Proceedings of IPACK2009
InterPACK'09
July 19-23, 2009, San Francisco, California, USA

IPACK2009-89032

INTEGRATED DESIGN AND MANAGEMENT OF A SUSTAINABLE DATA CENTER

Brian J. Watson, Amip J. Shah, Manish Marwah, Cullen E. Bash, Ratnesh K. Sharma,
Christopher E. Hoover, Tom W. Christian, Chandrakant D. Patel
Sustainable IT Ecosystem Laboratory
Hewlett-Packard Laboratories
Palo Alto, California, USA

ABSTRACT

The environmental impact of data centers is significant and is growing rapidly. However, there are many opportunities for greater efficiency through integrated design and management of data center components. To that end, we propose a sustainable data center that replaces conventional services in the physical infrastructures with more environmentally friendly IT services. We have identified five principles for achieving this vision: data center scale lifecycle design, flexible and configurable building blocks, pervasive sensing, knowledge discovery and visualization, and autonomous control. We describe these principles and present specific use cases for their application. Successful implementation of the sustainable data center vision will require multi-disciplinary collaboration across various research and industry communities.

NOMENCLATURE

AVT	Adaptive Vent Tile
CFD	Computational Fluid Dynamics
CRAC	Computer Room Air Conditioner
DCS	Data Center Synthesizer
DfC	Design for Change
DfE	Design for Environment
DfM	Design for Manageability
DfX	Design for X
IT	Information Technology
LCA	Life Cycle Assessment
LWPI	Local Workload Placement Index
PCA	Principal Component Analysis
PDU	Power Distribution Unit
SDC	Sustainable Data Center
SLA	Service Level Agreement
SLO	Service Level Objective
TCI	Thermal Correlation Index
TCO	Total Cost of Ownership
VFD	Variable Frequency Drive
VM	Virtual Machine

INTRODUCTION

A recent study found that IT is responsible for about 2% of global greenhouse gas emissions [1], about as much as the aviation industry. Furthermore, it projected that that this share would double by the year 2020. Increasing environmental concern and regulatory action will soon force a paradigm shift in how IT solutions are designed and managed across their lifecycles. Data centers are a prominent component of this impact, as well as the fastest growing.

To turn this crisis into an opportunity, we propose the development of a suite of technologies for a *sustainable data center* (SDC). The goal is to reduce the environmental footprint of a data center to such an extent that the services it offers are more environmentally friendly than conventional services offered within physical infrastructures. Thus, a SDC would have net positive effects on the environment.

Developing and demonstrating a SDC requires the multi-disciplinary collaboration of mechanical engineers, electrical engineers, computer scientists, and others. The *compute* infrastructure of the data center consists of thousands of servers hosting revenue-generating services, interconnected with each other and the outside world via networking equipment, and relying on storage devices for persistent data. The data center also has a *power* infrastructure that feeds electricity to all of the equipment, and a *cooling* infrastructure that removes heat from the equipment. The economic and environmental burden of the latter two infrastructures often equals or exceeds that of the compute infrastructure [2].

All of these infrastructures, and many of their components, have traditionally been designed and managed independently, resulting in unnecessary redundancy and waste. For example, CRAC units are often provisioned in tier 4 data centers at 2N redundancy [3]. This is to ensure sufficient backup cooling capacity in the event of CRAC unit failure. However, we have previously shown how to map the thermal zones of influence for each CRAC, and how to identify regions of the data center that naturally have 2x, 3x, or greater levels of cooling redundancy [4]. Using such thermal zones, hardware can be provisioned to services based on their availability and reliability requirements. For example, critical workloads can be placed within regions of the data center that are served by multiple thermal zones. The economic and environmental benefit of eliminating a redundant standby CRAC unit illustrates the type of advantage that can be garnered through integration of the compute and cooling infrastructures in the data center. Similarly integrating design and management of an entire data center, both within and across its compute, power, and cooling infrastructures, is crucial to our vision for a SDC.

As shown in Fig. 1, we have identified five principles for integrated design and management that cut across all three infrastructures, and indeed cut across the multiple disciplines that are needed to achieve this goal.

1) Data Center Scale Lifecycle Design. Existing data center design approaches are focused on assimilation of discrete components into an operational infrastructure that meets runtime objectives, such as performance and cost targets. From a sustainability perspective, however, the environmental impacts are distributed across the lifecycle of the data center – including the extraction of raw materials, manufacturing of the components and building, transportation, operation, and end-of-life. We propose an integrated approach that incorporates Design for Environment (DfE) principles across the lifecycle while allowing the data center designer to evaluate the necessary runtime objectives.

2) Flexible and Configurable Building Blocks. In the spirit of DfX, we utilize two additional concepts: *design for change* (DfC) and *design for manageability* (DfM). A DfC component can remain in service longer than its sub-components with shorter life-spans, and operators can *change* its functionality and operability over the lifetime of the data center in response to technological and business shifts. Similarly, DfM components can adjust to shifts on shorter timescales, like seasonal workload fluctuations, by exposing configurable features that administrators and other agents dynamically tune. DfC and DfM features may add complexity, cost, and environmental impact to an individual component, so data center scale lifecycle design is essential to assessing whether the holistic benefits of a given feature compensates for its increased impact.

3) Pervasive Sensing. To leverage the manageability of DfM components, sensing the current state of the system and identifying when the new state has been appropriately tuned are essential. This is accomplished via a pervasive sensing layer that continuously monitors the entire data center. There are many metrics of interest in a data center, and sensors must periodically sample each metric at different locations, both physical and logical. A scalable architecture for aggregating and disseminating this data to multiple consumers is essential, and the impact of each additional sensor must be weighed against its marginal benefit at the holistic level.

4) Knowledge Discovery and Visualization. What should be done with all of this data once it is collected? Knowledge discovery is a family of mathematical tools for detecting and predicting patterns and anomalies, and can provide better insight to the administrators and agents that manage a data center. Knowledge discovery techniques may also close the loop between integrated management and integrated design, by informing the data center scale lifecycle design process for infrastructure upgrades. Applying domain-specific human knowledge can augment these techniques, and we show how visualization technologies enable operators to assimilate a large quantity of data for better decision-making.

5) Autonomous Control. There are some decisions that must be made too frequently or that are too complex for humans to be effective arbiters. What could fill this gap is an ensemble of autonomous controllers that process information from the pervasive sensing and knowledge discovery layers, and adjust DfM building blocks in accordance with management policies defined by administrators, end user, and higher-level controllers. Coordinating this ensemble of controllers is necessary to ensure that they work together in harmony, and can holistically optimize a data center to the desired mix of operational cost and environmental impact.

This paper describes each principle in more detail, and quantifies their impact on specific use cases.

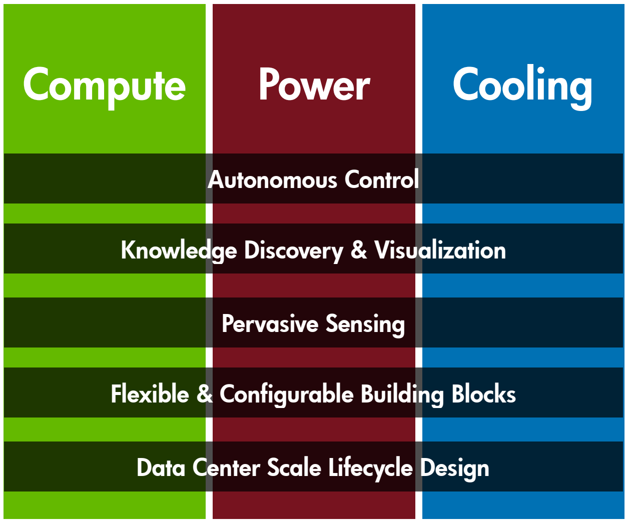

FIG. 1. PRINCIPLES FOR A SUSTAINABLE DATA CENTER

DATA CENTER ARCHITECTURE

Three infrastructures comprise a data center: information technology (compute), power distribution, and thermal management (cooling). Although their design and operation are often independent, their interactions affect the overall environmental impact of a data center over its lifetime.

Fig. 2 is a holistic diagram of a data center. The compute infrastructure typically consists of rows of racked IT equipment that can be separately categorized into server, storage, and networking components. Servers may run the IT services that generate value for a data center, and they rely on storage devices for persistent data and networking equipment for communication with each other and the outside world. IT equipment must operate within specified ranges of temperature and power criteria, and thus depend on high-quality power and cooling infrastructures.

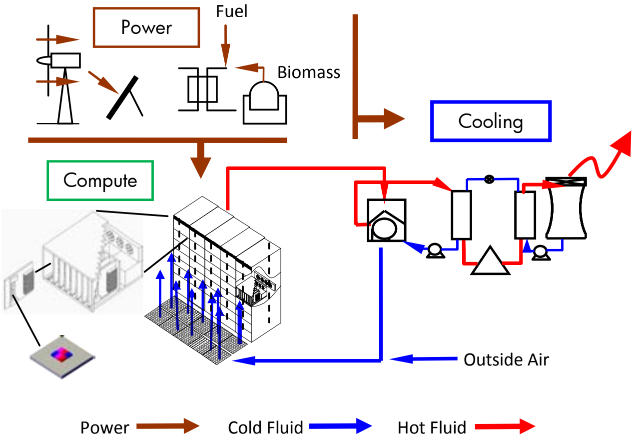

FIG. 2. DATA CENTER COMPONENTS

A variety of sources generate facility power, which local utility grids supply to most data centers. However, alternative approaches that utilize renewable and more sustainable energy sources – like solar-electric, wind, and fuel cells, among others – could be viable alternatives or supplements to grid-tied infrastructures. The source distributes power to transformers that manipulate the supply voltage. In the event of a failure, transfer switches enable the re-routing of power from a secondary source, such as on-site generators. Uninterruptible power supplies (UPSs) provide temporary power during the switching process. PDUs on the floor of the data center further adjust supply voltage and provide circuit-level distribution of power. Individual racks also contain PDUs that provide power to individual IT components in the rack.

Compressor-based systems commonly generate cooling, which either cool data center air directly or are coupled to a chilled water distribution system. Cooling towers or air-cooled condensers reject the resulting heat to the environment. Within the data center, CRAC units cool the air via heat exchange with the chilled water loop or via direct refrigeration. Blowers in the CRACs distribute the cool air into a plenum beneath a raised floor. Ventilation tiles in the raised floor near racks of IT equipment deliver the cool air, which absorbs heat from the IT equipment and returns to the CRACs. Alternatives that take better advantage of the local environment can replace or supplement parts of this standard infrastructure, including air- or water-side economization, direct use of outside air (i.e., “free cooling”), and absorption and adsorption units that run off of waste heat or combined heat and power systems.

INTEGRATED DESIGN

This section presents an integrated approach for data center scale lifecycle design, followed by a discussion about designing data center components for change and manageability.

Data Center Scale Lifecycle Design

Existing DfE principles are focused on component- or system-level implementation. Moreover, most DfE approaches are independent; practitioners generally consider other system design aspects (e.g., electrical, mechanical, thermal) separately and iterate for optimality. Such an iterative approach at the scale of a data center is not feasible. For example, many process-based LCA approaches that involve individual accounting of the different processes across the entire lifecycle – from extraction of raw materials and manufacturing to operation and end-of-life – are too laborious and costly to implement for each data center. Simultaneously, recursive design at the data center scale can be too complex to implement: if a particular compute design choice is found to be sub-optimal from an environmental perspective, the time and cost associated with manually redesigning the compute infrastructure (and its subsequent dependencies, including thermal and power delivery infrastructures) may be prohibitive.

Therefore, a SDC requires an integrated approach that systematically identifies environmental issues associated with different design choices, while ensuring that choices within one design domain do not violate design boundaries in another domain. This section describes such an approach.

Object-Based Environmental Design. First, a linked graph of "objects" describes the data center [5]. Each of these objects may possess static attributes (such as the mass of the system) as well as dynamic attributes (such as the runtime power consumption). The lifecycle of the object prior to its installation within the data center, also referred to as the "cradle to gate" portion of its lifecycle, usually determines the object’s static attributes. For example, a fixed amount of carbon is embodied in a server during the extraction of raw materials, manufacturing, assembly, and transportation to customer. This embodied carbon will not change regardless of how a customer uses the system, thus it is a "static" attribute of the system. On the other hand, while the design of a system might fix its peak power (another static attribute), its actual power consumption may change continuously depending on the applications it supports, the state of the system (peak versus idle), and even whether the system is operating or no longer in service. In this sense, the runtime power consumption of the system is a "dynamic" attribute that continuously changes as the server progresses through its lifecycle.

In addition to the embedded properties of the data center component that the static attributes represent, the spatial location of each object within the data center – and the flows via which the object links to other data center components – defines the scope of the object’s influence within the data center [5]. As an example, a microprocessor within an industry-standard server has the potential to impact everything from the chip package through the system up to the performance of the entire data center. On the other hand, the microprocessor is almost entirely independent of the construction materials being used in the physical shell of the building. Thus, the construction materials of the building will be out of the scope of influence of the microprocessor.

Thus, it becomes possible to describe the complete data center in spatio-temporal terms. The implementation of this framework should enable a variety of tools to simulate the performance of the data center. For example, lifecycle assessment methodologies [6,7] can describe most component-level static attributes, with particular emphases on cradle-to-gate aspects of the lifecycle. These component-level descriptions can then be aggregated into higher level indicators of sub-systems, or ultimately, entire systems and the data center at large. Performance models appropriate for each data center component may represent its dynamic attributes, such as IT system models (e.g., [8]) that can predict the dynamic behavior of particular aspects of the data center's compute infrastructure. The key output from this object-based approach is a library of components, systems, or infrastructures with static attributes and dynamic behavioral models in the data center environment.

The combination of objects from such a library enables the optimization of performance along additional design variables. We refer to this process of combining objects as 'synthesis'.

Data Center Synthesis. We are creating a *data center synthesizer* (DCS) that will automate and optimize the process of designing a SDC that meets its SLA and TCO requirements. The synthesis process begins with a description of the services to be provided and any constraints on the design. We have developed an SLA-driven approach that can transform SLOs for multi-tier applications into the computing resources required to deliver those services. The synthesizer next translates this equipment list into specific vendor and model numbers, using a component object library.

Power distribution and cooling equipment are selected to meet the operating characteristics of the computing, storage and networking equipment. Using workload and growth profiles from the services specifications, combined with knowledge about available real-time performance management for software service instances and data center hardware, a DCS can specify a minimal solution that meets requirements and allows for future expansion.

A projection of a candidate data center solution is the next step, which results in a complete model and physical layout of the data center, incorporating all computing, networking, and storage devices, as well as the complete power and cooling infrastructure required to support them. Following that is an analysis of the candidate data center to determine its operating characteristics. The analysis considers the integrated management capabilities to be discussed later in this paper, and its criteria include sustainability, performance, availability, thermal characteristics, and TCO. It uses existing CFD tools to perform a complete thermal analysis of data center operation.

The synthesizer then compares the analysis results of the candidate solution against the customer-specified goals for the data center and against other solutions from previous iterations. Provided that the solution meets the SLOs for services – or is reasonably close to those goals – it becomes a candidate for consideration as the recommended solution. If the data center design does not meet key customer requirements, or if there are other potential design solutions that might represent improvements, a re-synthesis process uses the results to modify the specifications and generate another candidate data center. Iteration continues in a fully-automated loop until an optimal solution is reached.

The DCS is still under development, but we manually applied its process to the use case of designing a hypothetical data center for a small city, which has a seaside location and a desert climate. Saltwater is abundant, but fresh water is scarce. Sunshine is plentiful all year long and rainfall is infrequent, an obvious opportunity for photovoltaic power generation. There is a predictable prevailing breeze that is reasonably consistent in direction, but lacks the velocity required for wind power generation. The mean daily low/high temperatures range from 12/24°C in the winter months to 29/42°C in the summer, making it possible to cool a data center with outside air during the evening for at least part of the year.

Data center requirements include e-mail services for 40,000 residents, a Human Resource Management system for 1500 businesses, real-time infrastructure monitoring for 10,500 sensors throughout the city, and a high-performance cloud computing environment for the local university. Using existing application sizers, we generated a baseline data center, and then we applied DCS techniques to derive more sustainable alternatives. The resulting design reduced TCO by over 11% and CO_2 emissions by 49% versus the baseline [9].

Flexible and Configurable Building Blocks

Data center administrators replace IT equipment on roughly three year cycles. In contrast, the cooling and power infrastructure is designed to last ten to fifteen years. This lifecycle variance can lead to reduced data center performance (efficiency and capacity) over time. Additionally, administrators manage IT equipment independently of facility infrastructures, further reducing efficiencies over the combined facility/IT lifecycle. In order to address this problem, the design of data center infrastructures must accommodate changes over time and thereby extend their overall useful lifetime. Additionally and related, infrastructure components must be flexible to accommodate the changing resource needs of the data center. This section provides examples for how to address these needs within the SDC context.

Design for Change (DfC). The evolution of computer architecture in recent years has significantly extended the useful lifetimes of chassis over past architectural generations. Fig. 3 shows a typical bladed chassis. This particular chassis houses 16 blades and cools them with 10 shared fans. Although the blades must accommodate certain chassis requirements, each blade can be different and can perform a wide variety of functions (e.g., processing, storage, etc.). Additionally, while the blades can undergo generational upgrades similar to traditional server architectures, they maintain compatibility with the original chassis architecture. This benefits the sustainability of the data center by reducing material use in the chassis over time, including metals, semiconductors, plastics, and ceramics, and by commensurately reducing the energy and other materials to manufacture the chassis and blades, versus an equivalent number of traditional rack-mounted servers. Following the work of Hannemann et al., we find that re-using the chassis of a rack-mounted industry standard server could potentially reduce the lifetime exergy consumption of the IT equipment by up to 17% [10].

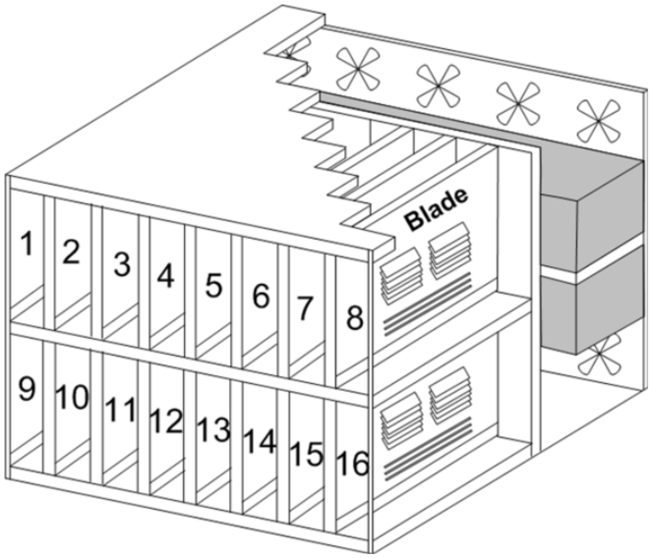

FIG. 3. BLADED CHASSIS

Design for Manageability (DfM). In addition to the need to design components that can accommodate change to improve their useful life, there is also a need to design components for manageability. In the bladed chassis of Fig. 3, the shared fans pull cool air from the data center through the chassis and over the blades. Each fan has a different thermal effectiveness for each blade, due to their proximity, and for each blade type, due to its design. One method of thermally managing this architecture is to operate the fans in unison with a uniform fan speed according to the thermal needs of the blades, where the hottest blade usually dictates the operation of the fans. If the blades have varying cooling requirements, this method can waste energy. In contrast, a DfM approach exposes independent control of each fan to external agents, which might consider the impact of each fan on each blade and adjust fan speeds commensurate with the needs of their associated blades. We have shown that such a methodology can reduce chassis fan power up to 30% over less flexible management schemes [11].

Fig. 4 shows another example of DfM: adaptive ventilation tiles for data centers. Traditional ventilation tiles are either fixed or equipped with manually operated damper mechanisms. As IT administrators replace equipment, they typically do not adjust or move the vent tiles accordingly, often because it is very difficult to make correct adjustments without a detailed analysis. This reduces the total heat load that the cooling infrastructure can accommodate, effectively reducing thermal capacity. To solve this problem, we have developed remotely-controllable adaptive ventilation tiles. Paired with an autonomous controller that considers the needs of nearby racks (i.e., the AVT Manager), such a DfM design can extend the life of the data center airflow management system by retaining thermal capacity that is often lost as IT equipment is replaced [12]. We have found that such a system can reduce power consumption in the cooling infrastructure by up to 49% over fixed flow tiles, the environmental benefit of which is likely to exceed any additional impact associated with manufacturing and operating an AVT.

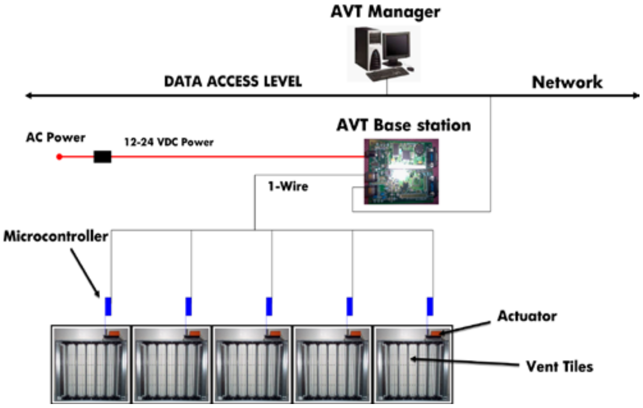

FIG. 4. ADAPTIVE VENT TILES FOR DATA CENTERS

INTEGRATED MANAGEMENT

Existing research on data center management focuses almost entirely on run-time optimization, ignoring the rest of the product lifecycle, and generally considers just a single technology vertical, like the compute infrastructure. By contrast, *integrated* management spans the power, cooling, and compute infrastructures within a data center to allow for finer-grained manageability and improved overall performance. In this section, we introduce three such principles for integrated management: pervasive sensing, knowledge discovery and visualization, and autonomous control.

Pervasive Sensing

Integrated management requires a pervasive sensing layer that continuously monitors the three infrastructures of a data center and the systems upon which the data center depends. There are many metrics of interest to periodically sample. This plethora of sensors generates a large volume of data that necessitates a scalable architecture for collecting, aggregating, and disseminating it to multiple consumers. In a SDC, data from the pervasive sensing layer feeds a number of subsystems, including historical data recorders, event detection and alarming mechanisms, data analytics, visualization tools, and autonomous controllers. Holistic data center management requires further extending sensing into the IT stack and even into business processes.

We previously described the thermal design of a typical raised-floor data center. Fig. 5 shows that there are significant sources of inefficiency in this design. Hot air ejected from IT equipment can *re-circulate* into equipment intakes, and cold air supplied by the CRACs can *short-circuit* back to the CRACs without cooling any IT equipment. Thermal management in a traditional data center relies on a temperature sensor in the return intake of each CRAC unit, which each CRAC maintains at a given set point. These adjustments happen independently. This architecture suffers from several problems. First, it cannot address the sources of cooling inefficiency in the data center, and thus requires over-provisioning of cooling to maintain reasonable thermal margins. Second, it is insensitive to local disturbances, like a vent obstruction or increased IT workload.

In prior work [13], we have shown that pervasive temperature sensing at the air intakes of IT equipment, rather than local sensing at the air intakes of CRAC units, combined with DfM enhancements to the CRAC units (i.e., installing VFDs in the blowers) and an autonomous control system, can reduce power consumption up to 35%, and can increase reliability by responding quickly to local disturbances. Five temperature sensors per IT equipment rack is sufficient to reap these benefits, and the incremental cost of materials and energy for such a system is miniscule compared to the cost of energy saved in cooling. Additional sensors on the IT equipment exhaust can help to detect a number of problematic and common modes of inefficient operation.

Besides temperature, other useful metrics include voltage, current, and frequency in various parts of the power infrastructure; pressure, mass flow, and velocity in the cooling hydronics system; and resource utilization, resource configuration, and application-level performance in the compute infrastructure. The latter provides examples of software-based sensors. The need for sensing is not confined to the environment within the data center. The performance of a chilled water hydronics system depends directly on site environmental conditions, as does free cooling and locally-sited energy sources, such as solar or wind. Measuring these conditions enables the sequencing of available resources for maximum efficiency. For example, a data center might only run low priority workloads when the sun is shining or when outside air is cool enough for the data center. However, designers must weigh the cost and environmental impact of each additional sensor against the opportunities for increased efficiency that they introduce.

As discussed above, the volume of this data and the size of many data centers require a high-performance, scalable architecture for collecting, aggregating, and disseminating the data to multiple consumers. Our architecture, code-named “Daffy”, allows a large number of loosely-coupled agents (“senders”) to publish sensor readings to its message broker. The message broker can scale horizontally as needed, and it sends the readings to any number of data consumers. The consumers indicate precisely what information they want (e.g., just temperature from specific racks). This reduces their load as well as that of the broker.

Daffy is based on a hybrid physical-logical-behavioral model of a data center that includes both facility and IT elements. From the facility domain, Daffy includes the location, name, and make-and-model of CRACs, PDUs, and racks, as well as facilities-level sensing capabilities, like thermal sensors and CRAC operating points. Daffy represents most of these objects as volumes in three-space geometry, and it can semi-automatically import much of this information from CAD drawings of the data center that are commonly available from facilities professionals.

From the IT domain, Daffy includes the location, name, make-and-model, and ownership (among other attributes) of IT equipment in the data center. Daffy represents the logical location of IT equipment by its particular rack and the number of U positions it occupies within its rack. There is an additional notion of an enclosure, which is necessary to model storage arrays and blades. As with the facilities data, Daffy can semi-automatically import much of this information from asset management tools used by IT professionals.

As an example of integrating facilities and IT domain information, Daffy provides real-time calculations of a *local workload placement index* (LWPI) for each IT device. This figure is measure of how efficient it is to cool a particular location in the data center [14,15]. LWPI is an example of *synthetic sensor* data, which is derived from measured values and other synthetic sensors, rather than measured directly itself.

Knowledge Discovery

Knowledge discovery refers to statistical, data mining and machine learning techniques that transform and analyze data collected from various sources. These techniques enable event and anomaly detection, causality inference, event and behavior

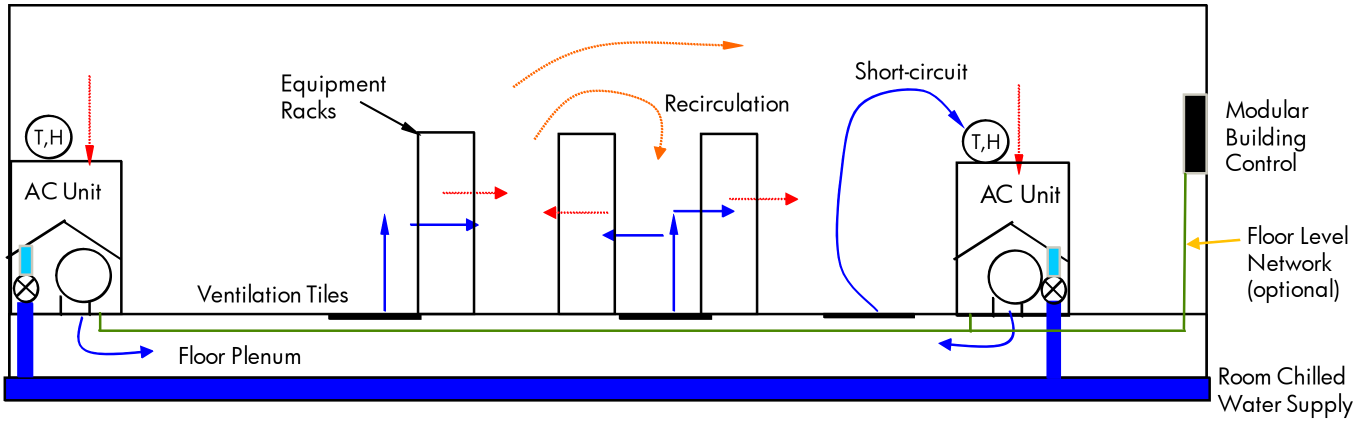

FIG. 5: DATA CENTER AIRFLOW

prediction, and discovery of patterns, rules, and associations. One of the goals of our research is to incorporate sustainability metrics (e.g., greenhouse gas footprint, energy consumption) into these techniques, facilitating operational optimizations from the perspective of sustainability.

In particular, these techniques further the goals of SDC integrated management in a number of ways. They can perform timely and specific detection, localization, and root cause analysis of anomalous behavior to determine corrective actions. For example, early detection of a CRAC unit failure allows redirection of alternate cooling resources to the affected racks. Knowledge discovery can also predict anomalous behavior and other significant events, facilitating pre-emptive resource re-allocation. This reduces the need to over-provision resources without sacrificing resiliency. Furthermore, sustainability metrics can drive device and process operational optimizations.

We are investigating a number of knowledge discovery techniques in the context of data centers. Here we briefly describe our experience with a few such techniques.

Principle component analysis (PCA). Temperature sensors continually monitor state-of-the-art data centers. Although simple threshold-based algorithms can easily detect thermal anomalies, like a large change or rate of change in a sensor temperature, they are inadequate for detecting thermal inefficiencies, such as increase in hot air recirculation (which decreases CRAC effectiveness, and hence increases power consumption). These anomalies routinely go undetected without human intervention. We use a technique based on PCA [16] for detecting events where correlation breaks between variables that are normally correlated. For example, if one sensor temperature increases while the temperature at a nearby sensor decreases, there might be a thermal inefficiency, which this technique would detect. Prompt detection of such anomalies provides an opportunity for correction.

Clustering. Clustering [17] is a technique that partitions a set of elements into groups (clusters), such that elements within a cluster are more similar than those across clusters. For example, applying a simple algorithm like k-means [17] to sensor data can aggregate similar operational points. Domain knowledge about the governing physical processes can enhance this algorithm through the use of constraints in the clustering algorithm. Estimated sustainability metrics can characterize each cluster, which indicate the desirability of operating in that cluster. For real-time assessment and active control of operational state, correlating sensor data streams to well-known clusters can obtain an estimate of the sustainability characteristics for the current state of operations.

Motif mining. Mining historical sensor data can reveal frequently occurring motifs (or patterns) [18]. Sustainability metrics can characterize these motifs and classify them as good or bad. Furthermore, sustainability metrics allow quantitative comparison of motifs. The ultimate objective is to constrain the operational state of the system to the most desirable motifs.

We applied this technique to manage an ensemble of five data center chiller units. We collected the utilization time series of the units over a period of about three weeks, and clustered the multivariate time series data using k-means. Next, we used the resulting cluster IDs to encode the five time series into a single sequence of symbols. We then mined this symbol stream for frequent motifs or patterns, which we characterized using sustainability metrics, such as energy consumption and carbon footprint. From these motifs, we observed that visually similar operational patterns differed significantly in their sustainability metrics. In fact, the most efficient motif provided a power savings of 41 KW over the least efficient motif [19]. This translates to an annual reduction of 287 tons of CO2.

Generational Knowledge Discovery. Knowledge gained during the operational phase can also benefit integrated design. Faults, inefficiencies, and limitations discovered during operations provide valuable insights for re-synthesis of the current data center and designing the next generation. Such operational information enables better re-design of operating regions for devices like CRAC blowers and chillers, and knowledge of thermal anomalies can aid in better layout of racks and CRAC units.

Visualization

Visual representation enables use of human cognitive abilities in the decision process. One use case on which we are focusing is to enable data center administrators to quickly identify and correct abnormal thermal states to improve cooling efficiency. These abnormalities often occur as a result of unexpected events in the cooling infrastructure, as well as changes to the IT equipment in the compute infrastructure, which alters airflow and heat loads and affects cooling efficiency in non-intuitive ways. We have applied a cell-based spatio-temporal view to simultaneously visualize complex thermal state changes over a high-resolution time series across hundreds of compute racks, and employed rich visualization techniques based on the detection of sensor temperature relations and events for identifying abnormal states and troubleshooting the root causes of problems [20,21].

As illustration, consider the example of chillers, which are a key hydronics component of the cooling infrastructure. They provide cold water to the CRAC units in the data center and extract heat from the warm return water. They eject the extracted heat into the environment through cooling towers or heat exchangers. Evaluating the efficiency of the chillers and estimating how long until the next maintenance downtime are crucial tasks. Data center administrators want to easily identify how each chiller is operating, examine its hourly usage patterns, and determine ways to improve its performance.

To monitor time-critical thermal states and to derive cooling policies for chillers in large scale production data centers, we have developed non-overlapped scatter plots, for which we applied a versatile method of displaying detailed data in a perspective that can be analyzed without overwhelming cognitive effort. Fig. 6 shows one such plot of chiller temperature and utilization. The x-axis represents the chilled water return temperature, and the y-axis represents the chilled water supply temperature. The color of a data point represents the chiller utilization (%U), from low (yellow) to medium (green to blue) to high (burgundy to red). A cell represents a data point *p(x, y, color)*, which our algorithm places into one of several temperature bins such that it does not overlap any other cells. Our placement algorithm defines the x- and y-value ranges for each bin according to the number of in- and out-temperatures for the chiller, and it determines the best location for each cell based on the available empty space in each bin. The algorithm sorts the cells by their utilization (color value) and clusters them together as shown in Fig. 6. This technique reveals interesting distributions and patterns.

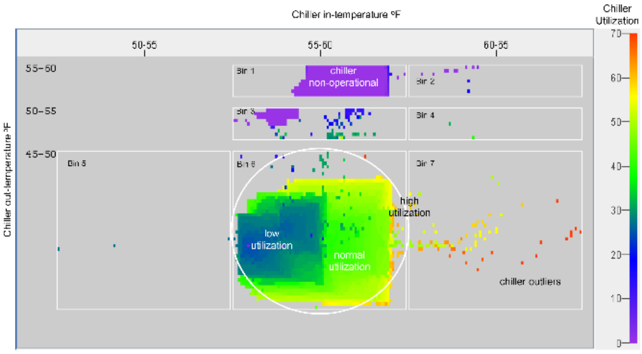

FIG. 6. NON-OVERLAPPING SCATTER PLOT OF CHILLER UTILIZATION

Through our user interface, an administrator can drill down into a bin to see the hourly usage patterns of a chiller, as shown in Fig. 7. This result leads the administrator to conclude that chiller utilization is higher during the day and early evening, as indicated by circles 1 and 2 (more red, orange, yellow, and green), than early morning and late evening (mostly blue). This may be due to increased server workloads generating high heat load. Interestingly, circles 1 and 2 indicate different operational patterns (more green in circle 2). Administrators can further drill down for more investigation, and can associate these patterns with cooling costs and environmental impact metrics. Identifying favorable modes of operation in clear and obvious terms will reduce uncertainty in infrastructure management.

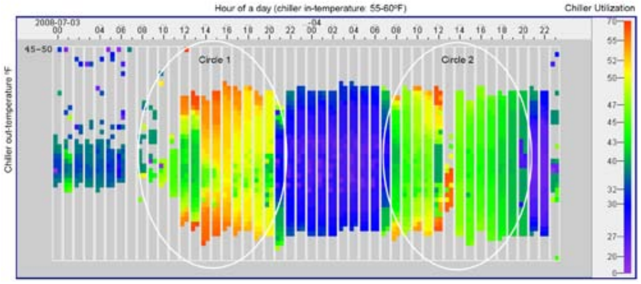

FIG. 7. CHILLER TEMPERATURE TIMESERIES

Autonomous Control

Some DfM modules operate on timescales too short for a human operator. Consider the example of the inkjet assisted spray cooling device in Fig. 8 [22,23]. In this example, an integrated circuit package incorporates an inkjet head, which supplies liquid coolant via hundreds of nozzles to hot spots on the chip die. Coolant flow rate can be varied per nozzle by adjusting tiny resistors in their firing chambers. The advantage of this solution is that it directs coolant on demand in a fine-grained manner according to the needs of the die, which are typically non-uniform across its surface. However, an administrator cannot manually control a device that operates on millisecond timescales. An autonomous controller is necessary.

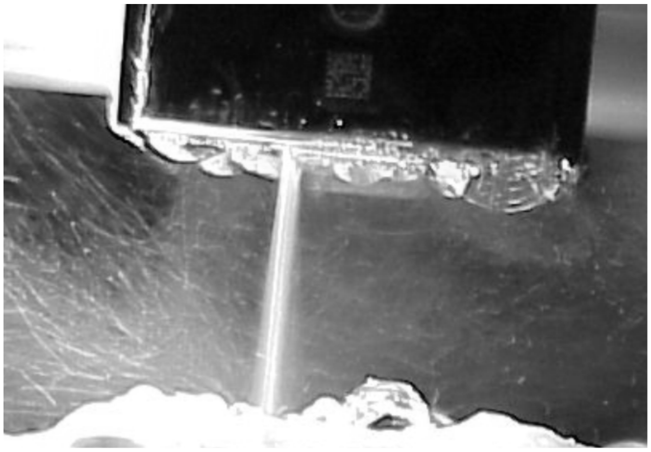

FIG. 8. INKJET ASSISTED COOLING

Other DfM modules require decision-making that is too complex for humans. An example of this is a set of CRACs outfitted with VFDs in their blowers, which theoretically enable the adjustment of data center cold air supply in response to time-varying heat loads in the data center. In practice, this is challenging to control. Continually measuring heat loads throughout an enterprise data center involves thousands of temperature sensors, generating too much data for a human operator to process. Furthermore, appropriately adjusting CRAC VFDs in response to this data requires an understanding of the *thermal correlation index* (TCI) between each CRAC and each temperature sensor, which can vary with changes in the data center airflow (e.g., adjusting vent tiles, adding / removing equipment) [13]. Knowledge discovery and visualization techniques could help an administrator to manage this complexity, but it is more straightforward and reliable to build an autonomous controller to manage these CRACs.

An even greater challenge is coordinating the control of multiple DfM components. An example in the compute infrastructure is virtualization, which adds many dimensions of flexibility to server management with its ability to dynamically adjust VM resource allocations and to move VMs between physical systems. When using the former to maintain low VM resource allocations and high utilization of server resources, demand spikes can quickly lead to response time violations unless a controller promptly allocates additional resources. The latter feature of VM migration is complex to control, since the number of possible mappings is equal to $(\textit{the number of physical systems})^{(\textit{the number of VMs})}$. Simultaneously controlling both DfM features introduces additional challenges [24].

In a cross-infrastructure example, conventional data centers typically allocate compute resources to incoming service requests without regard for cooling efficiency. The distribution of cooling resources in data centers is often variable across the floor space. Fig. 9 displays a CFD analysis of the temperature above the raised floor of a typical data center. The hot spot marked on the figure identifies a particularly inefficient location for thermal management. A poor allocation of IT resources can, therefore, lead to inefficient operation of the cooling infrastructure [14]. We have built a coordinated set of controllers, in which a VM migration controller leverages LWPI values from Daffy to place VMs according to cooling efficiency. By migrating as little as 0.5% of the total IT workload to efficient locations in our research data center, we reduced CRAC blower power by 27% [15]. A similar approach could provision batch workloads.

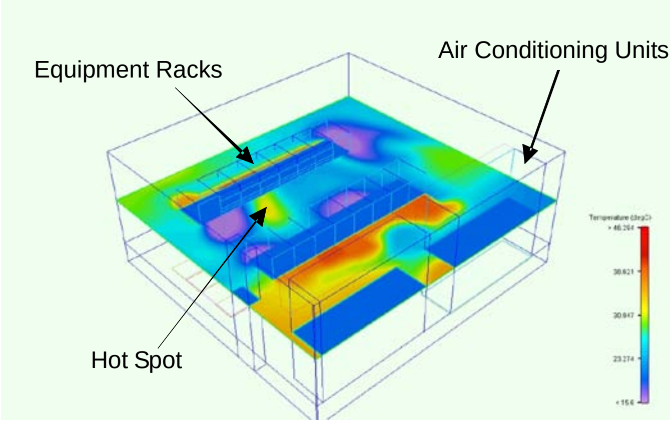

FIG. 9. DATA CENTER TEMPERATURE VARIATIONS

CONCLUSION AND NEXT STEPS

We propose an integrated design and management approach for developing a sustainable data center. The design of this data center consists of flexible and configurable building blocks that an automated synthesizer combines together in a manner that reduces the lifecycle environmental footprint of IT services below that of the conventional services that they replace. Integrating the management of this data center across its compute, power, and cooling infrastructures are the principles of pervasive sensing, knowledge discovery and visualization, and autonomous control. The synthesis process considers these management capabilities during its analysis of candidate data center designs, and knowledge discovery during operation provides feedback that improves the synthesis of future upgrades and subsequent generations of data centers.

Going forward, we are creating CAD-compatible tools for analyzing the impact of component design on the lifecycle sustainability of data centers, an automated data center synthesis tool for assembling these components together in the most sustainable manner, and a scalable data aggregation architecture for pervasive sensing of operational data. One of the flexible and configurable building blocks that we are developing is the adaptive ventilation tile, and there are likely to be many more opportunities for DfC and DfM components that the broader IT research community can investigate. Likewise, techniques for knowledge discovery and visualization of data center metrics, for controlling new DfM capabilities, and for coordinating the resulting ensemble of controllers are excellent areas for collaborative research in the pursuit of creating a sustainable data center.

REFERENCES

[1] Mingay, S., 2008, “IT’s Role in a Low Carbon Economy,” Keynote Address, *Greening the Enterprise 2.0 Conf.*

[2] Belady, C. L., 2007, "In the Data Center, Power and Cooling Costs More Than the IT Equipment It Supports," Electronics Cooling Mag., **13**(1), http://www.electronics-cooling.com/articles/2007/feb/a3/

[3] 2005, “Telecommunications Infrastructure Standard for Data Centers,” TIA-942, Telecommunications Industry Association, Arlington, VA.

[4] Bonilla, C. J., Ferrer, E., and Bash, C., 2007, "Thermal Zone Mapping: Data Center Design and Assessment Automated Visualization Tool for Thermal Metric Analysis," *IMAPS Adv. Technol. Workshop Therm. Manage.*

[5] Shah, A. J., Lugo, W., and Patel, C. D., 2009, "Thermodynamics-Based Object-Oriented Environmental Analysis of an Industrial Air Conditioner," *Proc. ASME Summer Heat Transf. Conf.*.

[6] Baumann, H., and Tillman, A.-M., 2004, *The Hitch Hiker’s Guide to LCA*, Studentlitteratur AB, Lund, Sweden.

[7] Hendrickson, C. T., Lave, L. B., and Matthews, H. S., 2006, *Environmental Life Cycle Assessment of Goods and Services: An Input-Output Approach*, Resources for the Future, Washington, DC.

[8] Mohamed, A. M., Lipsky, L., and Ammar, R., 2004, "Transient Performance Model for Parallel and Distributed Systems," *Proc. 10th Int. Conf. Parallel Distrib. Syst. (ICPADS’04)*, pp. 513-520.

[9] Christian, T. W., Chen, Y., Shih, R., Sharma, R.K., Hoover, C. E., Marwah, M., Shah, A. J., and Gmach, D., 2009, "Automated Synthesis of Sustainable Data Centers," *IEEE Int. Symp. on Sust. Syst. Technol. (ISSST’09).*

[10] Hannemann, C., Carey, V., Shah, A. J., and Patel, C. D., 2008, “Lifetime Exergy Consumption as a Sustainability Metric for Enterprise Servers,” *Proc. ASME 2nd Int. Conf. Energy Sustain. (ES2008).*

[11] Tolia, N., Wang, Z., Marwah, M., Bash, C. E., Ranganathan, P., and Zhu, X., 2009, "Unified Thermal and Power Management in Server Enclosures," *Proc. ASME InterPACK.*

[12] Beitelmal, M. H., Wang, Z., Felix, C., Bash, C. E., Hoover, C. E., and McReynolds, A., 2009, “Local Cooling Control of Data Centers with Adaptive Vent Tiles,” *Proc. ASME InterPACK.*

[13] Bash, C. E., Patel, C. D., and Sharma, R. K., 2006, "Dynamic Thermal Management of Air Cooled Data Centers," *10th Intersociety Conf. Therm. and Thermomechanical Phenom. Electr. Syst. (ITherm'06)*, pp. 445-452.

[14] Bash, C. E., and Forman, G., 2007, "Data Center Workload Placement for Energy Efficiency," *Proc. ASME InterPACK.*

[15] Bash, C. E., Hyser, C. D., and Hoover, C. E., 2009, "Thermal Policies and Active Workload Migration within Data Centers," *Proc. ASME InterPACK.*

[16] Jollife, I., 2002, *Principal Component Analysis, Second Edition*, Springer, New York, NY.

[17] Hastie, T., Tibshirani, R., and Friedman, J., 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, Springer, New York, NY.

[18] Lin, J., Keogh, E., Lonardi, S., and Patel, P., 2002, "Finding Motifs in Time Series," 2nd Workshop on Temporal Data Mining at 8th ACM SIGKDD Int. Conf. Knowl. Discovery and Data Mining.

[19] Patnaik, D., Marwah, M., Sharma, R. K., and Ramakrishnan, N., 2009, "Sustainable Operation and Management of Data Center Chillers using Temporal Data Mining," *14th ACM SIGKDD Int. Conf. Knowl. Discovery and Data Mining*.

[20] Hao, M., Dayal, U., Keim, D. A., Morent, D., and Schneidewind, J., 2007, "Intelligent Visual Analytics Queries," *IEEE Symp. Vis. Analytics Sci. Technol*, pp. 91-98.

[21] Hao, M., Dayal, U., Keim, D. A., and Schreck, T., 2007, "Multi-Resolution Techniques for Visual Exploration of Large Time-Series Data," *Proc. IEEE VGTC Symp. Vis. (EuroVis'07).*

[22] Bash, C. E., Patel, C. D., Shah, A. J., and Sharma, R. K., 2008, "The Sustainable Information Technology Ecosystem," *11th Intersociety Conf. Therm. and Thermomechanical Phenom. Electr. Syst. (ITherm'08)*, pp. 1126-1131.

[23] Escobar-Vargas, S., Fabris, D., Gonzalez, J. E., Sharma, R. K., and Bash, C. E., 2007, "High Power Density Dissipations by Spray Cooling," *Proc. ASME/JSME Therm. Eng. and Summer Heat Transf. Conf.*

[24] Zhu, X., Young, D., Watson, B. J., Wang, Z., Rolia, J., Singhal, S., McKee, B., Hyser, C. D., Gmach, D., Gardner, R., Christian, T. W., and Cherkasova, L., 2008, "1000 Islands: Integrated Capacity and Workload Management for the Next Generation Data Center," *5th IEEE Int. Conf. Autonomic Comput. (ICAC'08).*