

How to Cool Internet-Scale Distributed Networks on the Cheap

Vani Gupta[†], Stephen Lee[†], Prashant Shenoy[†], Ramesh K. Sitaraman[‡], Rahul Urgaonkar^{¶*}
[†]University of Massachusetts, Amherst [‡]Akamai Technologies [¶]Amazon
{vani, stephenlee, shenoy, ramesh}@cs.umass.edu urgaonka@amazon.com

ABSTRACT

Internet-scale Distributed Networks (IDNs) are large distributed systems that comprise hundreds of thousands of servers located in hundreds of data centers around the world. A canonical example of an IDN is a content delivery network (CDN) that delivers content to users from a large global deployment of servers around the world. IDNs consume significant amounts of energy to power their deployed server infrastructure, and nearly as much energy to cool that infrastructure. We study the potential benefits of using two new cooling technologies—open air cooling (OAC) and thermal energy storage (TES)—to reduce the energy usage as well as the operational and capital costs incurred by an IDN for cooling. We develop novel algorithms to incorporate both technologies into the IDN architecture and empirically evaluate their efficacy using extensive work load traces from Akamai’s global CDN and global weather data from NOAA. Our results show that both technologies hold great promise for the future sustainability of Internet-scale distributed networks. Our algorithm for power management of TES is provably near-optimal, is the first to incorporate storage efficiency, and is broadly applicable to other storage devices such as batteries.

Keywords

Internet-Scale Distributed Systems, Load Balancing, Renewable Energy, Energy Optimization, Data Centers

1. INTRODUCTION

Modern Internet services are delivered by Internet-scale distributed networks (IDNs) that consists of hundreds of thousands of servers deployed in a large number of data centers around the world. IDNs include cloud and Internet services such as content delivery networks (CDNs) that deliver web content, applications, and streaming media to clients on the web via hundreds of thousands of servers located in thousands of data center locations throughout the world [23]. An IDN consumes a significant amount of energy to power its servers and to cool them. It is not uncommon for a large IDN

*This work was performed when Rahul Urgaonkar was at IBM Research

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

e-Energy'16, June 21 - 24, 2016, Waterloo, ON, Canada

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4393-0/16/06...\$15.00

DOI: <http://dx.doi.org/10.1145/2934328.2934337>

to incur energy bills that run into millions of dollars per month. The environmental impact of Internet-scale networks, and data centers generally, is also considerable with carbon emissions from data centers growing by 15% each year [14]. Thus, it is imperative to re-design IDNs with energy as a key design consideration to ensure the sustainable growth of these networks.

Recently there has been considerable work in reducing the power consumption of servers in a data center and IDNs by using advanced cluster-wide power management techniques [3] or by simply shutting down servers or clusters during off-peak periods [16, 21]. While such techniques yield significant cost benefits, they only address the energy costs of powering the servers and do not *directly* address the energy cost of cooling them. Many data centers and IDNs still consume around 0.8 Watts of power to cool a server for every 1 Watt used to power that server—hence the energy bill spent in cooling an IDN is nearly as large as the energy used to power its servers. Rather than focus on server energy reduction, we focus explicitly on the complementary problem of reduction in cooling costs using two novel technologies: *renewable open air cooling (OAC)* and *thermal energy storage (TES)*.

Both OAC and TES are emerging technologies that are gaining traction within data centers. Open air cooling (OAC), also known as “free cooling”, uses outside air to cool servers whenever the climate permits, e.g. when the outside air is cool or dry. OAC can decrease, or even eliminate, the use of chillers used to chill the air for cooling the servers. Service providers who employ large data centers such as Facebook and Google have begun to use OAC, in part by building new data centers in carefully chosen (cold) locations where the climate permits the outside air to be used to cool the data center for the majority of the year. Thermal energy storage (TES) provides a mechanism for storing energy in the form of chilled water or ice that can be used for cooling the servers at a later time. Using TES, one can exploit electricity price fluctuations by storing thermal energy when prices are cheap and using stored thermal energy to cool servers when prices peak. While TES is a mature technology used in building and factory applications for decades, their use in data centers has been fairly recent [8].

Our focus is on understanding the efficacy of using OAC and TES for IDNs. To our knowledge, neither technology has been studied in an IDN context that is distinctive for the following reasons. IDNs such as CDNs have two defining characteristics: *a global deployment* of servers in multiple data centers around the world, and a *replication* of services across these data centers. The global deployment is often driven by the need for an IDN to have servers “proximal” to the end-users. For instance, Akamai’s CDN is deployed in hundreds of data centers in over 100 countries around the world, with users accessing content from servers in “proximal” data centers [23]. A corollary of this deployment model is that it

is not possible for IDNs to deploy *only* in places where weather is cold most of the year, or where electricity is cheap. They need to be deployed near where the users are. However, IDNs often replicate their services across their data centers, so that the workload of serving users can be easily shifted from one data center to another, albeit with a potential for performance degradation. These two characteristics provide an IDN the flexibility to move its workload across data centers to exploit climatic variations to optimize the use of OAC, a flexibility that services employing a single or a few data centers do not possess. The energy cost also varies across locations and across time. TES allows for “shifting” the cooling energy requirement in time, by drawing from the grid when energy is cheap during off-peak hours and using it during peak hours. Our algorithms orchestrate both the movement of load (i.e., load balancing) and grid energy (i.e., power management) to decrease an IDN’s cooling costs.

Contributions: Our work is the first scientific evaluation of the potential of OAC and TES in the context of an IDN. Our specific contributions follow.

- *Algorithm Design:* We design algorithms to integrate both OAC and TES, individually and in combination, with an IDN’s network-wide load balancing and cluster-wide power management systems. Of particular algorithmic interest is the TES power management described in Section 5.2 that uses the Lyapunov technique to decide when to use TES versus the chillers to minimize overall cost. Our algorithm is *online* in that it has no knowledge of the future workload or energy prices, but is provably within an additive factor of the minimum cost achievable by an offline algorithm that has full knowledge of both. Our TES algorithm is generic and can be used for a broad range of storage devices, including TES and batteries. While prior Lyapunov-based online algorithms [9, 19, 11, 28] are known, they assume that the efficiency of the storage device is 100%. To our knowledge, our work is the first to explicitly incorporate the notion of efficiency within the Lyapunov framework that is key to modeling storage devices. All real-world devices have various types of conversion losses that could range from 10-50% depending on the type of device. Our approach also better utilizes the available storage, resulting in a provably closer approximation of the optimal than prior work in [9, 19, 28, 11]. Further, our work also provides an evaluation of the cost achieved by the online Lyapunov approach *vis-à-vis* the offline optimal baseline for a range of workload, conversion loss, and energy prices.

- *Extensive Trace-based Evaluation:* We evaluate both OAC and TES using extensive load traces from the world’s largest CDN. The traces were collected from Akamai’s CDN every five minutes for the period of a month from over 115,000 servers deployed in over 973 data centers in 102 countries. Using these CDN traces, weather data from over 650 locations around the world, and energy price information we evaluate the ability of OAC and TES to reduce operational and capital expenditures via reductions in energy usage and costs.

- *Benefits of OAC:* We study the potential for deploying OAC at scale in a IDN. We design greedy “weather-aware” load balancing algorithms that direct load to the closest data center where the current weather permits “free” cooling using outside air. Our result shows that even during summer, a global IDN can extract more than 51% reduction in the energy spent for cooling using OAC. During winter when OAC is more plentiful, a 92% reduction in system-wide energy can be had. Further, these savings can be achieved without degrading the performance experienced by users. However, important exceptions remain. We find that Singapore has small or no potential for OAC throughout the year using current technology, while Japan is not conducive to OAC during the sum-

mer months. However, with newer (class A4) data center technology, even such places can use OAC, with energy savings for Japan rising from 0% to 83% in August.

- *Benefits of TES:* We study the potential for deploying TES-based cooling within an IDN to flatten peaks in cooling power requirements and reducing energy costs by exploiting electricity price differentials. We show that TES with a small storage capacity that can power a data center for 30 minutes is sufficient to reduce the global installed chiller capacity of the IDN by 41%, resulting in a capital expenditure (CAPEX) saving. For individual major metros in the world, the chiller capacity reduction for 30 minutes of TES capacity varied from 25-45%. Exploiting electricity price differentials require a larger TES capacity of about 10 hours and can result in nearly 18-43% energy cost reduction for major US metros, resulting in operating expenditure (OPEX) savings.

- *Benefits of OAC and TES together:* We study how to combine OAC and TES at the scale of a global IDN to extract even more CAPEX and OPEX savings. Combining OAC and TES produces greater cost savings than TES alone. With OAC and a 30-minute TES capacity, we can reduce the global installed chiller capacity of the IDN by 47%. In places such as San Jose, California where OAC is plentiful, combining OAC and TES can practically eliminate the need for chillers. The energy cost savings can also be significant.

Roadmap: We present background and methodology in Sections 2 and 3. We present our algorithms and results for OAC, TES and their combination in Sections 4, 5 and 6. We present related work in Section 7 and conclude in Section 8.

2. BACKGROUND

Internet-scale Distributed Networks: An IDN is a large distributed network of servers where the servers are deployed in data centers throughout the world. A canonical example is a content delivery network (CDN) that we discuss as an illustrative example. Content providers—such as media sites and e-commerce sites—use CDNs to host and deliver their content. When a user accesses content hosted by the CDN, the CDN’s *load balancing* system assigns the user to a server in a “nearby” data center, so as to minimize server-client latency and loss [23]. For this reason, major CDNs deploy their servers widely in hundreds of data centers around the world. A corollary of this deployment model that CDNs share with other highly-replicated Internet-scale services is that users have a redundant choice of data centers that each provide acceptable performance—an architectural feature we exploit extensively in this paper.

A typical IDN deploys cluster of servers in data centers that it owns or in rented co-location (colo) facilities. While an IDN has complete control over what (cooling) technologies are deployed in its own data centers, it typically has little control over cooling technologies used in colo facilities. However, there is an increasing trend towards containerized clusters, i.e. modular self-contained containers, that permit design flexibility in how the container is built and cooled, even in colo facilities. Given these future trends, we assume that an IDN (or, its proxy) has the ability to deploy or exploit advanced cooling technologies in its data centers, regardless of type.

A recent study of data center energy consumption [24] showed that servers and cooling consumed 56% and 30% of the total energy respectively, while power conditioning (8%), networks (5%) and lighting (1%) accounted for the rest. Thus, most of the energy consumed by a data center is spent in powering servers or cooling them; we refer to these components as *server energy* and *cooling energy*, respectively. Since cooling energy is a significant portion of the total energy consumption, we examine the potential for em-

ploying two new cooling technologies to reduce either the energy usage or the energy cost incurred by an IDN, or both.

Renewable Cooling using Open Air: A promising approach for reducing the cooling energy of a data center is to use the outside air, instead of HVAC chillers, to cool servers within the data center. OAC has the potential to reduce or even fully eliminate the need for chillers that consume much of the electrical energy used for cooling.¹ OAC has recently been successfully installed in a few facilities such as Facebook’s data center in Forest City, North Carolina [26]. However, OAC is feasible only when the air is “cold” and “dry” enough. As a result, OAC may not be possible everywhere. Further, even where OAC is possible, it may not be possible during all times of the day, or all seasons of the year.

In our work, we use (i) the instantaneous weather outside each of the IDN’s data center locations, and (ii) the recommendations of the American Society of Heating, Refrigeration, and Air-conditioning Engineers (ASHRAE) [10] to determine whether the outside air can be used to cool the data center at any point in time. Table 1 specifies ASHRAE’s temperature and humidity ranges for four different classes of data centers, where each class represents the type of server and other IT equipment used in the data center. The lowest class A1 represents the most basic equipment that allows the smallest operating ranges of temperature and humidity and as such represent the widest deployment of data centers today. The highest class of A4 represents the most advanced equipment that can function at a much large operating ranges of temperature and humidity. Given the wide deployment of IDNs, we conservatively assume that our data centers belong to class A1, but also consider what-if scenarios if data centers of higher classes become commonly prevalent. Note that assuming class A1 places a lower bound on the potential savings from OAC.

Class	Dry-Bulb Temp (° C)	Humidity Range	Max Dew Point (° C)
A1	15 to 32	20% to 80%	17
A2	10 to 35	20% to 80%	21
A3	5 to 40	8% to 85%	24
A4	5 to 45	8% to 90%	24

Table 1: ASHRAE’s allowable ranges for dry bulb temperature, relative humidity and the maximum dew point of the air that make it suitable for cooling different classes of data centers [10]. Higher data center classes correspond to newer technology allowing for broader ranges of tolerance.

Although ASHRAE standards do not specify the cooling technology to be used by a data center, the specified ranges enable us to determine the upper limits on *outside* air temperature, humidity and dew point that permits OAC to safely cool a data center of a particular class. To determine whether a data center can employ OAC at any point in time, we use the weather data for that location to determine the dry-bulb temperature, relative humidity, and dew point. Following the methodology presented in the GreenGrid consortium whitepaper [10], if the outside air is below the ASHRAE range, it can be rectified by mixing the outside air with the return air that is warmer. Thus, it suffices to compare the measured dry-bulb temperature, relative humidity, and dew point of the outside air with the upper limits of the allowable ranges in Table 1 to ascertain OAC feasibility.

¹Different forms of OAC include evaporative cooling, which uses a combination of water and outside air, and direct air cooling, which directly uses outside air, to cool servers. Our analysis is agnostic to the exact form of OAC employed by the data center.

Thermal Energy Storage: TES is a mechanism for storing thermal energy by either heating or cooling a medium so that it can be used later. TES systems create and store chilled water or ice that can subsequently be used to cool air (e.g. by blowing warm air through chilled water). TES can be used in conjunction with HVAC chillers, or perhaps even in lieu of them. The two common use cases of TES that we examine are (i) under-provisioning the chillers and using TES to bridge the gap when the cooling demand exceeds the capacity of the under-powered chiller, and (ii) exploiting energy price differentials by charging the TES when prices are low and using the stored energy for cooling when the prices are higher. The former provide capital cost (CAPEX) savings, while the latter reduces operational costs (OPEX).

TES systems are varied in their technology and design. We model key TES parameters as follows. At any given time, we assume that the TES is in one of two states. It is either in a recharge state where more chilled water or ice is produced and stored in its tank. Or, it is in a discharge state where the chilled water or ice is released and used for cooling. The maximum discharge rate of a TES represents the maximum amount of cooling (in KWs) it can dispense. Typically, TES is designed in a manner that it can provide the peak cooling demand of the data center. The charge rate is typically comparable to the discharge rate. The capacity of the TES is the amount of energy (in KWHs) it can store and relates to the size of the tank used to store the chilled water or ice. While the capacity can be expressed in KWHs, a more intuitive unit is the numbers of hours (or, minutes) a TES can be used to cool the data center at peak, i.e., TES capacity in hours is simply its capacity in KWHs divided by the max discharge rate. Finally, like any energy storage, TES systems incur energy losses in storage, charge and discharge cycles. TES efficiency varies from 50% to 90% [12] for water-based systems. Other types of TES that use Phase Change Materials (PCM) or chemical reactions to store the energy may have even higher efficiencies in the 75% to 100% range [12]. We assume TES efficiency to be 90%, but also show how lower TES efficiencies impact our results.

Energy Consumption Model for an IDN: The primary consumers of energy in an IDN are the widely deployed servers. The energy consumed (and dissipated as heat) by a server can be modeled as a function of its (normalized) load λ , $0 \leq \lambda \leq 1$, where λ is the load served by the server as a fraction of its capacity. However, deployed servers are not energy proportional and still use roughly 60% of the peak energy when idle. Hence, we use the standard linear model of server power consumption [2] where the power (in Watts) consumed by a server serving load λ equals $P_{idle} + (P_{peak} - P_{idle})\lambda$, where P_{idle} is the power consumed by an idle server, and P_{peak} is the power consumed by the server under peak load. While the servers themselves are not energy proportional, one can consolidate the server load and turn-off idle servers to save energy [16]. We assume that such an optimization is performed in each of the IDN’s data centers in calculating the energy dissipated by the clusters.

Clearly, the more the energy dissipated by the servers, the more the energy required by the cooling system to cool them. Thus, cooling energy is roughly proportional to the amount of heat dissipated in the data center, which in turn is governed largely by server energy. The proportionality factor is a measure of the effectiveness of the cooling system deployed at the data center and is captured by its power usage effectiveness (PUE) that is ratio of the total energy consumed by the facility to the energy consumed by the servers and other IT equipment. While specialized data centers with low PUE are known, recent studies have shown that the average PUE of a data center is 1.8 [27]. Thus, cooling energy still accounts for a

large fraction (say, 30 to 40%) of the total energy consumed by the IDN. Our work focuses on reducing the cooling component of the IDN’s energy consumption.

Energy Pricing Models: To incentivize consumers to reduce their energy use during peak hours, it is common place for utilities to have *time-of-use (TOU) pricing* that are lower during hours of low demand. Examples include Wisconsin electric [31] that we use for our experiments as being typical of TOU contracts. It is worth noting that our results are not sensitive to the precise dollar values of the electric costs, but rather only relative values of how much cheaper off-peak prices are to on-peak ones. Another form of time-dependent variable pricing that is gaining prevalence is *market pricing* [25] in which power can be bought on the “real-time” spot market at each data center location of the IDN. We study TES in both pricing models.

Energy Efficiency Metrics: We consider two aspects of efficiency in the context of cooling IDNs: energy *usage* and energy *cost*. Reducing the energy usage reduces carbon emissions, and also reduces energy cost, but not necessarily vice versa. For instance, OAC reduces both energy usage and energy cost. However, TES may increase energy usage slightly due to conversion and storage losses, but reduces energy cost. We consider two aspects of energy cost: (i) capital expenditure (CAPEX) of cooling equipment to be installed at a data center, and (ii) operational expenditure (OPEX) of running the cooling system. Both OAC and TES have the potential to reduce CAPEX and OPEX of an IDN.

3. EXPERIMENTAL METHODOLOGY

To derive the potential for using OAC and TES in an IDN, we performed trace-driven simulations using a combination of IDN workloads and weather data. We used extensive load data from across Akamai’s CDN for the period of one month. The trace includes load information from 115,246 servers deployed in 973 data centers locations in 102 countries around the world (see Table 2). Collectively, the traces account for 43.9 trillion requests served by the CDN that resulted in content served to users around the world with a peak traffic of 24.8 million requests/second. The dataset includes the load, requests served, and bytes served by each server every five minutes over the month-long trace. Further, the trace has detailed information about every data center including the number of deployed servers, total server capacity, and the location of the data center including its latitude, longitude, city, state, and country.

Our experimental evaluation also employs global weather traces provided by the National Oceanic and Atmospheric Administration (NOAA) for the year 2012. The dataset contains year-long weather data from 13,497 weather stations across the globe that record a large number of metrics including the hourly dry-bulb temperature and dew point. Since the exact location of each weather station and data center are known, we can compute the weather station that is closest to each CDN data center and use its weather data to represent the ambient weather conditions at that data center. Given the extensive network of NOAA weather stations, we were able to find a nearby weather station within 10km for the majority of data centers, including all of the “large” data centers near major population centers. We found a weather station within 40km for most of the remaining locations. The matching process yielded 651 weather stations that were mapped onto the 973 data centers (major cities have multiple data centers mapped to the same “nearby” weather station). The weather data was used to determine if the outside air at each data center was suitable for cooling at that time.

To compute the cooling energy required by the CDN, we first compute the server energy consumed (and dissipated) for each data center of the CDN for each 5-minute window using the load and

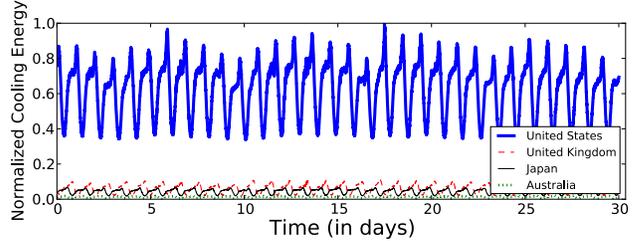


Figure 1: Normalized cooling energy required by Akamai’s CDN in the US, UK, Japan and Australia. Notice the diurnal variation as cooling energy is proportional to the load induced by users accessing content from those locations.

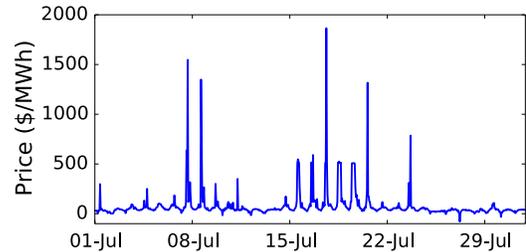


Figure 2: NYISO real-time prices.

server information in the traces and the server and cluster energy model presented in Section 2. Cooling energy is proportional to the server energy, where the proportionality factor is related to the PUE. Figure 1 shows the cooling energy required by Akamai’s CDN as computed from the traces for four major countries.

We make two simplifying assumptions in our analysis. We had load data every five minutes but weather data once an hour. We assumed that the weather parameters do not significantly change during the hour. Further, we had weather data for a whole year but comprehensive CDN load data only for a month. We assumed that the measured monthly CDN load pattern repeats through the year.

To evaluate the reduction in energy costs due to OAC and TES, we used Wisconsin electric’s TOU pricing [31] as being typical of TOU contracts. We obtained the real-time spot market prices available for several major US metros from public web sites (cf. Table 4), e.g. Figure 2 shows the real-time electricity price in the New York metro.

4. OPEN AIR COOLING

To integrate OAC into an IDN’s architecture, its global load balancer must be made “weather-aware”. The load balancer of an IDN assigns each user request to a “nearby” data center to optimize user-perceived performance. To evaluate the benefit of OAC, we propose a simple greedy algorithm that modifies the load assignments made by the (non-weather-aware) load balancer as reflected in our Akamai load traces by moving user load from data centers that have no OAC to nearby data centers that do, subject to performance constraints. Our greedy algorithm does the following for each of the IDN’s data centers at each time step: if the weather conditions at a data center location permits OAC, then user load mapped to that location is unchanged; however, if the weather conditions at a data center location do not permit OAC, the load balancer attempts to greedily re-assign the load destined for that location to other nearby data centers with spare server capacity where

Duration	1 month
Data centers	973 data centers, 102 countries
No. of servers	115,246
Total requests	43.9 trillion
Peak traffic	24.8 million requests/second

Table 2: Load data from Akamai CDN

Duration	12 months
Weather stations	13,497
Mapped stations	651
Resolution	1 hour

Table 3: NOAA weather data.

Dataset	Pricing type
Wisconsin Energy	Time of use (TOU)
New York ISO, ERCOT	Real time, spot
Texas, Midcontinent ISO	
California ISO	

Table 4: Electricity pricing data.

OAC may be available. The premise is that weather patterns exhibit sufficient regional variations so that OAC may be possible at a location even when it is not be feasible at another nearby location. We exploit these geographic variations by searching for alternate “nearby” locations where OAC is still feasible.

The primary performance impact to the user from the remapping is that a user may get mapped to a data center that is “farther” away, increasing response times. We can limit how far a user can be remapped by stipulating that our algorithm can only remap load to data centers within a radius r kms of the data center to where it was originally mapped. Specifically, the greedy algorithm reassigns the load of each data center without OAC to alternate data centers with OAC that have spare capacity to accommodate all or part of the load and are within radius r . The alternate data centers are examined in the increasing order of distance and any load left unassigned by this process is not remapped and must be cooled in its original data center using traditional HVAC chillers. The radius r represents a tradeoff between network performance and energy savings. The greater the r , the greater the user-perceived response times, but greater are the chances that there will be sufficient geographic weather variations such that OAC is possible at these alternate locations. In general, significant savings are possible with *no performance degradation at all* (i.e., for $r = 0$). Even for moderate values of r (e.g., $r \leq 1000\text{km}$), we expect the vast majority of the load to be served locally, while moving the residual load only by a small distance, limiting the performance degradation (cf. Fig. 5).

Our evaluation uses geographic distance as a proxy for latency and response times. This is because our load traces only include client locations and mask client IP addresses for privacy, allowing us to compute geographic distance but not network distance. Prior work has shown that network latency increases with increasing geographic distance, and so distance is a coarse measure of latency (cf. Table 1 of [23], or, Figure 4 of [13] that posits a marginal increase of 1 msec of network latency for every 50km of distance, or [25] that uses distance as a proxy for latency in a similar context).

4.1 Empirical Results

We evaluate the potential for OAC using our greedy algorithm outlined above on our IDN load and weather traces for a full year. In our simulations, we assume that each data center belongs to the most conservative ASHRAE A1 class.

4.1.1 Reduction in chiller capacity

We examine whether OAC can yield CAPEX savings for an IDN. Intuitively, if OAC reduces the worst-case peak demand on HVAC chillers, either by absorbing a portion of the peak demand locally using OAC or by redirecting a portion of the peak load to other nearby data centers that can be open air cooled, then the IDN can deploy lower capacity (and less expensive) chillers to cool the reduced peak load. However, it is not evident *a priori* whether OAC can reduce the worst-case peak demand on chillers, e.g. the worst-case peak load could occur on hot summer days where OAC is infeasible. Figure 3 depicts the average capital cost (CAPEX) savings across all global IDN locations due to a reduction in chiller

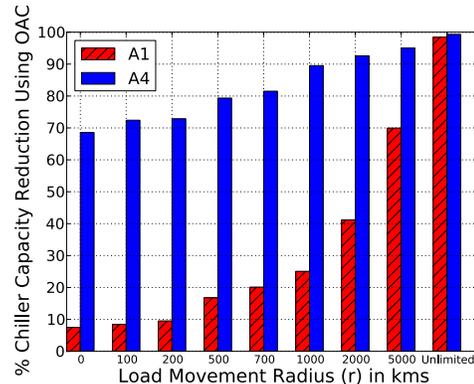


Figure 3: Savings in capital costs of chillers with OAC.

capacity. The figure shows that for A1 class data centers, OAC yields only a 7.5% savings when $r = 0\text{km}$, implying that peak load does coincide with hot or humid days when OAC cannot be used. Further, allowing the load to be redirected to locations within a 1000km radius yields 25% CAPEX savings. The CAPEX savings are significantly higher for the newer A4 class data centers with a mean reduction of 68.6% in cooling capex with $r = 0\text{km}$ to as much as 89.5% capex reduction when $r = 1000\text{km}$.

4.1.2 Reduction in energy usage

Global Savings: The energy savings can be computed by comparing the energy used with OAC to the energy used to cool the original load entirely with chillers without OAC. Figure 4(a) depicts the average percentage energy savings obtained across the entire CDN for different months of the year and for different values of distance r . The savings from OAC is generally higher during the cooler winter, early spring and late fall months of the northern hemisphere, with lower savings during the warmer summer months (May to September). Note that our analysis includes savings from data centers in both the northern and southern hemispheres. However Internet traffic from North America, Europe and Asia dominate the global Internet traffic, hence the seasonal benefits from the northern hemisphere dominate the global trends. Overall, our result shows that even during summer, a global IDN can extract significant cooling energy reduction of more than 51% even during summer *with no performance impact* ($r = 0$) and the savings due to OAC increase to over 92% during winter months; the savings increase as the performance constraints are relaxed by permitting $r = 1000\text{km}$ yielding an additional 13% savings during the warmest month of July. For $r = 5000\text{km}$ which allows for trans-continental load redirection, the savings increase to over 92% throughout the year, including summers.

Regional and Seasonal Variations in Savings: Figure 4(b) and (c) depict the energy savings seen in two major countries, USA

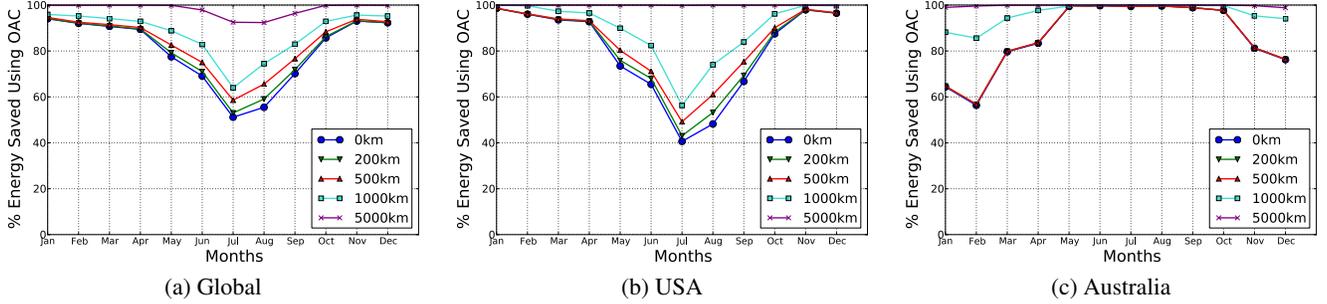


Figure 4: Energy savings for the entire global IDN and for major countries in each of the two hemispheres.

and Australia, in the northern and southern hemispheres, respectively. Energy savings in USA broadly follow the global trends, indicating that USA not only contributes a significant portion of the global traffic, but also has its seasons aligned with the dominant northern hemisphere. Further, USA being a large country in terms of geographic area, exhibits significant regional variations. Fig 4(b) shows high energy savings of over 96% in winter months, it indicates that most cities, regardless of location, see uniformly high energy savings. In the summer, however, there are considerable differences: southern cities such as LA see low OAC savings (for $r = 0$), while northern cities such as Seattle see higher than average OAC savings. Australia (in Fig 4(c)) sees similar differences between summer and winter, with OAC savings of 64% in the summer month of January for $r = 0$ and nearly 100% savings in the winter months of May to September. Further, allowing the load re-direction to a data center within a 1000km radius increases the summer savings to above 85.6%.

Impact of Newer Data Center Technologies: Thus far, we assumed that all IDN data centers belong to the most conservative A1 class in terms of server and cooling equipment. However commodity servers built in recent years are engineered to withstand higher temperatures without impacting reliability. Further, the latest cooling equipment can deal with a larger range of humidity scenarios. Consequently, we repeat the previous analysis by assuming all data centers are built for ASHRAE’s most aggressive A4 class, which permits the inside temperatures in the data center to be maintained as high as 45°C with relative humidity of 90 (cf. Table 1). Our experiment sheds light on the additional benefits from having A4 class data center, since OAC now becomes feasible even in warmer or more humid climates. With A4 class, we observe 95% energy savings from OAC year-round for $r = 0km$, with a slight decrease in the summer. With $r = 1000km$, the savings rise to about 98% even in the summer.

Network latency impact: Figure 5 shows the impact on performance (i.e., latency increase) due to OAC-driven load movements is likely to be small even when we allow our algorithm to move load to data centers that are up to 1000km away. This is because over 90% of the load is served locally due to OAC on an average day; even on the worst day of the year, over 68% of the load is not moved at all. On an average, 92.6% (resp. 97.5%) of the user load gets served by a data center within 300km for $r = 1000km$ (resp. $r = 500km$). Even on the worst day of the year that requires the most load movement, 75.5% (resp., 88.7%) of the load is served by a data center within 300km for $r = 1000km$ (resp., $r = 500km$). These results indicate that only a few users see a modest increase in latency due to OAC while most see no impact.

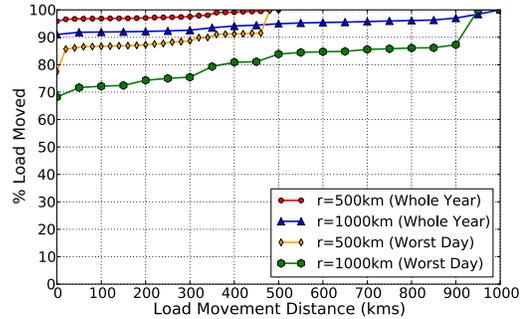


Figure 5: Average distance the load is moved by our algorithm for the whole year and on the worst day for $r = 500$ and 1000.

5. THERMAL ENERGY STORAGE

We explore two different benefits of TES. First, TES can reduce the need for chillers. With TES, a chiller with smaller capacity might suffice. Downsizing the chillers, or in some cases eliminating it, will result in a CAPEX savings that we quantify. Further, TES can also be used to reduce the OPEX by reducing the cost of energy needed to cool the IDN. In each of the IDN’s data centers, TES can be used to produce chilled water (or ice) when electricity prices are lower and used for cooling the servers when prices are higher.

5.1 Reduction in chiller capacity

Chillers are provisioned at a data center by first computing the energy that needs to be provided for cooling the data center over a long period (say, a year) and sizing the chiller to provide the peak cooling energy required over that period². Typically, provisioning is done in an offline fashion using historical or predicted values for the cooling energy. Our methodology follows this approach closely. We compute the cooling energy required for each IDN’s data centers over a year and take the peak requirement at each location as the chiller capacity (in KWs) required at that location. We then incorporate TES at each data center and recompute the chiller capacity required at each location. Figure 6 shows an example of downsizing the peak cooling power required using TES.

5.1.1 Computing the chiller capacity reduction:

To compute the smallest capacity chillers needed at each data

²In practice, one would inflate this peak requirement by a safety factor to over-provision the chiller. However, our results are independent of such a factor since we only use relative capacity values to compute savings.

center with TES, we minimize the chiller capacity required (chiller_cap) subject to the constraints below:

The total cooling power required is provided by either the chiller or the TES at each time step t .

$$\text{chiller_power}(t) + \text{tes_discharge}(t) = \text{cooling_power}(t), \forall t \quad (1)$$

The amount of power drawn from chiller at any time step is non-negative and cannot exceed its provisioned capacity.

$$0 \leq \text{chiller_power}(t) \leq \text{chiller_cap}, \forall t. \quad (2)$$

The TES discharge (resp., charge) variable is positive while discharging (resp., charging) and zero when charging (resp., discharging). The charge and discharge rates cannot exceed their respective maximum values.

$$0 \leq \text{tes_discharge}(t) \leq \text{max_discharge}, \forall t. \quad (3)$$

$$0 \leq \text{tes_charge}(t) \leq \text{max_charge}, \forall t. \quad (4)$$

Energy stored in the TES increases (resp., decreases) when it is charged (resp., discharged). Some energy is lost in the process as captured by the TES's efficiency.

$$\begin{aligned} \text{energy_stored}(t) + \text{tes_efficiency} \cdot \text{tes_charge}(t) \\ - \text{tes_discharge}(t) = \text{energy_stored}(t+1), \forall t. \end{aligned} \quad (5)$$

Energy stored cannot exceed the TES capacity.

$$0 \leq \text{energy_stored}(t) \leq \text{tes_cap}, \forall t. \quad (6)$$

Given cooling requirements, $\text{cooling_power}(t)$, $1 \leq t \leq T$ and TES parameters, we compute the smallest chiller capacity required at each data center using the algorithm below.

Step 1. Fix a value for chiller_cap and check if a chiller of that capacity would suffice in providing the cooling requirements for the data center as follows. Compute the required sequence of TES discharges by observing that when $\text{cooling_power}(t) > \text{chiller_cap}$ the TES should discharge to provide $\text{cooling_power}(t) - \text{chiller_cap}$. To check if this sequence of discharges is feasible, we use a greedy TES charging algorithm that charges the maximum allowable amount at each time step where $\text{cooling_power}(t) < \text{chiller_cap}$, subject to constraints 3, 4, 5, and 6. It is clear that the discharge sequence is feasible iff the greedy charging algorithm can provide those discharges without violating the constraints.

Step 2. Using binary search, fix values for chiller_cap and test for feasibility as above, till the smallest value is found.

Step 3. The chiller capacity reduction is simply $100 \times (1 - (\text{chiller_cap with TES}) / (\text{chiller_cap without TES}))$.

5.1.2 Empirical Results:

We analyze the reduction in chiller capacity as a function of TES capacity across the IDN on a global basis, in addition focusing on five major metro areas around the world (cf. Figure 7 (a)). Note that a small TES capacity of 30 minutes gives a substantial saving in chiller capacity of 25-45%, because even a small amount of TES is able to absorb and flatten the peaks in the cooling energy requirement. Without TES, those peaks would have to be provided by the chillers. Observe that with about 3.5 hours of TES capacity, one can reduce the chiller capacity by at least 50% globally and in all the five locations. Further, the law of diminishing returns hold where more TES capacity yields progressively smaller reductions in chiller capacity. Finally, while TES reduces chiller capacity it does not eliminate chillers all together.

Thus far, we assumed that the TES efficiency was 90%. To gauge the impact of efficiency on our results, we varied the TES efficiency from 50% to 90%, reflecting the many different TES technologies

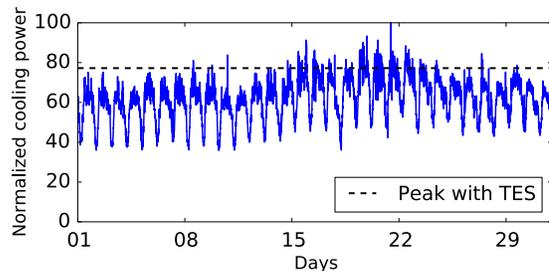


Figure 6: The peak cooling power required in Akamai's data center in New York can be reduced to 77% of its original value with 10 mins of TES capacity. TES discharges when cooling power required is above the 77% threshold value and charges when it is below the threshold, flattening of the power peaks.

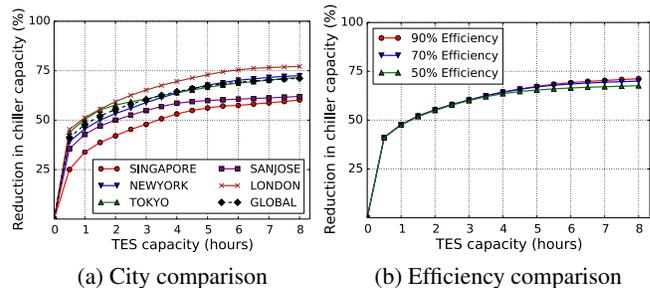


Figure 7: Impact of TES on chiller reduction

that are available today [12]. Figure 7(b) shows that the chiller reduction CAPEX decreases only a small amount as efficiency becomes smaller, since even at 50% efficiency the TES has sufficient time between the discharge intervals to recharge to nearly full capacity.

5.2 Reduction in energy cost

In addition to chiller capacity reduction (CAPEX saving), we evaluate the ability of TES to provide a reduction in the energy procurement cost (OPEX saving). The savings is due to shifting the electric power drawn from the grid from time periods when energy is expensive to when energy is cheap.

5.2.1 A near-optimal online algorithm for TES

A TES power management algorithm meets the cooling power requirements of each data center at each time step by either using the chillers or the TES. The algorithm decides when and how much to charge or discharge the TES, accounting for the workload and time-varying energy prices. Using the Lyapunov technique, we derive an online power management algorithm that achieves a cost that is provably within an additive constant of the optimal offline algorithm. An online algorithm makes decisions without knowing the future workload and energy prices, while an offline algorithm has full knowledge of both. Thus, the cost achieved by an offline optimal algorithm is a lower bound on what any online algorithm can achieve.

Our TES algorithm is generic and can be used for a broad range of energy storage devices, including TES and batteries. While prior Lyapunov-based online algorithms [9, 19, 11, 28] are known, they assume that the efficiency of the storage device is 100%. To our knowledge, our work is the first to explicitly incorporate the notion

of efficiency within the Lyapunov framework that is key to modeling storage devices. All real-world devices have various types of conversion losses that could range from 10-50% depending on the type of device. Our approach also better utilizes the available storage, resulting in a provably closer approximation of the optimal than prior work in [9, 19, 28, 11].

Given the time-varying price of electricity, energy cost is simply the sum over all time of the product of the energy price and the power drawn from the grid, i.e.,

$$\text{energy_cost} = \sum_{t=1}^T \text{grid_power}(t) \cdot \text{energy_price}(t). \quad (7)$$

We minimize the energy cost subject to the following constraints. The cooling energy is provided using power drawn from the grid and the power drawn from TES. Thus,

$$\begin{aligned} \text{grid_power}(t) + \text{tes_discharge}(t) - \text{tes_charge}(t) \\ = \text{cooling_power}(t), \forall t. \end{aligned} \quad (8)$$

Note above that when TES is charged, the grid provides the energy for both charging the TES and for cooling the data center via the chillers. However, when the TES is discharged, the grid provides only the difference between the cooling power requirement and the TES discharge amount. Finally, we enforce the condition that the power drawn from the grid should lie between 0 and max_grid , the maximum power that can be drawn from the grid, i.e.,

$$\text{max_grid} \geq \text{grid_power}(t) \geq 0, \forall t. \quad (9)$$

The energy cost optimization problem is simply minimizing the objective function in Equation 7 subject to the constraints that capture TES properties in Equations 5, 6, 3, and 4 and the constraints in Equations 8 and 9 that capture the manner in which power is drawn from the grid.

To design an online algorithm for the above optimization, we use an optimal control technique called Lyapunov optimization [22]. The resulting online algorithm works in real-time knowing only the inputs up and until the current time t , without knowledge of inputs greater than t . Specifically, it is a *myopic algorithm* whose control decisions in any slot t depend only on the system state in that slot.

To describe our online algorithm, we assume that the control horizon T is divided into intervals of size S and there are I such intervals, i.e., $T = S \cdot I$. Let max_price_i and min_price_i denote the maximum and minimum unit energy price in interval i . For each interval i , the Lyapunov-based algorithm makes use of a control parameter V_i which is defined as

$$V_i = \frac{\text{tes_cap} - \text{max_charge} - \text{max_discharge}}{\gamma_i - \text{min_price}_i} \quad (10)$$

where $\gamma_i \triangleq \text{max_price}_i / \text{tes_efficiency}$. The algorithm uses the queueing variable $X(t)$ in making its control decisions:

$$X(t) = \text{energy_stored}(t) - V_i \cdot \gamma_i - \text{max_discharge}. \quad (11)$$

where t_i denotes the interval containing slot t . The $X(t)$ variable is simply a shifted version of the TES energy level $\text{energy_stored}(t)$. Given $X(t)$ and current price $\text{energy_price}(t)$, the algorithm makes charge/discharge decisions by minimizing the following objective function every slot subject to the constraints 3, 4, and 8.

$$\begin{aligned} \text{tes_charge}(t)(\text{tes_efficiency} \cdot X(t) + V_i \cdot \text{energy_price}(t)) \\ - \text{tes_discharge}(t)(X(t) + V_i \cdot \text{energy_price}(t)) \end{aligned} \quad (12)$$

The optimal solution for above has the simple structure below:

1. If $X(t) \geq -V_i \cdot \text{energy_price}(t)$, then $\text{tes_charge}(t) = 0$ and the optimal discharge value is given by $\text{tes_discharge}(t) = \min[\text{max_discharge}, \text{cooling_power}(t)]$. Essentially, we discharge from TES as much as possible while drawing from the grid for remainder of the cooling power required.
2. If $X(t) \leq -V_i \cdot \text{energy_price}(t) / \text{tes_efficiency}$, then we have $\text{tes_discharge}(t) = 0$ and the optimal charge value is given by $\text{tes_charge}(t) = \min[\text{max_charge}, \text{max_grid} - \text{cooling_power}(t)]$. In this case, we charge the TES as much as possible in addition to drawing from the grid for all of the cooling power required.
3. If $-V_i \cdot \text{energy_price}(t) / \text{tes_efficiency} < X(t) < -V_i \cdot \text{energy_price}(t)$, then we do not charge or discharge and the cooling power is drawn from the grid.

The above structure follows by comparing the terms multiplying $\text{tes_charge}(t)$ and $\text{tes_discharge}(t)$ in (12) and choosing the strategy that minimizes the objective. It can be seen that this algorithm jointly makes use of the current energy level (reflected by $X(t)$) and the current price as well as the tes_efficiency is making its decisions and has a *threshold structure*. Our control algorithm differs from prior work derived in the context of batteries in [28] with the following key differences: (i) We explicitly consider storage efficiency, and (ii) unlike [28] that uses a fixed V for all t , we vary this control parameter dynamically. As shown in the theorem below, this results in a superior performance guarantee.

THEOREM 1. *Suppose that our online algorithm described above is implemented over T slots containing I intervals each of size S and suppose in each interval i , it uses control parameter V_i as defined in (10). Then, the following hold:*

1. *The queue $X(t)$ is deterministically upper and lower bounded for all t as follows:*

$$\begin{aligned} -V_i \cdot \gamma_i - \text{max_discharge} \leq X(t) \leq \text{tes_cap} - V_i \cdot \gamma_i \\ - \text{max_discharge} \end{aligned} \quad (13)$$

2. *The TES energy level $\text{energy_stored}(t)$ satisfies for all t :*

$$0 \leq \text{energy_stored}(t) \leq \text{tes_cap} \quad (14)$$

3. *Suppose the processes $\text{energy_price}(t)$ and $\text{cooling_power}(t)$ are i.i.d. over slots. Then the expected per slot cost under the online algorithm is within $\hat{B}/V_{hm}(I)$ of the optimal offline value.*

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \mathbb{E} \{ \text{grid_power_online}(t) \cdot \text{energy_price}(t) \} \leq \\ \frac{\hat{B}}{V_{hm}(I)} + \frac{1}{T} \sum_{t=1}^T \mathbb{E} \{ \text{grid_power_offline}(t) \cdot \text{energy_price}(t) \} \end{aligned} \quad (15)$$

where \hat{B} is a constant (independent of V) defined as

$$\hat{B} = \frac{(\text{max_discharge})^2 + (\text{max_charge})^2 + \frac{(\text{tes_cap})^2}{S}}{2} \quad (16)$$

and $V_{hm}(I)$ is the harmonic mean of V_i , i.e., $V_{hm}(I) = \frac{I}{\sum_{i=0}^{I-1} \frac{1}{V_i}}$.

PROOF. See Appendix A. \square

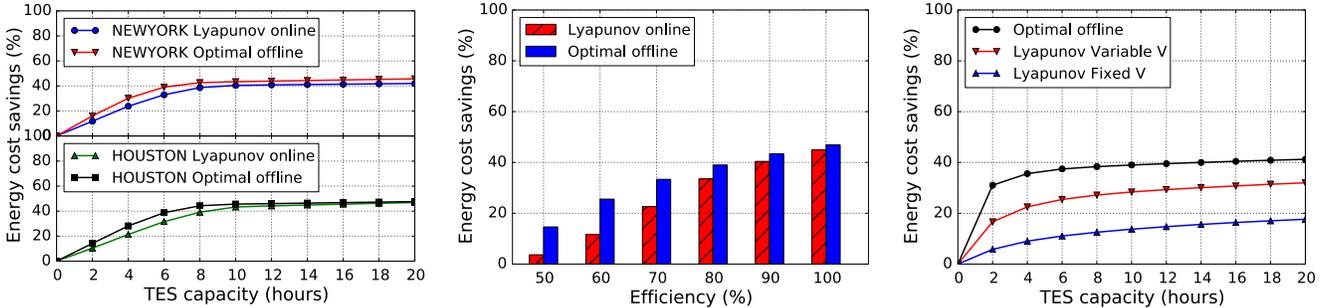


Figure 8: Lyapunov versus offline optimal for Figure 9: Impact of different TES efficiencies Figure 10: Comparing different Lyapunov algorithms with offline optimal. NY and Houston (TOU pricing). (New York with TOU pricing).

The performance bound (15) should be contrasted with that of [28] which uses a fixed V for all t . Indeed, the fixed V case corresponds to using a single interval. In this case, the V used by the algorithm will be $\min_i V_i$. Since $V_{\text{hm}}(I) \geq \min_i V_i$, the performance bound provided by our algorithm is superior. Also, note that while the theorem above assumes i.i.d inputs, that assumption can be relaxed to include markovian inputs using the framework in [22] to get similar bounds.

5.2.2 Empirical comparison with the offline optimal

In this section, we use real-world workload and pricing traces to evaluate the performance of our Lyapunov algorithm. We also compare it to the offline optimal solution that is obtained by solving a linear program as discussed before.

Theorem 1 guarantees that the Lyapunov algorithm is within an additive factor of the optimal offline and that this factor decreases as TES capacity is increased. To validate this, we calculated the cost savings under the Lyapunov algorithm for increasing values of TES capacity for two cities (New York and Houston) with the TOU pricing model. We also computed the offline optimal value using the LP formulation. Figure 8 shows that the gap between the offline optimal and Lyapunov becomes small as TES capacity is increased. In both cases, the gap is less than 3.4% for TES capacity of 10 hours or more.

So far, we assumed that the TES efficiency was 90%. Next we compare the impact of lower TES efficiencies on the cost savings under Lyapunov and offline optimal as shown in Figure 9. As TES efficiency increases from 50% to 100% for a 10 hour capacity TES, the gap between them reduces from 11% to 2%. This is because lower efficiency makes the $V_{\text{hm}}(I)$ smaller, thereby increasing the additive factor.

Finally, we compare the cost savings under Lyapunov and offline optimal for the real-time market pricing model. Unlike TOU pricing, it exhibits a large variation and degree of variation itself can change over the course of the control horizon. In Figure 10, we compare the performance of two variants of our Lyapunov algorithm: the more standard approach of using a fixed V and our newer approach of using a variable V . The interval size for the variable V algorithm was chosen to be 1 hour. As predicted by the bound (15), our variable- V algorithm significantly outperforms the fixed- V algorithm. Intuitively, this is because a dynamic V allows more efficient use of the available TES capacity. Note that the gap between offline optimal and Lyapunov remains non-negligible suggesting that for market pricing, there is room for improvement.

5.2.3 Empirical Results

We study the energy costs that can be saved using TES by us-

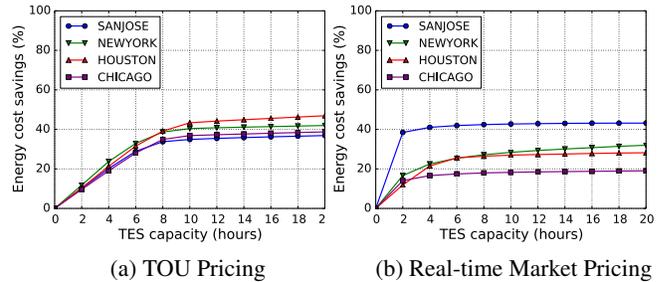


Figure 11: Annual energy cost savings for 4 major US cities.

ing both the TOU pricing model and real-time market pricing from four major metros in the US. Our choice of metros for market pricing were restricted to the US since we were unable to get real-time market prices for other locales. For each metro, we used our Lyapunov algorithm to simulate charge/discharge TES decisions to minimize the total energy cost for the entire year. The energy cost reduction is computed in comparison with using no TES at all. Specifically, energy cost reduction equals $100 \times (1 - (\text{cost with TES})/(\text{cost without TES}))$.

The energy cost reductions from TES arise from drawing energy from the grid when prices are lower and discharging when prices are higher. Figure 11 shows energy cost reduction as a function of TES capacity for TOU and market pricing. Unsurprisingly, as capacity increases so does the energy cost saving. Though returns diminish with increasing capacity as there are fewer energy price differentials that TES can exploit.

Now, compare the use of TES for chiller capacity reduction (Figure 7 (a)) versus its use for energy cost reduction (Figure 11). To achieve a 25-45% chiller capacity reduction a TES capacity of just 30 minutes was sufficient. While several hours of capacity is required for comparable energy cost savings. That is because chiller capacity reduction is obtained by flattening the peaks in cooling power, which requires only the small amount of energy in the “narrow” peaks to be stored. Whereas energy cost saving involves storing large amounts of energy drawn from the grid during low-price periods for use during high-price periods, requiring a larger capacity.

6. COMBINING OAC AND TES

We evaluate using both OAC and TES and study its impact on both CAPEX and OPEX. Our strategy is to first use as much OAC

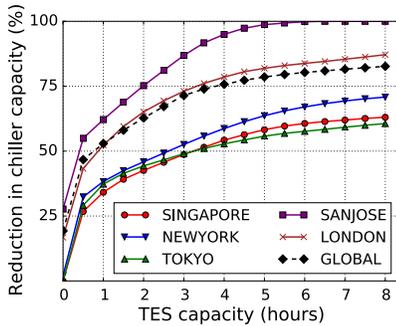


Figure 12: OAC and TES provide more chiller capacity reduction than just TES.

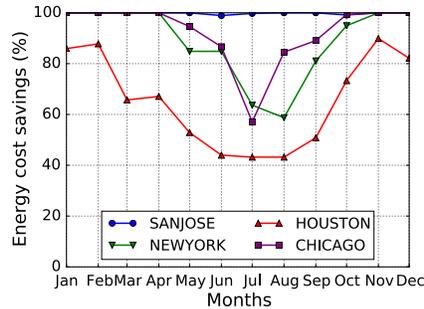


Figure 13: Energy cost savings for four major US metros with TOU pricing.

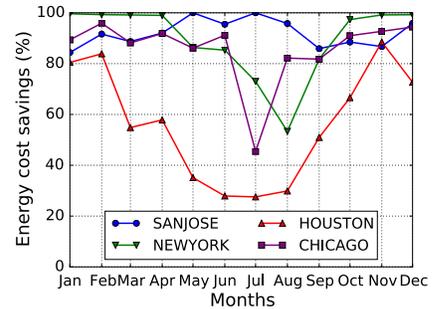


Figure 14: Energy cost savings for four major US metros with market pricing.

as possible, potentially by moving load across data centers of the IDN using the algorithms presented in Section 4. Next, we use TES within each data center using the algorithms presented in Section 5. The impact of OAC is to reduce the cooling energy supplied from the TES and chillers. TES further adds to the benefit by “flattening” the peaks of the cooling energy resulting in a chiller capacity reduction. Also, the TES can “move” the energy drawn from the grid from more expensive hours to cheaper hours. *For our evaluation, we use OAC with no performance impact ($r = 0km$), as even that provided favorable results in combination with TES.*

6.1 Reduction in chiller capacity

To determine the smallest chiller capacity required at each data center of the IDN we first use OAC whenever possible and then use the chiller reduction algorithm for TES as described in Section 5.1.1. Figure 12 shows the chiller capacity reduction as a function of TES capacity for the global IDN and five major metros around the world. In comparison with using TES only (cf. Figure 7(a)), the addition of OAC results in higher reductions, especially in San Jose and London where the chillers can be reduced by more than 50% with just 1 hour of TES capacity. Further, chillers can be *practically eliminated* in San Jose with OAC and TES capacity of about 5 hours providing all the required cooling energy. The other three cities each have periods of the year where OAC is unavailable resulting in less significant additional reductions due to OAC.

6.2 Reduction in energy cost

We study the reduction in energy cost that comes both from a decrease in the required cooling power from the grid due to OAC and shifting of the grid energy usage from more expensive periods to cheaper periods using TES. In Figures 13 and 14, we fix the TES capacity to 10 hours and show the energy cost reduction in the four major US metros through the whole year for TOU and market pricing respectively. San Jose has at least 84% savings all year round. The other metros show great seasonal variations with summer months having much less savings due to unavailability of OAC. OAC gives a bulk of the savings. But, in July and August when OAC is scarcer, TES provides more additional value.

7. RELATED WORK

Data centers and IDNs consume a significant amount of energy and techniques to reduce the energy consumption have been studied extensively. Much of this effort has focused on reducing the power consumption of server clusters through advanced cluster-wide power management techniques [3, 16]. In the IDN context

in particular, techniques such as server- and cluster shutdown have been proposed to make server clusters more energy proportional to IDN traffic [21]. Moving load across data centers of an IDN to exploit variation in energy market prices [25] and for increasing the use of renewables [18] has been studied. However, we explore load movement for evaluating the potential for OAC using extensive IDN load and global weather traces. Separately, the use of renewable energy to power and cool data center servers [17] and techniques to minimize server carbon footprint [6] have also been studied. In the context of reducing the cooling energy, thermal engineering techniques have been studied to optimize temperature and air flow through server racks or perform temperature-aware scheduling of workloads on “cool” racks [4]. However, the use of new cooling technologies in data centers has only recently begun to gain attention [1, 17]. In the context of OAC, recent groundbreaking work has focused either on the systems aspect of incorporating OAC into a modular data center [5] or on provisioning [20] and temperature management [7] within a single data center. Companies such as Facebook have begun to employ open air cooling in their data centers in recent years [26, 30]. Similarly, while TES has long been used in building and industrial applications, its use in data centers has only gained recent recognition [29].

To the best of our knowledge, prior scientific studies have not examined the potential benefits of exploiting OAC and TES at a global scale by a distributed network deployed across a large number of data centers, that is the focus of our work.

8. CONCLUSIONS

In this paper we studied the potential benefits of using OAC and TES to reduce the energy usage as well as the operational and capital costs incurred by an IDN. We presented algorithms to incorporate both technologies into the IDN architecture and empirically evaluated their efficacy using extensive traces from Akamai’s global CDN and global weather data from NOAA. We showed that, by using OAC, a global IDN can extract a 51% cooling energy reduction during summers and a 92% reduction in the winter. Further, a IDN can reduce its global chiller capacity by 41% with as little as 30 minutes of TES and reduce its energy costs by 18-43% with 10 hours of TES in major metros. Combining OAC and TES enables a IDN to reduce its global chiller capacity by 47% or even eliminate them entirely in places such as San Jose, California. Overall, we show that OAC and TES hold great promise for the future sustainable IDN design.

Acknowledgements: This research is supported in part by NSF grants CNS-1422245, CNS-1229059 and CNS-1413998 and by Massachusetts Department of Energy Resources.

9. REFERENCES

- [1] M. Arlitt, C. Bash, S. Blagodurov, Y. Chen, T. Christian, D. Gmach, C. Hyser, N. Kumari, Z. Liu, M. Marwah, A. McReynolds, C. Patel, A. Shah, Z. Wang, and R. Zhou. Towards the design and operation of net-zero energy data centers. In *Proc. 16th IEEE Conf. on Thermal and Thermomechanical Phenomena in Electronic Systems*, 2012.
- [2] L. Barroso and U. Holzle. The case for energy-proportional computing. *IEEE Computer*, 2007.
- [3] A. Chen, W. Das, A. Qin, A. Sivasubramaniam, Q. Wang, and N. Gautam. Managing server energy and operational costs in hosting centers. In *Proc. of ACM Sigmetrics*, 2005.
- [4] N. El-Sayed, I. Stefanovici, G. Amvrosiadis, A. Hwang, and B. Schroeder. Temperature management in data centers: Why some (might) like it hot. In *Proc. of ACM Sigmetrics*, 2012.
- [5] H. Endo, H. Kodama, H. Fukuda, T. Sugimoto, T. Horie, and M. Kondo. Effect of climatic conditions on energy consumption in direct fresh-air container data centers. In *Proc. of IEEE Intl. Green Computing Conference (IGCC)*, 2013.
- [6] P. Gao, A. Curtis, B. Wong, and S. Keshav. It's not easy being green. In *Proceedings of the ACM SIGCOMM*, 2012.
- [7] Í Goiri, T. D. Nguyen, and R. Bianchini. Coolair: Temperature-and variation-aware management for free-cooled datacenters. In *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2015.
- [8] *Google Embraces Thermal Storage in Taiwan, Data Center Knowledge, April 2012*. <http://bit.ly/LyhctC>.
- [9] Y. Guo and Y. Fang. Electricity cost saving strategy in data centers by using energy storage. In *Proc. of the IEEE Transactions on Parallel and Distributed Systems*, June 2013.
- [10] T. Harvey, M. Patterson, and J. Bean. Updated air-side free cooling maps: The impact of ASHRAE 2011 allowable ranges. *The Green Grid*, 2012. <http://bit.ly/Lf114Q>.
- [11] L. Huang, J. Walrand, and K. Ramchandran. Optimal demand response with energy storage management. In *Proc. of the Third IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Nov 2012.
- [12] Thermal energy storage technology brief. IRENA Technology Brief E17, January 2013.
- [13] S. Kaune, K. Pussep, C. Leng, A. Kovacevic, G. Tyson, and R. Steinmetz. Modelling the internet delay space based on geographical locations. In *Proc. of the 17th Euromicro International Conference on Parallel, Distributed and Network-based Processing*, Feb 2009.
- [14] J. Koomey. Growth in data center electricity use 2005 to 2010. *Analytics Press*, 2010.
- [15] S. Lee, R. Urgaonkar, R. Sitaraman, P. Shenoy. Cost minimization using renewable cooling and thermal energy storage in cdns. In *Proc. of the IEEE International Conference on Autonomic Computing (ICAC)*, July 2015.
- [16] M. Lin, A. Wierman, L. Andrew, and E. Thereska. Dynamic right-sizing for power-proportional data centers. In *Proc. IEEE INFOCOM*, 2011.
- [17] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyser. Renewable and cooling aware workload management for sustainable data centers. In *Proc. of ACM Sigmetrics*, 2012.
- [18] Z. Liu, M. Lin, A. Wierman, S. Low, and L. Andrew. Greening geographical load balancing. In *Proc of ACM Sigmetrics*, 2011.
- [19] A. Hasan Mahmud and S. Ren. Online capacity provisioning for carbon-neutral data center with demand-responsive electricity prices. *SIGMETRICS Perform. Eval. Rev.*, August 2013.
- [20] I. Manousakis, Í Goiri, S. Sankar, T. D. Nguyen, and R. Bianchini. Coolprovision: Underprovisioning datacenter cooling. In *ACM Symposium on Cloud Computing (SoCC)*, 2015.
- [21] V. Mathew, R. Sitaraman, and P. Shenoy. Energy-efficient content delivery networks using cluster shutdown. In *Proceedings of IEEE IGCC*, 2013.
- [22] M. J. Neely. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan and Claypool Publishers, 2010.
- [23] E. Nygren, R.K. Sitaraman, and J. Sun. The Akamai Network: A platform for high-performance Internet applications. *ACM SIGOPS Operating Systems Review*, 2010.
- [24] S. Pelley, D. Meisner, T. Wenisch, and J. VanGilder. Understanding and abstracting total data center power. In *Workshop on Energy-Efficient Design*, 2009.
- [25] A. Qureshi, R. Weber, H. Balakrishnan, J. Gutttag, and B. Maggs. Cutting the electric bill for internet-scale systems. In *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, 2009.
- [26] M. Treacy. Facebook Uses Open-Air Cooling in Super Efficient N.C. Data Center, 2012. <http://goo.gl/ZM1FdJ>.
- [27] *The Average PUE is 1.8: Uptime Institute*. <http://bit.ly/1eEqN9d>.
- [28] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. Optimal power cost management using stored energy in data centers. In *SIGMETRICS*, 2011.
- [29] Y. Wang, X. Wang, and Y. Zhang. Leveraging thermal storage to cut the electricity bill for datacenter cooling. In *Proceedings of the 4th Workshop on Power-Aware Computing and Systems*, 2011.
- [30] B. Weerts, D. Gallaher, R. Weaver, and O. Van Geet. Green data center cooling: Achieving 90% reduction: Airside economization and unique indirect evaporative cooling. In *Proc. of IEEE Green Technologies Conference*, 2012.
- [31] *Wisconsin Electric Rates*. <http://www.we-energies.com/pdfs/etariffs/wisconsin/elecrateswi.pdf>.

APPENDIX

A. PROOF OF THEOREM

Here, we prove Theorem 1. For notational convenience, we make use of the following shortened variables:

$$\begin{aligned} \text{cooling_power}(t) &= C(t), \text{energy_price}(t) = P(t) \\ \text{grid_power}(t) &= G(t), \text{energy_stored}(t) = Y(t) \\ \text{tes_charge}(t) &= R(t), \text{tes_discharge}(t) = D(t) \\ \text{max_charge} &= R_{\max}, \text{max_discharge} = D_{\max} \\ \text{tes_efficiency} &= \alpha \end{aligned}$$

PROOF. (Theorem 1 part 1) We omit the subscript i for notational convenience. We first show that (13) holds for $t = 0$. Using the definition of $X(t)$ from (11), we have that $Y(0) = \text{tes_cap} = X(0) + V\gamma + D_{\max}$. It follows that $X(0) = \text{tes_cap} - V\gamma - D_{\max} \geq -V\gamma - D_{\max}$. Also, $X(0) = \text{tes_cap} - V\gamma - D_{\max} \leq \text{tes_cap} - V\gamma - D_{\max}$. Thus, (13) holds for $t = 0$.

Now suppose (13) holds for slot t . We will show that it also holds for slot $t + 1$. First, suppose $X(t) \geq -V \cdot \text{min_price} \geq -VP(t)$. Then, by case 1 of the optimal solution to (12), we have that $R(t) = 0$ and $X(t + 1)$ cannot increase. Next, if $X(t) < -V \cdot \text{min_price}$, the most it can increase is R_{\max} . Thus $X(t + 1) \leq -V \cdot \text{min_price} + R_{\max} \leq \text{tes_cap} - V\gamma - D_{\max}$ (using the definition of V_{\max}). This shows the bound on the right hand side of (13).

Next, suppose $X(t) \leq -V\gamma \leq -VP(t)/\text{tes_efficiency}$. Then, by case 2 of the optimal solution to (12), we have that $D(t) = 0$ and $X(t + 1)$ cannot decrease. Next, if $X(t) > -V\gamma$, the most it can decrease is D_{\max} . Thus, $X(t + 1) \geq -V\gamma - D_{\max}$. This shows the bound on the left hand side of (13). \square

PROOF. (Theorem 1 part 2) Substituting the definition of $X(t)$ from (11) in (13), we have

$$-V\gamma - D_{\max} \leq Y(t) - V\gamma - D_{\max} \leq \text{tes_cap} - V\gamma - D_{\max}$$

Thus, we have that $0 \leq Y(t) \leq \text{tes_cap}$. \square

PROOF. (Theorem 1 part 3) We make use of the technique of Lyapunov optimization to show (15). We start by defining the following Lyapunov function: $L(X(t)) = \frac{1}{2}X^2(t)$. Define the conditional 1-slot Lyapunov drift as follows:

$$\Delta(X(t)) = \mathbb{E}\{L(X(t+1)) - L(X(t))|X(t)\}$$

Using (11), $\Delta(X(t))$ can be bounded as follows:

$$\Delta(X(t)) \leq B - X(t)\mathbb{E}\{D(t) - \alpha R(t)|X(t)\} \quad (17)$$

where $B = (R_{\max}^2 + D_{\max}^2)/2$. Following the Lyapunov optimization framework, we add to both sides of (17) the penalty term

$V_i\mathbb{E}\{G(t)P(t)|X(t)\}$ to get the following:

$$\begin{aligned} \Delta(X(t)) + V_i\mathbb{E}\{G(t)P(t)|X(t)\} &\leq B \\ &\quad - X(t)\mathbb{E}\{D(t) - \alpha R(t)|X(t)\} + V_i\mathbb{E}\{G(t)P(t)|X(t)\} \end{aligned}$$

Using the relation $C(t) = G(t) - R(t) + D(t)$, we can rewrite the above as:

$$\begin{aligned} \Delta(X(t)) + V_i\mathbb{E}\{G(t)P(t)|X(t)\} &\leq B + V_i\mathbb{E}\{C(t)P(t)|X(t)\} \\ &\quad - X(t)\mathbb{E}\{D(t) - \alpha R(t)|X(t)\} - V_i\mathbb{E}\{D(t)P(t)|X(t)\} \\ &\quad + V_i \times \mathbb{E}\{R(t)P(t)|X(t)\} = \\ B + V_i\mathbb{E}\{C(t)P(t)|X(t)\} &- \mathbb{E}\{D(t)(X(t) + V_iP(t))|X(t)\} \\ &+ \mathbb{E}\{R(t)(\alpha X(t) + V_iP(t))|X(t)\} \end{aligned} \quad (18)$$

Observing the last two terms, it can be seen that the online algorithm (12) minimizes the right hand side of (18) over all possible feasible control policies. This includes the offline solution to the problem. Indicating these by $G'(t)$, we have

$$\Delta(X(t)) + V_i\mathbb{E}\{G(t)P(t)|X(t)\} \leq B + V_i\mathbb{E}\{G'(t)P(t)|X(t)\}$$

Taking the expectation of both sides, summing over all slots in interval i , and dividing both sides by V_i , we get

$$\begin{aligned} \sum_{t=(i-1)S}^{iS-1} \mathbb{E}\{G(t)P(t)\} &\leq \frac{BS}{V_i} + \sum_{t=(i-1)S}^{iS-1} \mathbb{E}\{G'(t)P(t)\} \\ &\quad - \frac{\mathbb{E}\{L(X(iS))\}}{V_i} + \frac{\mathbb{E}\{L(X((i-1)S))\}}{V_i} \end{aligned}$$

Now summing over all intervals $i \in \{0, 1, \dots, I-1\}$, we get

$$\begin{aligned} \sum_{t=0}^{T-1} \mathbb{E}\{G(t)P(t)\} &\leq \sum_{t=0}^{T-1} \mathbb{E}\{G'(t)P(t)\} + \sum_{i=0}^{I-1} \frac{BS}{V_i} \\ &\quad + \sum_{i=0}^{I-1} \frac{\mathbb{E}\{L(X((i-1)S))\}}{V_i} \end{aligned}$$

Dividing both sides by $T = SI$ and using the fact that $L(X((i-1)S)) \leq \frac{(\text{tes_cap})^2}{2}$ for all i , we have

$$\begin{aligned} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{G(t)P(t)\} &\leq \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{G'(t)P(t)\} + \frac{B}{I} \sum_{i=0}^{I-1} \frac{1}{V_i} \\ &\quad + \frac{(\text{tes_cap})^2}{2SI} \sum_{i=0}^{I-1} \frac{1}{V_i} \end{aligned}$$

Note that $\frac{1}{I} \sum_{i=0}^{I-1} \frac{1}{V_i}$ is the inverse of the harmonic mean of V_i for $i \in \{0, 1, \dots, I-1\}$. This shows (15). \square