Environmentally Opportunistic Computing
Transforming the Data Center for Economic and Environmental Sustainability

Paul Brenner and Ryan Jansen
Center for Research Computing
University of Notre Dame
Notre Dame, IN, USA
Email: pbrenne1@nd.edu

David Go
Dept. of Aerospace and Mechanical Eng.
University of Notre Dame
Notre Dame, IN, USA
Email: dgo@nd.edu

Doug Thain
Dept. of Computer Science and Eng.
University of Notre Dame
Notre Dame, IN, USA
Email: dthain@nd.edu

***Abstract*—The United States Environmental Protection Agency forecasts the 2011 national IT electric energy expenditure will grow toward $7.4 billion [1]. In parallel to economic IT energy concerns, the general public and environmental advocacy groups are demanding proactive steps toward sustainable green processes. Our contribution to the solution of this problem is Environmentally Opportunistic Computing (EOC). Our Green Cloud EOC prototype serves as an operational demonstration that IT resources can be integrated with the dominate energy footprint of existing facilities and dynamically controlled to balance process throughput, thermal transfer, and available cooling via process management and migration. The Green Cloud is a sustainable computing technology that complements existing efficiency improvements at the application, operating system and hardware levels. Exhaust heat energy is transferred directly to an adjacent greenhouse facility and cooling is provided by free cooling methods. We will describe the architecture and operation of this successful prototype that has led to its growing use in our production environments.**

***Keywords*-sustainability, distributed systems, controls, grid, cloud, utility computing, energy, scheduling**

I. Introduction

The United States Environmental Protection Agency estimates the US spent $4.5 billion on electrical power to operate and cool information and communication technology (ICT) and high performance computing (HPC) servers in 2006 [1]. The same report forecasts the 2011 national IT electric energy expenditure will grow to $7.4 billion. The energy demand is already three percent of total US electricity consumption and growing rapidly. In parallel to economic ICT energy concerns, environmental stewardship and sustainability have come center stage. The general public and environmental advocacy groups are demanding proactive steps toward conservation and green processes.

In response, the technology industry has launched collaboratives such as The Green Grid http://www.thegreengrid.org and multiple individual corporate initiatives. Corporations, universities, and government labs are trying to reduce their environmental footprint and manage their energy consumption in a world economy where the environmentally sound production of energy is under intense scrutiny. Multiple organizations have made great strides in the optimization of traditional centralized data centers; such as the High-Performance Buildings for High-Tech Industries Team at LBNL [2], the ASHRAE Technical Committee 9.9 for Mission Critical Facilities, Technology Spaces, and Electronic Equipment [3], and the Uptime Institute [4]. One further example of this environmental emphasis in HPC is demonstrated by the emergence of a Green500 [5] to complement the Top500 [6] list of Supercomputers.

Our contribution to the solution of this problem is **Environmentally Opportunistic Computing** (EOC), which engages sustainable computing at the macro scale while complementing current scheduler, operating system and hardware level efficiency [7]–[12] and consolidation efforts in the virtualized data center. EOC integrates computing hardware with existing facilities to create heat where it is already needed, to exploit cooling where it is already available, to utilize energy when and where it is least expensive, and to minimize the overall energy consumption of an organization. EOC capitalizes on the dynamic mobility of virtualized services to exploit energy volatility for cost savings and lower environmental impact.

At Notre Dame, our latest operational EOC prototype is the Green Cloud, a bridge between traditional data center and containerized data center, located at the City of South Bend Botanical Conservatory and Greenhouse. During cold weather, the heat generated by the data center is vented into the greenhouse, saving both cooling costs for the data center and heating costs for the greenhouse. During hot weather, heat production and delivery is balanced by services migration. This prototype is used for high-throughput batch computing, allowing us to gain operational experience with the experimental system without placing critical services at risk.

In this paper we will first review our evolution of EOC methodologies with an emphasis on our most recent prototype, the Green Cloud framework implementation. In the Operation and Measures section we provide more detailed operational considerations of the system in production use and lay the ground work for the next prototype modifications discussed in Future and Related Work. Finally we conclude

978-1-4244-7614-5/10/$26.00 ©2010 IEEE

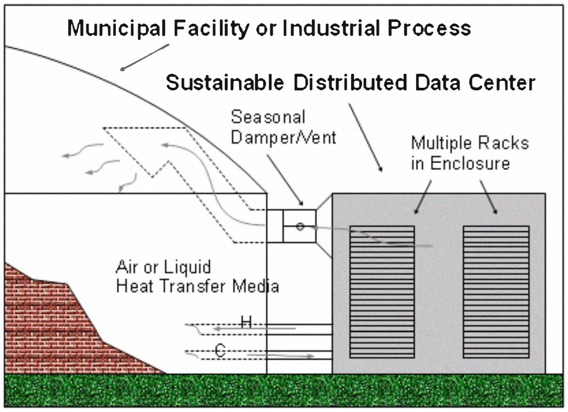

Figure 1. Sustainable Distributed Data Center Concept.

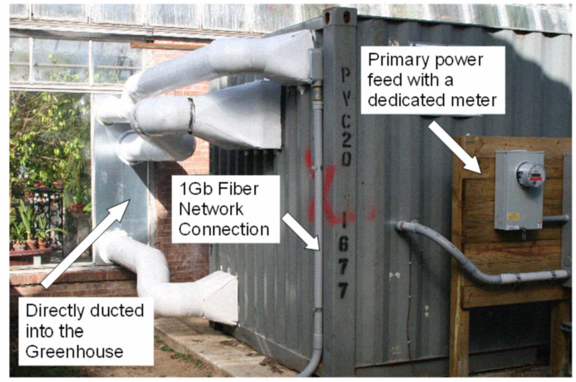

Figure 2. Sustainable Distributed Data Center Ducted Into the Greenhouse.

with a summary of the practical utility of our prototype and an acknowledgment of the many individuals and organizations who have come together to make EOC a success.

II. Methods

Environmentally Opportunistic Computing is a broad macro scale sustainable computing concept that capitalizes on the physical and temporal mobility of modern computer processes. EOC's economic and environmental benefits are realized through the integration of computing hardware into both individual and multiple facility infrastructures. The viability of this integration has grown steadily with the evolution of grid, utility, and cloud infrastructures bolstered by more robust data networks, enterprise virtualization solutions, and intelligent hardware. EOC architectures improve sustainability by providing the control mechanisms to migrate processes between geographically separated physical systems based on temporal considerations such as price of power fluctuations and available free cooling. Most importantly EOC benefits are direct additions to the wealth of data center, operating system, and hardware level efficiency improvements currently undertaken by research teams throughout the industry.

The authors have evolved toward the EOC concept based on numerous related smaller scale research prototypes. Our first model framework called Grid Heating (GC) [13] specifically focused on utilization and control of server exhaust heat within individual controlled human work centers. We were able to successfully demonstrate dynamic energy based ICT services migration in response to environmental stimuli [14]. The Grid Heating work proposed larger scale container based possibilities as show in Figure 1. We then investigated CPU core level energy utilization characteristics for benchmark grid loads to shape our macro scale migration policies [15], finding dominant efficiency benefits of hibernation states over voltage scaling and disabling individual cores.

Most recently the authors have constructed a heterogeneous, geographically distributed, multi-institutional grid infrastructure for executing service migration in production ICT environments. We have named this EOC prototype the Green Cloud (GC) http://greencloud.crc.nd.edu. The GC merges the Grid Heating concept with a macro scale Facility Integrated ICT (FIICT) [16] energy model in a distributed system architecture. Currently the core infrastructure is based on the University of Notre Dame Condor [17] pool serving a wide variety of high throughput research computing needs. Jobs can migrate from personal workstations to the traditional data center to dedicated Condor servers at the City of South Bend's Botanical Conservatory and Greenhouse; where the authors have designed and deployed a Sustainable Distributed Data Center (SDDC) container based prototype. This SDDC serves as a dedicated resource hub for the Green Cloud prototype and is shown in Figure 2.

III. Operation and Measures

Due to the breadth of the EOC concept and the numerous acronyms we have introduced thus far (EOC, GH, GC, ICT, FIICT, SDDC, etc..) in this section we will refrain from further topical discussions and focus only on the practical implementation details of our operational Green Cloud infrastructure.

The GC hardware resources are fully integrated into the general access Notre Dame Condor pool appearing no different than other resources to end user jobs. The differentiating factor lies in the environmentally aware controls system set in place for job management and scheduling. The controls system currently has two primary components: Condor and xCAT [18]. The Condor component handles the entire scientific workload management and each server's response to the workload based on environment (such as system temperature). The xCAT component handles real time system vitals monitoring through interface with the hardware's service processor BMC and IPMI. The interface between the two is handled with short Python scripts. It is rationale to consider a direct interface between Condor and

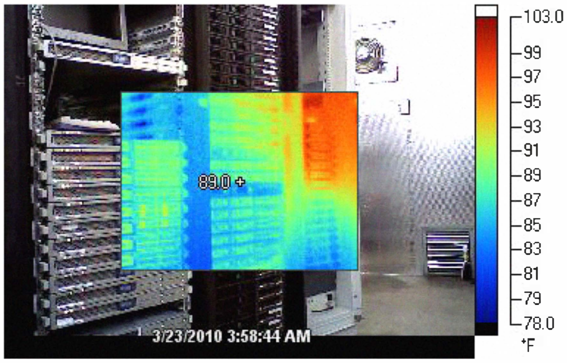

Figure 3. Sample Server Inlet Temperatures via Infrared

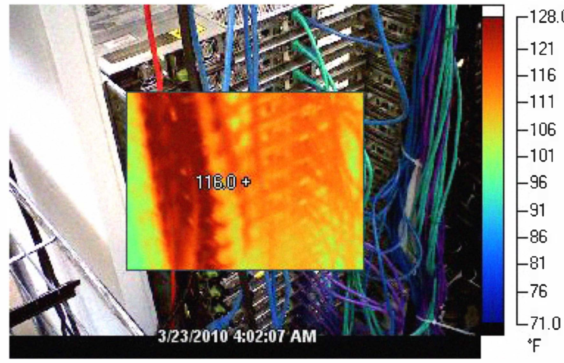

Figure 4. Sample Server Outlet Temperatures via Infrared

the hardware level diagnostics, however the robust existing capabilities of xCAT in this regard for a variety of hardware models has made the two component option the most readily viable (xCAT is also our standard cluster management tool for server installation and administration).

The SDDC was designed with a chief target to minimize cost while providing a suitably secure facility for use outdoors in a publicly accessible venue. The container based solution is standard 20' long by 8' wide retrofitted with the following additions: a 40KW capacity power panel with 208V power supplies to each rack, lighting, internally insulated walls, man and cargo door access, ventilation louvers, small fans, and duct work connecting the SDDC to the greenhouse; for a total cost under \$20,000. Exterior power infrastructure to include the transformer, underground conduit, panel, and meter were coordinated by AEP and the City of South Bend. The slab foundation was provided by the City of South Bend. The high bandwidth network connectivity critical to viable scaling is possible via 1Gb fiber network connectivity to the Notre Dame campus on the St. Joe County MetroNet backbone.

From the outset the system was designed for operation utilizing only direct free cooling via outdoor or greenhouse air. It was also determined that system performance and mean time to failure will be evaluated when pushing systems beyond ASHRAE [19] and industry specified [20] limits for temperature and particulate. By allowing a larger window for thermal fluctuation we are working to provide variable exhaust heat densities for delivery to the conservatory facility. System level temperatures are provided by the hardware IPMI and validated by occasional infrared camera measurements as shown in Figures 3 and 4 which indicate average inlet and outlet temperatures near 90 and 115F respectively. The system has multiple inlet, outlet, and fan options which will allow for hands on education of mechanical engineering undergraduates studying heat transfer mechanisms.

The eBay Corporation graciously provided over 100 servers for use in this rigorous prototype environment. Per our agreement with eBay we are not able to provide the reader with specific details on the server models utilized. We can however summarize that they are commercially available multicore systems. The servers began accepting jobs from the Notre Dame Condor pool in December 2009. Since that point seasonal temperature variations and continued tuning has allowed numbers of concurrently running jobs varying from 0 to 250. Machine utilization is posted publicly on the GC website. As the afternoon temperatures warm up at the greenhouse machines can be observed idling or migrating jobs; the jobs return in the evening when the environment is more suitable. In the same temporal cycle the greenhouse will become a priority service location in the evening as a power tariff under negotiation with the local utility will provide much lower costs at night.

One of the largest challenges in constructing the initial prototype has been heat management. To combat this problem, the authors developed a suite of temperature management scripts, designed to both efficiently run jobs in the pool as well as regulate the overall temperature of the SDDC. The primary script, the manager, monitors temperatures and updates configurations of machines in the pool at set intervals throughout the day. When initializing the manager, the user specifies a target temperature, which varies depending on the temperature thresholds deemed safe by each machine's hardware. For current operations, the target temperature is initialized to 95F, which is the upper limit that our hardware allows without entering an unsafe state. This target temperature, however, is not constant, and can change significantly depending on the status of the machines in the pool.

A dynamic target temperature allows the manager to, at any given iteration, control each machine's heat with regards to the status of the entire pool. For example, if a single machine is running in dangerous temperature ranges, idling or hibernating that single machine may not necessarily lower its temperature to a safe state. Instead, by incrementally lowering the overall target temperature of the SDDC over time, multiple machines in the pool will begin to idle and hibernate, and the machine in an unsafe state will eventually drop to an acceptable temperature. Conversely, by increasing the target temperature when all machines are running in a safe state, the manager ensures that the pool is always running with the greatest possible thermal output.

Example pseudo code logic for the control system is shown in Algorithm 1. Once the target temperature is calculated, the manager writes a configuration file for each machine, which includes that machine's current temperature, as well as several possible command expressions. These command expressions cause each machine to either continue accepting jobs or take appropriate action (suspend jobs, idle, hibernate, etc.) depending on the difference between the recorded machine temperature and the target. Once all of the configurations are written properly, the manager loads these new configurations on to the machines through Condor, and, upon updating, each machine takes its specified course of action.

Algorithm 1 Pseudo Logic for Macro Thermal Control

```
repeat

  Read all temperatures

  if machine is in the dangerous threshold range then
    Lower target temperature by 2 degrees
  else if machine is in the critical threshold range then
    Lower target temperature by 1 degree
  else
    Raise target temperature by 1 degree
  end if

  for all machines do
    Rewrite machine configuration file with updated
    inlet temperature and target temperature:
    Start if temp <target - start_offset
    Suspend if temp >target
    Continue if temp <target
    Preempt if temp >target + 1 - evict_offset
    Hibernate if temp >danger temp {Hibernate not
    currently automated}
    Unhibernate if temp <target + 1
  end for

  Send reconfigure command to all machines

until Incrementally update controls for the period of
operation
```

Recent tests to maximize heat density and server up time have provided positive results. Without the heat management software in place, during warm spring afternoons (75F outdoor temp), the GC was able to sustain a maximum of 24 active machines, running with an average inlet temperature of 85F. During this time, just over 18% of machines ran at maximum thermal capacity. With the manager running the cloud was able to sustain up to 44 machines for several hours, and then up to 28 machines after some of the hotter machines began to power off. During this time, the average inlet temperature for each machine was measured to be 108.11 degrees Farenheit, with 67% of machines running at maximum thermal capacity.

IV. Future and Related Work

Since December 2009 when the GC servers began to accept jobs the system has been both dynamically and statically adjusted to study and improve thermal delivery while maximizing throughput. The path forward promises no greater durations of steady state operations as modifications to the prototype present a wealth of interdisciplinary analysis opportunities. We do however look to formalize the system modification schedule to gather more robust data sets between changes. We are also currently evaluating alternative cloud administration software frameworks to compare and contrast with Condor. New cloud frameworks providing fully virtualized services may grow the number of production and development applications running in the GC via additional capabilities and lower technical barriers to entry.

It is important to note evolving commercial applications in grid, utility, and cloud computing [21]–[23] that will directly benefit from this technology. As the computational infrastructure configuration and locality is obscured from the end user, the flexibility to distribute and configure grows, allowing for additional economic and environmental optimizations. Along these lines, the growing acceptance of virtualization [24]–[26] in commercial applications will also allow greater flexibility in design and deployment.

While these grid frameworks are evolving, a large body of work has studied the problem of managing energy, heat, and load in large centralized data centers. Schmidt et al [27], [28] provide a good overview of the mechanical issues of cooling units, heat sinks, fans, and so forth. Server management techniques can also be applied to reduce energy costs. For example, inactive servers can be shut down, or loads migrated as more energy efficient hardware become idle/available. Chase [29] and Bradley [30] describe techniques for balancing performance, cost, and energy in this situation. To avoid hot spots, it is necessary to map the relation between components and heat [31], and then shape loads so as to evenly distribute the heat [32]. Further, large institutions such as the University of Illinois and NCSA are taking a holistic look at their entire campus utilities infrastructure to efficiently operate their data center. Despite the new efficiency benefits, Patel et al. [33] report that a typical data center still consumes about as much energy for cooling as it does for productive work. The advances toward more efficient traditional infrastructures and EOC frameworks will serve in tandem to provide greater computational capability while reducing economic and environmental costs.

V. Conclusion

Environmentally opportunistic computing provides economic and environmental benefits to improve the sustainability of information and communication technology infrastructure. With the growing enterprise utilization of cloud computing and virtualized services EOC becomes more viable in across the range of ICT services. Our Green Cloud prototype serves as an operational demonstration that ICT resources can be integrated with the dominate energy footprint of existing facilities and dynamically controlled to balance process throughput, thermal transfer, and available cooling via process management and migration.

With a modest budget and a supportive cast of community partners the authors have constructed a Sustainable Distributed Data Center prototype that anchors a Green Cloud infrastructure. Although the prototype is embedded within the community at a geographic separation from end user scientists, the grid infrastructure allows full integration into the primary Notre Dame Condor pool. While processing production scientific loads, the prototype has also become a tool for analysis of elevated operational temperatures when restricting systems to available free cooling.

The Green Cloud is a sustainable computing technology that complements existing efficiency improvements at the application, operating system and hardware levels. The exhaust heat energy is transferred directly to the adjacent greenhouse facility. Free cooling via outside or greenhouse air are the only cooling options permitted to minimize the cooling energy overhead. The measured average and maximum throughput continues to improve as we tune our Condor based controls system and modify the physical layout of the prototype. The initial success will stimulate continued project growth to the economic and environmental benefit of our organization (and thus user base).

Acknowledgment

The authors would like to recognize contributions to this work from multiple organizations and individuals. The City of South Bend and the South Bend Botanical Society have been close partners contributing funds and facility expertise. The eBay corporation generously donated over 100 servers for use in the Green Cloud prototype. Jarek Nabrzyski and the Center for Research Computing operations group contributed essential financial resources and system administration expertise to keep the production servers operational in very demanding environmental conditions. We would also like to thank professors Aimee Buccellato and Michael Lemmon for discussions of architectural and control system considerations respectively.

References

[1] U.S. Environmental Protection Agency, “Report to congress on server and data center energy efficiency public law 109-431,” U.S. E.P.A, Tech. Rep., Aug 2007.

[2] E. Mills, G. Shamshoian, M. Blazek, P. Naughton, R. Seese, W. Tschudi, and D. Sartor, “The business case for energy management in high-tech industries,” *Journal of Energy Efficiency*, vol. 1, pp. 1–16, 2007.

[3] T. Committee 9.9, *High Density Data Centers - Case Studies and Best Practices*. Atlanta, GA: American Society for Heating Refrigeration and Airconditioning Engineers, 2008.

[4] K. G. Brill, “Special report: Energy efficiency strategies survey results,” The Uptime Institute, Tech. Rep., April 2008.

[5] S. Sharma, C.-H. Hsu, and W. chun Feng, “Making a case for a green500 list.” in *IEEE International Parallel and Distributed Processing Symposium (IPDPS 2006)/ Workshop on High Performance - Power Aware Computing*, 2006.

[6] J. J. Dongarra, H. W. Meuer, and E. Strohmaier, “29th top500 Supercomputer Sites,” Top500.org, Tech. Rep., June 2007.

[7] I. Ahmad, S. Ranka, and S. U. Khan, “Using game theory for scheduling tasks on multi-core processors for simultaneous optimization of performance and energy,” in *The Next Generation Software (NGS) Workshop 2008*, April 2008.

[8] J. Kang and S. Ranka, “Assignment algorithm for energy minimization on parallel machines,” in *2009 International Conference on Parallel Processing Workshops*, 2009.

[9] R. Jejurikar and R. Gupta, “Dynamic voltage scaling for systemwide energy minimization in real-time embedded systems,” in *Proceedings of the 2004 International Symposium on Low Power Electronics and Design (ISLPED’04)*, 2004.

[10] V. W. Freeh, N. Kappiah, D. K. Lowenthal, , and T. K. Bletsch, “Just-in-time dynamic voltage scaling: Exploiting inter-node slack to save energy in mpi programs,” in *Journal of Parallel and Distributed Computing, Volume 68, Issue 9*, May 2008.

[11] Y. Chen, L. Keys, and R. H. Katz, “Towards energy efficient mapreduce,” in *Technical Report No. UCB/EECS-2009-109*, August 2009.

[12] C. Gunaratne, K. Christensen, and B. Nordman, “Managing energy consumption costs in desktop pcs and lan switches with proxying, split tcp connections, and scaling of link speed,” *Int. J. Netw. Manag.*, vol. 15, no. 5, pp. 297–310, 2005.

[13] P. Brenner, D. Thain, and D. Latimer, “Grid Heating Clusters: Transforming Cooling Constraints into Thermal Benefits,” in *Uptime Institute - IT Lean, Clean, and Green Symposium*, 2009.

[14] ——, “Grid heating: Transforming cooling constraints into thermal benefits,” University of Notre Dame, Computer Science and Engineering Department, Tech. Rep. 2008-30, April 2008.

[15] M. Lammie, P. Brenner, and D. Thain, “Scheduling Grid Workloads on Multicore Clusters to Minimize Energy and Maximize Performance,” in *10th IEEE/ACM International Conference on Grid Computing*, 2009.

[16] P. Brenner and D. Latimer, "Facility integrated information and communication technology," Invited Presentation:Data Center Dynamics Conference, Chicago IL, October 2009.

[17] M. Litzkow, M. Livny, and M. Mutka, "Condor - a hunter of idle workstations," in *Eighth International Conference of Distributed Computing Systems*, June 1988.

[18] E. Ford, "xcat extreme cloud administration toolkit," IBM internal software development. Open Sourced 2007, 1999.

[19] A. T. C. 9.9, *Best Practices for Datacom Facility Energy Efficiency*. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2008.

[20] D. Alger, *Grow a Greener Data Center*. Cisco Systems, 2010.

[21] T. Eilam, K. Appleby, J. Breh, G. Breiter, H. Daur, S. A. Fakhouri, G. D. H. Hunt, T. Lu, S. D. Miller, L. B. Mummert, J. A. Pershing, and H. Wagner, "Using a utility computing framework to develop utility systems," *IBM Systems Journal*, vol. 43, pp. 97–120, 2004.

[22] B. Raghavan, K. Vishwanath, S. Ramabhadran, K. Yocum, and A. C. Snoeren, "Cloud control with distributed rate limiting," in *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*. New York, NY, USA: ACM, 2007, pp. 337–348.

[23] W. Mitchell, "Data center in a BOX," *Scientific American*, vol. 297, pp. 90–93, Aug 2007.

[24] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," in *SOSP '03: Proceedings of the nineteenth ACM symposium on Operating systems principles*. New York, NY, USA: ACM, 2003, pp. 164–177.

[25] R. Figueiredo, P. Dinda, and J. Fortes, "Resource virtualization renaissance," *Computer*, vol. 38, pp. 28–31, 2005.

[26] R. Uhlig, G. Neiger, D. Rodgers, A. Santoni, F. Martins, A. Anderson, S. Bennett, A. Kagi, F. Leung, and L. Smith, "Intel virtualization technology," *Computer*, vol. 38, pp. 48–56, 2005.

[27] R. Schmidt, E. Cruz, and M. Iyengar, "Challenges of data center thermal management," *IBM Journal of Research and Development*, vol. 49, July 2005.

[28] R. Schmidt and M. Iyengar, "Best practices for data center thermal and energy managment - review of literature," *ASHRAE Transactions*, vol. 113, Jan 2007.

[29] J. Chase, D. Anderson, P. Thakar, A. Vahdat, and R. Doyle, "Managing energy and server resources in hosting centers," in *Symposium on Operating Systems Principles*, 2001.

[30] D. Bradley, R. Harper, and S. Hunter, "Workload based power management for parallel computer systems," *IBM Journal of Research and Development*, vol. 47, pp. 703–718, 2003.

[31] J. Moore, J. Chase, and P. Ranganathan, "Weatherman: Automated, online, and predictive thermal mapping and management for data centers," in *IEEE International Conference on Autonomic Computing*, June 2006.

[32] ——, "Making scheduling cool: Temperature-aware workload placement in data centers," in *USENIX Annual Technical Conference*, April 2005.

[33] C. D. Patel, C. E. Bash, R. Sharma, and M. Beitelmal, "Smart cooling of data centers," in *Proceedings of IPACK*, July 2003.