Water Efficiency Management in Datacenters: Metrics and Methodology

Ratnesh Sharma, Amip Shah, Cullen Bash, Tom Christian, Chandrakant Patel, *Fellow, IEEE*

***Abstract*— The demand for data center solutions with lower total cost of ownership and lower complexity of management is driving the creation of next generation datacenters. The information technology industry is in the midst of a transformation to lower the cost of operation through consolidation and better utilization of critical data center resources. Successful consolidation necessitates increasing utilization of capital intensive "always-on" data center infrastructure, reduction in the recurring cost of power and management of physical resources like water. A 1MW data center operating with water-cooled chillers and cooling towers can consume 18,000 gallons per day to dissipate heat generated by IT equipment. However, this water demand can be mitigated by appropriate use of air-cooled chillers or free cooling strategies that rely on local weather patterns. Water demand can also fluctuate with seasons and vary across geographies.**

Water efficiency, like energy efficiency is a key metric to evaluate sustainability of the IT ecosystem. In this paper, we propose a procedure for calculation of water efficiency of a datacenter and provide guidance for a management system that can optimize IT performance while managing the tradeoffs between water and energy efficiency in conventional datacenters.

***Index Terms*—water resource management, data centers, sustainability, water efficiency**

I. INTRODUCTION

THE design and operation of the data center infrastructure is one of the primary challenges facing IT organizations. Unprecedented growth in demand for IT services has led to development of large, complex, resource-intensive IT infrastructure to support pervasive computing. Emerging high-density computer systems and consolidation of IT resources into fewer data centers are stretching the limits of data center capacity [1] in terms of power and resource utilization. The large number of components in a data center including cooling systems, power systems, and computer systems and the diversity of these components make data center design and operation complex.

A 1MW data center operating with water-cooled chillers and cooling towers can consume 18,000 gallons per day to power and cool IT equipment [2]. However, this water demand can be mitigated by appropriate use of air-cooled chillers or free cooling strategies that rely on local weather patterns. Water demand can also fluctuate with seasons and vary across geographies. To improve customers' RoIT (Return on Information Technology) [3], it is critical to maximize the resource utilization efficiency of the data center and simplify its management.

II. BACKGROUND

A. Overview

Water is a key component in the physical resource mix that powers datacenters. Water is consumed not only for direct dissipation of heat generated by IT equipment but also indirectly, for generation of electrical power for data center operation. Water withdrawal for power generation accounts for 48% of total water withdrawal in the US [4]. Water withdrawal rates vary based on power generation technology with a large fraction of the withdrawal often occurring to meet cooling requirements related to condensation of steam in the power plant. Additional water may be needed for carbon sequestration processes and combined heat and power applications. Water consumption rates also depend on the power generation and heat rejection technology in use in the datacenter.

In addition to being used in power generation, water may also be used in datacenters in wet cooling towers. Air-cooling systems are generally considered as a standard replacement for reducing water consumption, but other options also exist based on geography and weather. For example, outside air cooling - if used at the right conditions - can also provide alternative means of dissipating heat from datacenters at reduced water consumption levels. Water usage can also be reduced by replacing wet cooling towers with ground-coupled heat rejection systems that involve dissipating heat from the datacenter directly into the ground without the need for a chilled water circulation loop.

B. Data Center Infrastructure

Fig. 1 shows the basic data center building blocks from utility grid to the cooling tower [5][6]. Switchgear comprised of transformers and static switches with associated panels distributes power to the cooling infrastructure and the IT infrastructure. The cooling infrastructure is comprised of chillers, cooling towers, computer room air conditioning (CRAC) units and primary/secondary pumps. IT infrastructure includes servers, network devices and storage devices housed in standard racks. UPS maintains power quality during normal

---

All authors are with Hewlett-Packard Laboratories, Sustainable IT Ecosystem Laboratory, Palo Alto, CA 94304 USA (e-mail: *Firstname.Lastname*@hp.com).

R. K. Sharma is corresponding author (phone: 650-857-3835; e-mail: ratnesh.sharma@hp.com).

operation and provides energy storage to operate the IT infrastructure during brown-outs or short power outages. Chillers provide chilled water to the data center room that houses the server racks and other IT equipment.

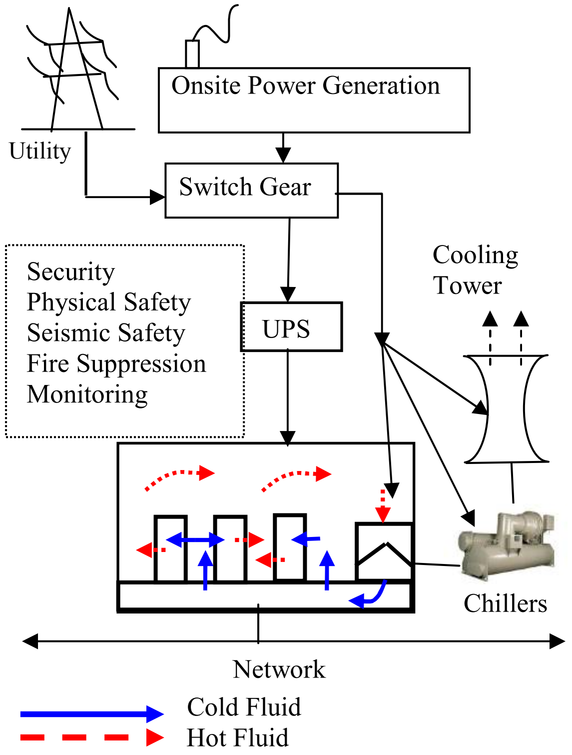

Fig. 1: Data Center Building Blocks

Fig. 2 shows the detail of the datacenter room including CRAC units, server racks and air flow paths [5]. Data centers are typically air-cooled inside the room, with a raised floor plenum to distribute cool air. Power and networking cables are also run through the plenum.

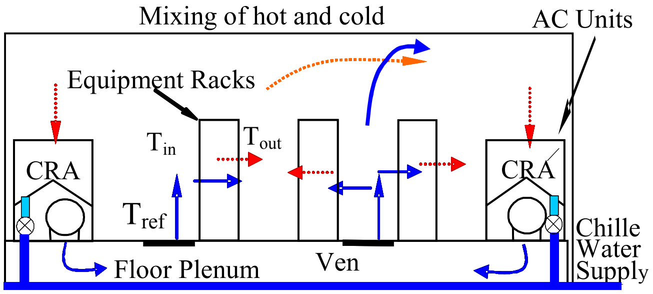

Fig. 2: Cross section of the datacenter

Fig. 2 depicts a typical state-of-the-art data center air-conditioning environment with under-floor cool air distribution. Computer room air conditioning (CRAC) units cool the exhaust hot air from the computer racks. Energy consumption in datacenter cooling inside the room comprises work done to distribute the cool air to the racks and work done to extract heat from the hot exhaust air. A refrigerated or chilled water cooling coil in the CRAC unit extracts the heat from the air and cools it within a typical range of 10°C to 18°C. The air movers in the CRAC units pressurize the plenum with cool air. The cold air enters the data center through vented tiles located on the raised floor close to the inlet of the racks. Typically the racks are laid out in rows separated by hot and cold aisles as shown in Fig. 2. This separation is done for thermal efficiency considerations. Air inlets for all racks face cold aisles while hot air is expelled to hot aisles. A number of other equipment layout configurations and non-raised floor infrastructures also exist.

Fig. 3 shows the details of the cooling infrastructure of a typical datacenter [5]. Typically, vapor compression based refrigeration chillers provide chilled water to the CRAC units. Exhaust air from the IT equipment (as described in Fig. 2) dissipates heat to the chilled water inside the CRAC units. The warm water is returned to the chiller for heat rejection. A condenser water loop carries this heat for subsequent rejection at the cooling tower..

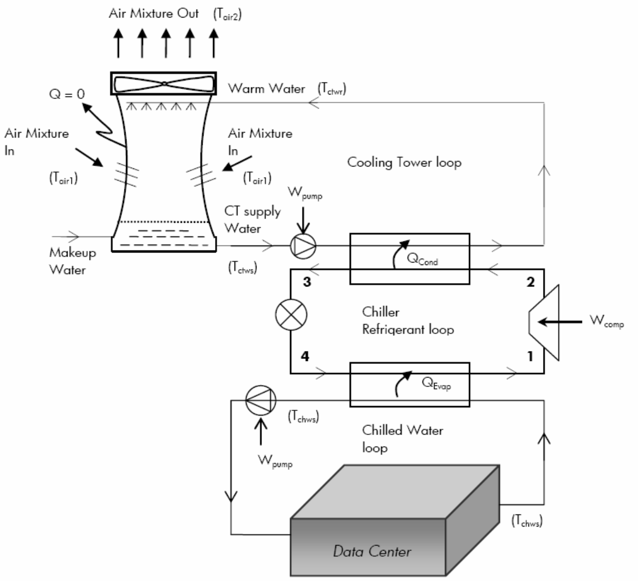

Fig. 3: Data Center Facility Cooling infrastructure

Water is lost by evaporation to the ambient environment during the process of heat rejection. Evaporation loss depends on the moisture content of the air and its temperature. In the case of air cooled chillers, the cooling tower is replaced with a heat exchanger, thus preventing loss of water to the environment. However, since air-to-liquid heat exchangers have lower effectiveness, air cooled chillers also have a lower coefficient of performance. As a result, more energy is required to reject the same amount of heat; and considering that water is required in the production of electricity, it is not obvious that the use of air-cooled chillers will necessarily result in a reduction of water usage. Note that even though we focus on a chiller-based cooling system with wet cooling towers in the present work, other cooling technologies do exist and can also be analyzed using the methodology described in this paper.

C. Water usage in Power generation

The generation of electricity usually requires available water for withdrawal and consumption, sometimes up to 30 gallons of freshwater for every kilowatt hour (kWh) generated in the case of some coal plants [7]. Water is mostly used as cooling water for the condensing of steam. Fig. 4 shows the schematic of a cooling tower with makeup water supply and a bleed system. The makeup water supply replenishes any losses due to evaporation and bleeds, while the bleed system is necessary to maintain water quality by prevention of fouling and microbial growth.

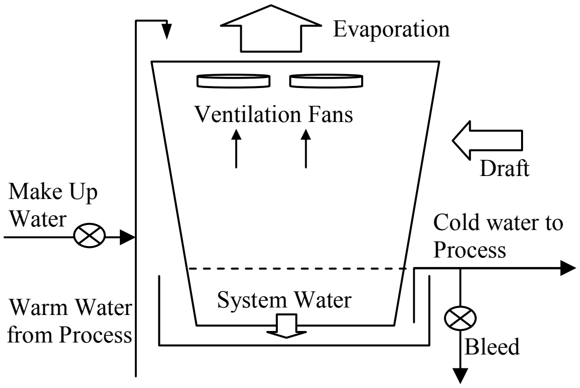

Fig. 4: Cooling tower schematic

Cooling occurs in a cooling tower by the mechanisms of evaporative cooling and the exchange of sensible heat. The loss of heat by evaporation lowers the remaining water temperature. The smaller amount of cooling also occurs when the remaining water transfers sensible heat to the air. Typically, the rate of evaporation is about 1.5% of the rate of flow of the re-circulating water passing through the tower for every 5.5^{o}C decrease in water temperature achieved by the tower. This evaporated water is replenished via the make-up water supply. The rate of replenishment will depend on the type of cooling system being used at the power plant. Open loop or once-through, cooling systems have relatively large water withdrawal rates and are widely used with fresh water. Power plants using closed loop cooling systems with cooling towers are designed to withdraw roughly the same amount of water that is consumed by direct evaporation within the power plant. Therefore, power plants with cooling towers require much lower withdrawal than open loop systems, but tend to consume nearly twice as much at the power plant. Typical water withdrawal rates for Rankine-cycle plants burning coal, oil, or natural gas are 20,000 to 50,000 gallons per MWh generated [7]. The lower end of the flow rate range corresponds to the higher temperature differential at the cooling tower, and vice versa. Air-cooling systems, which have negligible cooling water demand, can be considered as a replacement where water is scarce. Essentially, water demand for electricity generation is expected to increase fairly proportionally to the amount of electricity generation

Apart from water usage in power generation, water usage also impacts the electricity demand. For example, water distribution systems and water treatment plants consume 1.3MWh and 0.5MWh, respectively, for every million gallon of water processed. Electricity use in the water sector could nearly double by 2015, far outpacing population growth [8]. Thus, judicious use of water is also of interest for improving sustainability within the power generating sector.

III. EXPERIMENTAL TESTBED

Energy and water consumption data from a production datacenter has been analyzed in the present study. The datacenter is cooled by air and water cooled chillers which operate on a rotational basis based on diurnal schedules. The chiller utilization and performance is compared with water consumption to calculate a water usage metric [9] and conduct further analysis. Power consumed by IT equipment is also aggregated to understand the impact of water consumption per unit of IT power consumption. Diurnal averages of water and energy consumption data are used in the present study.

As discussed earlier, a key component in consideration of water usage will be the electricity requirements of the datacenter. Fig. 5 shows the typical cooling load for the IT facility being analyzed over a 24 hour period.

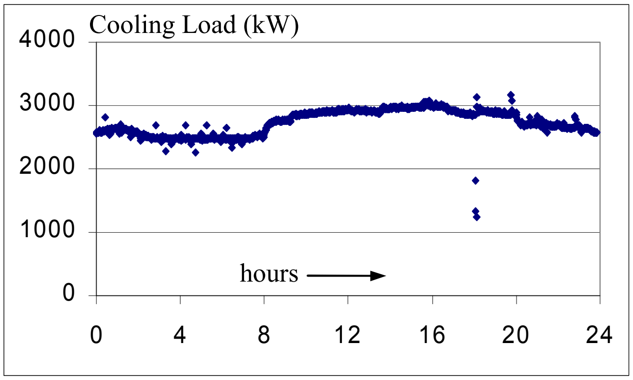

Fig. 5: Typical demand profile in an IT facility

Note that the facility consists of a datacenter and office space; the cooling load thus includes power consumed by IT equipment like servers, storage arrays, network devices etc. as well as office space cooling requirements. The power consumption of the cooling equipment itself is discussed later and therefore not included in Fig. 5

Both air cooled and water cooled chillers are used to provide the chilled water for cooling. The utilization of water cooled chillers varies with time and demand. Operational schedules of the chillers affect both the power consumption and water consumption of the facility. Fig. 6 shows the utilization of the water cooled chillers as a fraction of the total cooling load for two typical days. Observe that the water cooled chillers accounted for a greater fraction of the cooling during the period “day 2” as compared to the period “day 1”. The corresponding energy used by the IT equipment during these days was 73.2MWh on Day 1 and and 42.5MWh Day 2. The energy consumed by the cooling equipment to dissipate the heat was 18.19MWh on Day 1 and 9.28MWh on Day 2. Water consumption from the chillers was continually monitored at the makeup water valve (shown in Fig.4), and data from several days was subsequently analyzed to understand the tradeoffs between water and energy consumption.

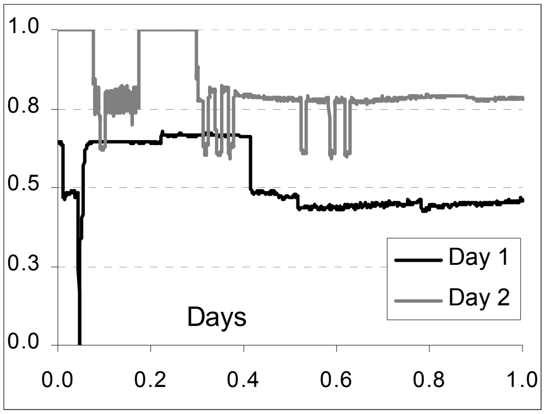

Fig. 6 Water-cooled chiller utilization as a fraction of datacenter cooling load

IV. RESULTS

A. *Water Consumption*

Fig. 7 shows the variation of daily water consumption (in thousands of gallons) for different water cooled chiller loads. Observe that the water consumption rises linearly with chiller utilization.

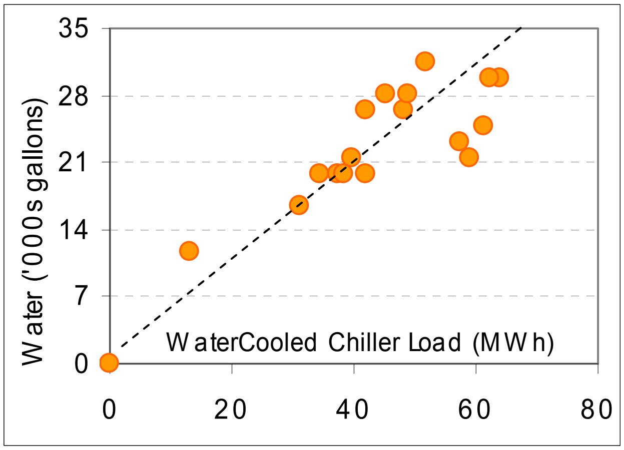

Fig. 7: Water consumption and water-cooled chiller utilization

In practical terms, the x-axis values represent the fraction of all IT energy consumption that is removed by one or more water cooled chillers in the facility. Tracking such variation can allow operators to throttle chiller utilization levels to meet water consumption targets. However, throttling chillers can also affect energy efficiency and indirectly increase the water consumption necessary to generate power.

Fig. 8 shows the scatter in water consumption values with IT energy consumption. The horizontal scatter is created by contributions from air cooled chillers (or water-less cooling technologies). Using air-cooled chillers allow for cooling of the IT load without increasing the *direct* water consumption in the paper. For these situations, higher IT energy consumption can be sustained without increasing water consumption. The vertical scatter is from contributions from additional load on water cooled chillers (i.e. situations where higher IT energy consumption can only be sustained with increased water consumption). The slope will depend on the effectiveness of the chiller (or combinations of chillers) in using water. In other words, an air-cooled (water-efficient) chiller may have a lower slope because it must consume additional water in order to bear a higher IT load. On the other hand, water-cooled (water-inefficient) chillers may have a fairly steep slope because they can continue to bear higher IT loads without consuming as much direct water consumption.

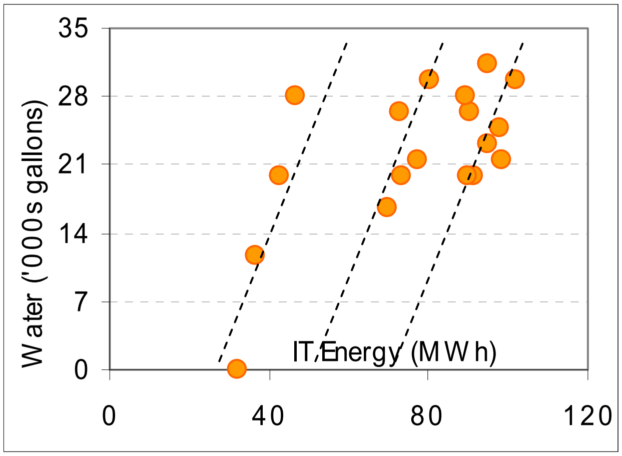

Fig.8: Scatter plot of water consumption and IT Energy Consumption

B. *Energy Consumption*

While reducing direct water consumption for cooling is important, increased energy consumption to do so can increase the indirect water usage at the point of power generation. Therefore, understanding the energy consumption component in cooling is important. Fig. 9 shows the linear rise in energy required to cool with IT energy consumption levels.

Observe clusters of data points based on IT energy consumption levels. Within each of these clusters there are instances with high and low water consumption. Key to management of water and energy consumption is to identify the optimal point for each of these IT Energy consumption scenarios and operate the cooling infrastructure with the right utilization mix.

It is also important to note that water-cooled chillers are more energy efficient than air-cooled chillers. Therefore, while increased utilization of water-cooled chillers increases water consumption at the facility, their higher energy efficiency may also reduce the power consumption of the cooling infrastructure. This reduced power consumption can indirectly reduce the water consumption at the power generation source, so that in aggregate, increasing the load on water-cooled chillers online may not change the *total* direct + indirect water consumption. This is precisely what we observe in our experimental testbed: even though the IT and cooling energy consumption is halved between day 1 and day 2, the total direct + indirect daily water consumption levels were similar at 19841 gallons (~75000 liters). In other words, even though the heat load changed and the direct water consumption changed, the *total* water use in datacenter cooling remained unchanged.

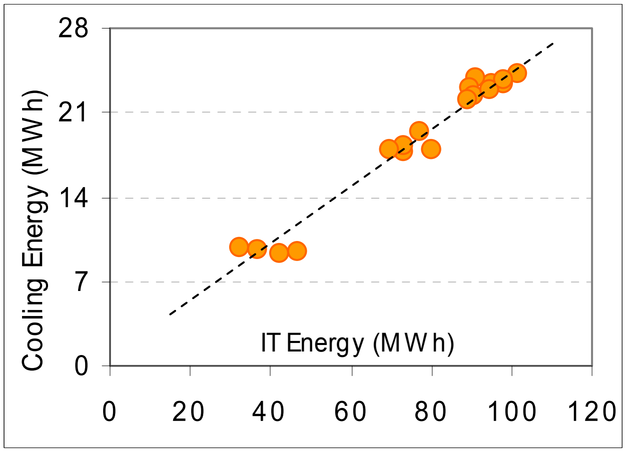

Fig.9: Cooling energy and IT energy consumption

Thus, there is a need for a coherent metric that combines water consumption and energy consumption in the datacenter. Essential to this is the transformation of water consumption to available energy consumed. In the next section, we discuss an approach and implementation for doing so.

V. DISCUSSIONS

A. *Water Usage Energy Metric*

Understandably, water and energy are interconnected in various ways. Electrical power is used in treatment and distribution of water. Power consumption for water treatment can vary greatly based on processes involved. Power consumption for water distribution can vary greatly based on the processes involved. Water distribution power consumption can vary greatly as well, based on the location of the datacenter. To explore the aggregate impact of these different processes, we analyze the impact of delivery of IT services on water usage in terms of the energy footprint of its water consumption

Datacenter water usage can be directly expressed as gallons of water consumed in performing useful IT services. If the power consumption of the IT equipment is used as a proxy for the IT services, then the data center water usage could effectively be represented as the water consumption per unit time, divided by the datacenter IT power consumption. But although easy to use and calculate, such a metric is not dimensionless and does not reflect the energy costs associated with obtaining fresh water in water-scarce regions of the world. Extraction of water from ground or desalination can be highly energy intensive processes. Any water usage metric should capture the energy impact of water usage as a whole.

To capture this available energy impact of water use, we propose a Datacenter Water Usage Energy Metric (ω), defined as the ratio of the embedded energy "footprint" of water consumption over the power consumption of IT equipment. The embedded energy footprint of water usage is the energy required to treat and distribute the water to the location of demand. Water usage is the rate of water consumption, including both direct (cooling) and indirect (power generation) usage. The embedded energy footprint of water usage is based on local utility or water supply district and the baseline for comparison is water drawn from a natural source. Datacenter water usage energy metric is defined as

$$\omega = \frac{\left[(E_d + E_n)\right]}{(\mathrm{E}_{\mathrm{IT}})} \times 10^3$$

E_d= Embedded Energy in direct water usage
E_n=Embedded Energy in indirect water usage
E_{IT}=Energy consumed by IT equipment

Fig. 10 shows the flow chart of the process for calculation of water usage energy metric.

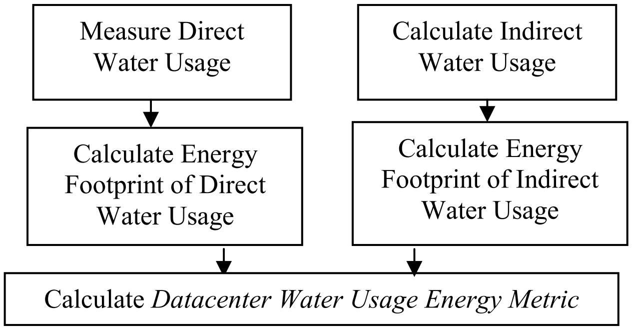

Fig. 10: Flowchart for calculation of data center water usage energy metric.

Embedded energy in direct-use water is calculated as the product of consumed water and the embedded energy multiplier. In order to calculate the embedded energy in the indirect-use water, we first obtain the water required to generate electrical energy for the total facility. Embedded energy in indirect-use water is calculated as the product of water consumed for generation and the embedded energy multiplier. For example, in our case study, based on published literature and site conditions, we estimated that the energy consumed to convey, treat and distribute water to the datacenter was approximately 1.5 MWh for every million gallons. This is the energy multiplier that can be used to obtain the embedded energy in water

To compare this embedded energy metric with energy efficiency, we calculate the coefficient of performance based on energy rather than power as the ratio of IT energy consumed over cooling energy consumption [6]. Fig. 11 shows a plot of the Water Usage Energy metric against the ratio of IT energy consumption over cooling energy consumption. The water usage energy metric at point A is 0.9, the minimum in the dataset. The reason for such a low water usage energy metric is that only air cooled chillers were active at Point A, so the direct water use was zero. However, this point also has the lowest coefficient of performance in the curve, because indirect water usage is still significant and accounts for the rest of the embedded energy. Point B, on the other hand, has a significant direct water use component because water cooled chillers dissipated a significant fraction (97%) of the IT energy consumption during that run. Hence, its water usage energy metric is 1.8 with a high coefficient of performance.

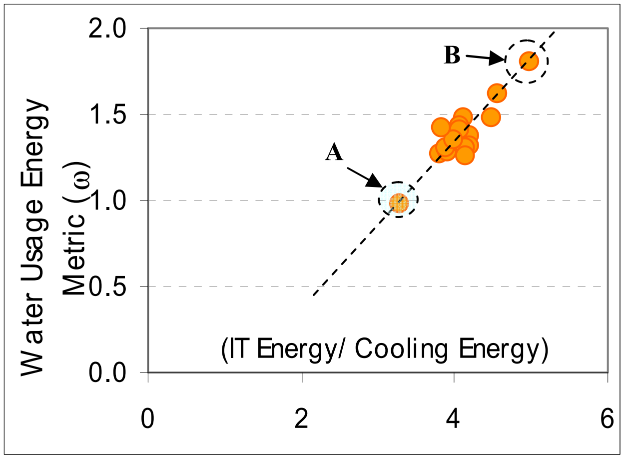

Fig.11: Water Usage Energy Metric for ensemble of chillers as compared with coefficient of performance

Each datacenter has different embedded energy content in its water supply. Such a curve can guide a datacenter operator in trade-off decisions between water and energy consumption. Curves can be generated automatically in real-time and an expert system could be developed to set appropriate targets for the water usage energy metric and corresponding coefficient of performance.

Data center water efficiency can be managed by minimizing the water usage energy metric for a given IT energy consumption. The metric can also be used to compare water efficiency across datacenters. There is a need to do seasonal benchmarking of this metric to capture the effect of regional weather patterns. As a preliminary approach, one could define a design curve for the metric over time to capture seasonal variations, then create operational policies for chiller and IT operation to meet the targets. Such a design curve could be a part of a service level agreement with local utilities or administration. As next steps, it may worthwhile to include the metric as a part of coefficient of performance of the datacenter ensemble. Management policies are another open area of research at the present moment.

VI. CONCLUSION

Water usage is closely coupled with energy usage. The proposed energy efficiency framework for datacenter can be extended to manage water efficiency. A water usage energy metric is proposed that provides a common framework for datacenters while capturing the diversity in water availability across geographies and seasons. Energy and water optimization can be achieved on a common basis through evaluation and subsequent application of water efficient technologies. Such a metric can identify "water hotspots" among a wider ecosystem, like a city/state-wide water management system, with other competing demands. The impact of water quality is also inherent in the definition of the metric.

Sustainable datacenters need to be designed around the local supply constraints of physical resources. Such constraints can be a function of service level agreements associated with IT services or with local administration/utilities.

ACKNOWLEDGMENT

The authors wish to thank Vaibhav Bhatia and P.V. Rajkumar for their support in data collection from the experimental test bed.

REFERENCES

[1] The Uptime Institute, "Heat density trends in Data Processing, Computer systems and Telecommunications Equipment", White Paper issued by The Uptime Institute, 2000
[2] Goldstein, R., Smith, W., "Water & Sustainability-U.S. Water Consumption for Power Production—The Next Half Century", EPRI Topical Report, March 2002
[3] Patel, C. D, Shah, A., "Cost Model for Planning, Development and Operation of a Data Center", Hewlett-Packard Laboratories Technical Report, HPL-2005-107(R.1), June 9, 2005
[4] S. Hutson et al., "Estimated use of water in the United States in 2000", US Geological Survey Circular 1268, USGS, 2004.
[5] R.K. Sharma, et al., "On building next generation datacenters: Energy Flow in the Information Technology Stack", ACM Compute 2008, Paper ID 129, Bangalore, India
[6] Patel et al., "Energy flow in the information technology Stack: Introducing the coefficient of performance of the ensemble", Proc. ASME Intl. Mech. Eng. Cong. Exp. (IMECE), IMECE2006-14830.
[7] King, C., Duncan, I., Webber, M., "Water Demand Projections for Power Generation in Texas", Bureau of Economic Geology, University of Texas at Austin, Apr 2008.
[8] Trask, M., "Water-Energy Relationship", Integrated Energy Policy Report 2005, California Energy Commission.
[9] R.K. Sharma, A. Shah, C.Bash, T. Christian, C.D. Patel, "Water Efficiency Management in Datacenters (Part I): Introducing a water usage metric based on available energy consumption", Hewlett-Packard Laboratories Technical Report, HPL-2008-206.