Cost Model for Planning, Development and Operation of a Data Center

Chandrakant D. Patel, Amip J. Shah[1]
Internet Systems and Storage Laboratory
HP Laboratories Palo Alto
HPL-2005-107(R.1)
June 9, 2005*

The emergence of the compute utility and growth in data center based compute services necessitates an examination of costs associated with housing and powering the compute, networking and storage equipment. The cost model must take into account the complexity in power delivery, cooling, and required levels of redundancies for a given service level agreement. The cost of maintenance and amortization of power delivery and cooling equipment must also be included. Furthermore, the advent of slim servers such as "blades" has led to immense physical compaction. While compaction can enable consolidation of multiple data centers into one, the resulting high thermal power density must be addressed from power delivery and cooling point of view. Indeed, the industry is abound with figures on sizing the total future data center capacity based on power loading per unit area, e.g. 500 W/m^2 today moving to 3000 W/m^2 in the future. However, in doing so, one needs to be mindful of the actual usage of the critical data center space with reference to the total rated and commissioned capacity. Therefore, a simple cost model that adequately captures the total cost of space and power based on utilization as well as recurring costs of "burdened" delivery is needed. This technical report shows a cost model that can be used in planning, developing and operating a data center. Techniques for "right" provisioning the data center are covered in the final chapters with reference to other reports published by the HP Labs team.

Approved for External Publication

TABLE OF CONTENTS

CHAPTER   1

Introduction

1.1 Objective

The objective of this report is to introduce a cost model for building and operating a data center.

1.2 The Need for a Cost Model

The data center of tomorrow is envisioned as one containing thousands of single board computing systems deployed in racks. The rack is an Electronics Industry Association (EIA) enclosure, 2 meter (78 in) high, 0.61 meter (24 in) wide and 0.76 meter (30 in) deep. A standard 2 meter rack accommodates forty to forty two complete single board computing units, and the dense rack configurations of the future, such as "blade" servers, will accommodate over 200 such computing units. The computing units include multiple microprocessors and dissipate approximately 250 W of power. The heat dissipation from a rack containing such computing units exceeds 10 KW today and the dense "blade" rack of tomorrow will reach over 30 KW. Thus, a data center with 1000 racks, over 30,000 square feet, will require over 10 MW of power for the computing infrastructure. The heat removal in such a high power density data center is quite complex. Conventional design and control approaches based on rules of thumb lead to gross over provisioning and inefficient design of cooling resources. It has been found through experiments at the HP Laboratories Smart Data Center that state-of-the-art control of cooling resources results in 0.8 W of power consumption by the data center cooling equipment for every 1 W of power dissipated by the compute hardware. A better approach using computational fluid dynamics (CFD) is necessary to design the air flow distribution and "right" provision the cooling resources. The thermo-mechanical research team at HP Labs has pioneered CFD modeling techniques and launched a service called "static" smart cooling of data centers for customers [1]. The intent of the service is to provide proper and efficient thermal management and reduce the power required by the cooling resources by 25%. Even with this static optimization, the recurring cost of power for the aforementioned 30,000 square foot data center is several million dollars per annum. Therefore, to avail greater savings, "dynamic" smart cooling of data centers is proposed through distributed sensing and control [1, 2]. Additionally, while the static and dynamic approaches enable savings in recurring cost of electricity, the maintenance and depreciation of the complex power delivery and cooling equipment must be captured to build adaptive and "right" provisioned data centers that can enable efficient data center based managed services.
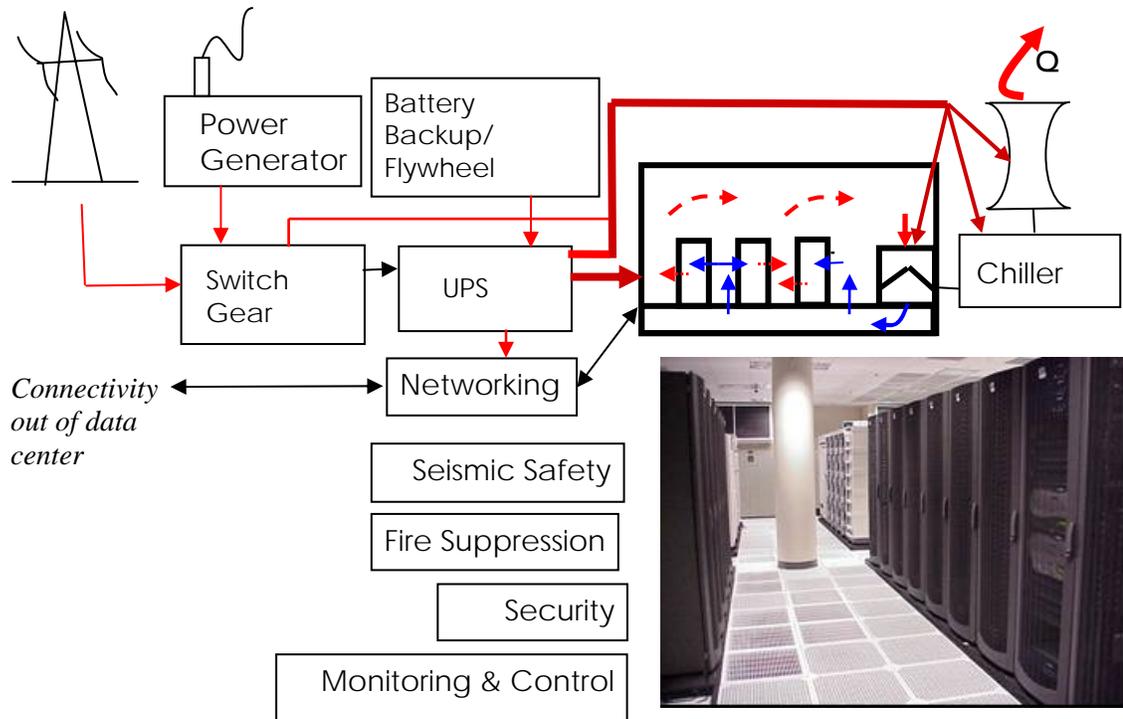
Figure 1.1. Key Elements of a Data Center

1.3 Salient Areas to Capture in the Cost Model

A cost model must capture the following salient points:

- o Cost of space. Real estate prices vary greatly according to the geographic location of the data center. For example, in early 2003, the commercial property prices in San Francisco were almost double those in other markets, such as Chicago [3]. A comprehensive data center cost model must account for such variance in real estate price.

- o Recurring cost of power. The electricity costs associated with continuous operation of a data center are substantial; a standard data center with a thousand racks spread over an area of 30,000 ft^2 requires about 10 MW of power for the computing infrastructure. The direct cost of drawing this power from the grid should be included.

- o Maintenance, amortization of the power delivery, conditioning and generation. Data centers are a critical resource with minimally affordable downtime. As a result, most data centers are equipped with back-up facilities, such as batteries/fly-wheel and on-site generators. Such back-up power incurs installation and maintenance costs. In addition, the equipment is monitored continuously, and costs associated with software and outsourced services must be included.

o <u>Recurring cost of power required by the cooling resources</u>. The environment internal to the data center must be maintained at a sufficiently low temperature for optimal rack operation. Since the electrical power consumed by the compute resources is dissipated as heat, an adequate cooling infrastructure is required for data center operation. State-of-the-art cooling resources tend to require almost as much power as that dissipated by compute resources to affect heat removal; this is a substantial addition that should be factored into a cost model.

o <u>Maintenance and amortization of the cooling resources</u>. Like the power system, data centers usually have backup chiller or air-conditioning units. The operation and maintenance (O&M) costs associated with the cooling infrastructure can be significant. In addition, the equipment is monitored continuously, and costs associated with software and outsourced services must be included.

o <u>Utilization of critical space</u>. The data center is a critical facility, and therefore care is required to "right" provision the data center resources. Sub-optimal operation (through "overprovisioning" or "underprovisioning" the available space) effectively raises the cost of ownership for a given data center.

Each of these factors is examined in further detail in the remainder of this report. Chapter 2 provides a brief overview of a typical data center design, thus introducing the various subsystems ("building blocks") that are used as a basis for the cost model formulated in Chapter 3 of the report. Chapter 4 discusses additional costs associated with data center operation, such as personnel and software expenses. Chapter 5 discusses how "smart" data center design can lower total data center operating costs, and the report concludes with a summary of key results in Chapter 6.

1.4 Nomenclature

A   area of the computer room (data center) or full property

Cap   capitalization rate, ratio of the net operating income (NOI) to the real estate property value. Lower cap rates reflect an appreciating real estate market.

AHU   Air Handling Unit served by chilled water unit (excludes direct expansion unit)

CRAC Computer Room Air Conditioning Unit (includes direct expansion units)

UPS   Uninterruptible Power Supply

PDU   Power Distribution Unit

COP   Coefficient of Performance, ratio of heat removed to work input (such as a compressor) in a refrigeration cycle.

IT_{dep} straight line monthly depreciation of IT equipment, typically over a 3-year period (per rack) [Eq. 4.2]

J_1 capacity utilization factor, i.e. ratio of actual data center power consumption to the maximum design (rated) power consumption [Eq. 3.6]

K_1 burdened power delivery factor, i.e. ratio of amortization and maintenance costs of the power delivery systems to the cost of grid power [Eq. 3.5]

K_2 burdened cooling cost factor, i.e. ratio of amortization and maintenance costs of the cooling equipment to the cost of grid power [Eq. 3.11]

L_1 cooling load factor, i.e. ratio of power consumed by cooling equipment to the power consumed by compute, storage and networking hardware [Eq. 3.9]

M_1 Ratio of the total number of information technology personnel servicing the data center to the number of racks

M_2 Ratio of the total number of facilities personnel servicing the data center to the number of racks in the data centers (excludes personnel costs associated with preventive maintenance contracts) to the number of racks

M_3 Ratio of the total number of administrative personnel servicing the data center to the number of racks in the data centers

M_{total} Ratio of the total number of all personnel servicing the data center to the number of racks in the data centers

NOI   Net Operating Income (in $), calculated as the difference between gross operating income and actual operating costs (not including amortization and maintenance)

P   power consumed by the hardware, networking or cooling equipment (Watts)

R   Number of racks utilized in a data center

S_{avg} Average salary of data center IT, facilities (not including preventive maintenance), administrative person per rack

SDC   Smart Data Center

σ_1 Software and licensing costs per rack (per month) [Eq. 4.3]

$U_{\$,grid}$ cost of grid power, in $/KWh or $/W per month

$U_{\$,A\&M}$ amortization and maintenance (depreciation) costs of either the power delivery or cooling equipment, in $/KWh or $/W per month

CHAPTER   2

Typical Data Center Design

2.1 Key Sub-systems – Power, Ping and Pipe

As shown earlier in Fig. 1.1, a data center is comprised of multiple subsystems. Electricity is supplied either from the grid or an on-site generator; this is then conditioned on-site before delivery to the computer room. Central plant chillers provide continuous supply of cold water for use in the computer room air-conditioning (CRAC) units. Practically all the electrical energy supplied to the computing equipment is dissipated in the computer room as heat, which is then removed by a cooling medium (usually air circulated by air handlers or fans). Additionally, apart from the rack switches, network connectivity must be provided for enabling data transmission within and outside of the data center.

Thus, there are three primary subsystems in a given data center: the power delivery system, which includes conditioning and backup equipment; the networking equipment, which includes all connectivity except the rack switches; and the cooling infrastructure, which includes both the central chillers and the computer room air-conditioners. These subsystems, often referred to as "power, ping and pipe" respectively, are the fundamental building blocks of a data center.

2.2 Building Blocks of a Data Center

2.2.1   Power Delivery, Conditioning and Backup System

Figure 2.1 shows a typical power delivery infrastructure. Electricity is drawn from the grid; in case of power failure, an on-site generator typically provides the input power to the data center. A backup battery bank is charged using this input electricity. These components help ensure continuous power availability to the racks and supporting infrastructure by compensating for fluctuations from the grid or temporary power loss. The degree of redundancy in the delivery system is an important consideration in the infrastructure, as systems with greater redundancy provide more contingency in case of failure, but the installation and maintenance costs of such a system are also higher.

Utility power is stepped down appropriately and conditioned in the UPS units. The conditioned power is utilized for charging the batteries and then delivered to data center PDUs. Rack-based PDUs are connected to breakers in the data center PDU. Dual connection to different PDUs improves reliability of the system. Static transfer switches provide graceful transition from utility power to battery or among dual sources at any stage of the delivery process. Three phase connections provide more power per unit core and are hence utilized extensively. The cooling infrastructure is more tolerant of

unconditioned power, but reliability and redundancy considerations usually decide the ultimate power delivery paths to these equipment. Additionally, even though uninterrupted cooling to the racks is required, the computer room air-conditioners can still be treated as a non-critical load to enable reduction of power delivery costs. Instead, to compensate for the downtime between a power failure and start-up of the generator, a chilled water storage tank can be used with the circulation pump being treated as a critical load. Based on these considerations, several competing power delivery architectures exist today. In all cases, however, the electricity reaches the computer room via PDUs, which step down the incoming 3 phase high-voltage electricity to more appropriate level based on equipment rating. The conditioned power is then made available to the racks, air-conditioners and auxiliary equipment as required. In terms of end-use, the racks and cooling infrastructure typically represent the largest loads on the power delivery system.
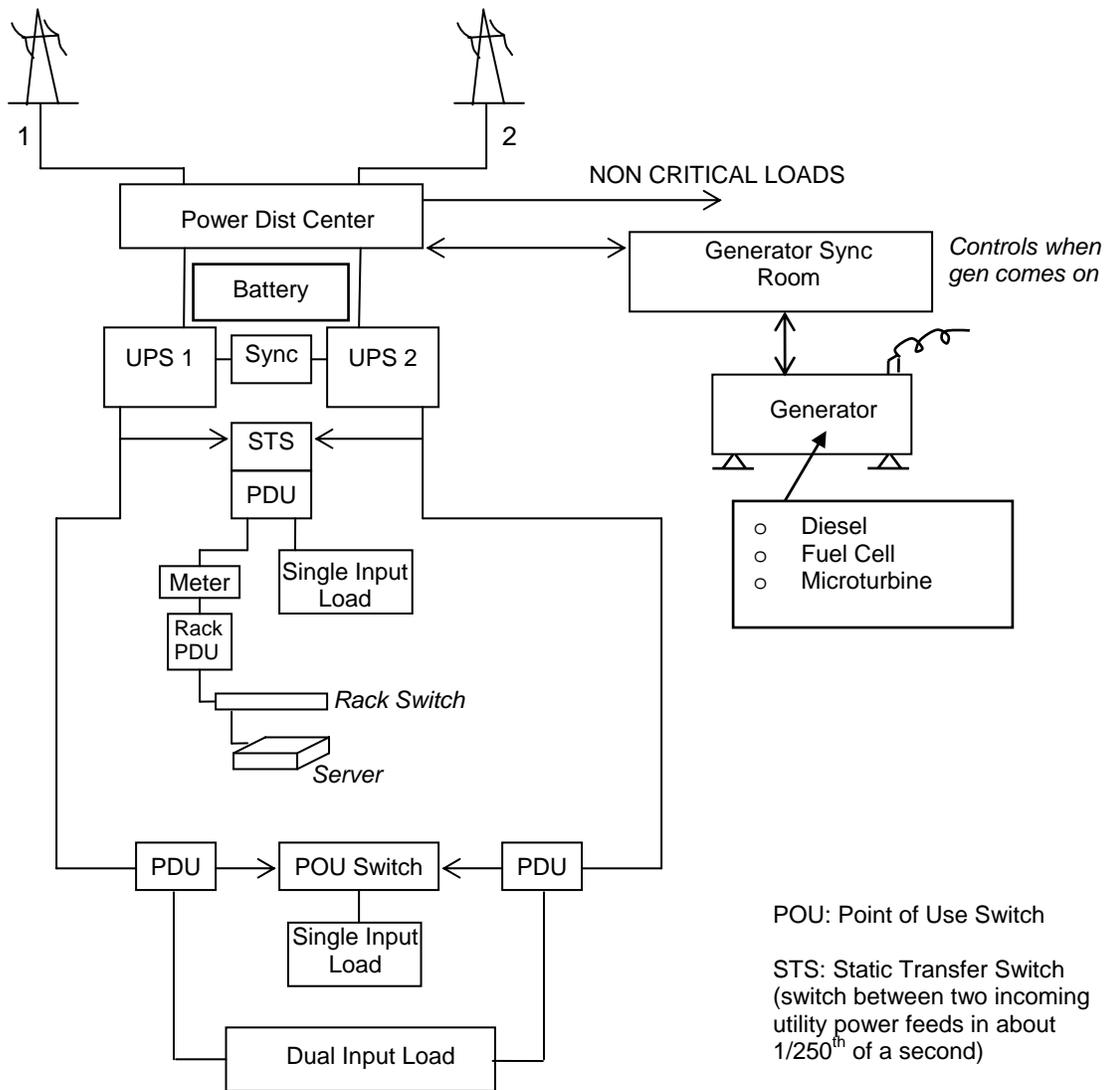
Figure 2.1. Typical power delivery system to a data center. The power delivery infrastructure includes the supply from the grid, the on-site generators, battery backups, and power conditioning equipment.

The figure 2.2 shown below is for a highly compacted
data center module shown in greater detail in Chapter 5
under modular construction of data centers.

COGENERATION

Gas-fired   Chiller

Backup

Boiler

Propane
Reserve

UPS
FLYWHEEL

STAGING

NET CONN

Service Area

800 sq ft

240 KW
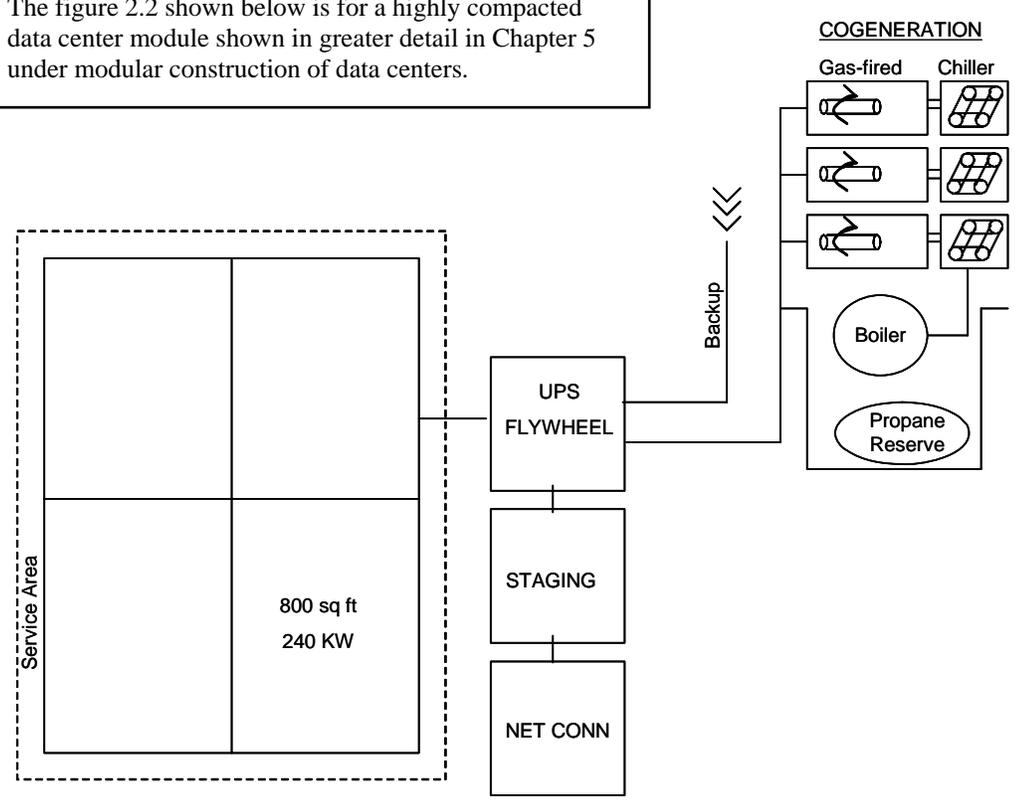
Figure 2.2. On-site power generation with absorption refrigeration cycle.

Server Rack   Hot Aisle   Cold Aisle   Temperature sense
                                                            point at return

                                                            Cooling
                                                            Coils

                                                      T

Raised
Floor

Plenum   CRAC unit

                                                            Air Mover

Cold Air   Hot Air   Vent Tiles

Figure 2.3. Typical raised-floor data center configuration.

2.2.1.1 Onsite Power Generation

Different techniques can be used for on-site power generation in data centers, which can eliminate dependence on grid power or be used for backup power in situations where the grid power is expensive or unreliable. On-site power production is most commonly achieved via diesel generators, although some smaller facilities have also utilized fuel cells [4]. Cog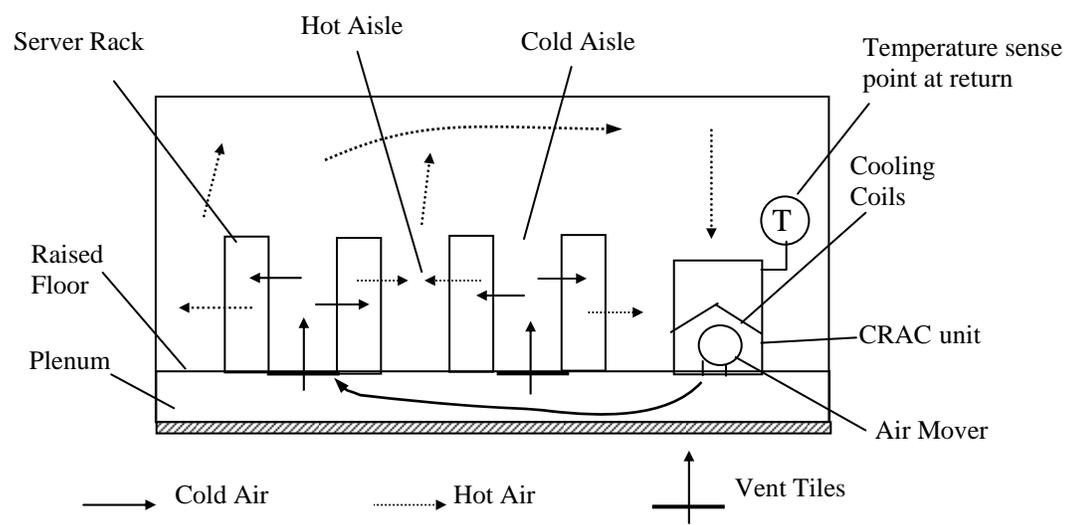eneration facilities, which can be utilized for both heating and electricity generation purposes, are also under consideration. Such facilities generally allow for more efficient use of resources, thus reducing overall operating costs. Figure 2.2 shows a conceptual design for a cogeneration system that runs on an absorption refrigeration cycle. In such a design, the waste heat generated in the electricity cycle (that is ordinarily simply rejected to the environment) augments as an input to a absorption refrigeration cycle running on propane. Such a system could impact the lifetime, fuel expenditure and maintenance costs of the power delivery.

2.2.2   Provisioning of Cooling Resources & Limitations of the State of the Art

Figure 2.3 displays an internal cross section of a typical state-of-the-art data center with modular air-conditioning units and under-floor cool air distribution. During the design phase, the computer room air conditioning (CRAC) units and associated back-end infrastructures, such as condensers, are sized based on projected heat loads in the room. The airflow is distributed qualitatively, and the CRAC units are controlled by temperature sensed at the return air inlet, typically set at 20 ^{o}C. Ideally, the rack exhaust should be directly returned to the CRAC units; however, the prevalent control methods promote mixing of the warm return air with the cool supply air – an inefficient phenomenon that prevents achievement of compaction goals. Patel et. al describe this limitation and propose new approaches based on computational fluid dynamics modeling in design of data centers [5, 6]. Sharma et. al introduce new metrics to assess the magnitude of recirculation and mixing of hot and cold streams [7]. These metrics, called Supply Heat Index and Return Heat Index, are based on non-dimensional parameters that describe the geometry of the data center. Bash et. al apply these design techniques to enable energy efficient data center design [8]. In order to adapt to rapid changes in data centers, Patel et. al introduce "smart" dynamic provisioning of cooling resources [1, 2]. Sharma et. al suggest thermal policies that can be used to allocate workloads [9]. Other work at HP Labs proposes global workload placement using a data center efficiency index based on external environment and internal thermo-fluids behavior [10]. Thus, a variety of modeling techniques and high-level control policies are available for assessing and improving cooling performance in data centers. In addition, the cost and performance of the cooling system in a data center can be impacted by considering the individual components of the infrastructure, such as the chillers and the CRAC units.

2.2.2.1 Air Handling Units with Central Chilled Water Supply

The cooling cycle in a data center begins with internally housed modular units, typically with ability to extract 100 KW of heat from the warm ambient intake. The heat extraction lowers the air temperature. This cold air is then supplied to the data center space, where it

picks up the heat dissipated by the compute equipment. The warm air is returned to the modular CRAC unit, where the heat is rejected to the chilled water stream. The chilled water loop is enabled by an external chiller – typically a screw-type compressor-based vapor compression unit with a 1 MW capacity.

2.2.2.2 Direct Expansion Units with Air Cooled Heat Exchangers

In a direct expansion CRAC unit, the refrigeration cycle takes place locally in the CRAC unit and the heat is rejected from individual condensers that serve each CRAC unit. The condensers are typically air cooled, but can also be water cooled using a cooling tower. This type of CRAC uses a vapor-compression refrigeration cycle (rather than chilled water) to directly cool the air with the help of an internal compressor. As a result, direct expansion units can be easier to monitor and control because the power consumption of the compressor is proportional to the cost of cooling the air, and the compressor power can easily be measured using a power sensor. However, to enable direct control, the direct expansion units need the ability to scale or flex its capacity, e.g. by changing compressor speed.

2.2.2.3 Common Deployment Model - AHU, Chiller with Cooling Tower

In a common deployment model, AHUs are coupled with appropriate hydronics to a central chiller supplying chilled water. Plant cooling tower(s) acts as a heat exchanger between the chiller and the environment. The implementation and monitoring costs of such a system are likely to be competitive with other options because such a system can be used to meet the chilled water requirements of the entire site, but redundancy for such a system may be more costly to implement. Further, because the refrigeration cycle takes place at a central chiller that may serve other thermal management infrastructures in the facility, the actual power consumption of each CRAC unit is difficult to measure. Instead, the heat load extracted by the CRAC is usually estimated by measuring the inlet and outlet chilled water temperatures and the chilled water flow rate. The actual cost of cooling is calculated by using the central chiller refrigeration cycle's coefficient of performance (COP), which can be measured directly or approximated using a simple heat transfer model. A technical report at Hewlett-Packard Laboratories [11] discusses such models in greater detail.

2.2.3   Security, Network Connectivity, Fire Protection, Security and Seismic Safety

While obvious, these costs can be significant in a data center, particularly in the event of a disaster. For example, even in the event of a contained fire, smoke particles may deposit on the disk and tape surfaces, rendering the data unrecoverable. Similarly, discharge of water or Halon fire retardants may cause further damage if the power supply to the computer systems is still active [12]. As a critical resource, protection and disaster recovery for data centers is essential.

CHAPTER   3

Data Center Cost Model:
Space, Power and Cooling

3.1 Key Drivers of the Cost Model

As discussed in Chapter 2, the three major building blocks required to enable data center operation are the "power, ping and pipe" (the electricity supply, networking infrastructure and cooling resources). In addition, physical space is required to host the building and equipment, and further operational expenditures are incurred related to personnel, software licenses, compute equipment depreciation etc. Thus, the true cost of ownership of a data center can be summarized as follows:

$$\text{Cost}_{\text{total}} = \text{Cost}_{\text{space}} + \text{Cost}_{\substack{\text{power} \\ \text{hardware}}} + \text{Cost}_{\text{cooling}} + \text{Cost}_{\text{operation}} \qquad (3.1)$$

The cost of power for hardware in Eq. 3.1 include both the compute and network resources. While the cost associated with cooling is also predominantly power consumption, these costs will be modeled separately. Each cost listed in Eq. 3.1 will now be discussed in further detail.

3.1.1   Appraising the Data Center Space

The real estate value for commercial properties is typically estimated as the ratio of two parameters: (i) the Net Operating Income (NOI) per month, which is usually calculated as the difference between gross income and operating expenses (without including debt service costs such as amortization, depreciation and interest), and (ii) the Capitalization Rate ("cap rate"), which is conceptually similar to the P/E (price/earnings) ratio in the stock market [13]. A P/E of 40 means the stock is selling at 40 times its current value, or it will take an investor 40 years to get his/her money back; similarly, the capitalization rate is a representation of the ratio of current valuation to expected return [14]. A cap rate of 10% is typical for most income property, although cap rates ranging from 9% to 13% may often be encountered [15].

Thus, the real estate value of a commercial property can be given as:

$$\text{Price}_{\substack{\text{real} \\ \text{estate}}} = \frac{\text{NOI}}{\text{Cap Rate}} \qquad (3.2)$$

Table 3.1. Typical leasing rates for different markets across the US.

Metropolitan area	Average rent (per sq ft)*
New York	$ 60-67
San Francisco	$ 50-60
San Jose / Santa Clara	$ 50-55
Boston	$ 36-43
Philadelphia	$ 30-36

* based on 2001 data from [16]; adjusted for 2005 prices assuming a 5% discount rate.

Table 3.1 lists typical rates for data center leasing in different US markets. It should be noted that these correspond to the Gross Operating Income (GOI); operating expenses would need to be subtracted from these values to obtain the NOI.

Equation 3.2 is based exclusively on the real estate value. For a data center, however, a distinction must be made between the total space, active space, and leased space [16, 17]. The total property value of a data center may include the real estate necessary for the plant chillers, power generation systems, and other auxiliary subsystems; however, operating income is typically only realized in that portion of the data center which is populated with computing equipment [18]. Thus, the level of occupancy in a given data center should be included in the calculation of Eq. 3.2:

$$\text{Cost}_{\text{space}} = \frac{\left(\text{NOI}/\text{ft}^2\right)\left(A_{\text{data center}}\right)\left(\% \text{ Occupancy}\right)}{\text{Cap Rate}} \qquad (3.3)$$

Example

Consider a 100,000 ft^2 property in the Bay Area with a 60,000 ft^2 data center (computer room) that is expected to be 50% occupied. Assuming an NOI of $50/ft^2, the estimated real estate value of the property using Eq. 3.3 will be:

$$\text{Cost}_{\substack{\text{real} \\ \text{estate}}} = \frac{\left(\$50/\text{ft}^2\right)\left(60{,}000\,\text{ft}^2\right)(0.50)}{0.10} = \$15 \text{ million}$$

That is, no more than $ 15 million (or $ 150 / ft^2) should be paid for the entire property.

Alternatively, suppose the model is to be applied to estimate leasing prices. That is, data center real estate was already purchased at $30 million, since initial estimates had assumed 100% occupancy, but we now want to find out how much should be charged for leasing out the real estate after occupancy has dropped to 50%. Additionally, assume that the real estate market has appreciated, so that lower cap rates (say 0.09) are in effect. Using Eq 3.3:

$$\frac{\text{NOI}}{\text{ft}^2} = \frac{\left(\text{Price}_{\substack{\text{real} \\ \text{estate}}}\right)(\text{Cap Rate})}{(\%\ \text{Occupancy})\left(\text{A}_{\substack{\text{data} \\ \text{center}}}\right)} = \frac{(\$30{,}000{,}000)(0.09)}{(0.50)(60{,}000\,\text{ft}^2)} = \frac{\$90}{\text{ft}^2}$$

Thus, to break even at 50% occupancy in a slightly appreciated real estate market, net operating income should be $90 / ft^2. The rent charged per unit area leased in the data center (GOI) would need to be higher to compensate for operational expenditures. In terms of a real estate cost estimate as shown in Eq. 3.1, however, the estimated NOI ($/ft^2) is the appropriate metric since this is the representative value of the space if used for other purposes.

3.1.2   Burdened Cost of Power Delivery

While an apartment or hotel can charge customers based on room space, a data center customer – a given system housed in a given rack – requires conditioned power on a continuous basis, and has special requirements for cooling. In addition to the electricity costs for power from the grid, as discussed in Section 2.1.1, the power delivery system in a data center includes conditioning, battery back-up, on-site power generation, and redundancy in both delivery as well as generation. Depreciation (amortization) and maintenance costs are associated with the infrastructure required to enable each of these functions. Indeed, data center experts often talk about conditioned power cost being three times that of grid power [18]. Thus, rather than just the direct cost of power delivery, a *burdened* cost of power delivery should be considered as follows:

$$\boxed{\text{Cost}_{\text{power}} = U_{\$,\text{grid}}\, P_{\substack{\text{consumed} \\ \text{hardware}}} + K_1 U_{\$,\text{grid}}\, P_{\substack{\text{consumed} \\ \text{hardware}}}} \qquad (3.4)$$

where the power burdening factor K_1 is a measure of the increased effective cost of power delivery due to amortization and maintenance costs associated with the power delivery system. It should be noted that depreciation is relevant to the maximum (rated) power capacity of the data center, i.e. if a data center is built and had equipment commissioned to operate at 10 MW, then the amortization costs will be derived from the capital necessary to enable a 10-MW delivery system. Thus, a data center being operated at lower capacities inherently fails to adequately recover initial investment. Therefore, a penalty for suboptimal operation should be inherently included in the K_1 factor, which can be given as follows:

$$K_1 = \frac{J_1\, U_{\substack{\$,\text{A\&M} \\ \text{power}}}}{U_{\$,\text{grid}}} \qquad (3.5)$$

where J_1 is the capacity utilization factor given as:

$$J_1 = \frac{\text{Installed (maximum) capacity [Watts]}}{\text{Utilized capacity [Watts]}} = \frac{P_{\text{rated}}}{P_{\substack{\text{consumed} \\ \text{hardware}}}} \tag{3.6}$$

From Eq. 3.5, K_1 is the ratio of the realized (weighed) amortization and maintenance costs per Watt of power delivered to the direct cost of electricity (per Watt) drawn from the grid. This can be calculated on a monthly basis. J_1 weighs the amortization costs by measuring the level of utilization of the conditioned space in the data center. For example, if a data center built at 100 W/ft^2 is actually being used at 10 W/ft^2 [i.e. the data center is "over-provisioned"], the space is not being used optimally, since the actual power usage is lower than anticipated (and therefore realized amortization costs are higher). Similarly, if a data center built at 100 W/ft^2 is operated at a local power density of 130 W/ft^2 [i.e. the data center is "under-provisioned"], then the cost of ownership will also increase because the supporting infrastructure is not designed to handle these high loads in local areas, and equipment performance will be degraded while potentially shortening system lifespan (and thus raising maintenance costs). Also note that the penalty for sub-optimal operation levied by the capacity utilization factor J_1 is different from the penalty related to a low occupancy rate in the data center (Eq. 3.3). The occupancy rate has to deal with unused floor space, and any related penalties are applicable because the real estate price includes a cost of expansion that has not yet been vested. The capacity utilization penalty, on the other hand, deals with unused compute resources, and is applicable because the initial capital assumed operation at full capacity (and failure to operate at maximum capacity results in reduced ability to meet amortization costs).

Combining Eqs. 3.4, 3.5 and 3.6:

$$\boxed{\text{Cost}_{\text{power}} = U_{\$,\text{grid}}\, P_{\substack{\text{consumed} \\ \text{hardware}}} + U_{\$,\text{A\&M}}\, P_{\substack{\text{rated} \\ \text{power}}}} \tag{3.7}$$

Equation 3.7 gives the costs associated with power delivery to the data center in terms of direct electricity costs ($U_{\$,\text{grid}}$) and indirect electricity costs (burdening factors, represented by $U_{\$,\text{A\&M}}$). In the absence of knowledge of specific amortization and maintenance costs, an approximate value for K_1 (nominally 2) can be substituted into Eq. 3.5.

Example

Based on personal communications data center operators [18], for a 10-MW data center with typical redundancies, amortization and maintenance costs often double the incurred cost of power delivery to the data center. That is, $U_{\$,\text{A\&M, power}} \sim U_{\$,\text{grid}}$. Using a typical value of $U_{\$,\text{grid}} = \$100/\text{MWh}$ over a one-month period,

$$U_{\substack{\$,\text{A\&M} \\ \text{power}}} = U_{\$,\text{grid}} = \left(\frac{\$100}{\text{MWh}} \left[\frac{24\,\text{hours}}{1\,\text{day}} \right] \left[\frac{30\,\text{days}}{1\,\text{month}} \right] \left[\frac{1\,\text{MW}}{10^6\,\text{Watt}} \right] \right) = \frac{\$0.072}{\text{Watt}} \quad \text{(per month)}$$

Further, if the data center is only 75% utilized (i.e. 7.5 MW of power consumed by the hardware), then the capacity utilization factor J_1 will be (using Eq. 3.6):

$$J_1 = \frac{P_{rated}}{P_{consumed \atop hardware}} = \frac{10}{7.5} = 1.33$$

From Eq. 3.5, the power burdening factor K_1 is then obtained as:

$$K_1 = \frac{J_1 \, U_{\$,A\&M \atop power}}{U_{\$,grid}} = 1.33$$

So the burdened cost of power delivery for this 10 MW (rated) data center can be estimated using Eq. 3.4 as:

$$Cost_{power} = (1 + K_1) U_{\$,grid} P_{consumed \atop hardware} = (2.33) \left(\frac{\$0.072}{Watt} \right) (7.5 \, MW) = \$1,260,000 \, / \, month$$

In terms of a cost per unit area, if the data center is constructed at an average room power density of 50 W/ft^2 (for a total room area of 200,000 ft^2), the cost per square foot of power delivery can be computed as:

$$Cost_{power} = \frac{\$1,260,000}{200,000 \, ft^2} = \$6.30 \, / \, ft^2 \text{ of room area}$$

3.1.3   Burdened Cost of Cooling Resources

In addition to the power delivered to the compute hardware, power is also consumed by the cooling resources. Amortization and maintenance costs are also associated with the cooling equipment, so that the cost of cooling can be given as:

$$Cost_{power} = U_{\$,grid} \, P_{cooling} + K_2 U_{\$,grid} \, P_{cooling} \qquad (3.8)$$

The load on the cooling equipment is directly proportional to the power consumed by the compute hardware. Representing this proportionality in terms of the load factor L_1, we get:

$$P_{cooling} = L_1 \, P_{consumed \atop hardware} \qquad (3.9)$$

Substituting Eq. 3.9 into Eq. 3.8,

$$\boxed{\text{Cost}_{\text{cooling}} = \left(1 + K_2\right) L_1 \, U_{\$,\text{grid}} P_{\substack{\text{consumed} \\ \text{hardware}}}} \tag{3.10}$$

where the cooling burdening factor K_2 is a representation of the amortization and maintenance costs associated with the cooling equipment per unit cost of power consumed by the cooling equipment. As with the power delivery system, the K_2 cost is based on the maximum (rated) power capacity of the data center, i.e. if a data center is built to operate at 10 MW, then the amortization costs will be derived from the capital necessary to enable the cooling system necessary for a 10-MW load. Thus, a data center being operated at lower capacities inherently fails to adequately recover initial investment. Therefore, a penalty for suboptimal operation should be included in the K_2 factor, which can be given as follows:

$$K_2 = \frac{J_1 \, U_{\substack{\$,\text{A\&M} \\ \text{cooling}}}}{U_{\$,\text{grid}}} \tag{3.11}$$

where J_1 is the capacity utilization factor given by Eq. 3.6.

It should be noted that the capacity utilization penalty of Eq. 3.11 is different from the load factor penalty of Eq. 3.9 in the sense that L_1 is non-optimal because of *design* inefficiencies in the cooling system, while J_1 is non-optimal because of *operational* inefficiencies in the data center.

Combining Eqs. 3.6, 3.10 and 3.11:

$$\boxed{\text{Cost}_{\text{cooling}} = U_{\$,\text{grid}} \, L_1 \, P_{\substack{\text{consumed} \\ \text{hardware}}} + U_{\substack{\$,\text{A\&M} \\ \text{cooling}}} J_1 L_1 \, P_{\substack{\text{consumed} \\ \text{hardware}}}} \tag{3.12}$$

Example

Based on study of cooling equipment purchase and preventive maintenance costs for a 10-MW data center with typical redundancies, amortization and maintenance costs for the cooling infrastructure are approximately one-half the costs incurred for the power delivery. That is, $U_{\$,\text{A\&M,cooling}} = 0.5 \, U_{\$,\text{A\&M,power}}$. Using the typical grid pricing of \$100 per MWh on a monthly basis of $U_{\$,\text{A\&M,power}} = \$ \, 0.072/\text{W}$ (as calculated in previous example),

$$U_{\substack{\$,\text{A\&M} \\ \text{cooling}}} = 0.5 \, U_{\substack{\$,\text{A\&M} \\ \text{power}}} = 0.5 \left(\frac{\$0.072}{\text{Watt}}\right) = \left(\frac{\$0.036}{\text{Watt}}\right) \quad \text{per month}$$

Assuming the data center is only 75% utilized (i.e. 7.5 MW of power consumed by the hardware), then the capacity utilization factor J_1 will be (using Eq. 3.6):

$$J_1 = \frac{P_{rated}}{P_{consumed \atop hardware}} = \frac{10}{7.5} = 1.33$$

From Eq. 3.11, the cooling burdening factor K_2 is then obtained as:

$$K_2 = \frac{J_1 \, U_{\$,A\&M \atop cooling}}{U_{\$,grid}} = (1.33)\left(\frac{\$0.036/W}{\$0.072/W}\right) = 0.667$$

Lastly, experiments at Hewlett-Packard Laboratories indicate that a typical state-of-the-art data center has a cooling load factor of 0.8, i.e. 0.8 W of power is consumed by the cooling equipment for 1 W of heat dissipation in the data center. That is,

$$L_1 = 0.8$$

Substituting the above factors into the burdened cost of cooling resources per Eq. 3.10:

$$Cost_{cooling} = (1 + K_2)L_1 U_{\$,grid} P_{consumed \atop hardware} = (1.67)(0.8)\left(\frac{\$0.072}{Watt}\right)(7.5\,MW) = \$720,000/month$$

In terms of a cost per unit area, if the data center is constructed at an average power density of 50 W/ft^2 (for a total area of 200,000 ft^2), the cost per square foot of cooling can be computed as:

$$Cost_{cooling} = \frac{\$720,000}{200,000\,ft^2} = \$3.6/ft^2 \quad \text{of room area}$$

3.2   Overall Model for Data Center Cost of Space, Power and Cooling

Thus far, the real estate value of the data center has been estimated [Eq. 3.3] as well as the recurring costs for burdened delivery of power [Eq. 3.4] and cooling [Eq. 3.10]. To determine the total cost of ownership for the data center, the real estate valuation must be translated into a per-month recurring cost estimate (to enable summation with the power + cooling costs). As a first-order estimate, we consider that the reason the data center space becomes more valuable than any other commercial property is because it is enabled by burdened power delivery and cooling resources. However, the costs for cooling and power are accounted for separately, i.e. if K_1, K_2, L_1 etc were all zero, then the value of the data center should become identical to that of any commercial real estate. Thus, the average $/ft^2 commercial rates for the market can be used as a basis for determining the value of the data center real estate. (The appropriate $/ft^2 value will depend on the cap rate and other considerations discussed earlier in Eq. 3.1 to Eq. 3.3.) That is,

$$\text{Cost}_{\text{space}} \approx \left(\$/\text{ft}^2\right)\left(A_{\text{data center}}\right)\left(\%\ \text{Occupancy}\right) \tag{3.13}$$

where the $\$/\text{ft}^2$ is the average income per unit area that might be realized each month for a commercial property specific to the geographic area under consideration.

Combining Eqs. 3.4, 3.10, and 3.13, the overall cost for data center space and power is obtained as:

$$\text{Cost}_{\substack{\text{space}\\\text{power}\\\text{cooling}}} = \left[\left(\$/\text{ft}^2\right)\left(A_{\text{data center}}\right)\left(\%\ \text{Occupancy}\right)\right] \\ + \left[\left(1 + K_1 + L_1 + K_2 L_1\right)U_{\$,\text{grid}}\ P_{\substack{\text{consumed}\\\text{hardware}}}\right] \tag{3.14}$$

The first term in the above represents the cost of the real estate, while the second term represents the combined costs of power for hardware and power for cooling equipment. Amortization costs as well as penalties for inefficient utilization of critical data center resources is captured in the multiplying factors K_1, K_2, and L_1. The remaining costs of operation, including personnel, licensing and software, are discussed in the next chapter.

The equation can be further simplified to cost out space in terms of critical space used in ft^2 as:

$$\boxed{\text{Cost}_{\substack{\text{space}\\\text{power}\\\text{cooling}}} = \left(\frac{\$}{\text{ft}^2}\right)\left(A_{\text{critical}},\text{ft}^2\right) + \left(1 + K_1 + L_1 + K_2 L_1\right)U_{\$,\text{grid}}\ P_{\substack{\text{consumed}\\\text{hardware}}}} \tag{3.15}$$

Scenarios that can have significant impact on cost:

Two likely scenarios where the features incorporated into the above cost model could have significant impact are:

o   The overall heat load from all the equipment, not just local heat load, in a data center far exceeds the total cooling capacity. From the first law of thermodynamics, as the capacity of the equipment to remove heat does not match the dissipated power, additional cooling equipment will be required. Hence, additional cooling equipment purchased and installed to enable higher heat loads must be factored into K_2.

o   In cases where overall capacity exists, but high local power density results in "over-temperature" on some systems, the customer may purchase a service, such as HP "static smart cooling" to optimize the layout. This will add to the cost, but will also result in savings in energy used by the cooling resources, and must be factored into both J_1 and L_1. The energy savings associated with the cooling resources lowers L_1 and results in a payback of about 9 months. The type of thermo-fluids provisioning studies conducted to accommodate local high power densities are shown in ref [1, 5, 6, 8, 11].

20

CHAPTER 4

Estimating the Total Cost of Operating and
Running a Data Center

4.1 Personnel and Software Costs

The objective of this report is to focus on the physical space, power and cooling costs. However, in order to estimate the total cost, an estimate of personnel and software costs is critical. The personnel costs associated with management of heterogeneous set of applications and physical compute, power and cooling resources across the data center needs to be captured. The emergence of utility models that address the heterogeneity at a centralized level enables one to normalize the personnel cost to the number of active racks in the data center. Furthermore, in standard instances, there is a clear segregation of installations resulting from the need for physical security or due to high power density deployment of uniformly fully loaded racks.

4.1.1   Personnel Costs

Personnel cost may be considered in terms of number of information technology (IT) facilities and administrative resources utilized per rack per month. These are internal personnel (excluding those covered by preventive maintenance contracts for power and cooling equipment) utilized per rack in a data center. Including costs for the average number of IT personnel per rack (M_1), the average number of facilities personnel per rack (M_2), and the average number of administrative personnel per rack (M_3), the total personnel costs are obtained as follows:

$$\text{Cost}_{\substack{\text{personnel} \\ \text{per rack}}} = (M_1 + M_2 + M_3) S_{avg} = M_{total} S_{avg} \qquad (4.1)$$

Example

Suppose the average fully loaded cost of personnel in a data center is $10,000 per month (i.e. S_{avg} = $10,000). Then, for a fraction of resource – say M_{total} = 0.05 (i.e. one person for every 20 racks), the personnel cost on a per rack basis is $500 per month.

4.1.2   Depreciation of IT Equipment

In a managed services contract, the equipment is provided by the data center operator, and thus the total cost of operation must account for IT equipment depreciation. Assuming a rack full of industry standard servers, the list cost of IT equipment can be determined using manufacturer's list price. Using a straight line depreciation, the cost of depreciation of a rack full of industry standard servers may be determined on a monthly basis as follows:

$$\text{Cost}_{\substack{\text{depreciation} \\ \text{per rack}}} = \text{IT}_{\text{dep}} = \frac{\text{Rack Purchase Cost}}{\text{Lifetime of rack (usu. 3 years)}} \tag{4.2}$$

Example

Assume the purchase cost of a rack full of approximately 40 1U (U is 44.4 mm high server in a standard rack) servers and a 2U rack switch is \$180,000. A straight line depreciation over a period of 3 years per Eq. 4.2 will yield:

$$\text{Cost}_{\substack{\text{depreciation} \\ \text{per rack}}} = \frac{\$180,000}{(1\,\text{rack})(3\ \text{years})} = \$5000\,/\,\text{month}$$

4.1.3   Software and Licensing Costs

In a managed services contract, data center personnel administer the installation of the operating system, patches and resources for load balancing. The user's only responsibility may be to manage the application on the server. The personnel and hardware cost associated with the data center side of the cost is captured in sections 4.1.1 and 4.1.2. However, the software and licensing costs must also be applied to determine the overall cost. A first order strategy is to look at the total licensing cost, and the number of racks being utilized for IT purposes, and determine the cost on a per rack basis as follows:

$$\text{Cost}_{\substack{\text{software} \\ \text{per rack}}} = \sigma_1 = \frac{\text{Total licensing costs}}{\text{R}} \tag{4.3}$$

4.2   Total Cost of Rack Ownership and Operation

Combining Eq. 4.1, 4.2 and 4.3, the total costs of data center operation can be estimated as:

$$\text{IT Operation Cost}_{\text{total}} = \text{R}\left(\text{M}_{\text{total}}\text{S}_{\text{avg}} + \text{IT}_{\text{dep}} + \sigma_1\right) \tag{4.4}$$

In conjunction with Eq. 3.15, the desired total data center cost model in terms of critical space used in ft^2 will be:

$$\boxed{\begin{aligned}\text{Cost}_{\text{total}} = &\left(\frac{\$}{\text{ft}^2}\right)\left(\text{A}_{\text{critical}},\text{ft}^2\right) + \left(1 + \text{K}_1 + \text{L}_1 + \text{K}_2\text{L}_1\right)\text{U}_{\$,\text{grid}}\text{P}_{\substack{\text{consumed} \\ \text{hardware}}} \\ &+ \text{R}\left(\text{M}_{\text{total}}\text{S}_{\text{avg}} + \text{IT}_{\text{dep}} + \sigma_1\right)\end{aligned}} \tag{4.5}$$

Example

Suppose a customer requires five 20 Amp/110 V circuits for servers to host an application. The customer asks the data center services provider to provision industry standard servers based on the power allocated. In addition, the services provider is asked to install and maintain the operating system, load balancers, etc. Once the configuration is ready, the customer will install a given application and expect a given performance based on an agreement.

Assume the following apply to the equipment provided to the customer:
- o Approximately 10 KW of power to the rack(s) occupying 20 ft^2.
  - o From a space point of view, the cost of space can be determined by using Eq. 3.3 and is estimated to be $50/ft^2
- o Further, based on ratio of power consumed by hardware to the rated capacity, the data center provider determines a J_1 of 2 for this installation using Eq. 3.6.
- o K_1 is determined to be 2 (Eq. 3.5) for the redundancies needed.
- o K_2 is assumed to be 1 (Eq. 3.11) for the redundancies needed.
- o L_1 is determined to be 1.5 (1.5 W of power required by cooling equipment for 1 W of heat load in the data center).
- o Monthly cost of power is $0.072 per W based on $0.10 per KWh.

With these assumptions, and using a modified form of equation 3.14, the first order monthly cost for power, space and cooling for the 10 KW installation is:

$$\text{Cost}_{\substack{\text{space} \\ \text{power} \\ \text{cooling}}} = \left(\frac{\$}{\text{ft}^2} \right) \left(\text{Area, ft}^2 \right) + \left(1 + K_1 + L_1 + K_2 L_1 \right) U_{\$,\text{grid}} P_{\substack{\text{consumed} \\ \text{hardware}}}$$

$$= \left(50 \frac{\$}{\text{ft}^2} \right) \left(20 \, \text{ft}^2 \right) + \left[1 + 2 + 1.5 + (1)(1.5) \right] \left(\frac{\$0.072}{W} \right) (10,000 \, \text{W})$$

$$= \underline{\$5320 \text{ per month for power, space and cooling}}$$

To approximate the total cost of housing a 10 KW installation, consider the purchase price for a full rack of servers – 40 1U industry standard servers at 250 W each as an example - to be $180,000, yielding a straight line depreciation over 3 years of $5,000 per month. The total cost for servers, space, power and cooling is thus $ 10,320 per month.

Suppose the cost for software, licensing and personnel adds another $10,000 per month. Then, to a first-order approximation, **the total monthly cost to host a fully managed set of IT resources with an aggregate power of 10 KW will be approximately $20,000**.

CHAPTER   5

Key to Cost Effective "Smart" Data Center
Design and Operation

5.1 Right Provisioning of the Data Center: Dynamic Cooling

Construction of new data center facilities is often driven by anecdotal information and rules of thumb preached by experts, and it is quite common to hear W per unit area as the kernel for discussion of data center design. While this approach is a good start, this thread leads to provisioning of internal data center equipment and layout without providing due considerations for power supply and cooling plant located outside the building. Instead, one must start from the cooling tower and switchgear down to the room level and vice versa. At the room level, the business end of the data center with all the hardware, the state of the art approaches of designing and controlling the environment are lacking as noted in Section 2.2.2. The biggest challenge is to go beyond current industry techniques of controlling and designing the environment with too much focus on the projected maximum power in terms of Watts-per-square-foot rule of thumb. An example of a good start is to examine the IT needs of the company over time, and create a flexible infrastructure that can be "right" provisioned based on the need.

A case study specifically emphasizing the importance of "right" provisioning has been documented by Beitelmal and Patel [11]. Figure 5.1 shows the data center used in this study, which is the HP Smart Data Center (SDC) in Palo Alto, California. The total data center (DC) area is 267 m^2 (2875 ft^2) divided into three sections. The first is the South DC or production section, 126.7 m^2 (1364 ft^2) in area, which houses 34 racks with a measured computational (heat) load value of 136.5 kW. A Northeast DC or research section, 63.8 m^2 (687 ft^2) in area, houses 17 additional racks with a measured computational load of 89.3 kW. A Northwest DC or high-density section, 76.6 m^2 (824 ft^2) in area, contains 15 racks with a measured computational load of 144 kW. These numbers represent the operational load at a given time. For modeling purposes, at this loading at a given point in time, the total heat load dissipated by the server racks is 369.8 kW and the total cooling capacity available from the CRAC units is up to 600 kW when all CRAC units are operational and running at full capacity (100 kW per CRAC unit). The total flow rate provided by the CRAC units is 38.18 m^3/sec (80900 CFM) based on 6.84 m^3/sec (14500 CFM) from CRAC 1 thru CRAC 3 and 6.14 m^3/sec (13000 CFM) from CRAC 4 and 5.76 m^3/sec (12200 CFM) from CRAC 5 and CRAC 6. A chilled water-cooling coil in the CRAC unit extracts the heat from the air and cools it down to a supply of 20°C. The cooled air enters the data center through vented tiles located on the raised floor near the inlet of the racks. The most notable difference in this infrastructure is the ability to scale the capacity of the CRACs by changing the blower speed to modulate the flow work, and adjusting the amount of chilled water entering the CRACs to modulate thermodynamic work.
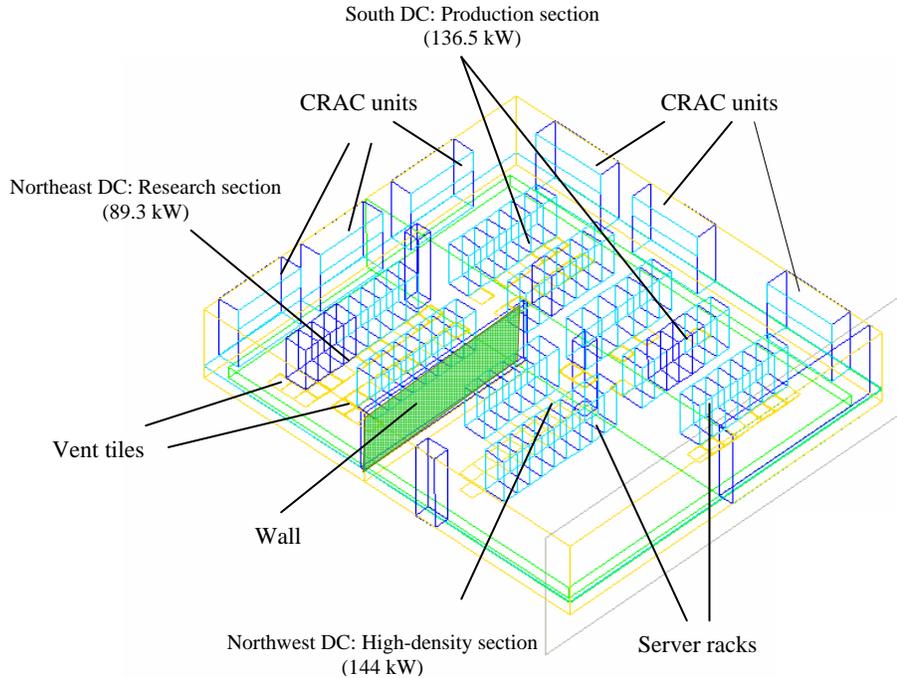
Figure 5.1. Data Center Layout

Computational fluid dynamics (CFD) simulations were used to analyze the heat load distribution and provisioning in the data center. Beitelmal and Patel [11] discuss the modeling techniques in greater detail. Fig. 5.2(a) shows the temperature distribution when all CRAC units are operating normally. The plot shows the capacity utilization per CRAC unit. CRAC units 1 and 2 are handling most of the load in the data center due to the load/racks layout relative to the CRAC units. In this case study, the contribution from CRAC units 5 and 6 to the high density area is minimal due to existence of a wall as shown in Fig. 5.1. It is clear from the results shown in Fig 5.2(a) that under the given heat load distribution, the CRACs will operate at various levels of capacity. And, to avail the requisite energy savings, one would need to change the CRAC flow and temperature settings to scale down to the correct levels – e.g. 27 % for CRAC 6. Use of distributed sensing and policy based control in conjunction with flexible infrastructure can enable this "smartness" – a proposition called "dynamic" smart cooling [1, 2].

Fig 5.2(b) shows the temperature distribution in the data center when CRAC unit 1 has failed. The plot shows the high heat load accumulating in the northwest high density section, causing the temperature to rise in that area. The plot also shows that the failed CRAC unit was handling the largest amount of load compared to the other five CRAC units in the model. After the failure of CRAC 1, previously separated warm air exhausted from the high density racks in the Northwest section travels across the room to reach CRAC units 2 through 4 warming up some of the cooled air along its path. The model also shows CRAC unit 1 failure leads to the mal-provisioning of CRAC unit 2 and CRAC unit 3, leading to severe temperature rises and unacceptable environmental conditions in some regions (Northwest DC section) of the data center. This can force system shutdown in those regions, resulting in large operational losses.
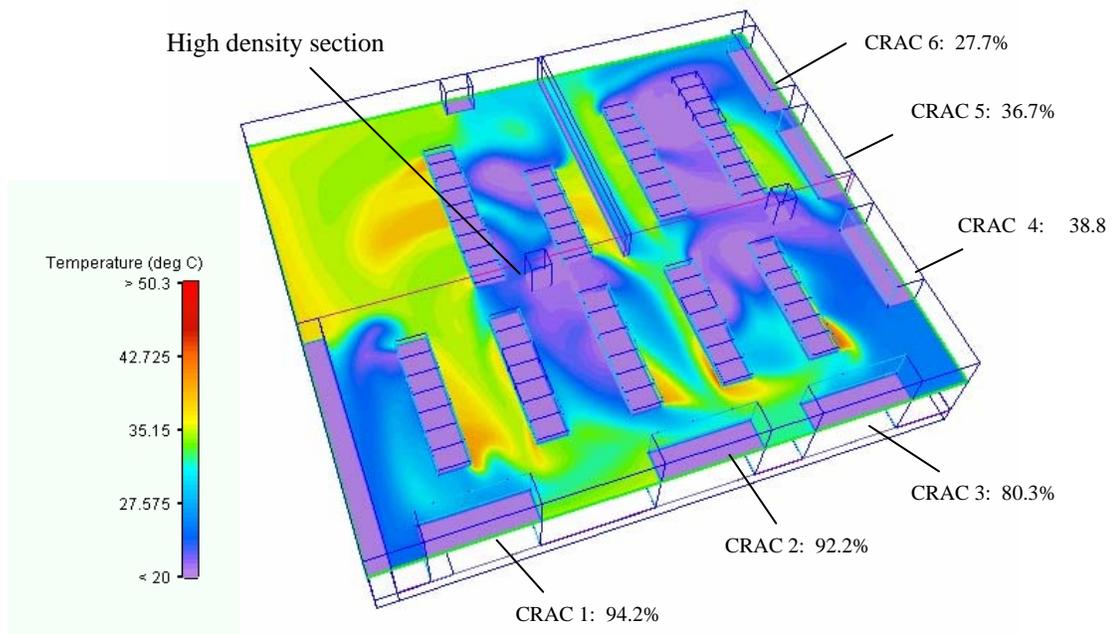
High density section

CRAC 6: 27.7%

CRAC 5: 36.7%

CRAC 4: 38.8

CRAC 3: 80.3%

CRAC 2: 92.2%

CRAC 1: 94.2%

Temperature (deg C)
> 50.3
42.725
35.15
27.575
< 20

Figure 5.2(a). Temperature distribution before CRAC unit failure

High density section

CRAC 6: 29.8%

CRAC 5: 44.6%

CRAC 4: 46.5 %

CRAC 3: 103%

CRAC 2: 146%

CRAC 1: FAILED

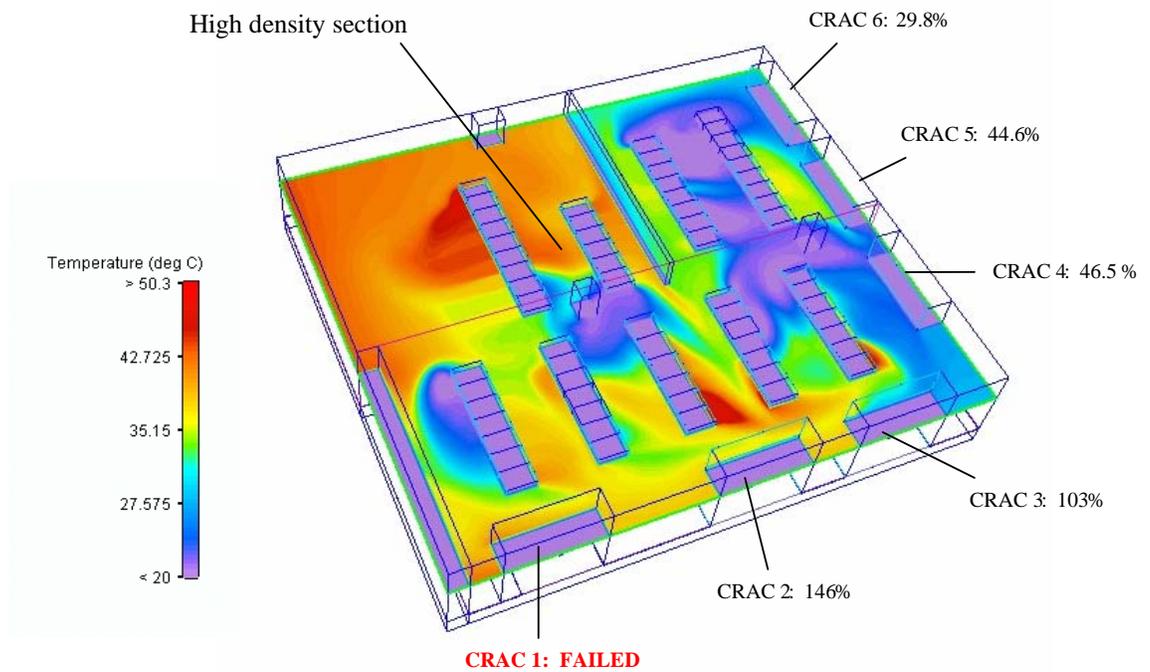Temperature (deg C)
> 50.3
42.725
35.15
27.575
< 20

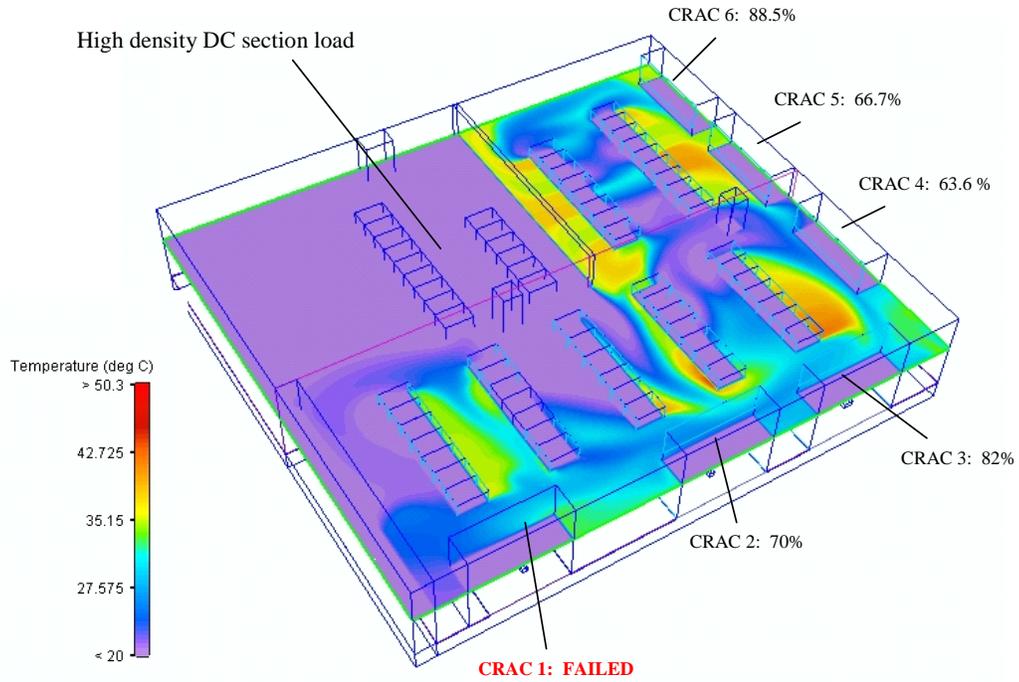Figure 5.2. (b) Temperature distribution after CRAC unit failure.

5.2 Introducing "Smart" Redundancy

Traditionally, data center designers and operators combat such a failure scenario by installing redundant back-up cooling capacity uniformly at the outset. For example, in a typical data center, one or more backup CRAC units might be found in each zone of the SDC (northeast, northwest and south DC). However, as discussed in Section 3.1.3, the realized costs of added cooling capacity are much higher than just the added value of equipment purchase. Under normal operation (i.e. no CRAC failure), the addition of backup CRAC units leads to a grossly overprovisioned system; this is manifested as an increase in the factors L_1 and K_2, which leads to a higher total cost of ownership per Eq. 4.5. To avoid these excess costs, the cooling resources in the data center can be "smartly" re-provisioned by:

o   designing physical redundancies against planned use. For example, if a given area is critical (such as the high density NW area in Fig 5.1), then additional cooling capacity can be provisioned specifically for this region without requiring redundancies across the length and breadth of the data center, or

o   alternately or as a complementary move, migration of the compute workload can also be used to mitigate failure, i.e. the entire compute load in the high-density area is redistributed within the rest of the data center. Figure 5.3 (a) shows the temperature distribution after workload distribution for failed CRAC operation. Compared to Fig. 5.2(b), much better thermal management – leading to lower CRAC energy consumption and reduced chance of equipment failure – is achieved by appropriately provisioning the CRAC resources for the workload in each section of the data center. The benefit of migrating workload away from an underprovisioned zone can be seen even more clearly in the comparison of Figure 5.3 (b) and (c), which shows the critical temperature distribution (above 34 $^{\circ}$C) before and after workload redistribution respectively.

With respect to workload migration, there is complementary research underway at HP Labs as described in the following references:

o   Sharma et. al suggest a thermo-fluids based policy that can be used to redeploy workloads [9].

o   Moore et. al build on a theoretic thermodynamic formulation to develop a scheduling algorithm [19].

o   Janakiraman et. al describe checkpoint-restart on standard operating systems [20] and Osman et. al show an implementation using a migration tool [21].

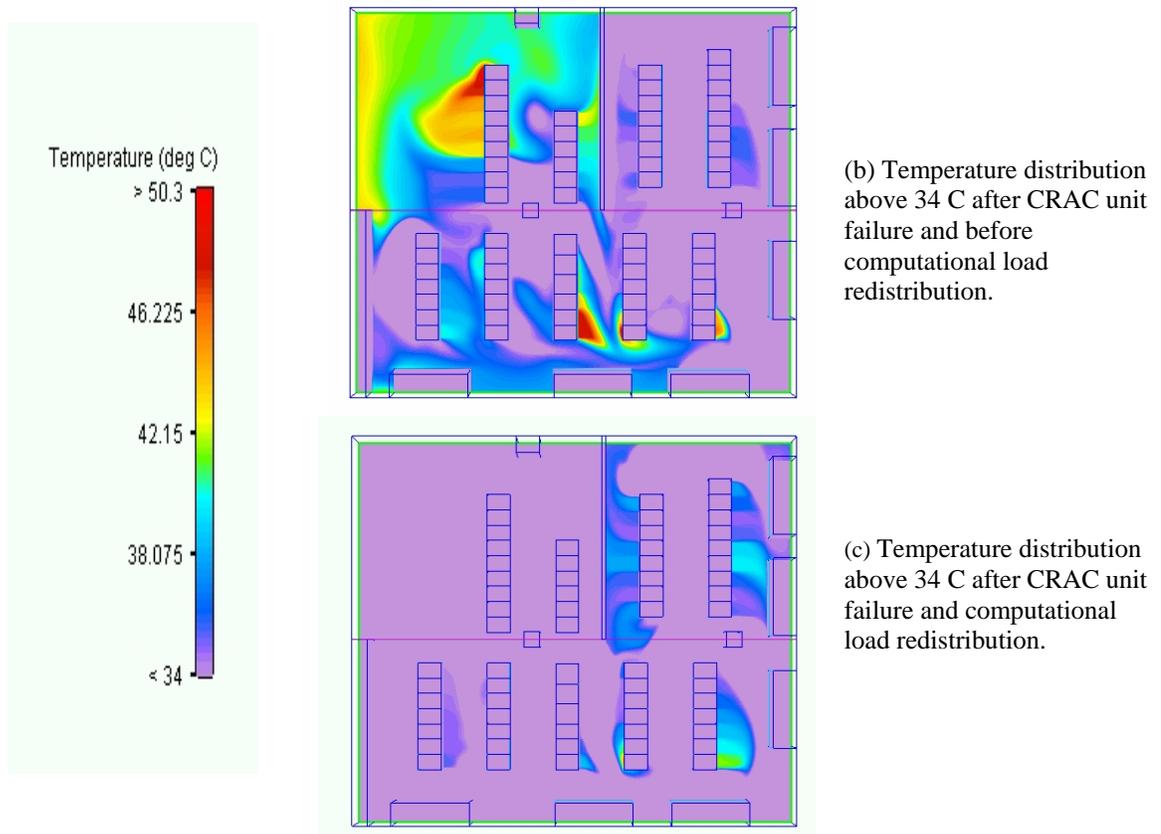(a) Temperature Plane Plot after CRAC unit failure with heat load redistributed.

(b) Temperature distribution above 34 C after CRAC unit failure and before computational load redistribution.

(c) Temperature distribution above 34 C after CRAC unit failure and computational load redistribution.

Figure 5.3. Temperature Plane Plots with load redistribution for failed CRAC operation.

5.3 Modular Construction and High-Level Control Policies

As seen in the previous example, a "smart" cooling system that can dynamically address each section of the data center is required for minimization of operating cost. However, such a system can be difficult to develop and implement for a large multi-component data center, since several dependent factors such as the geometric layout, physical system boundaries, CRAC settings and rack heat loads must be taken into account. Instead, smaller systems with fewer variables are more easily (and accurately) modeled; such "modules" also allow for improved control and commissioning at greater granularities. The data center should hence be viewed in a modular fashion. Fig. 5.4 shows such a 4 MW facility, which in itself can be one of many 4 MW facilities at the main switch gear level. The 4 MW facility is shown to service two 2 MW data center enclosures (~ 1 MW rack power + 1 MW cooling). Each data center enclosure is divided into four 240 KW modules containing 16 racks with maximum power of 15 KW each, served by adequate number of AC units. Fig. 5.4 (b) shows 2 AHU units with a capacity of 120 KW each. Additional AHU capacity maybe added to increase the level of redundancy, and additional modules can be added for expanded capacity based on projected growth.

One of the great advantages of such a modular approach is that it can uniformly be applied along each of the two vectors shown below:
1. systems, racks, rows and room(s) and working out to the cooling tower, and
2. cooling tower, cooling plant, hydronics, power feed, power conditioning, room(s), rows, racks, systems

The cost model developed in earlier sections of the report is then applicable for each module from systems out to the cooling tower(s) (1), and from cooling tower(s) to the systems (2). The cost model, while simple in nature, can be applied effectively by a person skilled in thermodynamics, heat transfer, fluid flow, mechanical design and experienced with practical equipment selection to create a "right" provisioned data center. The complexity and cost requires that the person(s) have field experience with data centers. This first order design can be followed by detailed modeling of the type shown in references [1, 5-9, 11].

4X Modules ~ 2 MW enclosure (compute & cooling power requirement)

20'

20'

GEN 1

~ 4 MW FACILITY

960 KW

GEN 2

CHILLER

PUMP ROOM

GEN 3

40'

800 sq ft MODULE

800 sq ft

POWER UPS FLYWHEEL

CENTRAL CONTROL

STAGING

800 sq ft
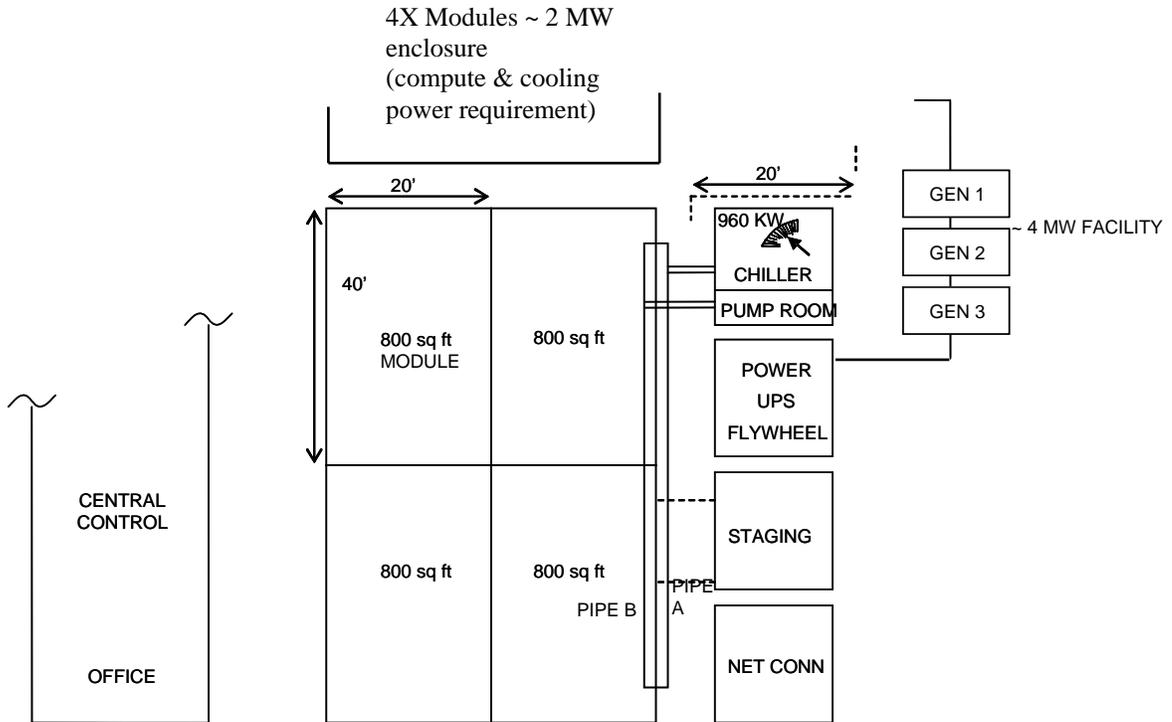
800 sq ft

PIPE A

PIPE B

OFFICE

NET CONN

Figure 5.4. (a) Plot plan: The compute area is assumed to be comprised of four 40' x 20' data center modules – 4 Modules form an enclosure, 4 enclosures form a facility around a Central Control Complex.

Module

4 Modules of 800 sq. ft form an Enclosure

Multiple Enclosures form a Facility

20'

Service Area

40'

Fan & Hex_DC   120 KW each

80-100 gpm

16 racks @ 15 KW
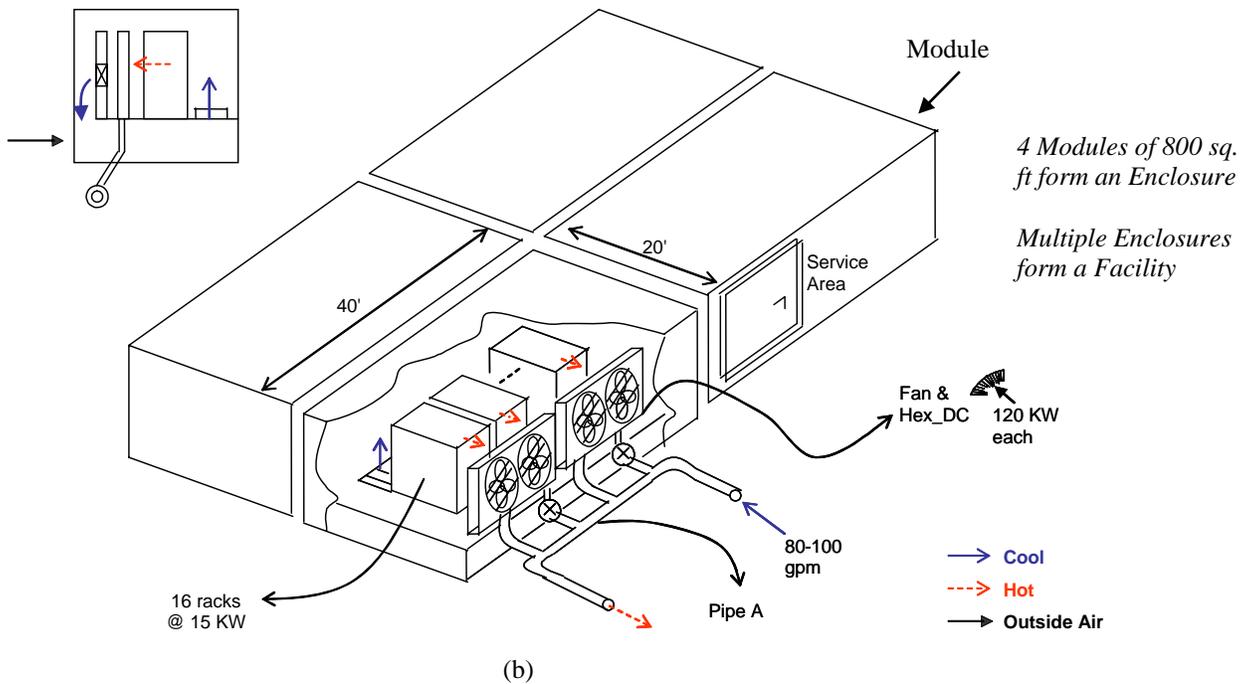
Pipe A

→ Cool
---→ Hot
→ Outside Air

(b)

Figure 5.4. (b) Schematic of a single 800 sq ft module, which contains the computer racks and air-conditioning units.

5.3.1 Exergy Destruction Method for Data Centers

In addition to these references, an exergy destruction (destruction of available energy) modeling technique based on the 2nd law of thermodynamics can pinpoint inefficiencies. As shown in references [22-27], quantification of exergy destruction is very powerful as it scales well along both the above vectors. Because exergy is defined with respect to a common ground state – which can be changed depending on the control volume being analyzed – an exergy-based approach offers a common platform for evaluation of different systems on a local as well as global basis. For example, by defining the external ambient as the ground state, the thermodynamic performance of a data center operating in New Delhi, India can be compared to a data center operating in Phoenix, Arizona. Alternatively, by defining the ground state as the plenum temperature inside a data center, the performance of the different components within the data center (racks, CRAC units and airflow patterns) can be comparatively evaluated. A detailed framework for using exergy in this manner to enable optimal component configuration inside the data center is being developed [25-27]. Additionally, by combining exergy loss, which measures thermodynamic performance of a system, with another metric – such as MIPS – that measures computational performance, it becomes possible to define a figure-of-merit based on MIPS/unit exergy loss that can be applied to the entire compute path from the chip to the data center. Such a metric would allow for identification of inefficiencies at each stage of computation, thus simplifying the decision-making process of how to best globally allocate resources for the i-th processor of the j-th system of the k-th rack of the m-th data center.

By relating the cost model developed in this paper to the above proposed figure-of-merit, a dollar estimate for inefficient operation in the data center can be achieved. An important realization in this work is that inefficient design or operation (as measured by the figure-of-merit) is likely to lead to higher personnel costs as well, i.e. the different components of the cost model in Eq. 4.5 are not necessarily independent. Thus, the cost model developed in this paper is an important step towards correlating economic inefficiencies in cost allocation with operational inefficiencies in resource allocation. A combined model that appropriately defines an exergo-economic metric would help to quantitatively identify those areas which require the most future research.

CHAPTER   6

Summary and Conclusions

As computing becomes a utility, a great deal of attention has centered around the cost of operation of a data center, with specific attention to power required to operate the data centers [28]. This report presents a comprehensive cost model for data center ownership and operation. Significant costs related to real estate, burdened cost of power delivery, personnel as well as software and licensing have been discussed together with examples throughout the paper to demonstrate application of the cost model for a variety of scenarios. Further, by defining the costs in terms of multiplication factors, the proposed model enables comparative evaluation of different data center practices and designs. The model penalizes inefficiencies in the data center infrastructure, including suboptimal usage of critical space, cooling resources and personnel allocation. Potential cost improvements from implementation of efficient cooling techniques, such as "smart" cooling, are explicitly addressed. Additionally, the model enables a direct comparison of capital expenditure (cost of equipment) versus operational expenditure. We find that both are of comparable magnitude; in fact, the operational expenditures associated with "burdened" cost of power delivery and cooling are typically larger than the expenses induced by depreciation of industry standard compute equipment.

The foundation proposed for cost- and energy-efficient design is a modular approach, where individual data center modules are designed and then integrated into a larger full-scale expansion plan. Optimization of the data center modules is contingent on optimal design of the power delivery and cooling subsystems, which is a subject of ongoing work at Hewlett Packard Laboratories. Of particular emphasis is the potential benefit that can be availed from rightly provisioning the cooling resources and enabling smart dynamic redundancy for minimal expense. As an example, a typical state of art data center with 100 fully loaded racks with current generation 1U servers (13 KW aggregate power) would require:

- o $1.2 million for cooling, $1.2 million for powering the servers per year in recurring cost of power, and
- o An additional $1.8 million per year burden would result from amortization and preventive maintenance cost of power and cooling equipment used to support the hardware in the data center.

These costs are shown in Fig. 6.1. For comparison, costs are also shown for a data center with dynamic smart cooling [1, 2] and a data center with both dynamic smart cooling and temperature-aware scheduling [9, 19]. Experiments at HP Labs have suggested that cooling costs can be cut in half with dynamic provisioning, i.e. $600,000 for the 1.2 MW data center. Additionally, we believe the amortization expenses on the cooling equipment would also decrease proportionally with proper application of modular construction and provisioning techniques. Thus, we estimate that a total of around $900,000 can be saved

annually per 1.2 MW module through optimization of cooling resources. In addition, the ability to scale processor and system power can add further flexibility in energy savings and cost reduction. As an example, based on available hooks such as multiple power states from industry standard processors, future systems deployment will result in a demand-based power dissipation profile in a data center. Dynamic smart cooling techniques coupled with temperature aware power scheduling algorithms provide immense opportunity to reduce cost – by as much as half, or a net savings of $2 million per annum. To realize such savings, high-level control policies for cost- and energy-efficient operation are under development at Hewlett Packard Laboratories [7-10].
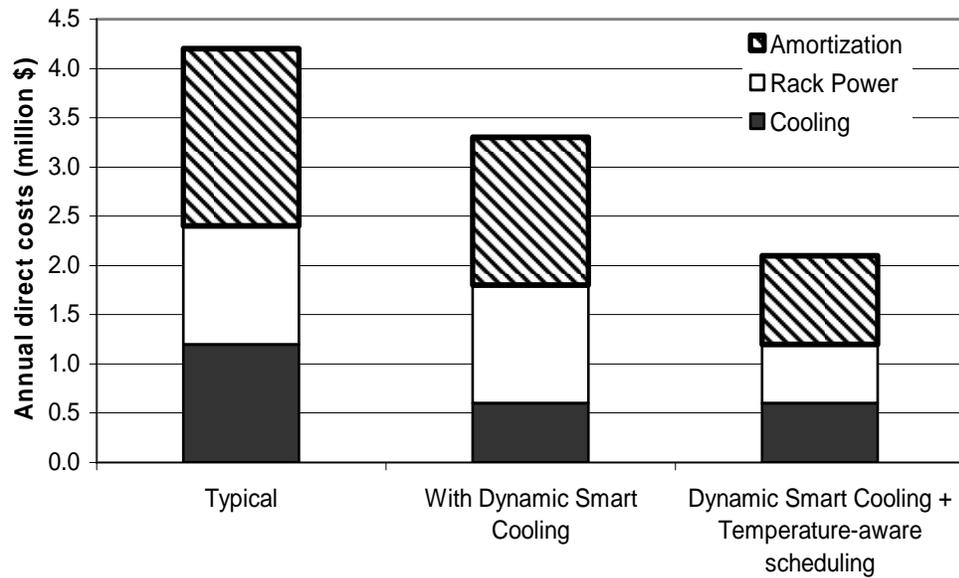

Figure 6.1. Distribution of costs in a 1-MW data center module as estimated from the cost model.

In conclusion, the cost model presented in this report suggests that the total annual cost of ownership for a fully optimized modular data center can be several million dollars lower than typical data centers found on the market today.

ACKNOWLEDGMENTS

REFERENCES

[1] Patel, C. D., Bash, C. E., Sharma, R. K, Beitelmal, A., Friedrich, R.J., 2003, "Smart Cooling of Data Centers", IPACK2003-35059, *The PacificRim/ASME International Electronics Packaging Technical Conference and Exhibition* (IPACK'03), Kauai, HI.

[2] Patel, C.D., Bash, C.E, Sharma, R.K., Beitelmal, A., Malone, C. G., "Smart Chip, System and Data Center enabled by Flexible Cooling Resources," *IEEE Semitherm, 2005*, San Jose, CA.

[3] Pienta, G. Van Ness, M., Trowbridge, E. A., Canter, T. A., Morrill, W. K., 2003, "Building a Power Portfolio," *Commercial Investment Real Estate* (a publication of the CCM Institute), March/April 2003.

[4] "Case Study: Fuel Cell Installation Improves Reliability and Saves Energy at a Corporate Data Center," 2002, *Energy Solutions for California Industry: Ways to improve Operations and Profitability*, Office of Industrial Technologies, U. S. Department of Energy and the California Energy Commission (May 2002).

[5] Patel, C. D., Sharma, R. K, Bash, C. E., Beitelmal, A., 2002, "Thermal Considerations in Cooling Large Scale High Compute Density Data Centers", *Eighth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems* (ITHERM '02), San Diego, CA.

[6] Patel, C. D., Bash, C. E., Belady, C., Stahl, L., Sullivan, D., 2001, "Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications," *The PacificRim/ASME International Electronics Packaging Technical Conference and Exhibition* (IPACK '01), Kauai, HI.

[7] Sharma R. K., Bash, C. E., Patel, C. D., 2002, "Dimensionless Parameters for Evaluation of Thermal Design and Performance of Large Scale Data Centers," AIAA2002-3091, *American Institute of Aeronautics and Astronautics Conference*, St. Louis, MO.

[8] Bash, C. E., Patel, C. D., Sharma, R. K., 2003, "Efficient Thermal Management of Data Centers – Immediate and Long-Term Research Needs," *International Journal of Heat, Ventilating, Air-Conditioning and Refrigeration Research*, April 2003

[9] Sharma, R. K., Bash, C. E., Patel, C. D., Friedrich, R. S., Chase, J., 2003, "Balance of Power: Dynamic Thermal Management for Internet Data Centers," *HPL-2003-5, Hewlett-Packard Laboratories Technical Report.*

[10] Patel, C. D., Sharma, R. K., Bash, C. E., Graupner, S., 2003, "Energy Aware Grid: Global Workload Placement based on Energy Efficiency," IMECE2003-41443, *ASME International Mechanical Engineering Congress and Exposition* (IMECE '03), Washington, DC.

[11] Beitelmal, M H., Patel, C. D., 2002 "Thermo-Fluids Provisioning of a High Performance, High Density Data Center," *HPL 2004-146 (R.1), Hewlett Packard Laboratories Technical Report.*

[12] "Natural Disasters in the Computer Room," 1997, *White Paper published by Intro Computer Inc,* available http://www.intracomp.com

[13] Sivitinades, P., Southard, J., Torto, R. G., Wheaton, W. C., 2001, "The Determinants of Appraisal-Based Capitalization Rates," *Technical report published by Torto Wheaton Research*, Boston, MA.

[14] "Special Report for those purchasing or Refinancing Income Property," *Report published by Barclay Associates*, available http://www.barclayassociates.com, last accessed online December 24, 2004

[15] Simpson, J., Simpson, E., 2003, "Cap Rate Follies," *Commercial Investment Real Estate* (a publication of the CCM Institute), July/August 2003. Available http://www.ccim.com/JrnlArt/20030811a.html, last accessed online January 4, 2005.

[16] Blount, H. E., Naah, H., Johnson, E. S., 2001, "Data Center and Carrier Hotel Real Estate: Refuting the Overcapacity Myth," *Technical Report published by Lehman Brothers and Cushman & Wakefield*, New York, NY.

[17] Evans, R., 2003, "The Right Kind of Room at the Data Center," Delta 2031, *White Paper published by MetaDelta (Meta Group)*, Stamford, CT.

[18] Personal communciation with Ron Kalich and Gautam Barua, Real Estate Workshop, February 2002, Rocky Mountain Institute Data Center Charette, San Jose, CA.

[19] Moore, J., Chase, J., Ranganathan, P., Sharma, R., 2005, "Making Scheduling "Cool": Temperature-Aware Workload Placement in Data Centers," *Usenix 2005*.

[20] Janakiraman, G.J, Santos J.R., Subhraveti, D., Turner, Y., 2005, "Cruz: Application-transparent distributed checkpoint-restart on standard operating systems," *Dependable Systems and Networks 2005*.

[21] Osman, S., Subhraveti, D., Su, G., Nieh, J., 2002, "The design and implementation of Zap: A system for migrating computing environments," *Fifth Symp. On Operating Systems Design and Implementation (OSDI)*, pp 361-376, Dec 2002

[22] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2003, "Exergy Analysis of Data Center Thermal Management Systems," *International Mechanical Engineering Congress and Exposition* (IMECE 2003-42527), Washington, DC.

[23] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2004, "Development and Experimental Validation of an Exergy-Based Computational Tool for Data Center Thermal Management," *Joint ASME/AIChE Heat Transfer/Fluids Engineering Summer Conference* (HTFED 2004-56528), Charlotte, NC.

[24] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2004, "An Open-System, Exergy-Based Analysis of Data Center Thermal Management Components," *International Microelectronics and Packaging Society (IMAPS) Advanced Technology Workshop on Thermal Management*, Palo Alto, CA.

[25] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2004, "An Exergy-Based Control Strategy for Computer Room Air-Conditioning Units in Data Centers," *International Mechanical Engineering Congress and Exposition* (IMECE 2004-61376), Anaheim, CA.

[26] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2005, "Exergy-Based Optimization Strategies for Multi-component Data Center Thermal Management: Part I, Analysis," *Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition* (IPACK2005-73137), San Francisco, CA.

[27] Shah, A. J., Carey, V. P., Bash, C. E., Patel, C. D., 2005, "Exergy-Based Optimization Strategies for Multi-component Data Center Thermal Management: Part II, Application and Validation," *Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition* (IPACK2005-73138), San Francisco, CA.

[28] LaMonica, M., "Google's secret of success? Dealing with failure," CNET News.com article, published March 3, 2005; accessed March 18, 2005. Available http://news.com.com/Googles+secret+of+success+Dealing+with+failure/2100-1032_3-5596811.html