RC24768 (W0903-078) March 20, 2009
Computer Science

IBM Research Report

True Value: Assessing and Optimizing the Cost of Computing at the Data Center Level

John Karidis, José E. Moreira, Jaime Moreno
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

True Value: Assessing and Optimizing the Cost of Computing at the Data Center Level

John Karidis
IBM T.J. Watson Research Center
Yorktown Heights NY 10598
karidis@us.ibm.com

José E. Moreira
IBM T.J. Watson Research Center
Yorktown Heights NY 10598
jmoreira@us.ibm.com

Jaime Moreno
IBM T.J. Watson Research Center
Yorktown Heights NY 10598
jhmoreno@us.ibm.com

ABSTRACT

There are five main components to the cost of delivering computing in a data center: (i) the construction of the data center building itself; (ii) the power and cooling infrastructure for the data center; (iii) the acquisition cost of the servers that populate the data center; (iv) the cost of electricity to power (and cool) the servers; and (v) the cost of managing those servers. We first study the fundamental economics of operating such a data center with a model that captures the first four costs. We call these the physical cost, as it does not include the labor cost. We show that it makes economic sense to design data centers for relatively low power densities, and that increasing server utilization is an efficient way to reduce total cost of computation. We then develop a cost/performance model that includes the management cost and allows the evaluation of the optimal server size for consolidation. We show that, for a broad range of operating and cost conditions, servers with 4 to 16 processor sockets result in the lowest total cost of computing.

Categories and Subject Descriptors

K.6.2 [**Management of Computing and Information Systems**]: Installation Management – *benchmarks, computer selection, computing equipment management, performance and usage measurement, pricing and resource allocation.*

General Terms

Performance, Design, Economics, Management.

Keywords

Data center, Servers, System management.

1. INTRODUCTION

An increasing amount of computing in the world is being performed in large data centers that aggregate thousands to tens of thousands of servers. These large data centers seek to achieve economies of scale by consolidating significant computing capacity in one location. In some cases, large data centers are used by a single company to deliver the computing power it needs for its business operations; examples include Google, eBay, or the data centers used internally by large enterprises. In other cases, a large data center is used by a service provider to host computing capacity for several different customers. Examples of the latter include data center outsourcing services offered by companies like IBM and Digital Realty Trust.

To design better data centers, it is important to have a quantitative understanding of their cost. There is an entire spectrum of cost models that ranges from simplistic views of cost proportional to square footage, to detailed studies that characterize the performance and cost of servers in such data centers [3],[4],[5]. Models that normalize costs by the area tend to obscure the true source of key costs, and spreadsheet type models that include every local cost item can obscure important trends, so in this paper we use a high-level analytical model that captures the key components of the cost, and supports sensitivity analysis to explore the design space. Our derivation follows a different approach from what is done in the existing literature [4], and, as far as we know, many of our observations and conclusions regarding workload consolidation and optimal node sizes are original.

We identify five major components to the cost of delivering computing in a data center: (1) The construction of the data center building itself. (2) The power and cooling infrastructure. (3) The acquisition cost of the servers that populate the data center. (4) The cost of electricity to power (and cool) the servers. (5) The labor cost of operating and managing those servers.

To better study the fundamental economics of doing computing in a data center, we start by creating a model, similar to those in [4],[5], that quantifies the contribution of the first four components to the total cost. We call this the "physical cost of computing," to differentiate it from the labor cost in operating and managing the servers. This model shows that a large fraction of the cost of computing comes just from having the infrastructure to turn the servers on, while the incremental cost for adding load to a server (i.e., increasing its utilization) is relatively small. A direct consequence of this last observation is that the cost per unit of computing delivered goes down with increased utilization of servers. This is an intuitive result and has been noted before [1]. In this paper, we present a quantitative study of the cost of computing as a function of server utilization and validate the intuition.

Since cost of computing goes down with server utilization, it makes sense to consider servers that can operate at high utilization even if those servers are more expensive to purchase. Computer system vendors often promote the concept of *workload consolidation*, by which workloads from many small servers are combined on a few bigger servers that run at higher utilization [6]. Thus, even when the larger servers have a higher acquisition cost (on a per processor basis), many data centers adopt this strategy because of the total cost reduction.

We then expand our cost/performance model to take into account the management cost and evaluate the optimal server size for consolidation. Our results show that, for a broad spectrum of cost parameters, the optimal server size is in the range of 4 to 16 processor sockets.

Our model and analysis is targeted at mainstream air-cooled data centers. In this paper, we do not analyze data centers with primarily liquid-cooled equipment, nor do we consider highly dedicated data centers that require an absolute minimum amount of management. Those data centers are still considered somewhat specialized, while our results apply primarily to modern large data centers of typical design.

The rest of this paper is organized as follows. Section 2 presents our quantitative cost model for the physical infrastructure of a data center. Section 3 presents our analysis for the physical cost of computation. Section 4 introduces our discussion of optimal server size for consolidation. Finally, Section 5 presents our conclusions.

2. PHYSICAL INFRASTRUCTURE COST

The first step in delivering an operational data center is creating the physical infrastructure; that is, the facility in which the IT equipment for the data center will reside. To compute the cost of such facility, we start with a model from the Uptime Institute [5]. That cost model uses two primary elements to compute the cost of the data center: (i) the size of the computer room, and (ii) the amount of provisioned power and cooling.

The size of the computer room, measured in area of raised floor, defines how much space is available for IT equipment. The Uptime Institute uses an average cost of \$300 per ft^2 of raised floor space. That figures covers the usual overhead associated with the computer room itself: offices, restrooms, lobby, building shell, etc.

The amount of provisioned power and cooling defines how much IT equipment can be installed. Power and cooling equipment includes all the power distribution equipment (cables, transformers, panels), power backup equipment, water chillers, heat exchangers, computer room air conditioners, etc. The complexity and cost of that infrastructure depends on the degree of fault tolerance and functionality desired. The Uptime Institute classifies the data centers in four tiers, from Tier I, the least sophisticated, to Tier IV, the most sophisticated. The contemporary cost per kW of capacity (power and cooling) for each tier is shown in Table 1.

Let us illustrate the computation of the cost of two 10 MW Tier I data centers. The first one has 100,000 ft^2 of raised floor space (100 W/ft^2 average power density). The second data center has 25,000 ft^2 of raised floor space (400 W/ft^2 average power density). The cost of building the first data center is 10 MW × \$11,500/kW + 100,000 ft^2 × \$300/ft^2 = \$115 million + \$30 million = \$145 million. The cost of building the second data center is 10 MW × \$11,500/kW + 25,000 ft^2 × \$300/ft^2 = \$115 million + \$7.5 million = \$122.5 million.

Table 1: Cost of power and cooling infrastructure per kW of capacity for different data center tiers, as defined by the Uptime Institute [5].

Tier	Cost per kW of capacity
Tier I	\$11,500.00
Tier II	\$12,500.00
Tier III	\$23,000.00
Tier IV	\$25,000.00

We note that in both cases the cost of the data center is dominated by the infrastructure cost. That would be even more pronounced in the case of higher tier data centers. Note that a savings of 75% in floor space results in only 15% of total savings. This example shows that trying to save on the cost of the data center by "squeezing" space is not a particularly efficient approach. In fact, we show that it can be a bad idea.

Let us take the following scenario comparing the two data centers above. In both cases, the data center is populated with IT equipment that has an average density of 200 W/ft^2. In one scenario, the first data center is populated by spreading out the equipment to yield an average density of 100W/ft^2, in which case the owner has perhaps wasted some money on higher-density equipment. Alternatively, if the data center can handle twice the design power density over half of it's design area, the first data center can populate 50,000 ft^2 of space before running out of power, effectively wasting 50,000 ft^2 of space (100,000 ft^2 total minus the 50,000 ft^2 populated), which cost \$15 million (50,000 ft^2 × \$300/ft^2). The second data center populates all 25,000 ft^2 of space but uses only 5 MW of power, effectively wasting 5 MW of infrastructure which cost \$57.5 million (5 MW × \$11,500/kW). This simple example shows that it is far more expensive to run out of space (and have power left over) that to run out of power (and have space left over).

The conclusion from this simple exercise is that it makes sense to design data centers with lower power density in mind. That is, with some extra space. Such a data center is far more flexible with respect to what equipment can be hosted than a data center designed with high average power density.

We now look at the problem from a different angle. Let us say we can choose the actual power density with which to operate the data center. (In the previous example, we stipulated a 200 W/ft^2 density.) Is there an optimal density that minimizes cost?

Let D be the target power density of the data center. Let T be the infrastructure cost per watt of capacity for the desired tier level (e.g., \$11.50/W for Tier I.) Let F be the cost per square foot of raised floor (\$300/ft^2 in the above model). Then, the normalized cost C (dollar per watt of capacity) of the data center can be computed as $C = T + F / D$.

Plots of the cost per watt of capacity for different values of power density and raised floor cost can be seen in Figure 1. The plots show that, although the cost per watt of capacity decreases with density, we clearly observe a diminishing return effect. There is little to be gained from operating the data center at densities above 100 W/ft^2 and almost nothing to be gained at densities above 200 W/ft^2. Moreover, the qualitative results do not change even if we triple the raised floor cost, from \$300/ft^2 to \$900/ft^2.

We note that Figure 1 represents an idealized scenario, in which power and cooling efficiencies are independent of power density. In reality, efficiencies decrease at high densities (or more expensive cooling solutions are needed) because the air gets hotter and/or more air needs to be moved. These inefficiencies can quickly erase any small reduction in cost from smaller floor space.

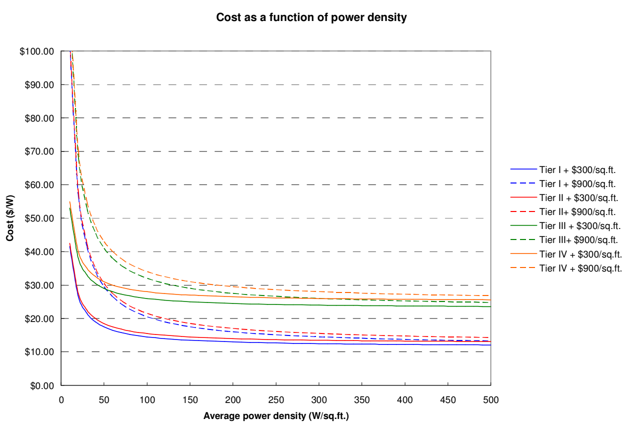

Figure 1: Cost per watt of capacity for different power densities. Solid lines are for the reference cost of \$300/ft^2 of raised floor. The dashed lines are for the hypothetical cost of \$900/ft^2 of raised floor.

Let us use this model to revisit the comparison between the two Tier I data centers that we performed earlier in this section. Figure 2 quantifies the impact of operating a data center at densities different from the design point. We include the plot for Tier I data centers (at \$300/ft^2), which shows the cost per watt if the data center is both built and operated at the indicated density. The other two lines indicate the cost per watt for operating the data center at 100W/ft^2 and 400W/ ft^2, assuming the data center was built for a specific density. (This is what actually happens.) These are the same parameters that we used in our previous example in this section.

We note that operating the data center at densities higher than designed results in a flat cost per watt, since the data center fills up to maximum power and cooling capacity. That cost is only marginally higher than the cost of a data center purpose built for a higher operating density. (The increased cost is the difference between the lines at densities greater than 100 W/ft^2 and 400 W/ft^2, respectively.) However, operating at lower than designed densities results in significantly higher cost. (The increased cost is the difference between the lines at densities smaller than 100 W/ft^2 and 400 W/ft^2, respectively.) Furthermore, we note that the 100 W/ft^2 data center is more cost effective than the 400 W/ft^2 data center for any operating point lower than (approximately) 340 W/ft^2. This analysis shows how constrained is the cost-effective range of operations for a high-density data center.

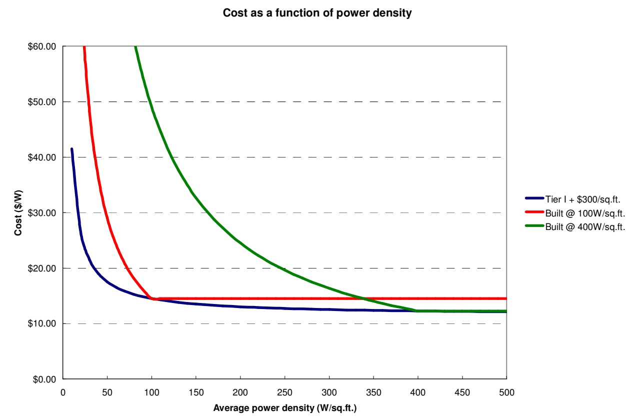

Figure 2: Cost per watt of capacity for different densities.

3. PHYSICAL COST OF OPERATION

We now derive an analytical model of the total physical cost of operation of a data center. In the physical cost we include only the mechanical and electrical infrastructure cost, the floor space cost, the cost of the IT equipment and the cost of electricity to operate the data center. We do not include the labor cost from operations, maintenance and system administration. (We will get to that in Section 4.) We also assume that our IT equipment consists primarily of servers.

We compute the cost in dollars (\$) per kW-compute-hour. This allows treating the depreciation costs on a per hour basis, instead of the depreciation period. A kW-compute-hour (kWch) is the amount of computing delivered by 1 kW of servers operating for one hour at full utilization. If the servers operate at lower utilization, then proportionally less computing is delivered.

We start the computation of the total physical cost with the model from the previous section. Let us use a Tier I data center and a density of 100 W/ft^2. Let I be the cost of the mechanical and electrical infrastructure for 1 kWh of computing. For Tier I data centers, 1 kW of capacity costs \$11,500. That infrastructure can typically be amortized over 20 years [Ref]. Since there is an average of 8766 hours in a year, the cost per hour per kW of capacity installed is

$$I = \frac{11500}{20 \times 8766} = \$0.066 / \text{h/kW installed.}$$

Similarly, we can compute F , the floor space cost per kWh of computing. At a density of 100 W/ft^2, it takes 10 ft^2 of raised floor space to host 1 kW of servers. Again, this can be amortized over 20 year and the cost comes to

$$F = \frac{3000}{20 \times 8766} = \$0.017 / \text{h/kW installed.}$$

For this exercise, let us populate the data center with servers that consume 500 W of power and cost \$4,000 each (a typical dual-socket server). Therefore, 1 kW of servers costs \$8,000. (Cheaper and more expensive servers exist, but the exact value does not change the qualitative results of our analysis.) Servers typically can be depreciated in approximately 5 years [Ref]. Therefore, the server cost S is

$$S = \frac{8{,}000}{5 \times 8766} = \$0.183 / \text{ h/kW installed.}$$

Finally, we have to account for the electricity used to power the data center, both to directly power the servers and to power the entire supporting infrastructure (backup, cooling, etc). A reasonably efficient data center will consume approximately 1.5 kWh of electricity for every kWh actually delivered to the servers [Ref]. At a cost of \$0.06/kWh, each kWh of server consumption costs \$0.09.

Let u be the average utilization of the servers in the data center. The actual power used by a server depends on its load. Measurements show that an idle server uses approximately 50% of the power it uses when fully loaded [2]. We adopt a linear model in which the power used by a server is proportional to the utilization. That is, the power used by 1 kW of servers is $(1+u)/2$ kW and the electricity cost $E(u)$ is

$$E(u) = \frac{1+u}{2} \times \$0.09 / \text{kWh.}$$

The total physical cost per hour for each kW of server installed $C(u)$ is calculated as a function of server utilization by summing the above four terms: $C(u) = I + F + S + E(u) =$

$$\left(\$0.066 + \$0.017 + \$0.183 + \frac{1+u}{2} \times \$0.09\right) / \text{ h/kW installed.}$$

Figure 3 plots this cost as a function of server utilization. We note that, except for electricity, the other cost components are independent of utilization. The electricity cost can be comparable to the infrastructure cost and the two together are comparable to the cost of the servers themselves. Raised floor space is a relatively small part of the cost, as we saw before.

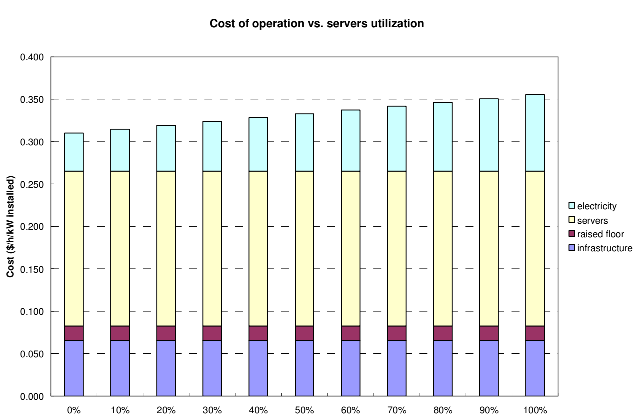

Figure 3: Total physical operation cost as a function of server utilization. The cost is per hour of a kW of servers in operation. Of the four main components of cost, infrastructure, floor space, servers and electricity, only the latter varies with utilization.

Since a kW of servers operating at utilization u produces u kWch, the cost per kWch is given simply by

$$\frac{C(u)}{u} = \frac{I + F + S + E(u)}{u} / \text{ kWch.}$$

We plot that cost in Figure 4(a). It is clear that the cost of computing decreases with the server utilization. In fact, the cost per computing at 10% utilization is almost 10 times the cost at 100% utilization. The more utilized the servers, the cheaper it is to produce a unit of computation. This property has several implications, but in this paper we focus specifically on the motivation for higher utilization servers.

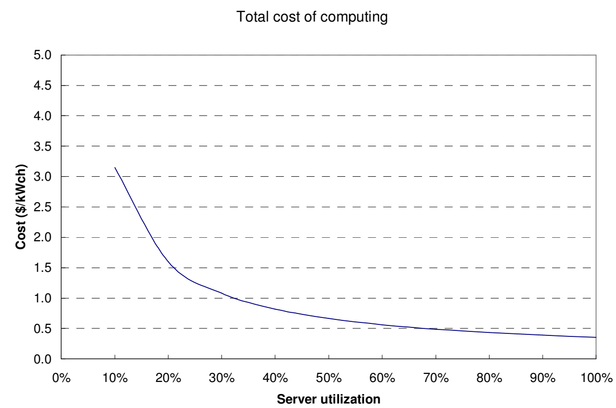

Figure 4: Total physical cost of computing per kWh of computing delivered, for the reference servers at \$8000/kW.

Let us consider what happens if we use more expensive servers. Figure 5 plots the cost of computing using the same equation above, but with servers arbitrarily assumed to cost from 2 to 5 times more than the reference server. The other costs are held constant. We note that a more expensive server (on a per capacity basis) can be more cost effective than a cheaper server if it can be operated at higher utilization. For example, the x5 server (costing \$40,000/kW) operating at 80% utilization delivers 1 kWch for \$1.30, whereas the reference server (\$8,000/kW) operating at 20% utilization delivers 1 kWch for \$1.60.

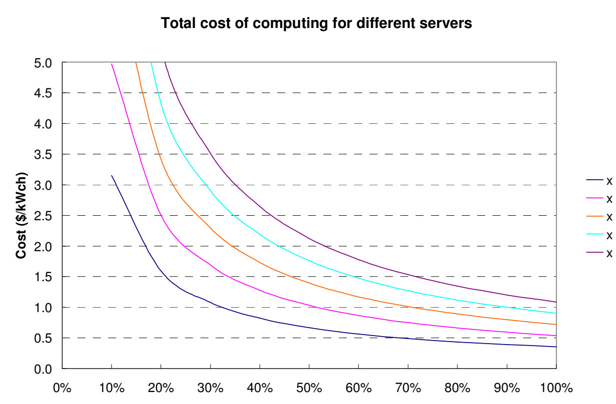

Figure 5: Total physical cost of computing per kWh of computing delivered. These results are for servers costing 2 to 5 times the reference of \$8000/kW.

4. OPTIMAL SERVER SIZE

In the previous section, we demonstrated how a more expensive server can be more cost effective than a cheaper one, if it can operate at higher utilization. There are several factors that can enable a server to operate at higher utilization, including the choice of operating system, application workload, usage policies, etc. In this paper we use results from queuing theory to show that, all other things being equal, a larger server (i.e., a server with more processors and more memory) can operate at higher utilization than a smaller server (i.e., a server with fewer processors and less memory).

To avoid confusion between processors and cores (as with multi-core processors), in this paper we do an analysis based on number of processor sockets (sockets, for short). A server can have 1 or more processor sockets (two being the most common) and in each socket we can plug a processor chip with one or more cores. Also associated with each socket is a certain amount of memory. In our model, the processor chip (that plugs into a socket) is the unitary engine of computing and, as the number of cores in a processor chip increases, the workloads are assumed to evolve to use those cores as a unit. We also assume that the workloads do not incur any additional overhead when they are consolidated onto a larger system.

To understand the claim that larger servers can run at higher utilization, let us consider two server configurations, both with a total of m sockets. The first configuration consists of m single-socket servers while the second configuration consists of one m-socket server. Let the memory per socket, and therefore the total memory, be the same in both configurations. We can model these configurations using the queuing models in Figure 6 and Figure 7. Each single-socket server is modeled by an M/M/1 queue while the m-socket server is modeled by an M/M/m queue. The work going through the queues corresponds to workloads that have to be processed by the processor sockets. For normalization purposes, we will adopt an average service time in each socket of 1. Therefore, the work arrival rate λ is equivalent to the utilization of each socket.

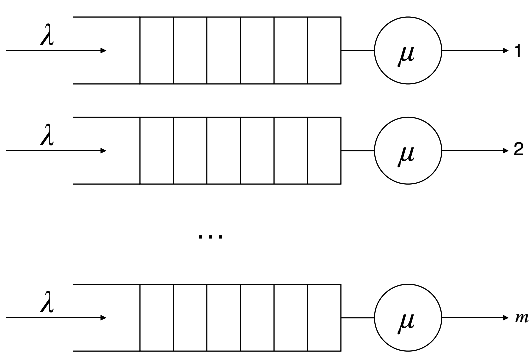

Figure 6: Queuing models for two server configurations: (a) model for m single-socket servers and (b) model for one m-socket server.

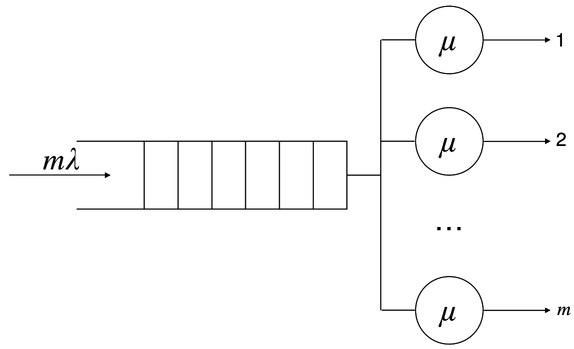

Figure 7: Queuing model for one m-socket server.

Figure 8 shows the response time as a function of utilization for the two queuing models of Figure 6, for different values of m. (We note that $m = 1$ corresponds to the M/M/1 model.) We observe that for any given utilization, the M/M/m queue, which models the m-socket server, has lower response time as m increases. Put in another way, for a given target response time, the m-socket server can operate at higher utilization, thus delivering more throughput, than m single-socket servers. The larger the m, the bigger the advantage. This simple model helps us understand why larger (and more expensive) servers can be more efficient than smaller servers when workloads are consolidated.

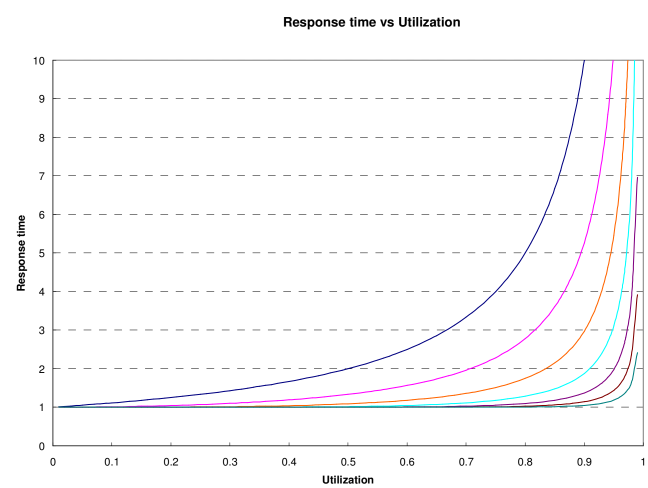

Figure 8: Response time as a function of utilization for the two queuing models of Figure 6. For any given utilization, the M/M/m queue has lower response time as m increases.

We assume that a constraint of consolidation is the preservation of the response time of the original unconsolidated systems. Figure 9 plots server utilization as a function of server size for a fixed response time. The plots are derived as follows. We use an M/M/1 model to compute the response time of a single-socket server operating at utilizations of 10, 20, 30, 40, 50, 60, 70, 80 and 90% (each of the plot lines in the figure). For each response time computed, we then use an M/M/m model to compute what is the utilization that an m-socket server can operate at and deliver the same response time. The plot shows that the lower the utilization in the single-socket server, the higher the benefit of going to an m-socket server. For example, if a single-socket server operates at 10% utilization, then an 8-socket server can operate at approximately 70% utilization with the same response time. In other words, an 8-socket server can produce seven times the throughput of 8 single-socket servers. Correspondingly, if the single-socket server already operates at 90% utilization, then there is little benefit from moving to larger servers.

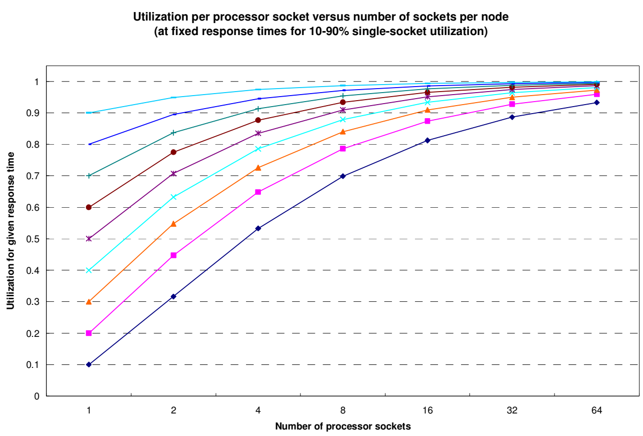

Figure 9: Server utilization as a function of server size for a fixed response time. We note that the lower the utilization of the single-socket server, the greater the benefit of moving to larger servers.

To compute price/performance metrics for the servers we need a cost model that takes into account the number of processor sockets in the server. The total cost of operation of a server includes the physical costs previously described (building infrastructure, electricity, IT equipment) plus the labor cost of managing and operating the servers.

Let us start by refining the server cost model that we previously used. In Section 3 we assumed a cost of $8,000/kW of server, based on a dual-socket server costing $4,000 and consuming 500 W. Let us adopt a consumption of 250 W/socket and model the equipment cost of an m-socket server $C_{\text{server}}(m)$ as having a fixed component, a component that scales linearly with the number of processor chips (e.g., memory) and a component that scales faster than linearly with the number of processor chips (e.g., the aggregate cost of those processor chips themselves), $C_{\text{server}}(m) = C_{\text{fix}} + C_{\text{mem}} \cdot m + C_{\text{proc}} \cdot m^{\alpha}$. Since an m-socket server corresponds to $250 \cdot m$ watts of power, the cost per kWh based on what we did in Section 3 comes to

$$S(m) = \frac{C_{\text{fix}} + C_{\text{mem}} \cdot m + C_{\text{proc}} \cdot m^{\alpha}}{0.250 \cdot m \cdot 5 \times 8766} \text{ / kWh.}$$

We can now plug this result in the formula for cost of kW-compute-hour. Figure 10 shows plots of cost per kWch for servers from 1 to 64 sockets, using the values of $C_{\text{fix}} = \$500$, $C_{\text{mem}} = \$1000$, $C_{\text{proc}} = \$500$, and $\alpha = 2$. The exponent α accounts for both the inherently increased wiring and complexity of large SMP servers and the common pricing models whereby per-processor prices rise with maximum supported SMP system size. We illustrate that pricing behavior by listing the prices for different models of AMD processor chips offered for sale at newegg.com, shown in Table 2. The processor chips are identical in computing capacity, but differ on maximum system size supported. We note that the larger the maximum system size, the higher the per-processor chip price.

Table 2: List price at newegg.com for different models of AMD processors. Each processor chip has the same computing capacity. The only difference is how large of a system can be assembled with them. (Fitting these costs to our model yields $\alpha \cong 1.8$.)

Processor model	Frequency	Maximum sockets	Price per chip (socket)
AMD Opteron 1356	2.3 GHz	1	$259
AMD Opteron 2356	2.3 GHz	2	$699
AMD Opteron 8356	2.3 GHz	8	$1509

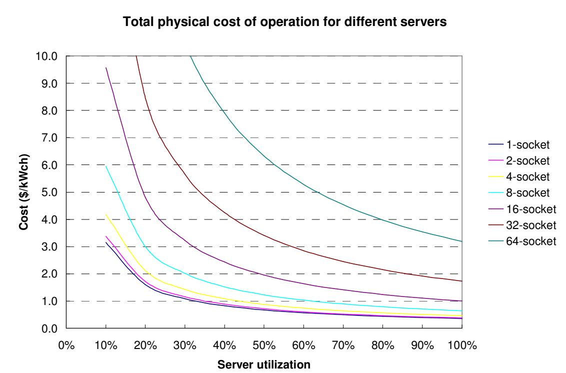

Figure 10: Cost of operation, without the management component, for servers of different sizes.

We can now add the labor cost of managing and operating the servers. We adopt a model where the labor cost of a server is a function of its size, $C_{\text{labor}}(m) = C_{\text{mgmt}} \cdot m^{\beta}$, where β is typically less than 1, since managing an m-socket server is easier than managing m single-socket servers. If C_{mgmt} is expressed in $/year, then we can compute the labor cost per kWh as

$$L(m) = \frac{C_{\text{mgmt}} \cdot m^{\beta}}{0.250 \cdot m \cdot 8766} \text{/kWh.}$$

Finally, we can compute the total (physical plus labor) cost of computing for an m-socket server operating at utilization u as

$$\frac{C(u)}{u} = \frac{I + F + S(m) + L(m) + E(u)}{u} \text{/kWmch.}$$

We introduce a new metric, a kW-of-managed-compute-hour (kWmch), which is a kWch of computing (at maximum utilization) but delivered from a managed server. This distinction makes it clearer if we are talking just about the physical cost of operation (without labor) or the total cost (with labor).

Figure 11 shows a plot of total cost of computing (in \$/kWmch) for servers of different sizes, with $\alpha = 2$, $C_{\text{mgmt}} = \$2000$ /year and $\beta = 0$ (That is, a server costs \$2000/year in management costs, irrespective of its size.) Comparing with Figure 10(a) we observe that the relative costs can change considerably when management costs are included. In particular, for the specific parameters we adopt, the lowest cost server for a fixed utilization is a four-socket server, as opposed to single-socket. Note that this is irrespective of any benefits from increased utilization at larger socket counts.

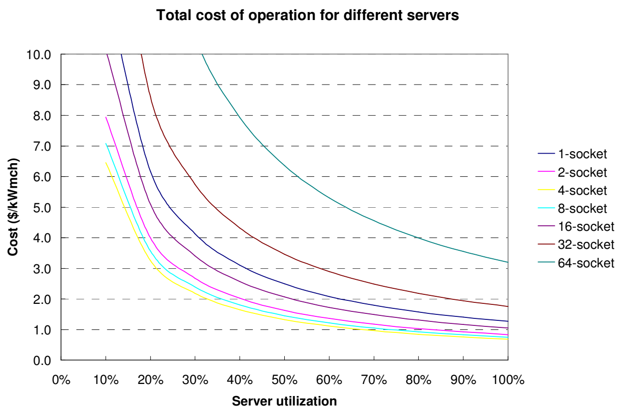

Figure 11: Cost of operation, with the management component, for servers of different sizes. Note that once cost of management is included, cost at a fixed utilization does not necessarily increase monotonically with server size.

Figure 12 plots the cost of computing as a function of server size for a constant response time, with the above values of $\alpha = 2$ and $\beta = 0$. Each of the lines is for a fixed response time – namely, the response time of a single-socket server operating at 10, 50 and 90% of utilization. (That is, 1.1, 2 and 10 times the service time, correspondingly.) It is computed by evaluating the cost at the utilization given by Figure 9. We note that the cost of computing first decreases with system size, as higher utilization more than offsets the increased cost of the server. Eventually, larger servers become too expensive. The most cost effective server size depends on the single-socket response time of the workloads, but it is in the 4- to 8-socket range for this particular cost/performance model. Smaller servers cannot sustain a high-enough utilization, whereas larger servers are just too expensive.

We conclude this section with a sensitivity analysis for different values of α and β. We vary α between 1 and 2 (i.e., aggregate processor cost from linear to quadratic with the system size) and β from 0 to 1 (i.e., management cost from fixed to linear with system size), and for each case we repeat the plot of Figure 12. We show those results in Figure 13 and make the following observations. First of all, if the cost of the processors is linear with the system size ($\alpha = 1$), then it always makes sense to use the largest server available. Second, if the management cost is flat with the system size ($\beta = 0$), then large servers are also favored, although the cost of the large servers can overwhelm that advantage ($\alpha = 2$). Finally, for most of the spectrum of configurations, the most cost effective server size is in the range of 4 to 16 sockets.

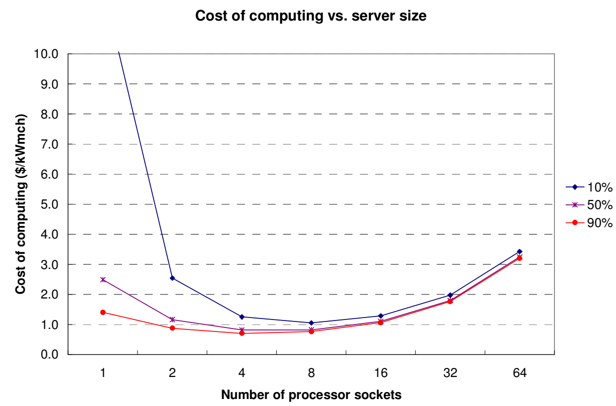

Figure 12: Cost/performance analysis for a server, as a function of its size (number of processor sockets).

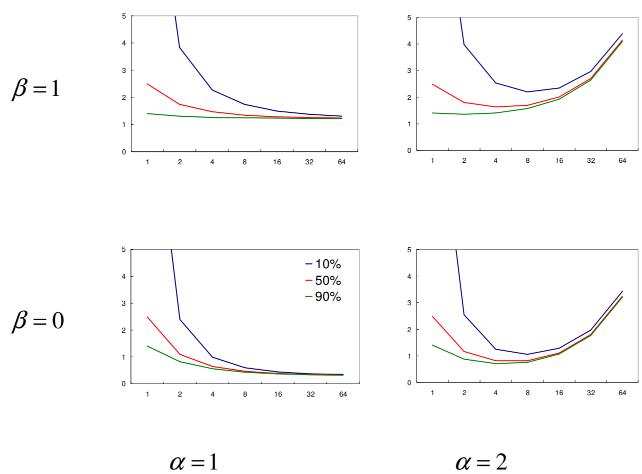

Figure 13: Results of a sensitivity analysis on cost parameters. Each plot is for one combination of α (the exponent for processor socket cost) and β (the exponent for management cost) values. We note that for a broad range of parameters, there is an optimal server size that usually varies between 4 and 16 processor sockets. For each plot, the x- and y-axis are as in Figure 12.

5. CONCLUSIONS

There are three main conclusions from our work: (1) Data centers should be designed for relatively low average power densities. (2) An efficient way to decrease the cost of computing is to increase server utilization. (3) Larger servers can operate at higher utilization than smaller servers and, because of non-linearity in the cost, there usually is an optimal server size for each situation.

The first conclusion derives from two observations. First, the cost of the mechanical and electrical infrastructure for a data center far outweighs the cost of the raised floor space. Second, fully populating a data center with equipment at lower power densities than the data center design point is very expensive. Therefore, one should design data centers with extra space, keeping the flexibility of populating them with different kinds of equipment until the more precious resource, power and cooling capacity, is exhausted.

The second conclusion derives from the observation that the physical cost of computing has three components that are independent of server utilization (infrastructure, floor space and the servers themselves), with only the electricity component being partially dependent on the utilization. Therefore, given the high fixed costs and low variable costs, increasing server utilization dramatically reduces the total cost of computing.

Our third conclusion, derived from a cost/performance model, shows that larger servers can indeed operate at higher utilization than smaller ones. This is a good justification for the consolidation play that the main U.S. server vendors (HP, IBM and Sun) execute with their high-end systems. These systems are often more expensive, on a per processor basis, than volume servers. But because they can be driven to higher utilization, they end up resulting in a lower total cost of computing.

Our cost/performance model also shows that whenever there are non-linear (with the number of sockets) components of the cost (e.g., processor and management), then there is an optimal server size for consolidation. Small servers have high management cost and low utilization, which offsets their low purchase cost. Servers bigger than the optimal size have high purchase costs that are not fully offset by the increase in utilization and reduction in management costs. The actual value of the optimum server size depends on the specific cost and performance parameters, but we show that for a broad spectrum of configuration parameters, servers with 4 to 16 processor sockets are optimal according to our models.

Finally, we should mention that our analysis and conclusions in this paper are valid for a certain spectrum of users and applications. Some users need extreme density and will resort to exotic cooling techniques to achieve it. Other users have applications that need the largest servers possible. In those cases, very large servers are the most effective (sometimes the only way) despite the additional cost. Correspondingly, some users have optimized their applications to obtain high utilization from small servers. In those cases, large servers have little or nothing to add. Nevertheless, we believe our analysis and results shed quantitative light on much anecdotal evidence that has been available and discussed for a while.

6. REFERENCES

[1] Luiz A. Barroso, Jeffrey Dean, Urs Hölzle. Web search for a planet: The Google cluster architecture. *In IEEE Micro Magazine*. April 2003.

[2] Xiabo Fan, Wolf-Dietrich Weber, Luiz A. Barroso. Power provisioning for a warehouse-size computer. *In Proceedings of the ACM International Symposium on Computer Architecture*. San Diego, CA, June 2007.

[3] Kevin Lim, Parthasarathy Ranganathan, Jichuam Chang, Chadrakant Patel, Trevor Mudge, Steven Reinhardt. Understanding and Designing New Server Architectures for Emerging Warehouse-Computing Environments. *In Proceeding of the ACM International Symposium on Computer Architecture*. Beijing, China, June 2008.

[4] Chandrakant D. Patel, Amip J. Shah. Cost Model for Planning, Development and Operation of a Data Center. Technical Report, HP Laboratories. HPL-2005-107R1. Palo Alto, CA, June 2005.

[5] W. Pitt Turner, Kenneth G. Brill. Cost model: dollars per kW plus dollars per square foot of compute floor. *White paper*. Uptime Institute. 2008.

[6] Werner Vogels. Beyond Server Consolidation. ACM Queue vol. 6, no. 1 - January/February 2008.