoWorld, 29*(9), 2. Retrieved November 12, 2007 from EBSCOhost database.

GE Energy. (2007). *Wind turbines*. Retrieved November 10, 2007 from http://www.gepower.com/prod_serv/products/wind_turbines/en/index.htm

Harris, R. (2007, July 27). *Google's warehouse-size power problem*. Retrieved November 15, 2007 from http://blogs.zdnet.com/storage/?p=165

Koomey, J. (2007, February 15). *Estimating total power consumption by servers in the U.S. and the world*. Retrieved November 12, 2007 from http://enterprise.amd.com/Downloads/svrpwrusecompletefinal.pdf

Lundquist, E. (2007, June 25). Google's mixed message. *eWeek, 24*(23), 8. Retrieved November 12, 2007 from EBSCOhost database.

Margulius, D. (2007, February 26). Gauging net consumption. *InfoWorld, 29*(9), 29. Retrieved November 12, 2007 from EBSCOhost database.

Mills, M. (2007, February 5). *Futurist: target Google*. Retrieved November 15, 2007 from http://public.cq.com/docs/cqw/weeklyreport110-000002443248.html

Ontario Power Generation. (2007). *Nanticoke generating station*. Retrieved November 10, 2007 from http://www.opg.com/power/fossil/nanticoke.asp

Pucher, J., Peng, Z., Mittal, N., Zhu, Y., & Korattyswaroopam, N. (2007). Urban transport trends and policies in China and India: impacts of rapid economic growth. *Transport Reviews, 27*(4), 379-410. Retrieved November 2, 2007 from http://www.chinaplanning.org/Publications/China%20and%20India_PUCHE.pdf

Taylor, C. (2007, October 1). Veritable virtualization: concept of server virtualization software evolves. *Electronic News, 52*(40), 26. Retrieved November 16, 2007 from EBSCOhost database.

VMware Inc. (2007). *VMware ESX Server datasheet*. Retrieved November 16, 2007 from http://www.vmware.com/files/pdf/esx_datasheet.pdf

VMware Inc. (2007). *VMware Server datasheet*. Retrieved November 16, 2007 from http://www.vmware.com/pdf/server_datasheet.pdf