Energy Flow in the Information Technology Stack: Coefficient of Performance of the Ensemble and its Impact on the Total Cost of Ownership

Chandrakant D. Patel, Ratnesh K. Sharma, Cullen E. Bash, Monem Beitelmal
Enterprise Software and Systems Laboratory
HP Laboratories Palo Alto
HPL-2006-55
March 21, 2006*

chip, system, data center, data center management, power, cooling, smart data center; sustainability, energy efficiency, data center metrics, TCO

The industry is in the midst of a transformation to lower the cost of ownership through consolidation and better utilization of critical data center resources. Successful consolidation necessitates increasing utilization of capital intensive "always-on" data center infrastructure, and reducing recurring cost of power. A need exists, therefore for an end to end physical model that can be used to design and manage dense data centers and determine the cost of operating a data center. The *chip core to the cooling tower* model must capture the power levels and thermo-fluids behavior of chips, systems, aggregation of systems in racks, rows of racks, room flow distribution, air conditioning equipment, hydronics, vapor compression systems, pumps and cooling towers or heat exchangers. As a first step in data center consolidation, the ensemble model must be able to characterize a given data center and its level of capacity utilization, controllability, and room for expansion. Secondly, the continuous operation of the data center management system demands that the ensemble model be programmable to create new "set points" for power and cooling based on current customer cost and performance needs. The overall data center management system, when bundled as a product, must result in a simple payback of 1 year by increasing data center utilization to 80% of rated capacity and through savings in recurring cost of power.

Therefore, economic ramifications drive a business need for the creation of an information technology management tool that can maximize the utilization of critical data center resources and minimize the power consumption. The creation of such an end to end management system that can sense and control a complex heat transfer stack requires a thermodynamics based evaluation model. Earlier work has outlined the foundation for creation of a "smart" data center through use of flexible cooling resources and a distributed sensing system that can provision the cooling resources based on the need. This paper shows a common thermodynamic platform which serves as an evaluation and basis for a policy based control engine for such a "smart" data center with much broader reach - *from chip core to the cooling tower*.

* Internal Accession Date Only
A portion of this paper will be presented at the International Mechanical Engineering Congress and Exposition, 5-10 November 2006, Chicago, IL, USA
Approved for External Publication

Energy Flow in the Information Technology Stack: Coefficient of Performance of the Ensemble and its Impact on the Total Cost of Ownership

Chandrakant D. Patel, Ratnesh K. Sharma, Cullen E. Bash, Monem Beitelmal
Hewlett-Packard Laboratories, Palo Alto, California
Phone: 650-857-7140; Fax: 650-857-7029
Email: {chandrakant.patel, ratnesh.sharma, cullen.bash, monem.beitelmal}@hp.com

Abstract

The industry is in the midst of a transformation to lower the cost of ownership through consolidation and better utilization of critical data center resources. Successful consolidation necessitates increasing utilization of capital intensive "always-on" data center infrastructure, and reducing recurring cost of power. A need exists, therefore for an end to end physical model that can be used to design and manage dense data centers and determine the cost of operating a data center. The *chip core to the cooling tower* model must capture the power levels and thermo-fluids behavior of chips, systems, aggregation of systems in racks, rows of racks, room flow distribution, air conditioning equipment, hydronics, vapor compression systems, pumps and cooling towers or heat exchangers. As a first step in data center consolidation, the ensemble model must be able to characterize a given data center and its level of capacity utilization, controllability, and room for expansion. Secondly, the continuous operation of the data center management system demands that the ensemble model be programmable to create new "set points" for power and cooling based on current customer cost and performance needs. The overall data center management system, when bundled as a product, must result in a simple payback of 1 year by increasing data center utilization to 80% of rated capacity and through savings in recurring cost of power.

Therefore, economic ramifications drive a business need for the creation of an information technology management tool that can maximize the utilization of critical data center resources and minimize the power consumption. The creation of such an end to end management system that can sense and control a complex heat transfer stack requires a thermodynamics based evaluation model. Earlier work has outlined the foundation for creation of a "smart" data center through use of flexible cooling resources and a distributed sensing system that can provision the cooling resources based on the need. This paper shows a common thermodynamic platform which serves as an evaluation and basis for a policy based control engine for such a "smart" data center with much broader reach – *from chip core to the cooling tower*.

Keywords

chip, system, data center, data center management, smart data center, control, energy, energy efficiency, sustainability, total cost of ownership, TCO, data center power consumption

Nomenclature

A	*Area (ft^2 or m^2)*
IT	*Information Technology Equipment Cost ($)*
J	*CRAC Capacity Utilization Factor (Rated heat extraction capacity/Actual heat extracted)*
K	*Amortization and Maintenance Factor for Power and Cooling*
L	*Cooling Load Factor (Power required by cooling resources/Power dissipated by compute hardware)*

M	*Number of personnel (typically a variety of IT personnel, and Facilities not part of maintenance)*
Pwr	*Power (W)*
Rk	*Number of Racks*
S	*Salary ($/person)*
U	*Cost of Power ($ per KWh or $ per KW-month for amortization estimate)*
$\dot{V}$	*Air Flow rate (m^3/s or CFM)*
σ	*Cost of Software Licenses ($)*
T	*Temperature (℃)*
Q	*heat load (W)*
$\dot{V}$	*Volume Flow (m^3/s)*
W	*Work (W)*
COP	*Coefficient of Performance*
$\dot{m}$	*mass flow, kg/s*
ζ	*efficiency, wire to fluid (flow work)*
η	*efficiency (thermodynamic work, motor, etc)*
C_p	*specific heat capacity of fluid (KJ/kg-K)*
T	*Temperature (°C)*
ρ	*density (kg/m^3)*
α	*Seeback Coefficient*
I	*Current (Amperes)*
R	*Resistance (Ohms)*
P	*Pressure (Pa)*
v	*Specific Volume (m^3/kg)*
Thr	*Thermal Resistance (°C/W)*

Subscripts

1,2	*Used to breakdown the costs such as power and cooing, or refer to instances*
n	*Polytropic Index*
Amort	*Amortization*
Maint	*Maintenance*
dep	*Depreciation*
critical	*Critical resource e.g. area occupied by racks in a data center*
blade	*Single blade module e.g. processor or storage blade*
cp	*chip scale*
sup-dev	*support devices on "blades", including system enclosure fans*
sys	*system enclosure level made up multiple blades*
r	*rack*
a	*air-side*
int	*interface*
b	*blower*
static	*static pressure in blower curve, enclosure, plenum*
cr	*Computer Room Air Conditioning Unit (CRAC)*
hydronics	*chilled water distribution system*
comp	*compressor*
p	*polytropic*
in	*inlet*

out	*outlet*
cond	*condenser*
evap	*evaporator*
ch	*chiller*
ref	*refrigerant*
ct	*cooling tower*
dc	*datacenter*
p	*pump*
$	*cost*
h	*hot side*
c	*cold side*

Motivation

Demands of greater processing speed and myriad functionalities have motivated the design of computer systems that enable the greater number of processors, and thus higher processing power, per rack volume. The standardization of compute equipment, together with the compaction of compute, networking and storage hardware, has fuelled the growth in data center based services. The demand for data center based computing services continues to grow as millions of new services become available, particularly in the field of mobile access devices. A modern data center is akin to a factory – the residents of the data center – the compute, networking, and storage hardware – require conditioned power and cooling to enable reliable operation. Fig. 1 highlights these key elements along with various ancillary components. The power delivery and cooling infrastructure in such a data center scales with the power demands of the compute hardware. The capital cost of power delivery and cooling resources, and the associated preventive maintenance costs, must be considered together with recurring cost of power to develop a complete understanding of the data center cost structure. One approach is to burden the recurring cost of power with the amortization and maintenance of the power delivery and cooling resources such as UPS, battery, backup power generation, chillers, pumps, etc shown in Figure 1 [1]. The monthly “burdened” cost of power so defined, and cost of critical data center space, can exceed the depreciation of the computing hardware found in the data center[1]. Therefore, the planning, design and operation of power delivery and cooling infrastructure, at required levels of redundancies to support a given service level agreement, is paramount from a cost point of view. Additionally, the rate of increase in power density in the data center is beginning to outpace the state-of-art in cooling technology [2]. This is exacerbated by the use of conventional design and operation techniques which reduce the effectiveness and efficiency of such technology at high heat loads. Consequently, computer manufacturers are faced with the choice of either limiting system performance in favor of reduced power consumption, or of providing customers with higher performance products that are impractical to deploy. A holistic model based approach, therefore, is necessary to efficiently provision the power and cooling resources to operate the data center in a cost effective manner. While earlier work has provided approaches in sensing and control, with underlying flexibilities in compute hardware power management and air conditioning resources, an overall model from chips to data centers is lacking [3]. Models that create a comprehensive thermofluids based description in the chip-system-datacenter stack, and that enable the expedient extraction of costs associated with various operational modes within this path, can become the foundation for an end to end management tool that can be used to make operational decisions.

Data Center Total Cost of Ownership

The total cost of ownership (TCO) of a data center and the potential savings in the recurring cost of power and better utilization of capital intensive resources can be quantified through the application of an appropriate cost model. Patel and Shah introduced a cost model that examines the burdened cost of power delivery that can be used to quantify the savings [1]. The model, as shown by equation 1, accounts for real estate expenses by use of standard appraisal techniques such as net operating income over capitalization rate in a given geography to determine cost per unit area ($A_{critical}$). It accounts for capital expense and maintenance of the redundant power and cooling infrastructure such as uninterruptible power supplies, generators, hydronics, chillers, pumps, air handling units, etc by applying "burdening" factors (K_1 and K_2 for power and cooling respectively) to the power consumed ($Pwr_{consumed\ hardware}$) by the compute, network and storage hardware in the data center. It applies the cooling load factor, L_1, to account for the amount of electricity used by the cooling equipment to remove a given watt of power from the compute hardware in the data center. The model uses standard electric grid pricing ($U_{\$grid}$). Furthermore, within the burdening factors K_1 and K_2, the model applies a factor (J) to account for utilization of expensive capital equipment. As an example a data center with 1000 KW of power and cooling capacity being used at a compute load of 100 KW would have a J factor of 10. Thus lowering of J results in savings by reducing the burdening factors K_1 and $K_{2.}$ and better realization of capital and maintenance costs.

$$Cost_{total} = \left(\frac{\$}{m^2}\right)\left(A_{critical}, m^2\right) + \left(1 + K_1 + L_1 + K_2 L_1\right) U_{\$,grid}\, Pwr_{consumed,hardware} + Rk\left(M_{total} S_{avg} + IT_{dep} + \sigma_1\right) \quad (1)$$

where K_1 and K_2 are function of J. And, J is the utilization of power and cooling resources and is represented as

$$J = \frac{Pwr_{rated}}{Pwr_{consumed}} \quad (2)$$

$$K_{1,2} = J\left(\frac{U_{amort-maint}}{U_{\$-grid}}\right) \text{ calculated on a monthly basis} \quad (3)$$

- *where $U_{\$-grid}$ is the cost of utility per month (at \$0.10 per KWh, it will be \$0.072 per Watt per month)*
- *and $U_{\$-amor-maint}$ is the ratio of amortization and maintenance cost per month to the rated power (so a data center with amortization and maintenance cost per month of \$10,000 with a rated capacity of 1,000,000 W will have a $U_{\$-amor-maint}$ of \$0.01 per month)*

The second half of the equation is used to determine software licensing and personnel cost, where M represents the number of personnel servicing a given rack Rk. S represents the salary, IT_{dep} represents the depreciation of compute equipment and σ_1 represents software licensing cost per rack.

Framework of the Model

In order to better understand the recurring cost of distributing cooling resources at all levels of the heat transfer path, the data center physical infrastructure requires a model that traces the energy flow path from chip core to the external ambient environment via the available heat transfer path. Furthermore, inefficiencies introduced due flow and thermodynamic irreversibilities along the flow path must be understood.

Energy Flow Path from Chip Core to the Cooling Tower

Figure 2 shows a schematic of energy transfer in a typical air-cooled data center through flow and thermodynamic processes. Figure 3 shows the cross section of a typical data center with modular air conditioning unit which extracts the heat from the data center hot air return.

At the chip scale, Figure 4 shows the heat transfer to the heat sink. A solid state active cooling scheme at the interface is examined to provide specificity to the example. The high power density at the chip level will likely require active heat transfer means at the heat sink interface. Alternatively, at the chip scale, as in spray cooling, heat may be transported to a fluid stream which may or may not undergo phase change. Heat is transferred from the heat sink from a variety of chips to the cooling fluid in the system e.g. air as a coolant enters the system and undergoes a temperature rise based on the mass flow driven by system fans and is exhausted out into the rack. Fluid streams from different servers undergo mixing and other thermodynamic and flow processes in the exhaust area of the rack. As an example, for air cooled servers and racks, the dominant irreversibilities arise from mixing and mechanical efficiency of air moving devices (blower power).

Referring back to Fig. 2, these fluid streams are subsequently exhausted out into the data center physical space and transfer heat among themselves by mixing or convection leading to further flow related irreversiblities. These streams (or some fraction of them) flow back to the modular computer room air conditioning units (CRACs) and transfer heat to the chilled water (or refrigerant) in the cooling coils. Heat transferred to the chilled water at the cooling coils is transported to the chillers through a hydronics network. The coolant in the hydronics network, water in this case, undergoes pressure drop and heat transfer until it loses heat to the expanding refrigerant in the evaporator coils of the chiller. Work is added at each stage to change the flow and thermodynamic state of the fluid. Development of a performance model at each stage in the heat flow path can enable energy efficient design, and subsequently, operation of the complete thermal management ensemble.

While, the example shown here refers to use of chilled water air conditioning units, the data center may have direct expansion CRAC units with air cooled or water cooled heat exchangers. Regardless, the performance model should be agnostic and be applicable to an ensemble of components for any environmental control infrastructure.

Model Framework

The main power consumers in the cooling infrastructure of Fig. 2 include compressors in the chillers (or CRAC units), chilled and cooling tower water pumps, cooling tower (or heat exchanger) blowers and air handling unit blowers. Total heat load of the datacenter is assumed to be a direct summation of the power delivered to the computational equipment via UPS and PDUs. The coefficient of performance (COP) of the datacenter, then, is the ratio of total heat load to the power consumed by the cooling infrastructure.

$$\text{COP} = \frac{\text{Total Heat Dissipation}}{(\text{Flow Work} + \text{Thermodynamic Work})\text{ of Cooling system}} = \frac{\text{Heat Extracted by Air Conditioners}}{\text{Net Work Input}} \qquad (4)$$

The data center COP can be separated into components that describe the operational efficiency of each level of the heat transfer path. An aggregate COP can be used to describe the ensemble. For example, the coefficient of performance at the system (or server level) is the ratio of power consumed by the server over the power required to drive the server-level cooling infrastructure. For refrigerated systems, power supplied to the cooling infrastructure will include compressor, blower and pump power. In case of liquid cooled systems, cooling infrastructure power demand will include pump and blower power.

A COP based model shown by Equation 2, has the advantage of being utilized with an instrumented ensemble (i.e. chip, system and data center). The heat load and subsequent work done by any element of the cooling system can be estimated or monitored in real-time during data center operation by an appropriate monitoring network. The use of distributed sensor networks to monitor environmental conditions and power consumption in a data center has been described by Patel et. al. [3]. Such a network has been used to effectively distribute cool air to equipment racks in a typical data center resulting in significant cost savings associated with CRAC operation [4]. Extension of such a network to include each element within the ensemble heat transfer path would reduce the dependence on potentially inaccurate estimates and result in further cost savings. The following section describes the proposed model and provides detail on the use of both a holistic sensing infrastructure along with appropriate relationships that can be utilized in absence of real-time data.

Description of the Ensemble

The overall model of the ensemble is comprised of several subsystems, namely, chips, systems, racks, datacenter, chiller and the cooling tower or heat exchanger. Figure 3 shows a cross section of a typical data center with rows of racks, CRACs and vent tiles. Each rack contains a collection of systems with processors. Each system contains a cooling infrastructure consisting primarily of heat exchangers and fluid movers. At current overall computer system heat loads, cooling by sensible heat gain with air as the working fluid is sufficient in most systems. In future high density systems, air may be replaced by water as a working fluid. At microprocessor level, based on a variety of future microprocessor functional floor plans, Patel discusses the inevitability of cooling means such as pumped liquid loop, spray, Rankine vapor compression cycles, etc., to overcome the temperature rise from chip to heat sink [5].

Chips:

The continuing trends of non-uniform power profile and high power density poses a severe challenge for traditional air cooling which is already plagued by large thermal interface and spreading resistance losses. As an example, a 100 W chip with local power density in excess of 200 W/cm^2 requires active micromechanical means to overcome the temperature rise from chip to heat sink. Patel et. al. examined various microprocessor organizations, and have shown an example 100W processor chip (see Fig.4) – containing a 50 W CPU core occupying an area of 5 mm by 5 mm on a 20 mm by 20 mm die – has a temperature rise of 46 °C from chip to heat sink given the state of art in chip to heat sink interfacing[]. Therefore, for a chip typical maximum temperature requirement of 85 ºC, the 46 °C rise from chip to heat sink necessitates a cooling solution that can maintain the heat sink at 40 °C in a 40 °C ambient – i.e. a "0" °C temperature rise from heat sink to air [6]. Hence, using this as an example, one can deduce that devices such as solid state thermo-electric modules or even vapor compression refrigeration modules will be required to move the heat from such a processor to a given heat sink i.e. active devices will be required to move the heat across the interface. Miniature thermoelectric coolers (TEC) can provide an effective negative thermal resistance with low heat spreading resistance. As shown in Figure 4, for a temperature difference of 15 ºC between the hot and the cold side of the TEC, the TEC module requires 30 W of power to remove the heat from the 100 W microprocessor [7].

Therefore, with reference to Fig. 4, the work required to remove the heat at chip scale, W_{cp}, is the sum of the power required at the interface W_{int} (TEC module) and the power required for a chip scale blower, W_b. As shown in Ref [7], the TEC module characteristics can be used to determine W_{int}. The work required to remove a given amount of heat for a TEC at the interface can be determined by:

$$W_{\text{int}} = N\left[(\alpha_h - \alpha_c)I(T_h - T_c) + I^2 R\right] \qquad (5)$$

where N is the number of TE couples, α is the Seebeck coefficient, I is the current and R is the resistance in each element. The subscripts h and c denote hot and cold side.

With respect to the active TEC interface, for a given thermoelectric geometry, an optimum current exists based on the hot and cold side temperatures[7]. The COP can be maximized by operating the TEC module at the optimum current level to satisfy given thermal management constraints. Cold side temperatures can also be controlled by traditional air cooling to improve COP of the package.

With respect to the chip heat sink, the product of chip scale heat sink pressure drop (ΔP_{cp}) and volume flow ($\dot{V}_{cp}$) determines the pumping power or fluid work required to move a given coolant (air) flow through the heat sink. The minimum coolant volume flow ($\dot{V}_{cp}$) for a given temperature rise required through the heat sink can be determined from the first law of thermodynamics (Eq. 6).

$$Q = \dot{m} C_p \left(T_{out} - T_{in}\right) \quad \text{where } \dot{m} = \rho \dot{V}_{cp} \tag{6}$$

As shown in Equation 7, the electrical power required by the blower is the ratio of the fluid work to the blower wire to air efficiency, ζ_b. The blower characteristic curve, showing static pressure (P_{static}) with respect to Volume flow ($\dot{V}$), can be superimposed with the efficiency curve and held in a "lookup table" to use in conjunction with a controller that strives to minimize W_b, and maximize the COP.

$$W_b = \frac{(\Delta P_{cp} \times \dot{V}_{cp})}{\varsigma_b} \tag{7}$$

In order to distinguish the blower used at chip level from one used at other levels, the heat sink blower work is represented as $W_{b\text{-}cp}$. The COP at chip level is then represented as

$$COP = \frac{Q_{cp}}{W_{cp}} = \frac{Q_{cp}}{W_{\text{int}} + W_{b-cp}} \tag{8}$$

An active chip interface driven by thermal inkjet assisted sprays, by contrast, can also be used to dissipate heat by phase change from specific high power density locations on the chip. Flow through micro-nozzles can be controlled via manipulation of the firing frequency of each chamber[8]. Micro-fluidic systems condense and recycle the vaporized fluid back to the reservoir to complete the spray cycle[6]. Such systems have been tested with water to perform at W_{int} of 6 W for the aforementioned 100 W simulated chip area in an open environment [8].

Systems

The state of art in systems design is centered around bladed architectures that consist of single board computers with one or more microprocessors, associated memory and power converters mounted in an enclosure as shown in Fig. 5a. The networking connections are hard wired on the back plane board, with redundant wired networking cables at the enclosure level. The enclosure contains multiple processor blades, storage blades, etc. The power supply is either part of the enclosure or the power is "bussed" to the blades enclosure from an independent rack mounted power supply. As evident from Fig. 5, a single blade can be 250 W resulting in a 2.5 KW system enclosure. The system blower at enclosure or blade level provides the required air flow to maintain approximately 15 °C caloric rise through the system for a given pressure drop. In some cases higher temperature rise, approximately 20 ºC, may be used in determining the volume flow rate through a system.

The power consumed by the blower, $W_{sys\text{-}b}$, can be used to define the COP for a system. The blower power can be determined by calculating the fluid work and using the blower efficiency curve as shown earlier for the chip heat sink. The pressure drop across various blades can be difficult to determine due to the complexity of the air flow path in a system enclosure. A representation that can approximate the flow resistance in each of the blade channels is necessary to create an expedient model. Blade modules, depending upon the functionality, have varying geometric layouts that effects the pressure drop (see Fig.5b).

The thermal performance of the blade module is a function of mass flow of coolant through the blade. The thermal performance can be defined as thermal resistance, Thr_{blade} ($^oC/W$), with respect to given mass or volume flow of coolant through the blade "volume".

In order to facilitate the determination of the blade thermal resistance:

1. Determine key point such as microprocessor ($T_{blade\text{-}cp}$), disc drive, etc as measures for temperature
2. Determine the heat dissipated by the blade, Q_{blade}, Watts
3. Determine the difference in temperature between key point(s) and inlet temperature to the blade ($T_{blade\text{-}inlet}$)
4. Thr_{blade} ($^oC/W$), then, is defined as:

$$Thr_{blade} = \frac{\left(T_{blade-cp} - T_{blade-inlet}\right)}{Q_{blade}} \tag{9}$$

The thermal resistance is a function of mass flow, and hence volume flow, of the coolant through the blade as shown in Fig 5c. The required volume flow, $(\dot{m}/\rho)_a$, determined based on the energy equation (Eq. 6) must be delivered by the blower against the given "volume" resistance of the blade.

The "volume" resistance is flow resistance due to the blade volume and is represented as pressure drop, $\Delta P_{blade,}$ in Pa. vs. Volume flow, $\dot{V}_{blade}$ in m^3/s. Thus, blade modules of different varieties can be represented as vivid "volume" resistances in an enclosure as shown by dashed lines in Fig. 5d.

The modeling and metrology associated with determination of "volume" and 'thermal" resistance for heat sinks is presented by Patel and Belady[9]. The authors have combined thermal resistance and volume resistance to introduce the "thermo-volume" resistance approach with the contention that the "thermo-volume" characteristics determined by the technique shown can be applied to system design. Bash and Patel have shown the application of "thermo-volume" resistance technique in system design [10]. The area is further taught in greater detail by Patel and Bash in a course on thermal management[11].

Thus, the thermo-mechanical attributes of blades can be shown as "thermo-volume" curves that characterize the flow resistance (N/m^2 or Pa), and thermal resistance ($^oC/W$) with respect to volume flow (m^3/s)[9][10][11]. These can be determined during the blade development using computational fluid dynamics tools or experimentally using calibrated air sources and heaters. The use of thermo-volume resistance curves (Fig. 5c. and 5d.) becomes the higher level abstraction at blade module level that can now be used to determine the volume flow needed (m^3 per sec) at a given pressure drop (Pa) and can be used to size air movers. The thermo-volume curves can be associated with variety of blade models, or as blades become standardized, a library of thermo-volumes curves can be selected to synthesize a system enclosure. The wire to air efficiency curves from the manufacturer of the air moving devices, ζ_b, in conjunction with the thermo-volume curves determine the operating point as shown in Figure 5e.

The operating point shows the volume flow required to remove a given amount of heat at a given pressure drop. The volume flow, $\dot{V}$, and pressure drop, ΔP, together with efficiency, ζ_b, applied to equation 5 determine blower power needed to remove heat from a system of blades. In this “thermo-volume” approximation, the minimum mass flow required to remove a given amount of heat from a blade was calculated using the energy equation. This does not negate the need to perform detailed convective analysis to determine the film resistance and conductive package analysis[11].

The COP of the system can then be determined as follows:

$$COP_{sys} = \frac{Q_{sys}}{W_{sys}} = \frac{\sum (Q_{cp} + Q_{\sup-dev})}{\sum (W_{cp} + W_{\sup-dev})} = \left(\dot{m}_a C_{p,sys}\left(T_{in,sys} - T_{out,sys}\right)\right) \Big/ \left(\sum_i \left(W_{cp} + W_{\sup-dev}\right)\right) \tag{10}$$

where Q_{sys} and W_{sys} denote the system heat load and the system power required for cooling, respectively. Q_{sys} is the sum total of heat dissipated from all the chips (Q_{cp}) and other support devices ($Q_{sup-dev}$) like memory, I/O, power converters and storage.

Depending on the system construction, as shown in Fig. 5a entitled “blade enclosure”., Q_{sys}, is the enclosure level power dissipation – the sum total of power dissipated by all the blades containing chips and support devices. $W_{sup-dev}$ is the blower(s) power required to transfer heat from all support devices and W_{cp} is the power required to transfer the heat from the chips. Therefore, W_{sys} can be thought of as the total power required to transfer heat generated in the blade system enclosure.

Racks:

Typically, racks do not have significant additional cooling. However, as the total rack power approaches 20 KW, augmentation in the form of additional air movers or liquid cooled heat exchangers is being introduced as shown in Fig. 6 [2]. Power consumed by locally mounted pumps, compressors and blowers can add up and reduce overall COP. At a general level, the same thermo-volume resistance abstraction used to characterize the system can be applied at rack level. The volume resistance can capture other elements of pressure drop, such as cables, outside the blade system enclosure. The power draw by air movers, used as augmentation, can be determined from the efficiency curves and the COP at rack level can be represented as:

$$COP_r = \sum_i Q_{sys} \Big/ \left(\sum_i W_{sys} + W_r\right) \tag{11}$$

where i indicates the i^{th} server. W_r represents the power consumed by the local pump or blower installed in the rack. In case of refrigerated racks, this could also include the power consumed to run the compressors.

Data Center

The datacenter model typically consists of an air distribution and chilled water distribution system but can also employ alternative wo rking fluid streams. The air distribution system distributes cold air to the racks through a system of blowers, CRAC unit heat exchanger coils (air-side), vent tiles and raised floor plenum (or other infrastructure). Figure 3 is a simplified representation of a conventional data center with under floor cool air distribution and room return infrastructure. The hot exhaust air from the EIA (Electronics Industries Association) racks is cooled by re-circulating the air through CRACs. Multiple CRAC units, sized to extract the total heat load in the data center, are located within the data center or in close proximity. A refrigerated or chilled water cooling coil in the CRAC unit cools the air to a temperature of approximately 10ºC to 17ºC. Other data center infrastructures with different air handling systems like ceiling supply/return also exist. Although CRAC units may also utilize direct expansion refrigeration, chilled water units have been used in this paper. The model developed here is applicable to CRAC units with vapor compression.

As shown in Fig. 7, the chilled water system includes a combination of secondary pumps, cooling coils (liquid-side), piping and valves. The hydronics system distributes chilled water to the cooling coils inside the various CRAC units to achieve the prescribed temperature setpoints. The chilled water circulation is modulated and controlled by a primary or a booster circulating pump depending on how far the data center is from the central chiller location. The pump is usually installed at the supply side of the chilled water. The chiller receives the warm water from the CRACs in the data center and transfers the heat to the circulating chiller refrigerant. The chiller refrigerant is compressed to a high pressure high temperature vapor where it condenses in the chiller condenser to a saturated liquid by transferring heat to the circulating condenser water. The warm condenser water circulates back to the cooling tower where the heat is transferred to the ambient air. The majority of the heat transfer in the cooling tower takes place by virtue of evaporative cooling.

Data Center Air Distribution:

Air flow distribution in data centers is complicated and non-trivial. Unbalanced flow patterns can lead to excessive energy consumption in the blowers and parasitic heat loads due to mixing of cold and hot air streams. The data center, with complex three dimensional air flow distribution, requires a combination of modeling and metrology to understand the utilization of CRAC units[12]. As a few CRAC units serve a multiplicity of asymmetrically distributed heat loads (racks) in the room, the level of utilization varies and has to be determined by air flow modeling. The modeling takes the form of computational fluid dynamics analysis[12]. The CFD analysis is used to assure system inlet temperature specifications[13], and determine the provisioning of the CRAC units 12. The analysis is used to initially deploy the systems in the data center, and can be followed by fine grained measurement using a distributed sensing network [3] for fine tuning the cooling system. The sensor data can be further used to upgrade the model, and to improve the COP of the CRACS resident in the data center. The CRAC COP can be calculated by determining the level of capacity utilization and manufacturer data on blower efficiency. An example of varying CRAC utilization level is shown in Fig. 8. This is a production data center at HP Laboratories in Palo Alto in a given "statically provisioned" state [14]. The optimum COP will be determined by the blower efficiency curve with respect to flow rate at the given utilization level in this example as the the CRAC is a chilled water unit and the only active heat removal element in the unit is the blower.

The COP of the air distribution system can be calculated on per CRAC unit basis to assess the impact of provisioning (utilization) variations and benefit from higher part load efficiencies.

$$\left(COP_{cr}\right)_l = Q_{cr}/W_{b-cr} \tag{12}$$

where Q_{cr} and $W_{b\text{-}cr}$ are cooling load and blower power consumption for l^{th} CRAC unit. As the data center dynamics changes so does the CRAC utilization and the sensor data is used to monitor the COP_{cr}. The analysis can be extended out to the hydronics, chillers and cooling towers or heat exchangers.

Data Center COP

In order to determine the COP of the data center, COP_{dc}, the work required to humidify and dehumidify must also be considered in addition to the heat added by CRAC blowers.

- $Q_{b\text{-}cr}$ represents the heat from the blower that must also be extracted in addition to the heat generated in the data center.
    - $Q_{b\text{-}cr}$ is calculated by subtracting flow work from the power consumed by the blower – the difference being the change in internal energy and frictional irreversibilities.
- $W_{humid\text{-}dehumid}$ represents the work required to periodically humidify or dehumidify the air in the data center

Therefore, to calculate the composite COP of the datacenter, total cooling load on all the units is divided by total power consumed by all the blowers in the datacenter.

$$COP_{dc} = \sum_{l} (Q_{cr} + Q_{cr-b}) \Big/ \sum_{l} (W_b + W_{humid-dehumid}) = Q_{dc} \Big/ \sum W_{dc} \tag{13}$$

Chilled water distribution (Hydronics)

Figure 7 shows the chilled water distribution system consisting of secondary pumps, cooling coil in the CRAC unit and associated system of piping and valves. The purpose of the system is to circulate chilled water from the chillers to the CRAC units. Maximum cooling should be provided at the CRAC units, with a minimum of pump power consumption.

The COP of the chilled water distribution system is given by

$$COP_{hydronics} = Q_{dc} \Big/ \sum \left(W_p\right)_{secondary} \tag{14}$$

where $\Sigma(W_p)_{\text{secondary}}$ represents the power consumed by all the secondary pumps (see Fig. 7)

Chillers

Chillers provide chilled water to the data center facility as shown in Fig. 7. Heat dissipated from the data center is rejected to the environment at the chillers through various refrigeration techniques like vapor compression, absorption, adsorption etc[15]. Neglecting heat gain by the room from the environment through walls, windows, etc and heat transfer gains and losses from pumps and piping, cooling load at the chiller (Q_{ch}) is identical to data center heat load (Q_d). Coefficient of performance of chillers is affected by the method of compression and heat rejection, load to capacity ratio, chilled water distribution system and the environmental conditions[15]. COP of the chiller is defined as

$$COP_{ch} = \frac{\text{Total Chiller Heat Extraction}}{\text{Power Consumed by Compressor}} \tag{15}$$

For chillers utilizing direct expansion, compressor maps can be used to determine COP as a function of critical operating parameters like evaporation and condensation temperatures and refrigerant mass flow rate [16]. Equation 7 describes the work done by a positive displacement compressor producing polytropic compression. Similar relationships exist for alternative compressor designs.

$$W_{comp} = \frac{\dot{m}_{ref} n P_2 v_2}{\eta_p \eta_{motor} (n-1)} \left[\left(\frac{P_3}{P_2} \right)^{(n-1)/n} - 1 \right] \tag{16}$$

The COP of the chiller is given by

$$COP_{ch} = \frac{Q_{ch}}{W_{comp}} \tag{17}$$

The total chilled water heat dissipation can calculated from the sensible heat load on the chiller.

- Mass flow sensor, and temperature sensors at the supply and return of the chilled water supply circuit, can be used with the caloric equation (Eq. 3) to determine the chiller heat load.
- Compressor, pump and blower power can be obtained from the chiller controls or manufacturer information on these devices. Power consumption for pumps and blowers can also be obtained from variable frequency drives, if available.

Cooling Tower

Proper selection and operation of cooling towers is key to improving chiller performance. Chiller cooling load combined with environmental and local factors dictate the operation and selection of cooling towers. Blowers account for a major portion of cooling tower power consumption, followed by circulation pumps. Power consumption to cooling towers can be obtained from blower and pump motor drives. Heat dissipated in the cooling tower can be obtained from the chiller controls or by measuring the sensible heat loss from the condenser cooling water.

Thus the total heat removed by the cooling tower can be determined by:

- Mass flow sensor, and temperature sensors at the supply and return of the cooling tower water supply circuit, can be used with the caloric equation (Eq. 3) to determine the cooling tower heat load.
- Power consumption for pumps and blowers can be obtained from manufacturer data or monitored from variable frequency drives, if available.

The COP formulation for the cooling tower can be represented as:

$$COP_{ct} = Q_{ct} / W_{ct} = \frac{Q_{dc} + W_{comp}}{W_{ct}} \tag{18}$$

where Q_{ct} is the heat loss from the condenser cooling water in the cooling tower and W_{ct} is the total power consumed by the cooling tower blowers and the condenser cooling water pumps. From first law of thermodynamics, Q_{ct} is equal to sum of Q_d and compressor work for chillers based on vapor compression refrigeration cycle along with blower and pump heat loads and other miscellaneous cooling components required to move the heat originating in the data center to the cooling tower per Fig. 7.

Representing the Ensemble

The grand COP for the overall ensemble based formulation for chips, systems and data center can be represented as:

$$COP_G = \frac{Q_{dc}}{\sum_k \left(\sum_j \left(\sum_i \left(W_{cp} + W_{\mathrm{sup}-dev} \right) \right) + W_r \right) + \sum_l W_b + \sum_m W_p + W_{comp} + W_{ct}} \qquad (19)$$

where

$$Q_{dc} = \sum_k Q_r \text{ for all racks and support equipment in the datacenter,} \qquad (19a)$$

$$Q_r = \sum_j Q_{sys} \text{ for all systems in a given rack,} \qquad (19b)$$

$$Q_{sys} = \sum_i \left(Q_{cp} + Q_{\mathrm{sup}-dev} \right) \text{ for all boards or subsystems} \qquad (19c)$$

Conventional COP is defined as Q_{dc} / W_{comp} ignores the power consumption from blowers, pumps and system blowers. Rewriting Eq. 13 we get:

$$COP_G = \frac{Q_{dc} / W_{comp}}{1 + \sum_k \left(\sum_j \left(\sum_i \left(W_{cp} + W_{\mathrm{sup}-dev} \right) \right) + W_r \right) \Big/ W_{comp} + \sum_l W_b \Big/ W_{comp} + W_p / W_{comp} + W_{ct} / W_{comp}}$$

$$COP_G = \frac{Q_{dc} / W_{comp}}{1 + A + B + C + D} \qquad (20d)$$

where A, B, C and D are ratios based on power consumed by the compressor at the chiller:

- "A" depends on power consumed by cooling infrastructure in chips, systems and racks,
- "B" depends on total power consumed by all the CRAC unit blowers,
- "C" depends on total power consumed by secondary and primary pumps, and
- "D" depends on power consumed by blowers and pumps at the cooling tower.

Minimization of A, B, C and D will lead to the overall COP of the ensemble being close to the conventional COP of the chiller. In most cases, due to high air flow needs of the systems and the consequent high air distribution demands in the datacenter, A and B can each be close to 1 if not more for datacenters. In this instance, the actual COP of the ensemble, becomes less than half of the conventional COP of the chiller. In case of small and dedicated datacenters in close proximity to chillers, C can be negligible. For large buildings and data centers, C can be significantly higher.

Applying the Ensemble Model in an Example Data Center

In order to apply the model, consider a simplified representation consisting of 10 rows of 10 racks with approximately 12 KW of maximum heat load each, as shown in Figure 6. The fully loaded racks are made up of dual IA-32 Intel Xeon 1U servers used for rendering application such as the Utility Rendering Service (URS) in HP Labs Smart Data Center. The 1U servers with integral hard drives are assumed to have two power states – idle and max – with idle being 60% of max power dissipation. A diurnal duty cycle assumes that rendering jobs are loaded at 6:00 P.M for an overnight 12 hour run at maximum load.

The rest of the time the systems are assumed to be idling. The racks are deployed front to front and back to back in the typical hot aisle-cold aisle layout, in symmetry with the CRAC units in the room. The total power consumption for computing has been assumed to vary between 1200kW and 700kW (~60% of full load) while rendering. Each CRAC unit has a capacity of 100 kW, and there are 20 units for a total heat extraction capacity of 2 MW. Individual CRAC capacity utilization varies depending on the distribution of heat load, flow asymmetry due to plenum blockages, etc.

- An example of this asymmetric loading is shown in Fig. 8 in the HP Labs Smart Data Center in Palo Alto [14]. The complex nature of fluid flow necessitates computational fluid dynamics modeling [11] to determine the thermofluids provisioning. Sensing and control can enable operation of CRACS in their optimal range[3]

The data center CRAC unit blowers consume a total of 200kW at full load. Figures 6 and 7 shows the a schematic of fluid delivery circuit and associated chillers and cooling tower assumed in the model. The chilled water delivery system comprising of primary and secondary pumps and piping consumes a maximum power of 56kW. The combined power consumption of cooling tower fans and condenser cooling pumps is 115kW at full load.

	Q_{sys} (W)	W_{sys} (W)	Q_{dc} (kW)	W_b (kW)	W_p (kW)	W_{comp} (kW)	Q_{ct} (kW)	W_{ct} (kW)
Idle	174	13	700	43	12	116	816	25
Max	300	60	1200	200	56	200	1400	115

* With respect to Q_{dc}, the heat addition from the blowers and work required for humidificiation and dehumidification has been ignored for simplicity

Three cases were studied:

1. **Base Case Data Center Management**: No variation in power consumed by system cooling system
    a. System cooling power consumption levels fixed at max load shown in Table 1 e.g. W_{sys} at max
    b. Blower , pump and cooling tower power consumption (W_b, W_p, W_{ct}) at max load setting shown in Table 1
    c. Compressor power, W_{comp}, in the chiller varied based on COP of the chiller which was assumed to be 6.

- The base case assumes the presence of a sophisticated compressor in the chiller with the ability to turn down its thermodynamic capacity as the heat load in the data center, Q_{dc}, varies from 1200 KW to 700 KW.

2. **Smart Data Center Management:** Variation in cooling power consumption at datacenter level.
    a. System cooling power consumption identical to that in Case 1 i.e. at max setting.
    b. Blower , pump and cooling tower power consumption varied with compute load
        i. Blower power consumption varied from 200kW to 43kW
        ii. Pump power consumption varied from 56kW to 12kW
        iii. Cooling Tower power consumption varied from 115kW to 25kW
    c. Compressor power consumption is identical to that in Case 1 i.e. varied with data center heat load.

3. **Smart Systems-to-Data Center Management**: Variation in cooling power consumption from chips to datacenter level.
    a. System cooling power consumption varied from 240kW to 52kW [6]
    b. Blower, pump and cooling tower power consumption varied with compute load as in Case 2.
    c. Compressor power consumption is identical to that in Case 1 i.e. varied with data center heat load.

The COP of the ensemble is calculated for all cases for varying computational loads between 1200kW and 700kW using Eq. 18d. Figure 9 shows the results of the case study for the three cases with changing compute load. Data centers with conventional control continue consuming power for cooling even when the computational load drops. With Smart Data Center control, the cooling infrastructure is modulated to reduce power consumption while maintaining appropriate thermal management levels. Point Q in fig. 9 shows that such measures can provide a power savings up to 291kW at an idle load of 700kW when compared to conventional case. However, this still does not modulate the cooling infrastructure within systems which still continues to consume significant power. Control of system cooling power can provide additional savings of 188kW at idle load for the case in study. This is shown by point P in Fig. 9.

- At full load the COP of the ensemble (COP_G) is 1.5 for a constant COP_{ch} (see eqn. 11) of 6.
- As the load reduces, cases 1, 2 and 3 lead to COP_G 0.9, 1.5 and 2.7 respectively.
    a. Case 1: At idle heat load (servers idling) of 700 KW, the power required by the cooling resources is 770 KW for the base case
    - Case 2: Upon controlling the data center blowers in CRACS, the power required by the cooling equipment is 420 KW
    - Case 3: Upon controlling the system cooling resources and the CRAC blowers, the power required by cooling equipment is 259 KW.

Therefore, full power management of cooling resources from chips to cooling tower triples the COP_G leading to a substantial savings in recurring cost of energy. In this example, 511 KW reduction in power use (difference between 770 KW and 259 KW) from applying chip to cooling tower control, will result in a savings of $440,000 per year based on recurring cost of power of $0.10 per KWh.

Impact on the Total Cost of Ownership

It is important to look at the impact on total cost of ownership as well to see if the expensive capital resources used to support the high power needs are being well utilized. The example requires 1200 KW of uninterruptible power for the hardware and an additional ~1200 KW of power for heat removal for a total data center power requirement of ~2.4 MW.

The example data center housing this rendering service has 4 MW of available power delivery, conditioning and back up equipment. Cooling equipment has been allocated 2 MW of power, and the hardware in the data center has been allocated 2 MW of power. The data center has two cooling towers rated at 300 tons (~1MW) each, and approximately 20 chilled water CRAC units at 100 KW each.

As an example, while cost of installation varies with degree of difficulty, etc, the following is a rough approximation for cooling resources for this 2 MW cooling need of the data center. The example is provided by way of reference and process, and the reader is encouraged to get new estimates:

Example Calculation

The capacity utilization, J in Equation 2, for power and cooling is:

$$J = \frac{Pwr_{rated}}{Pwr_{consumed}} \tag{2}$$

- J for cooling equipment is 1.7 (2 MW divided by 1.2 MW)
- J for power equipment varies with efficiency of the cooling system
    - At L_1 of 1 (1W of cooling for 1W of power dissipated), J is 1.7

Assuming \$1.6 million for capital cost for the 2 MW cooling capacity in the example. The amortization cost of the cooling resources per month for a 10 year period will be approximately \$13,300. Including maintenance(including on-call facilities personnel contract) of \$11,000 per month, the amortization and maintenance will be \$24,300 per month for the cooling resources.

For a 2 MW cooling infrastructure, the $U_{\$amort\text{-}maint}$ is \$0.012 per month. It is determined by dividing monthly amortization and maintenance cost by rated capacity i.e. \$24300 divided by 2,000,000 W as described in Eq. 3. Next, the ratio of amortization and maintenance of cooling equipment normalized to grid power cost on a monthly basis i.e. $U_{\$\text{-}amort\text{-}maint}$ of \$0.012 per month to $U_{\$\text{-}grid}$ of \$0.072 per W-month is

$$\frac{U_{\$-amort-maint}}{U_{\$-grid}} = \frac{0.012}{0.072} \approx 0.17$$

The 4 MW power distribution, including conditioning equipment, diesel generators, is estimated to cost \$4 million resulting in monthly amortization of \$31,000. Maintenance and including power monitoring as part of building management system, etc is estimated at \$25,000 per month. The total amortization and maintenance on a monthly basis is \$56,000 for this 4 MW power delivery and conditioning infrastructure.

For a 4 MW infrastructure, the $U_{\$amort\text{-}maint}$ is \$0.014/W per month (\$56000 divided by 4,000,000 W as described in Eq. 3). Next, the ratio of amortization and maintenance of cooling equipment normalized to grid power cost on a monthly basis i.e. $U_{\$\text{-}amort\text{-}maint}$ of \$0.014 per month to $U_{\$\text{-}grid}$ of \$0.072 per W-month is

$$\frac{U_{\$-amort-maint}}{U_{\$-grid}} = \frac{0.014}{0.072} \approx 0.2$$

Applying to Equations 2 and 3, and using a J of 1.7:

$$K_{1,2} = J\left(\frac{U_{\$-grid}}{U_{\$-amortization-maintenance}}\right) \textit{ calculated on a monthly basis} \tag{3}$$

$$K_1 = 1.7(0.2) = 0.34 \text{ (power)}$$

$$K_2 = 1.7(0.17) = 0.29 \text{ (cooling)}$$

The TCO is determined by Equation 2:

$$Cost_{total} = \left(\frac{\$}{m^2}\right)\left(A_{critical}, m^2\right) + \left(1 + K_1 + L_1 + K_2 L_1\right)U_{\$,grid}\, Pwr_{consumed,hardware} + Ra\left(M_{total} S_{avg} + IT_{dep} + \sigma\right)$$

The *burdened cost of power and cooling* per month is then represented by the second component of the TCO equation:

$$Cost_{power\&cooling} = \left(1 + K_1 + L_1 + K_2 L_1\right)U_{\$,grid}\, Pwr_{consumed,hardware} \tag{21}$$

In the power and cooling cost equation, L_1 is a function of COP_G

$$L_1 = f(COP_G) = \frac{1}{COP_G} \tag{22}$$

and the burdened cost of power and cooling may be approximately represented as:

$$Cost_{power\&cooling} = \left(1 + K_1 + \frac{1}{COP_G} + K_2 \frac{1}{COP_G}\right)U_{\$,grid}\, Pwr_{consumed,hardware} \tag{23}$$

In equation 21, applying K_1 and K_2, at different COP_G and utilization, shows that on a monthly basis:

- At L_1 of 1, and J of 1.7, burdened cost is 2.5 times that of power consumed
- At L_1 of 0.3, and J of 1.7, burdened cost is approximately 1.7 times the power consumed

Increasing the data center resource utilization, by adding 30 additional racks totaling 360 KW, will result in J of 1.28, and K_1 and K_2 of 0.26 and 0,22 respectively from Eq. 3.

Thus,

- At L_1 of 0.3, and J of 1.28, burdened cost is approximately 1.6 times the power consumed.

While the fluctuation of power on a monthly basis is not factored into Eq. 21, the approximation and the equation serves as a good structure for understanding the impact of COP_G on TCO. The improvement in space utilization, resulting in improved J, has a positive effect as it leads to better utilization of amortization and maintenance expenses.

Summary and Conclusions

The paper has tracked the energy flow path in an IT data center, from chips to heat exchanger or cooling tower, and created a simple formulation to evaluate the performance. The performance evaluation criteria is based on the coefficient of performance (COP) metric commonly used in thermodynamics textbooks. The authors have extended the COP definition, and created a unified model called COP_G - or coefficient of performance of the ensemble. The COP_G formulation enables one to track the performance by having adequate instrumentation along the energy flow path. The COP_G can now be used as a basis to operate the data center in a variety of modes and to compare data centers. In addition to the COP_G representation as shown, one normalized to chiller COP in a given data center, other algebraic representations may also be derived and applied. The COP_G representation was used to study a hypothetical example data center with 100 racks. The racks were assumed to be idling at 60% of maximum power for half a day and running at maximum power for the other half. The chips to data center level control resulted in a COPG approximately a factor of 3 better than the one which relied on "chiller" turn down alone. The net energy savings was found to be approximately 511 KW, a savings of $440,000 per year based on recurring cost of power of $0.010 per KWh. Electricity pricing of $0.010 per KWh is conservative. In locations such as Hawaii, the electricity rates of $0.20 per KWh could yield a savings of $880,000 per year.

An application of COP_G to total cost of ownership has also emerged in this paper. The inverse of COP_G is found to be the "cooling load factor – L_1" in the TCO equation (Eq.3) developed by Patel and Shah. And, thus, the application of COP_G enables the TCO equation to be extended further and rewritten as shown below.

$$Cost_{total} = \left(\frac{\$}{m^2}\right)\left(A_{critical}, m^2\right) + \left(1 + K_1 + \frac{1}{COP_G} + K_2 \frac{1}{COP_G}\right) U_{\$,grid} Pwr_{consumed,hardware} + Ra\left(M_{total} S_{avg} + IT_{dep} + \sigma\right)$$

In very simple terms, using the same hypothetical example data center, better utilization resulted in approximately 30% savings from the burdened cost of power point of view – the second component of the TCO equation. The savings resulted in better utilization of critical resources and improvement in K_1 and K_2.

There are many inter-dependencies in the TCO equation. As an example, improvement in COP_G enables one to improve the power utilization as the cooling resources require less power. This will be explored in subsequent papers. Lastly, this model can be the basis of a high level evaluation engine used to enable control, a proposition of enabling Smart Chip, Smart System and Smart Data Center using flexible power and cooling resources[18].

References

[1] Patel, C.D., Shah, A., “Cost Model for Planning, Development and Operation of a Data Center”, HP Laboratories Technical Report, HPL-2005-107R1, 2005, http://www.hpl.hp.com/techreports/2005/HPL-2005-107R1.html

[2] Schmidt, R., Kamath, V., Chu, R., Lehman, B., Ellsworth, M., Porter, D., Madhu Iyengar, “Maintaining Datacom Rack Inlet Air Temperatures with Water Cooled Heat Exchangers”, ASME International Electronics Packaging Technical Conference and Exhibition, July 2005.

[3] 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.

[4] Bash, C.E, Patel, C.D., Sharma, R.K. “Dynamic Thermal Management of Air Cooled Data Centers” HP Laboratories Technical Report, HPL-2006-11, http://www.hpl.hp.com/techreports/2006/HPL-2006-11.html

[5] Patel, C.D., 2000 “Enabling Pumped Liquid Loop Cooling: Justification and the Key Technology and Cost Barriers”, International Conference and Exhibition on High Density and Systems Packaging, Denver, CO, pp. 145-152.

[6] Bash, C.E, Patel, C.D., Beitelmal, A. Burr,R., “Acoustic Compression for the Thermal Management of Multi-Load Electronic Systems” Itherm 2002, San Diego, California, USA

[7] Bierschenk, J., Johnson, D., “Extending Limits of air cooling with thermoelectrically enhanced heat sinks”, 2004 Inter Society Conference on Thermal Phenomena, pp 679-684, 2004

[8] Bash, C., Patel, C.D., Sharma, R., “Inkjet assisted spray cooling of electronics, Advances in Electronic Packaging, American Society of Mechanical Engineers, Maui, HI, Vol. 2, pp. 119-127, 2003.

[9] Patel, C.D. and Belady, C.L., "Modeling and Metrology in High Performance Heat Sink Design"' 1997 47th IEEE Electronics and Components Conference, San Jose

[10] Bash, C. E. and Patel, C. D., "Modeling and Metrology for Expedient Analysis and Design of Computer Systems", The International Journal of Microcircuits and Electronic Packaging, Volume 22, No 2, 1999"

[11] Patel, C.D., Bash, C.E., “Thermal Management of Electronic Systems and Packages”, Santa Clara University, Course Notes, ME 310, Spring 2006

[12] Patel, C.D., Bash, C.E., Sharma, R.K. and Beitelmal, A., “Thermal Considerations in Cooling Large Scale High Compute Density Data Centers”, 8th Intersociety Conf. Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), San Diego, CA, May 2002

[13] Patel, C.D., Bash, C.E., Belady, C., Stahl, L., Sullivan, D., “Computational Fluid Dynamics Modeling of High Compute Density Data Centers to Assure System Inlet Air Specifications”, July 2001, Proceedings of IPACK’01 – The PacificRim/ASME International Electronics Packaging Technical Conference and Exhibition, Kauai, Hawaii.

[14] Beitelmal M. H. and Patel, C. D., "Thermo-Fluid Provisioning of a High Performance High Density Data Center," Accepted for publication in the Distributed and Parallel Databases Journal-Special issue on high density data centers, 2005.

[15] Patel, C.D., Sharma, R.K, Bash, C.E., Graupner, S., “Energy Aware Grid: Global Workload Placement based on Energy Efficiency”, *ASME Intl. Mech. Eng. Cong. RD&D Expo (IMECE),* IMECE2003-41443, Washington, D.C., Nov 2003.

[16] Bash, C. E., 2001, “Analysis of Refrigerated Loops for Electronics Cooling”, IPACK, July 2001, Kauai, HI.

[17] Saad, S.A., *Thermodynamics: Principles and Practice*, Prentice Hall, 1997

[18] 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.

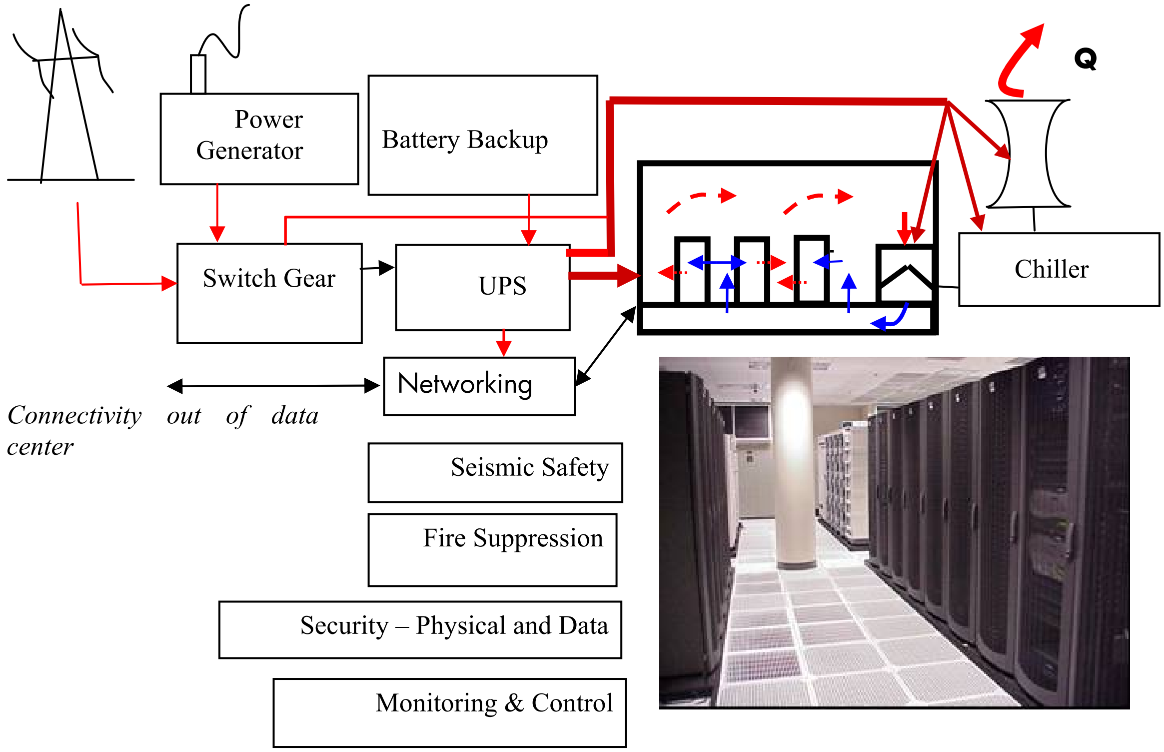


Figure 1. Key Elements of a Data Center

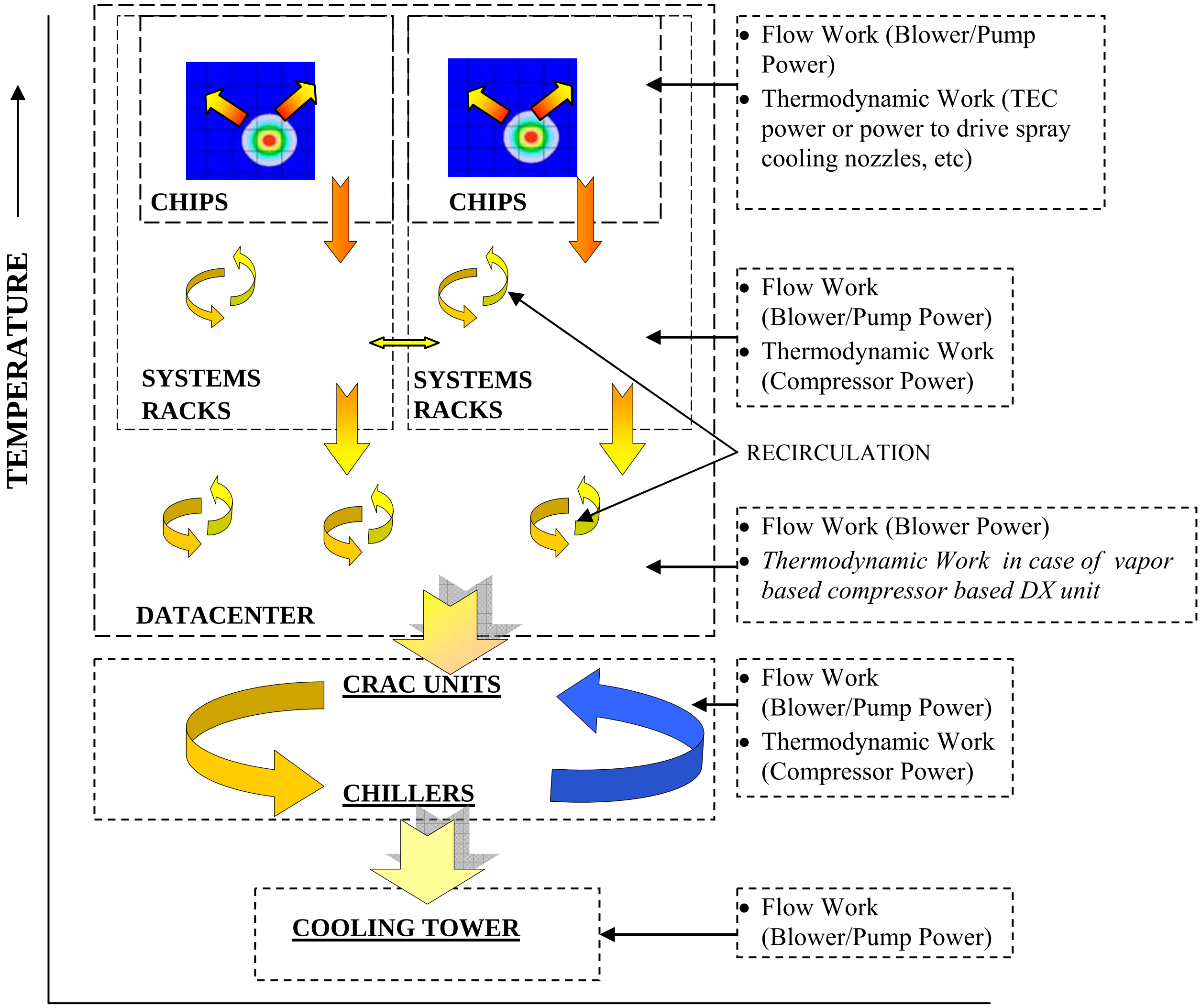


Figure 2: Schematic of Energy Flow from chips to data center with work done at each stage.

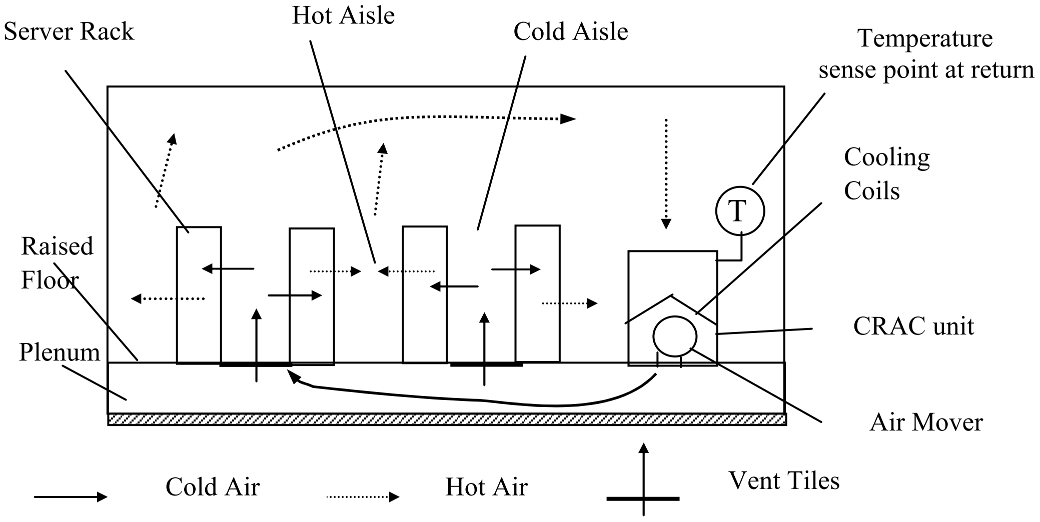
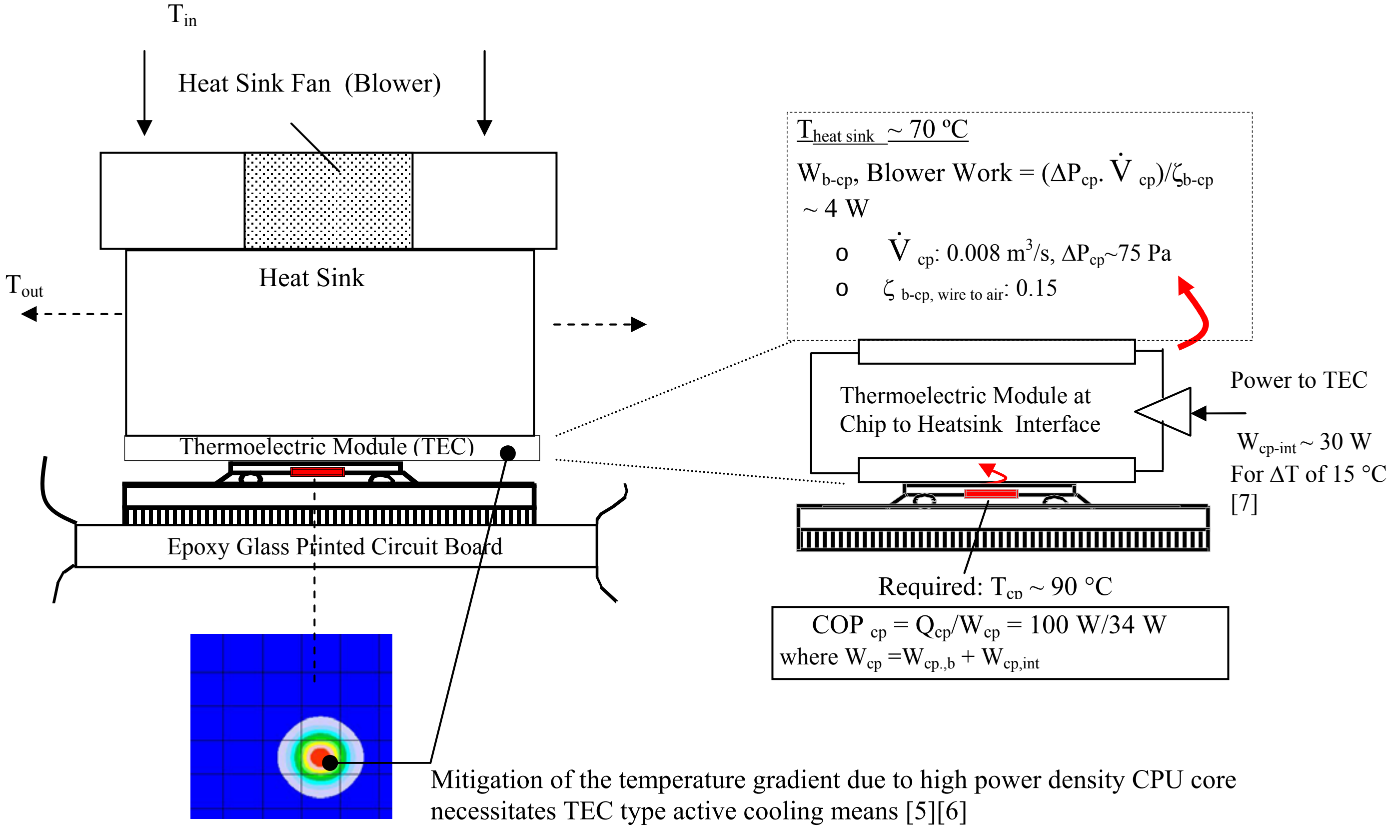
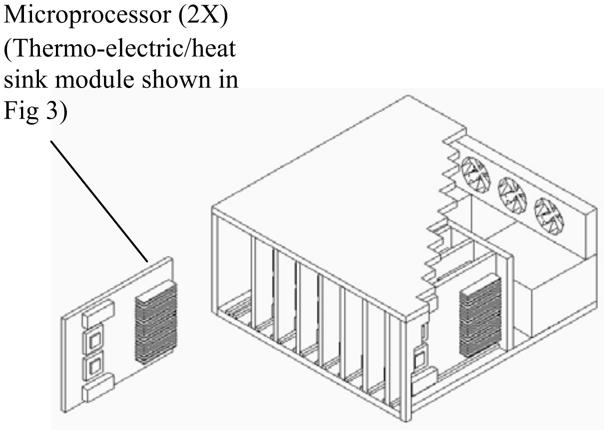


Figure 3. Internal Cross Section of a Data Center with under floor air plenum



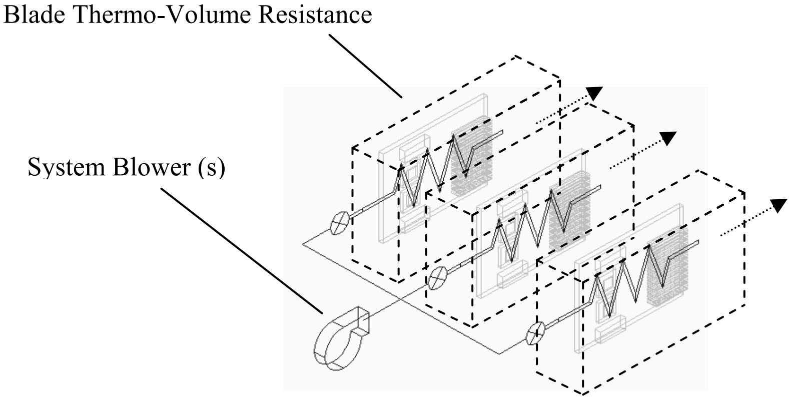
Mitigation of the temperature gradient due to high power density CPU core necessitates TEC type active cooling means [5][6]

Figure 4. Example Chip Scale Cooling Solution and Power Required to Remove Dissipated Heat



- 10 Blades at 250 W each
- 2.5 KW enclosure
- Enclosure level volume flow of approximately 1.25 m^3/s for a 15 °C rise at sea level

Fig 5a. Blade Enclosure (System)



- Equivalent Thermo-volume resistance used with blower curve to determine the volume flow, V’
- Blower Power: $(\Delta P.V')/\zeta_b$
    - where ζ_b is blower wire to air efficiency from the blower characteristic curve

Fig 5b. Blade Thermo-mechanical Model

Figure 5. System “Blade” Scale Cooling Solution and Blade Thermo-Volume Resistance

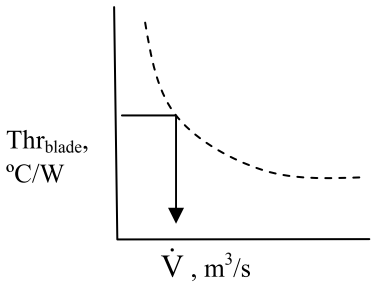


Fig 5c. Thermal Resistance

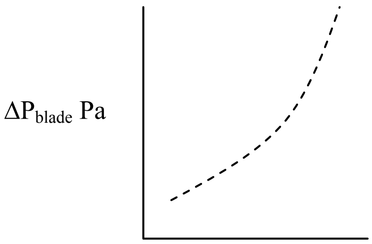


Fig 5d. Volume Resistance

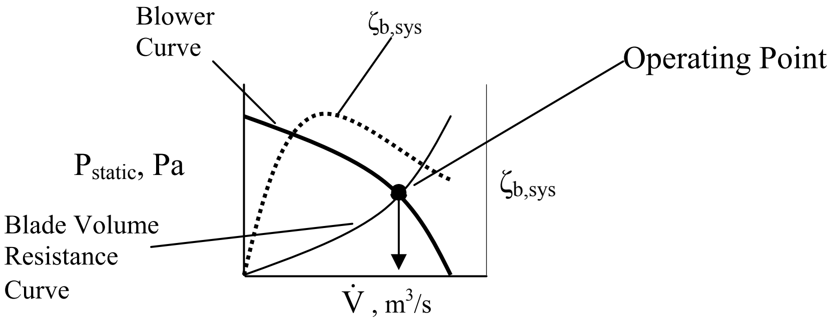


Fig 5e. Blower Characteristic Curve

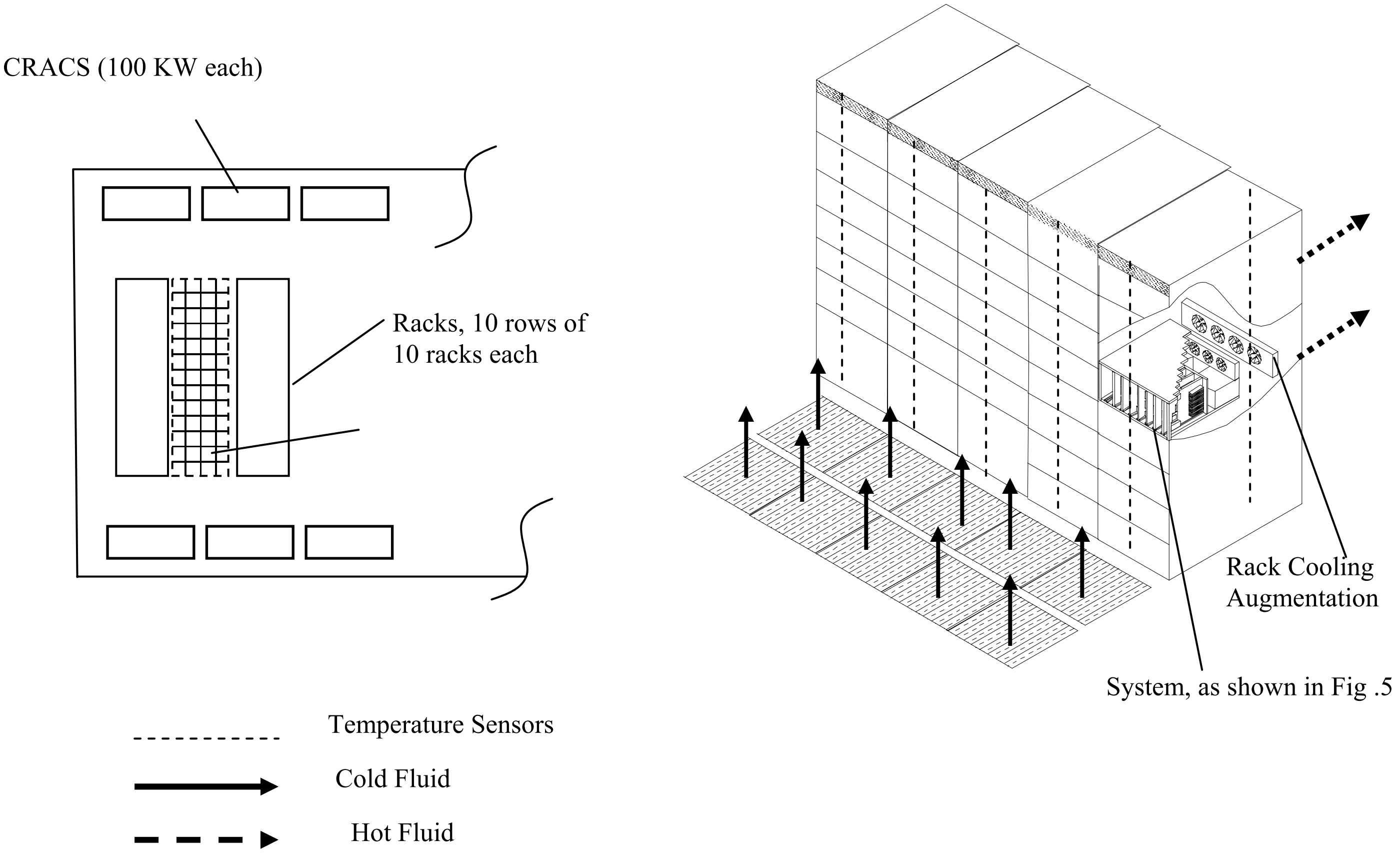


Figure 6: Example Case Study

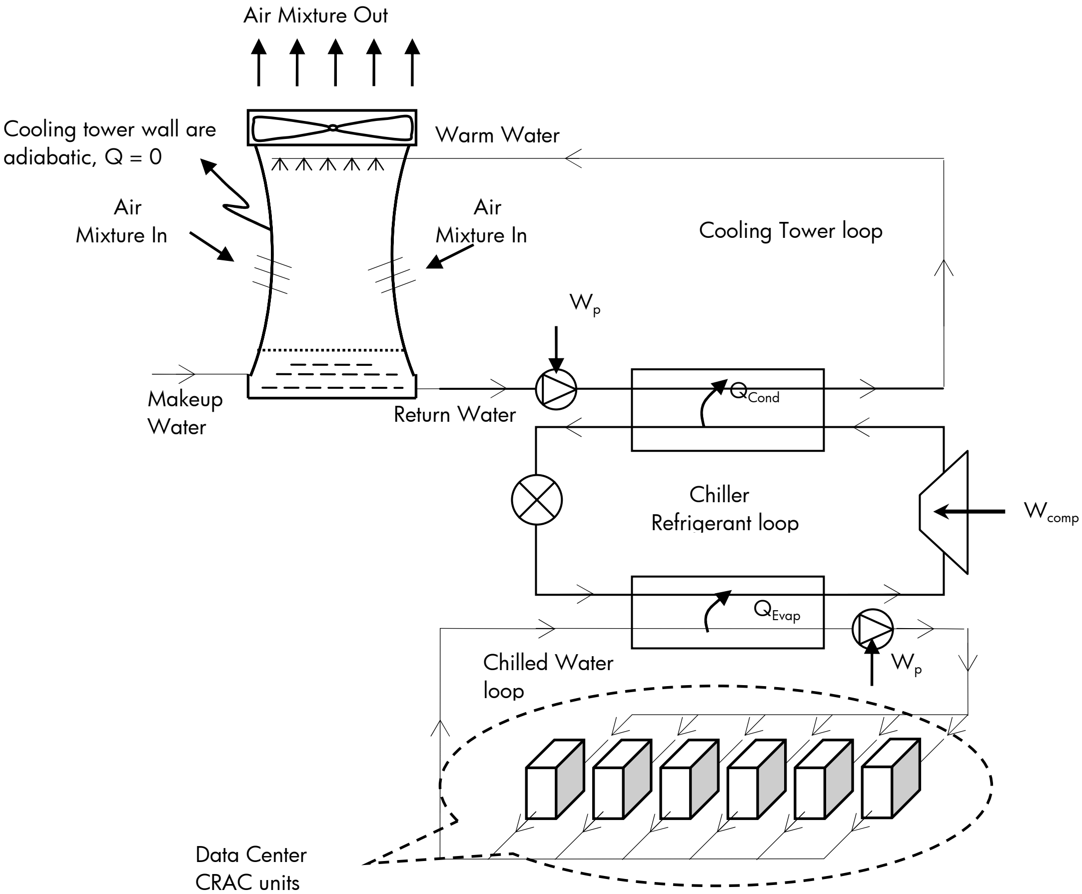


Figure 7. Data Center – Room CRACs to Cooling Tower

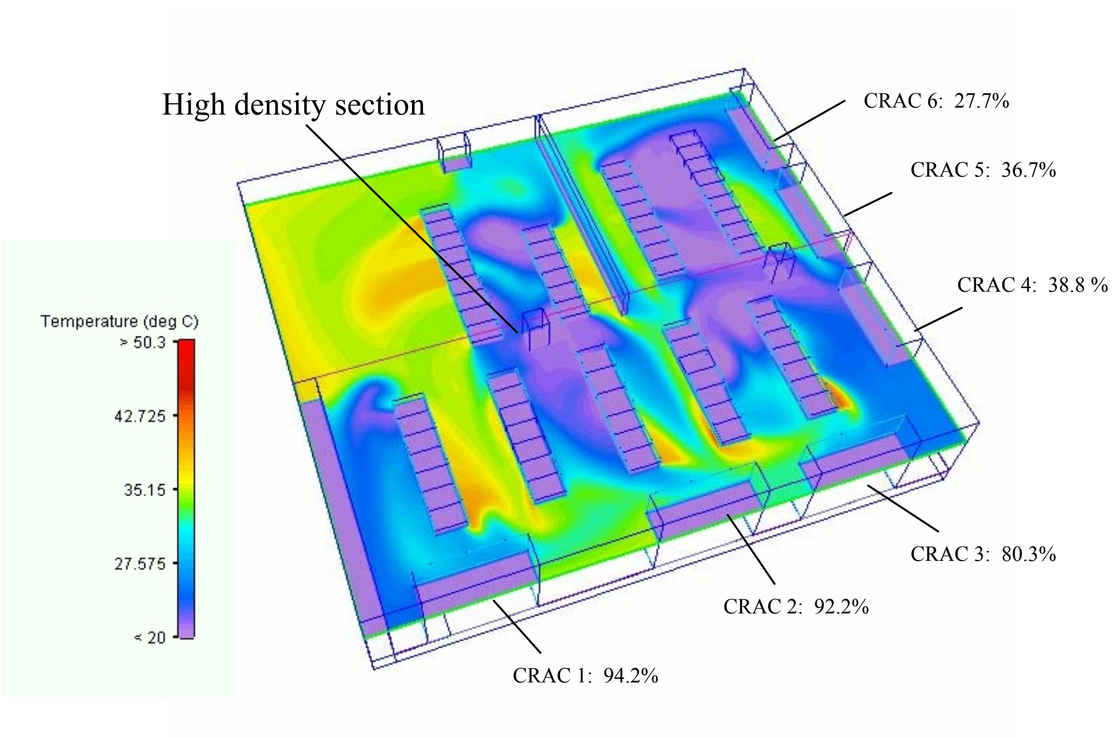


Figure 8. Given Provisioned State in a Data Center showing Temperature and CRAC Utilization States[14]

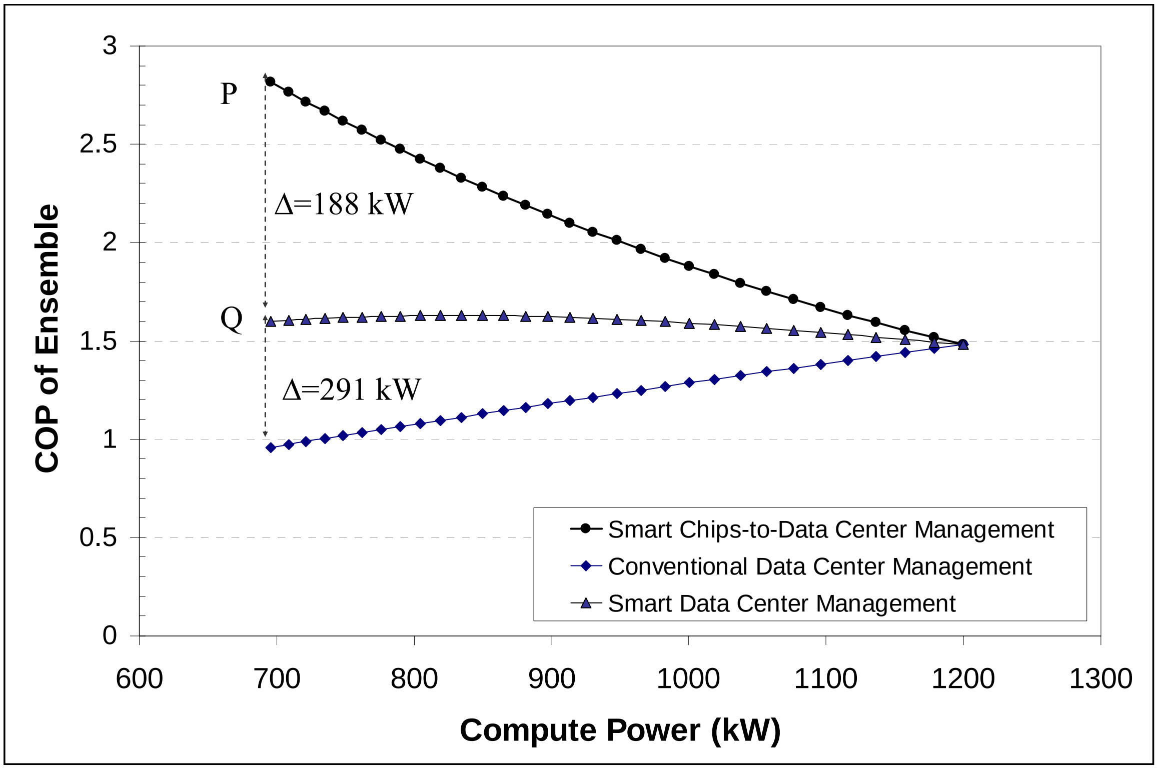


Figure 9: COP_G vs. Compute Power