Lawrence Berkeley National Laboratory

High-performance data centers: A research roadmap

Authors
Tschudi, William
Xu, Tengfang
Sartor, Dale
et al.

Publication Date
2004-03-30

LBNL-53483

HIGH-PERFORMANCE DATA CENTERS

A RESEARCH ROADMAP

Developed by:
Lawrence Berkeley National Laboratory with input from industry partners representing data center facility design and operation firms, industry associations, research organizations, energy consultants, and suppliers to data centers

William Tschudi, Tengfang Xu, Dale Sartor
And
Jay Stein – E Source

Sponsored by:
The California Energy Commission through the Public Interest Energy Research (PIER) Program

DISCLAIMER

ROADMAP FOR PUBLIC INTEREST RESEARCH

For

HIGH-PERFORMANCE DATA CENTERS

William Tschudi, Tengfang Xu, Dale Sartor, and

Jay Stein – E Source

Building Technologies Department
Environmental Energy Technologies Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
1 Cyclotron Road
Berkeley, California 94720-8134 USA

Table of Contents

Data Center Energy Research and Deployment Roadmap

Executive Summary

When California's electric utilities began receiving requests for huge electrical demands for data center facilities, it became evident that little information existed to validate actual data center electrical performance, or to see how the energy performance could be improved. As a result, California utilities and the California Energy Commission became interested in learning more about the data center market. Utility case studies and preliminary investigations confirmed that research with the objective of reducing the large, continuous electrical loads in data centers was clearly merited, however the role of public interest research for these types of facilities was not clear.

To tackle this problem, the Public Interest Energy Research (PIER) Industrial Program set out to define and prioritize energy efficiency research areas by engaging Data Center Industry professionals. In preparation of this roadmap, researchers from Lawrence Berkeley National Lab (LBNL) facilitated workshops, participated in industry forums, and researched energy issues related to data centers. As the topics in the roadmap were developed, opportunities for California public interest research and market transformation activities were the primary focus. Other research and standardization activities by others were noted, and it will be important to keep abreast of their progress as the California research agenda is advanced. In addition, data center professionals identified other parts of the energy efficiency puzzle that must be solved by the industry itself due to the highly specialized nature of much of the equipment in data centers. Even though the research in these areas will proceed through industry efforts, public interest encouragement may accelerate the development and adoption of new innovations.

There are many types of data centers in California ranging from corporate data centers in a wide range of industries, banks, telecommunication facilities, and Internet hosting facilities. Data centers are also found in other institutions such as research organizations, universities, National Laboratories, and government facilities. The crosscutting nature of the market involves many California industries that directly or indirectly rely on data centers, as well as the suppliers of goods and services for data centers.

Figure 1 The National Energy Research Scientific Computing Center (NERSC) operated by LBNL

The data center market when defined broadly represents a significant and important component of the California economy. Regardless of their mission and make-up, most data centers are much more energy

intensive than other buildings due to the high power requirements of the computing equipment and the infrastructure needed to support the computing equipment. Based on their energy density, large data centers more closely resemble industrial facilities than commercial buildings.

The roadmap development identified many areas where significant efficiency gains could be achieved through adoption of current best practices, better application of existing technology, and research into new technological solutions. The roadmap organizes these areas as follows:

1. Activities aimed at understanding the Data Center Market – The size and growth rate of the market as well as local concentrations of data centers is of interest to planners and implementers of electrical power generation and distribution.
2. The benefits of obtaining energy benchmarks – By monitoring and comparing the energy consumption of a variety of data centers, operators and designers will be able to learn what is possible to achieve.
3. Identification and promotion of best practices – Adopting current best practices in existing or new data centers will provide significant improvement in the short term.
4. Improving data center facility systems' efficiency - Facility systems containing both conventional equipment such as chillers, and specialty equipment such as Uninterruptible Power Supplies are far from optimal.
5. Improving the interface between building systems and IT Equipment-The systems that house and support electronic equipment in data centers are typically not designed to optimize the efficiency of the building infrastructure systems they interface with.
6. Improving the efficiency of IT Equipment - Energy use in data centers is dominated by the servers, hard drives, routers, and switches that are used to process, store, and transmit data. Efficiency improvements in IT equipment are compounded by secondary effects in HVAC and power supply facility systems

During the course of the roadmap development, data center experts suggested several promising research topics that were outside the scope of this roadmap, which is focused exclusively on improving the energy efficiency of data centers. Most of these topics, such as thermal storage for peak demand reduction and distributed generation, have the potential to provide other societal benefits, and as such, are under investigation by other parties outside of the data center industry so they were not included in the roadmap. Furthermore, it was recognized that research efforts were underway by various organizations and industry associations, such as iTherm, CEETHERM, (a collaboration between the University of Maryland and Georgia Tech) and major electronics companies. In some cases the roadmap cites these efforts, observing that there are other research efforts outside those included in this roadmap that will make important contributions toward solving the overall problem. Collaboration and awareness of developments by others will be important to make sure that the research undertaken in California is headed in the right direction.

Lastly, while developing the roadmap, we uncovered an important difference of opinion within the data center industry. The electronic equipment used by this industry is continually evolving. Some industry observers have noted that the energy intensity exhibited by this equipment (measured in Watts per square foot) is increasing. Others, noting the recent availability of more efficient microprocessors, have proposed that at some point in the future, the trend towards increasing energy intensity will either slow down, level off, or decline. In other cases, more powerful computing equipment, although itself more energy intensive, has replaced many other pieces of equipment resulting in a net decrease in energy use. We did not take a position in the debate over whether intensities will rise or fall. However, given such uncertainty regarding future electrical and cooling demands, our efforts were directed at identifying strategies that would allow for efficient data center operation regardless of how technology evolution and business conditions play out.

Introduction and Background

Data Center Definition

The market addressed by this roadmap employs a broad definition of the term ”data center”. Generally, we use the term data center to be a facility that contains concentrated equipment to perform one or more of the following functions: Store, manage, process, and exchange digital data and information. Such digital data and information is typically applied in one of two ways:

- ♦ Support the informational needs of large institutions, such as corporations and educational institutions.
- ♦ Provide application services or management for various types of data processing, such as web hosting, Internet, intranet, telecommunication, and information technology.

We do not consider spaces that primarily house office computers, including individual workstations, servers associated with workstations, or small server rooms, to be data centers. Generally, the data centers we include are designed to accommodate the unique needs of energy intensive computing equipment along with specially designed infrastructure to accommodate high electrical power consumption, redundant supporting equipment, and the heat dissipated in the process.

To accommodate these needs, data centers typically exhibit these characteristics:

- ♦ Physically house various types of IT equipment, such as computers, servers (e.g., web servers, application servers, database servers), main frame computers, switches, routers, data storage devices, load balancers, wire cages or closets, vaults, racks, and related equipment.
- ♦ Exhibit critical requirements for security and reliability.
- ♦ Most, but not all, data centers utilize raised floors or other specialized computer room air conditioning systems.
- ♦ Provide for redundant and uninterruptible power.

Figure 2 Supercomputers in LBNL’s NERSC Center

Other terms used to refer to facilities that meet the definition of data center as used in this roadmap include: computer center, data storage and hosting facility, server farm, data farm, data warehouse, co-location facility, co-located server hosting facility (CoLo), corporate data center, managed data centers, internet hotel, internet service provider (ISP), application service provider (ASP), full service provider (FSP), wireless application service provider (WASP), telecommunication hotel (or telco hotel), carrier hotel, internet hotel, telecommunications carriers, or other data networks.

Case Studies and Prior Investigations

A number of case studies have been performed to characterize the energy end use in various types of data centers[1]. A large variation in energy intensity and energy efficiency of key systems was observed in the various facilities that were studied. Design features of the better performing systems were noted yet every facility had the potential for energy efficiency improvement. Recommendations for efficiency improvements were provided as part of the case studies. The recommendations and findings often identified common issues. Many of the issues noted in the case studies suggested areas where further research could lead to much better performance.

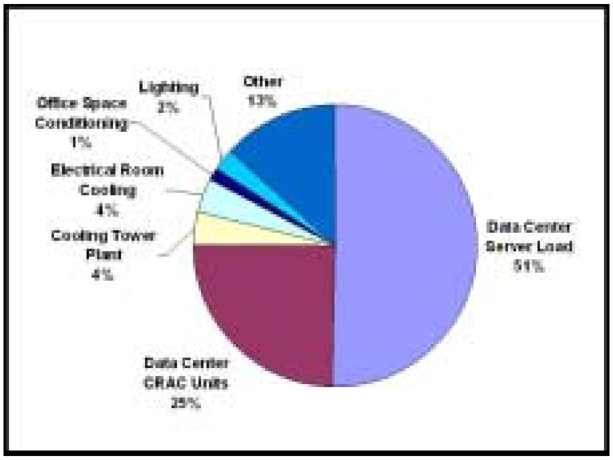

Figure 3 Typical Data Center Energy End Use

Workshops to Develop the Roadmap

There are many stakeholders interested in data center energy consumption. These stakeholders include data center industry associations, research organizations, public interest groups (i.e. utilities, federal and state agencies), as well as the individual firms that operate data centers, and manufacture IT or data center specialized equipment.

Input was obtained from data center professionals and experts in order to develop this research roadmap for high-performance, energy-efficient data centers, and to validate research issues and possible actions identified through case studies. Their input was obtained throughout the project by conducting workshops, attending data center conferences and meetings, and interviewing data center professionals. Leading data center designers, specialty equipment suppliers, computer manufacturers, energy consultants, and industry associations were contacted to solicit input concerning the state of the data center industry, and for help in defining where public interest research could make a difference. Industry associations such as the 7 X 24 Exchange Organization (www.7x24exchange.org) and the Uptime Institute (www.upsite.com) participated and provided valuable input including research topics and possible actions to address them. These professionals also helped to prioritize the issues and possible actions.

Generally, two types of recommendations emerged:

- **Recommendations to Improve performance through better use of existing technologies**

  Through case studies, a wide variation in energy performance using today's technologies was observed. As a result, the identification of current best practices and efforts to

[1] Case studies and summary information are available on LBNL's website: http://datacenters.lbl.gov .

influence the market to adopt best practices should be a high priority. Some roadmap recommendations involve better use of existing strategies and technologies.

- **Research and Development for new solutions**

To advance beyond the current best practices, research and development will be needed in a number of areas. Industry participants identified many areas and issues where new solutions are needed, both for improving efficiency and for handling expected increases in heat intensity.

RMI Data Center Charrette

The California Energy Commission, along with other public interest organizations, sponsored a data center "charrette[2]". This event, hosted and coordinated by the Rocky Mountain Institute (RMI- www.rmi.org), assembled over 75 leading stakeholders in the data center market for three days of brainstorming and sharing their expertise and information. The charrette participants identified many of the same issues previously identified during the case study work and other workshops, but also introduced additional issues. Many ideas for needed research were presented from the chip level up through the facility power source. Some research topics are more appropriate for various areas of the data center industry to take the lead (e.g. Efficiency of "blade servers"), and others where public interest research is needed and appropriate (e.g. benchmarking energy use). The public interest research topics along with some topics for possible joint participation are included in this roadmap. RMI will report charrette findings separately.

Figure 4 RMI Charrette Announcement

[2] The term charrette describes a process widely used by architects to critique a design and brainstorm new solutions. Normally the charrette occurs early enough in the design process to allow improvements to be incorporated.

Development of the Roadmap

In addition to the workshops and design charrette described above, an extensive literature review further helped to identify trends, current practice, and also suggested areas where improvement is needed. References that were reviewed were annotated and included as appendix A. Through these activities, a list of issues with a bearing on energy efficiency was compiled and forms the basis of this roadmap. For each of the issues one or more suggested research and/or market transformation actions were developed. Some of the actions are intended to provide near term improvement by determining current best practices, creating new ways to use existing technology, overcoming barriers, and helping the market adopt energy efficient concepts. Other actions are longer term but have the potential to bring further dramatic efficiency improvement to the market. The participants at the RMI data center charrette (RMI 2003), for example, felt that an order of magnitude energy reduction was possible.

To achieve this level of improvement, it is likely that all elements in the data center from chip level through building systems, to building shell would need to be optimized. This level of improvement would require simultaneous and coordinated RD&D efforts involving all energy using devices and systems – for example; improving the efficiency of computer chips, computer power supplies, heat transfer through cabinets, HVAC systems, UPS systems, standby power reduction, more efficient computer code, etc. This lofty goal is unlikely to occur without strategic guidance given the fragmented nature of market, the number and variety of data center suppliers, and the evolving nature of the wide assortment of IT equipment. In addition there are numerous barriers to change. Issues such as fast track design and construction schedules, reliability at any cost, inertia to maintain proven (although inefficient) design, etc. are preventing advances in this market. However, large efficiency gains are possible in the areas identified in this roadmap. An integrated strategy that leverages public interest funding has the potential to achieve a dramatic efficiency gain.

Current understanding of data center energy efficiency in the industries and institutions that rely on them is very limited. Typically, data center professionals have a thorough understanding of issues related to power quality, reliability, and availability, but energy efficiency is not a high priority. The general lack of benchmark information, various definitions of energy intensity, together with traditional barriers limiting efficiency improvements in data centers immediately suggests areas for further research, development and market transformation. In addition, case study recommendations and other industry input point to many areas where large energy savings are possible.

Organization of this Roadmap

The topic descriptions in the roadmap are organized into the following categories:

- Collecting, analyzing, and applying data center market data
- Improving Facility Infrastructure Efficiency
- Improving Building Systems and IT Equipment Interfaces
- Improving IT Equipment Efficiency

For each roadmap activity, industry participants attempted to identify the activities most suited for public interest research, development, and demonstration (RD&D) actions recognizing that some areas in need of research are not appropriate for public interest efforts. Research and advancement in some areas can best be accomplished by industry efforts, such as improving idle state performance of computing equipment. Other activities are good candidates for public interest involvement since they would not otherwise be accomplished such as benchmarking performance across the various industries that operate data centers. Still other actions may be better accomplished through collaboration or setting standards of performance with data center IT or equipment suppliers. Examples of this category are potentially to develop more efficient power supplies for IT equipment or to develop more efficient specialty infrastructure systems.

Roadmap

Collecting, Analyzing, and Applying Data Center Market Information

Understanding the Data Center Market

Data centers house computers and IT equipment to provide functions such as data information storage, data processing, and information dissemination. Many definitions have been used to describe "data centers" [ACEEE and CECS 2001; Aebischer et al. 2002b; Blount et al. 2001; Brown et al. 2001; Callsen 2000; Elliot 2001; Gruener 2000; Intel 2002; Mitchell-Jackson 2001; Planet-TECH 2002; Robertson and Romm 2002]. With the boom associated with the Internet in the late '90's came new names for data centers including "server farms", "collocation facilities" and "telecommunication hotels". Although data centers have been important to industries, institutions, and governmental agencies for some time, it was the Internet and mission-critical facilities that brought energy consumption in data centers to the forefront.

Electric power requirements for data centers became an important issue for three very different reasons. First, computer technology, primarily chip technology, was creating higher heat density in smaller and smaller geometries. The simultaneous compaction and increase in electrical power caused concern over the ability to cool future generations of IT equipment. Secondly, the facilities that support the Internet were requesting unrealistic levels of electrical power from utilities. That requested power, if it materialized, would have required major changes in electrical utility generation and distribution infrastructure. Third, IT professionals, data center operators, and facility designers aggravated both situations by predicting huge increases in electrical demand for future computing equipment. Limited energy benchmarks in operating data centers confirm that present day energy use is much lower than predicted. When those high-expected loads did not materialize, over sizing of data center infrastructure resulted in inefficient operation in many data centers. If criteria are not developed to improve the understanding of near and longer-term electric load requirements, such inefficient operation is likely to continue into the future.

To first come to grips with the extent of this problem, the place to begin is to characterize the stock of data centers and their load intensity in California. These characteristics have turned out to be difficult to estimate. The market is characterized by constant change and there is no reliable source of market data covering all of the various types of data centers. Load intensity for data centers supporting the Internet fluctuates greatly with the rise and decline of dot COM companies but data center load intensity is also affected by the trends in computing capability and energy intensity within IT equipment. One scenario suggests that the total computing electrical load is increasing at a modest pace and being compacted into a smaller number of data centers. Anecdotal evidence indicates that recently completed data center facilities are being converted to other uses. Other scenarios suggest that computing electrical load may actually decrease as the computational capability of future generations of IT equipment will outstrip the computational need and allow older equipment to be retired. [Anonymous 2001; Baer; Bors 2000; Mandel 2001]. For these reasons, identifying and tracking energy trends in the industry is a prerequisite for coping with increasing energy intensity within data center facilities and to predict the impact on electric utility infrastructure.

Possible Public Interest Actions:

- Update the California data center market assessment and develop a better understanding of the market by surveying industries that provide specialized goods or services for data centers such as manufacturers of raised floor, or UPS systems.
- Monitor trends in the data center market, such as space availability and processor heat intensity, through collaboration with industry associations such as iTherm, 7x24 Exchange, and the Uptime Institute
- Project future data center market and energy demand by working with industry associations
- Develop market data at the utility level to facilitate system planning and identification of potential bottlenecks. Monitor Utility load requests for new projects

Benchmarking Energy Use

Understanding Data Center Computer Load Density

A comprehensive literature review revealed that there is limited information available for understanding data center energy performance. This situation has created considerable confusion and conflicting information concerning the true electrical power demand in data centers [Baer; Bors 2000; Hellmann 2002; Mandel 2001; Mitchell-Jackson et al. 2001; Roth et al. 2002; The

Uptime Institute 2000; Thompson 2002; Wood 2002]. Appropriate PIER involvement would be to provide an overview of the current energy use through benchmarking a diverse sampling of the state's data centers. This would establish a baseline to develop an understanding of current operation and enable comparison to similar facilities. The benchmarking framework could then be used to track energy performance over time using a consistent set of metrics. As has been demonstrated with other building types and equipment ratings (i.e. EnergyStar), benchmarking will lead to improved energy efficiency through identification and use of best practices in the case of building systems, and improved component efficiency for items such as computer room air conditioners, or computer power supplies. It is also likely that areas requiring research to get over other technological or institutional barriers will be identified.

During our research IT professionals and data center designers expressed a good deal of confusion regarding data center load densities. There is a wide variety of computing and communication equipment each characterized by varying energy demand and intensity. There currently is little measured benchmark data for energy end-use taking into account load diversity and other operational factors. Hence IT professionals and data center designers frequently overstate energy requirements by relying on nameplate ratings or other conservative estimates. IT equipment includes many types of devices from mainframe computers to "blade" servers to disc storage devices yet the problem of identifying the composite true composite electrical load is a common theme.

Even after current load density benchmarks are established, they will likely require continuous maintenance as industry conditions rapidly evolve. The trend in processors and storage media has been to provide exponential improvement in computing capability as predicted by "Moore's Law", and processors have exhibited corresponding increases in heat density. This trend produces locally intense heat at the processor and when servers are stacked together - in the data center. Many data centers are constantly adding and/or removing processing equipment due to growth, changes of occupants, or technology improvements. While these changes have relatively little impact in the short term, they can lead to load growth for the data center over time. This situation leads to difficulty in understanding the operational state for the current collection of IT equipment and the situation becomes even less clear when trying to predict future trends. In one scenario, processor heat load is expected to rise exponentially as it has in the recent past. In another, processors and related components are expected to become more thermally efficient. And in yet another, computing capability is theorized to outstrip computing needs resulting in fewer IT devices. There are also load uncertainty issues due to electrical load diversity such as occurs within computing equipment due to various operational states (sleep mode, full processing, data storage, etc.) and on a macro level for all electrical systems (various operating combinations of IT and infrastructure equipment, fans or compressors on or off, etc.).

By understanding the current heat producing electrical loads, and trending their changes over time, the industry can better design systems to operate efficiently today, and make them adaptable for efficient operation in the future. Limited benchmarking and case studies to date provide insight into the actual range of energy intensity in California data centers (figure 6). However, the load densities exhibited by facilities studied varies widely, and further work remains to characterize these facilities so that the data collected can be used to predict the load density of future facilities. Additional benchmarking will help to provide comparative data for various types of data centers and are very likely to lead to identification of best practices.

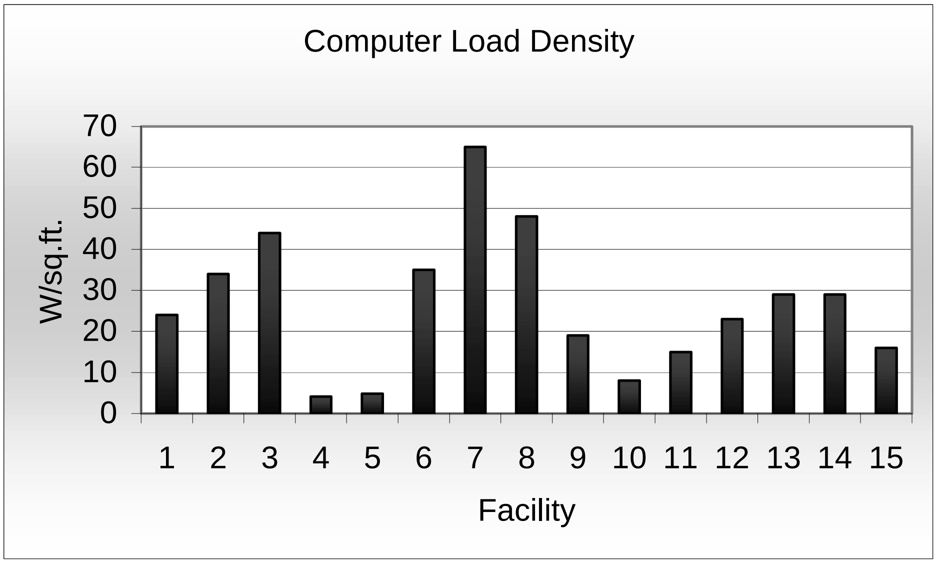

Figure 5 Computer Load Density Benchmark Results

Possible Public Interest Actions:

- Develop robust benchmark data by compiling available end use data and adding new benchmark data through case studies or industry self- benchmarking.
- Encourage sub-metering and instrumentation to facilitate monitoring energy end use in data centers.
- Through case studies, illustrate the margin between actual loads and original design loads.
- Benchmark actual operating temperature and humidity ranges in data centers.
- Develop and deploy a benchmarking protocol to enable data center designers, owners and operators, commissioning agents, and other energy engineers to perform benchmarking in a standard manner.
- Develop a database of energy benchmarks using standard benchmark data collected through case studies or self-benchmarking (by use of standard protocol).

Developing Energy Metrics

Benchmarking, using consistent metrics provides data center industry professionals with valuable data to enable performance comparisons much the way that commercial buildings are compared. Facility designers and operators have traditionally used the metric, kW/sq. ft., as a basis to define computer, lighting, HVAC, and other electrical loads. However, computer equipment load intensity has been expressed in many different ways based upon:

- Area under server rack
- Area of raised floor
- Total building area
- Raised floor area, less HVAC equipment occupying space on the raised floor
- Other variations

Additionally, some data center professionals have abandoned the kW/sq. ft. metric in favor of W/rack, with the number of racks determined from physical space available. Providing a consistent metric to define IT equipment load intensity is important for a consistent understanding of design capacity and actual performance.

Other metrics that quantify computational efficiency such as millions of instructions per second per kilowatt (MIPS/kW) are also being proposed[3]. Key facility systems' efficiencies can be evaluated through the use of other metrics such as kW/ton of chilled water, or cfm/kW of air, both of which provide direct system level efficiency comparisons.

To develop metrics most useful to the data center market, the first step may involve an examination of the pros and cons of existing metrics used by engineers and researchers from different disciplines [Aebischer et al. 2002a; Aebischer et al. 2002b; Beck 2001; Feng et al. 2002; Mitchell-Jackson 2001; PG&E 2001]. The second step would be to further refine and get consensus on the metrics that can be of most use. One data center industry association, the Uptime Institute (www.upsite.com), has attempted to standardize the definitions of kW/sf in data center spaces, however their target constituency represents only a portion of the data center market, and other important facility system efficiency metrics (such as chilled water system efficiency in kW/ton) are not addressed. Nonetheless their data helps to characterize the current computing equipment load density in data centers (figure 6). On average this data correlates well with case studies performed to date.

[3] MIPS is defined as "million instructions per second" and is a measure of the rate that computations are occurring in a computer

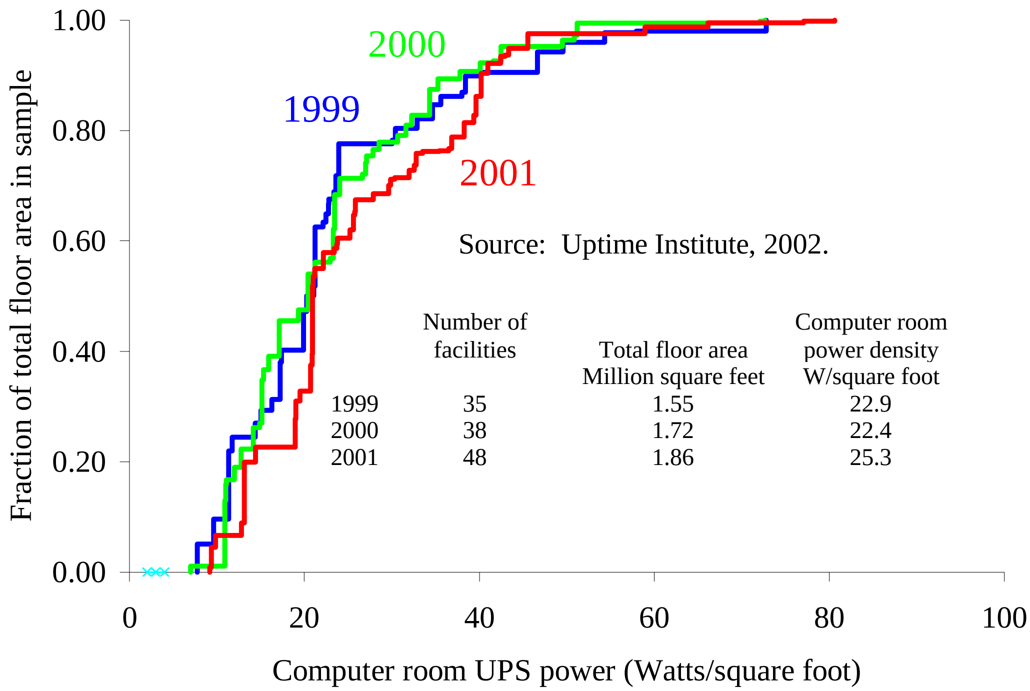

Figure 6. Uptime Institute Energy Density Data

Possible Public Interest Actions:

- Work with industry associations such as 7X24 exchange (www.7x24exchange.org), Uptime Institute (www.upsite.org), Silicon Valley Manufacturers Group (www.svmg.org), AFCOM (www.afcom.com), Advanced Computing Systems Association - USENIX (www.usenix.org) etc. to develop metrics of most use for the data center community.
- Provide workshops with industry associations and public utilities to promote use of standard metrics
- Develop protocols that include and integrate energy efficiency into data center performance metrics (e.g., availability, reliability). Develop case studies to evaluate the related construction and operational cost implications of these key areas.

Benchmarking IT Equipment - Actual vs. Nameplate

Predictions of electrical requirements for IT equipment are often determined by use of "nameplate" values. Common nameplate information for most pieces of computer or network equipment usually provide electrical values designed with a "safety factor" to ensure that the equipment will energize and run safely. Typically the values specified by the manufacturer are conservatively set with little correlation to normal operational conditions. When equipment nameplate information is used directly to develop facility power consumption and resulting cooling requirements, the facility systems are often oversized by factors of four or more. Obtaining and publicizing true power demand for IT equipment would provide a much needed, rational basis for determining real power requirements. Comparing actual and nameplate values will provide important insight for IT and facility professionals and can lead to improved sizing of electrical and mechanical systems. One professional described the need to determine an electrical "Expected Maximum Load" (EML) and resulting "Expected Maximum Heat" (EMH) rejected for each piece of equipment as an alternative to nameplate ratings.

Possible Public Interest Actions:

- Benchmark actual loads of various types of IT equipment typically found in data centers and compare to nameplate values. Publicize findings and develop training guidance to deal with nameplate values in data center facility design.
- Deploy guidance through workshops and training sessions.
- Develop testing protocols to characterize the Expected Maximum Load (EML) and Expected Maximum Heat (EMH) by working with manufacturers of IT equipment.
- Monitor projections from industry organizations such as iTherm (www.itherm.org), Intel's Developer's Forum, etc.
- Investigate characteristics of new and emerging computer technologies, such as blade servers.

Identifying and Developing Best Practices

By reviewing case studies and benchmark data, the better performing facility systems will become evident. To investigate how better energy performance is achieved, the design features and operating procedures deployed in the top energy performing systems can be examined and documented. In this way, best practices for the data center infrastructure will be identified. For example, superior equipment used in interfacing IT equipment with facility systems (racks, cabinets, etc.) also affects energy performance and similar to the facility systems, better energy efficiency performance in certain configurations can be identified. Examples of best practices in the operating procedures area might include selections of the range of allowable temperature and humidity. . Once best practices are identified, existing and new data centers' performance can be improved by actively promoting the better performing technologies and strategies.

Figure 7 Server Cabinets and Racks

Possible Public Interest Actions:

- From benchmarking activities, identify the top performing data centers from an energy perspective.
- Conduct investigations at these facilities to determine which practices contributed to such performance.
- Confirm the cost-effectiveness of these practices.
- Disseminate information about the identified best practices to major industry stakeholders.
- Add to, or modify the energy research roadmap as gaps in available solutions are identified.
- Determine widest observed temperature and humidity set point operating range in data center spaces. Work with industry associations to establish broader ranges yet maintain reliability.
- Research available modeling tools and provide designers with comparative data
- Survey available energy storage/un-interruptible power technologies and their relative efficiencies

Improving Facility Infrastructure Efficiency

Develop Better Tools including Planning and Design Guides

Data Center designers and owners confirm that critical facets of the design and operation of data centers are often accomplished through rough approximations or trial and error. There are many tools currently available for modeling energy, airflow, and heat transfer, however they are not routinely used in data centers. Apparently, the complexity, technical limitations, perceived inaccuracy, and cost of these tools are barring them from widespread use in the data center sector. To improve the accessibility of available tools, research will be required to provide data center designers and operators with a comprehensive base of existing resources. Areas where better design tools are needed could then be identified and targeted for future development. For example, the use of Computational Fluid Dynamics (CFD) software to design HVAC systems for complex arrangements of server racks may be desirable but available software packages are unwieldy in their current state. Research to improve such software may well result in efficiency gains for this industry.

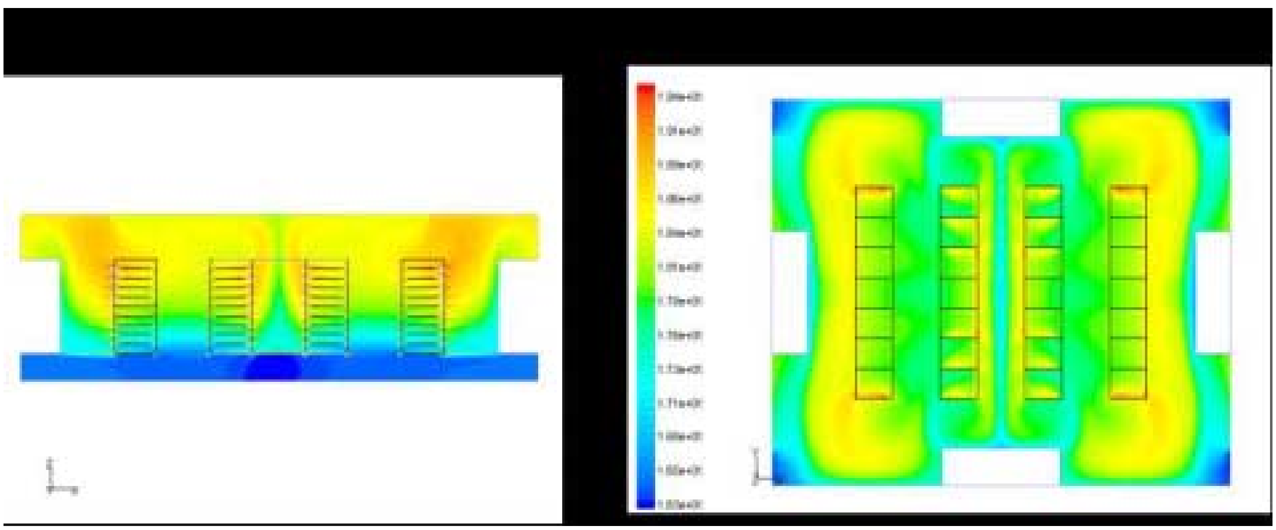

Figure 8 Computational Fluid Dynamics Models

Another example of where better tools are needed is in sizing and placement of floor tiles to provide air-cooling to racks of computers through raised floor systems. Design professionals describe current practice as far from an exact science, and accomplished through judgment, experience, or trial and error. Examining best practices may yield some clues, however the industry needs a simplified, yet accurate method to assist in energy efficient computer room HVAC design.

Over-sizing of electrical infrastructure is epidemic in the data center industry. Accounts of installed infrastructure capable of serving power densities ranging from 100 to 300 Watts per square foot (W/ft^2) have been routinely cited in media reports about these facilities (Stein 2002). Yet, both published and unpublished studies of actual data center power demand suggest that on average most of these facilities actually exhibit energy intensities between 30 to 55 W/ft^2 (Mitchell-Jackson 2001).

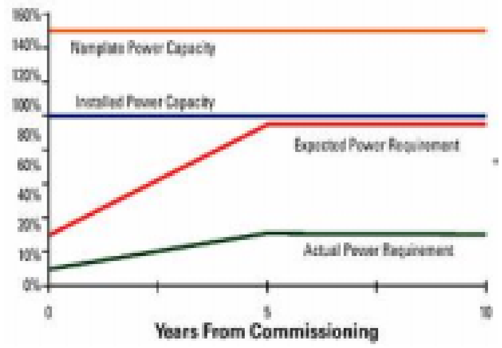

Figure 9 Representative Loading in Data Centers

Mitchell-Jackson (2001) and others have identified numerous reasons why data center operators and designers oversize electrical infrastructure. An article in Energy User News illustrated this issue graphically as shown in Figure 9.

The following reasons that designers over-size electrical systems are cited:

- Actual power requirements for computing equipment are often much less than nameplate data. Nameplate ratings are sometimes used as the basis to size electrical equipment and then result in oversized mechanical HVAC equipment for cooling.
- They have little guidance regarding how the power demand of electronic equipment varies depending on whether that equipment is in active mode, or is idling. (i.e. diversity of equipment load)
- They sometimes, inappropriately, apply power densities based on small areas to much broader areas such as power density for a computer rack being applied to the entire raised floor area.

In the last decade, many data centers were hastily planned since IT equipment had to be quickly installed on fast-track schedules to meet business objectives. One Case Study participant related that it was impossible to find people with data center design or operational experience during that time, and that the industry has been learning as they go. Often, undesired consequences induced by the lack of planning and experience were not apparent until after serious reliability problems or poor environmental control had already occurred [Sullivan 2002; Thompson 2002].

Designers are tasked with accurately predicting space, energy requirements, and cooling needs to ensure data center reliability. Due to the fast-track nature of data center projects, the lack of experienced technical expertise, and the myriad of design decisions encountered, it is appropriate to consider a data center planning and design guide addressing efficiency in key areas.   Such a tool could include guidance on thermal trends, growth in the amount and intensity of servers, incremental build-out, and flexibility [Anonymous 2001]. Efficiency gains could be maximized by incorporating green building principles and by integrating more efficient IT and facility equipment as they become commercially available. To ensure adoption, energy efficiency, reliability, and security would need equal consideration and evaluation. [Anonymous 2002a; Beck 2001].

Many in the Data Center Industry expect the power density of electronic equipment commonly used in data centers to rise rapidly in the near future (Uptime Institute 2000). As a result, some types of data centers believe that they must demonstrate to their customers that they have surplus capacity and redundant systems - resulting in a "more is better" philosophy - even where additional electrical power is not reasonably going to be needed.

Regardless of the reasons, the over-sizing of electrical infrastructure has several consequences, most of which are negative for both the data center industry and society at large. Some oversized electrical equipment operates inefficiently at small part loads, which wastes energy. Excessive capital costs and possible delays in obtaining power from the local utility are likely outcomes when electrical equipment is oversized and this becomes a barrier to development in the data center sector. Lastly, many electric utilities responding to power requests based upon exaggerated power demand estimates may over invest in transmission and distribution infrastructure, or as has happened in some California locations, may deny the request for service based upon transmission constraints, forcing the data center to be located elsewhere - possibly outside of the state. In addition, some utilities are contemplating rate schedules that include provisions for recapturing capital cost of new transmission and distribution infrastructure.

Possible Public Interest Actions:

- Research available data center design and analysis tools and summarize their features on a website.
- Independently confirm adequacy of CFD modeling tools to accurately predict thermal performance.
- Identify systems, components, and issues for which guidance is needed. Develop a guide (or guides) incorporating current best practices along with any new ideas.
- Develop advanced design and modeling tools
- Develop and implement modular (scalable) system concepts to improve part load efficiency
- Develop mechanical and electrical system sizing guidelines including use of benchmark results to account for load diversity to allow efficient operation initially and as IT equipment loads change. Consider the relationships between reliability, availability, and energy efficiency.
- Use benchmark data to establish a correlation between the nameplate ratings and actual loads associated with IT equipment. Work with industry associations to influence manufacturers to establish and publish realistic nameplate values for various operational states (i.e. sleep mode etc.)
- Use benchmark data to develop guidelines to account for the role of equipment diversity (active vs. idle) when estimating data center electrical loads.

Developing Better Monitoring and Control Tools

Although building management systems are currently used to monitor and control energy intensive systems operating cost in the data center sector, they are rarely used to optimize energy performance Research is needed to develop and deploy improved building monitoring systems which are able to evaluate and correct energy performance as well as protect critical computing equipment.

Possible Public Interest Actions:

- Survey a sampling of Data Center operators to determine current practice, existing system capability, and barriers to more extensive monitoring and control.

Maintaining Efficiency as IT Equipment Electrical Requirements Change

In order to improve the efficiency of data centers, HVAC and electrical systems are needed that are capable of providing efficient operation over the range of operating loads that occur as data centers are incrementally built-out, or as computing equipment and electrical loads vary with changing technology. Currently, electrical and HVAC systems in most data centers operate considerably below their design basis. Over the life of the facility, many different loading conditions will be present based upon the changing conditions discussed above under benchmarking. The challenge that designers face is to provide design options that satisfy both fully loaded and partially loaded conditions while achieving high efficiency. Design professionals have identified that design options traditionally have been limited by capital budget constraints, unrealistic owner demands, or other reasons, but often are due to lack of good planning guidance. Planning for incremental build-out, for example, could defer the cost of capital equipment until needed in future years, while selecting energy efficient designs matched to smaller initial electrical loads.

Clearly, systems are needed that can maintain efficient operation over extremely wide load variations. Actual electrical loads may vary from design values for many reasons - overly conservative design requirements, change in computing equipment, or simply changes in the mission of the data center. Changes in technology in the future, such as use of smaller, more efficient servers, or use of direct liquid cooling instead of air, may also result in part load operation of conventional cooling systems. Strategies may include incremental connection of UPS systems, chillers, pumps, and fans while using low-pressure drop distribution. Research into optimization strategies and promotion of best practices will increase the likelihood of industry adopting more efficient approaches using current technologies. The same strategies may also enable demand response reductions for emergencies or rate relief. The desire to have more reserve capability may be overcome by demonstrating the ability to economically and quickly increase the capability of infrastructure systems. Future data center business should be more cost-competitive, and designs that can deliver major savings in both capital cost (correct sizing) and operating cost (high efficiency) should provide their owners and operators with a competitive advantage [RMI and DR International 2002].

Possible Public Interest Actions:

- Develop case studies to demonstrate how modular design of facility systems can improve efficiency and reliability.
- Develop a model design criteria that facility owners could use to specify efficiency goals.

Electrical System Issues

Power Reliability and Uninterruptible Power Supplies

An uninterruptible power supply (UPS) is used to protect sensitive loads (computers, servers, critical electronic equipment, etc.) from various forms of power disturbances during transmission or distribution that can affect their operation or service life. UPS's are installed between the utility (or distributed) power supply and end-use loads. These systems usually contain both electronic control circuitry and electrical storage. Often, electrical storage is provided by lead-acid batteries, although flywheels seem to be picking up market share.

Figure 10 An older UPS system

Prior benchmarking by LBNL has identified that UPS systems represent a large opportunity for energy savings. UPS systems are continuously energized to provide standby power and power conditioning for IT equipment and varying amounts of facility infrastructure depending upon the critical nature of the center. The UPS energy losses occur due to electrical power conversion and to charge batteries or maintain inertial systems. The efficiency of UPS systems drops off significantly at partial load conditions, which is typically how most data centers operate. See figure 11 below displaying actual measured values from previous case studies. To determine the loss in the UPS, electrical load was directly measured at the input and output of the UPS.

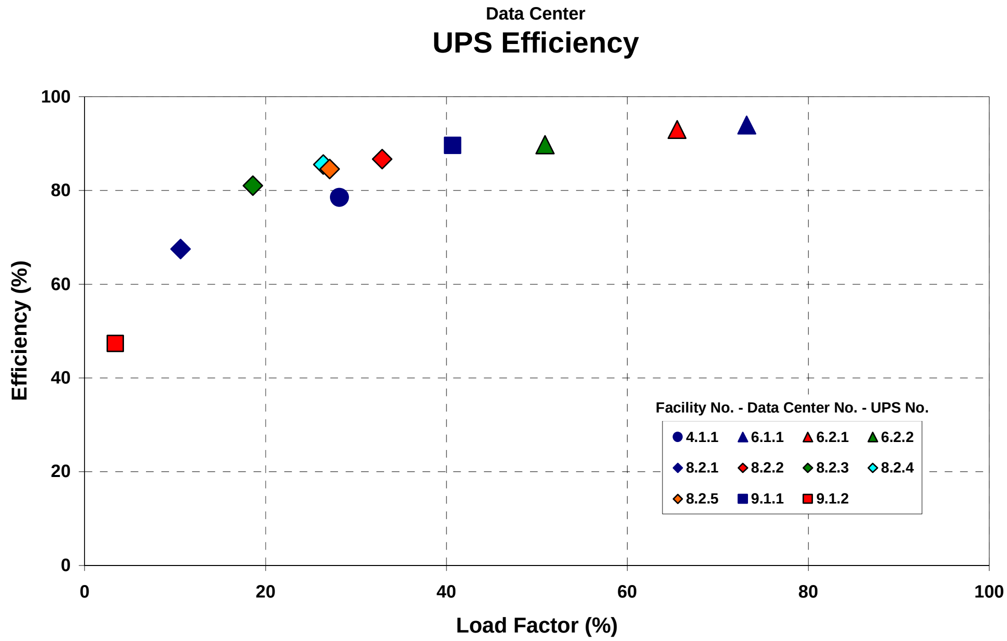

Figure 11 Measured UPS Performance From Case Studies

To compound the inherent inefficiency of UPS systems, redundancy strategies often call for use of multiple UPS's where each may be lightly loaded.

Even small savings in UPS efficiency can yield large on-going savings. To investigate the efficiency opportunity, research into the available UPS systems' to determine efficiencies over varying percentages of full load is necessary. This research could lead to development of better labeling, more useful ratings, and could be used to develop financial incentives for implementing more efficient systems. Results will provide owners and designers useful information on the efficiency of available systems for various loading conditions. In addition, efficiencies related to common redundancy strategies can be studied to determine best practices for achieving both desired redundancy and energy efficiency. This research has a broader application since UPS systems are also prevalent in other building types such as cleanrooms and laboratories with hazardous environments or where other life safety equipment is involved.

Once best practices utilizing current technology are developed, it should be possible to collaborate with UPS manufacturers and other researchers to develop next generation systems. For example, many researchers are working on improved battery systems, and inertial UPS systems are becoming more prevalent. Collaboration with those developing new battery or inertial technology or may yield breakthroughs in both efficiency and capacity. Other technologies such as fuel cells may also play a role transforming the UPS market.

For facility operators or designers who would like to select UPS based, at least in part, on energy efficiency, the web of options, claims, and counterclaims is confusing at best. Impartial reporting of the efficiency of various systems could provide owners and designers with valuable selection criteria. For example, many major UPS manufacturers claim energy savings associated with their UPS when compared to their competitors. This situation might be improved if there were an independent party that could offer credible advice regarding integrating energy efficiency considerations into the UPS selection process.

Possible Public Interest Actions:

- Survey available UPS and energy storage technologies. Evaluate controls and control strategies. Determine losses versus load for each UPS
- Hold workshops involving facility electrical design professionals, and other researchers to investigate UPS system design concepts and configurations to achieve desired redundancy
- Develop more efficient UPS solutions for various levels of reliability (N+1, N+2, 2N, etc.)
- Provide training workshops summarizing comparison of various UPS equipment and efficient methods of achieving redundancy.
- Utilize UPS ratings to develop rebate programs through public utilities
- Develop a model of data center UPS to evaluate a variety of storage technology combinations. Analyze the lifecycle costs and benefits of the different combinations. Distribute the results of this analysis to data center industry stakeholders.
- Perform research to develop more efficient energy storage technologies.

Reduce losses due to harmonics

Single-phase electronic loads, such as computers, servers, hard drives, and other devices used in data centers, are susceptible to harmonics, which can cause overheated neutrals in 3-phase/4-wire systems and overheated transformers due to circulating currents. Such overheating can lead to shortened transformer life, as well as potential safety concerns (fires, explosions, etc.) and there are electrical power losses caused by resistance to the large neutral currents. Many types of electronic equipment can generate harmonics. The heating caused by harmonic currents also wastes energy required to provide cooling [Gilleskie 2002; Greenberg 2001; Howe et al. 2001]. Today, commodity servers containing inexpensive power supplies often introduce significant harmonics. When large numbers of servers generating harmonics are present, high transformer operating temperatures result along with high neutral conductor current in three phase electrical systems.

Determining to what extent harmonic currents are induced by clusters computing equipment in data centers could be a worthwhile first step. Then, existing technologies for mitigating these currents could be evaluated. Based upon these findings, new technologies or strategies to eliminate the source of the problems or to mitigate the effects could be developed.

Possible Public Interest Actions:

- Research available technologies and develop improved cost effective techniques to diagnose and mitigate harmonics effects
- Benchmark actual power factor vs. IT equipment manufacturer's claimed power factors

Optimize electrical distribution

There are often several levels of power conversion occurring in data center facilities and within the IT equipment contained therein. Most data centers experience significant electrical power losses in their facility systems' supply and distribution including losses in transformers, power line conditioners, UPS, line loses, etc. Add to these the power losses due to inefficient power conversion (power supplies and/or voltage changes) within IT equipment, and the result is a large loss in useful electrical power coupled with energy use for HVAC systems to remove the heat produced in conditioned spaces.

An optimal system might integrate the IT equipment with the facility in such a way as to minimize power conversions. For example, the individual power supplies in servers could be eliminated if the correct voltages of DC power could be supplied efficiently from a central system, or in the case of fuel cells, directly from the power source. One industry expert envisions the data center of the future similar to a computer in its case. Taking this idea a step further, the electrical system could be thought of as an integrated system from where it enters the data center to the ultimate end use. When viewed in this manner, optimized systems could be designed so as to optimize energy (distribution and conversion losses), reliability, power quality, and potentially provide additional benefits such as elimination of harmful harmonics.

Possible Public Interest Actions:

- Workshop involving computer/server manufacturers, facility electrical design professionals, and other researchers to investigate system design concepts from the electric meter through to the computing device - whole system approach.
- Benchmark losses in power conversion throughout the electrical system including transformers losses, UPS losses, power supply losses etc.

Develop integrated solutions to meet demands for power generation, distribution, reliability, and efficiency

Although the focus of this roadmap is on the energy efficiency of the data center and not distributed generation (DG), there are areas where distributed generation may hold promise for efficiency improvements also. Based upon business needs, power reliability and power quality often are the primary driving force in data center design. Mission critical data centers strive to avoid any sort of power outage because an interruption of power could cost millions of dollars per occurrence.

Grid-connected DG systems and in particular, fuel cell-based systems could be important solutions [Planet-TECH 2002]. One option would be to have the primary power supply be an on-site generation system achieving a minimum of double redundancy, with the grid as backup. For example, use of fuel cells with the grid as backup, could simplify power conversions, be a ready source of uninterruptible power, and could eliminate costly and inefficient equipment.

Strategies can be developed to improve power reliability, power conditioning, uninterruptible power supply (UPS), and improved energy efficiency through use of DG systems [Anonymous 2001]. High availability can also be achieved with Combined Heat and Power (CHP) for Mission Critical Applications [Cratty and Allen 2001].

Possible Public Interest Actions:

- Demonstrate more efficient power distribution. To simplify the path of power to the computers, the dual redundant, on-line UPS systems could be replaced with simpler self-stabilizing buffer/transient technology systems (flywheels, new high power batteries or super-capacitors), powered by a clean, reliable on-site power source (*e.g.,* turbines, fuel cells, etc.) Part or all of this strategy could be demonstrated in an operating data center.
- Demonstrate a thermal-based cooling system that uses an on-site generator's waste heat to drive absorption, desiccant or other cooling cycle technology
- Accelerate the development of reliable, clean, and economically feasible distributed generation technologies (such as fuel cells) for critical power applications

Improve lighting efficiency

Energy used for lighting in data centers represents a small fraction of the overall energy use, yet the opportunity for efficiency savings is great – much more than for commercial office space – due to the amount of time when unoccupied. Modern Internet hosting facilities take advantage of this fact by providing a "lights out" philosophy where lighting is provided only when needed. Standard lighting controls in combination with more sophisticated building management systems can easily achieve a 50% reduction in lighting electrical energy use.

Possible Public Interest Actions:

- Demonstrate saving potential through cases studies and demonstration projects using existing lighting controls.
- Develop energy efficient task maintenance lighting to avoid lighting large data center areas for locally small maintenance or installation activity.

HVAC System Issues

Improve the performance of Data Center Air Conditioning

Computer Room Air Conditioners (CRAC units) are the most widely used specialty HVAC components in data centers. The efficiency performance of today's CRAC units operating with raised floor systems, however, is far from optimal. Some examples from case studies of where efficiency was impacted by inefficient CRAC unit deployment and operation illustrate the extent of such problems:

- It was not uncommon to find some CRAC units humidifying while others were de-humidifying the same space.
- Often, CRAC units were not placed in optimal locations.
- In one case, CRAC units were found not to be providing any cooling and could have been turned off, relying on a more efficient central house system.
- Often, more CRAC units were operating than needed.
- Air return to CRAC units did not take advantage of thermal stratification in the data center.
- CRAC units were manually turned on and off.
- Humidification methods were extremely inefficient.
- Openings in raised floors allowed air to bypass its intended use.
- Placement of raised floor tiles with openings was subjective and/or based upon experience with less than optimal results.
- Areas under raised floors were blocked, preventing airflow to where it was needed.

There appear to be many efficiency improvement opportunities with CRAC units and raised floor systems in data centers. These opportunities span design, operations, and maintenance involving the air conditioning units and their controls, the raised floor system, and interaction with other building air systems. Many measures could benefit existing data centers suggesting a public interest role to encourage and provide guidance for energy efficiency commissioning and retrofits. Best practices discovered through case studies and benchmarking, can be applied to both existing data centers and in new data center facilities.

Figure 12 Typical CRAC unit

Possible Public Interest Actions:

- ♦ Develop incentives to apply Best Practices in new or retrofit data center HVAC systems
- ♦ Hold workshops with Industry Associations such as the 7X24 Exchange Organization to develop additional improvements, disseminate best practice information, provide

information on public interest incentive programs provided by the California Energy Commission or Utilities.

- Develop a demonstration project to illustrate efficiency improvement opportunities, in addition to other benefits such as improved thermal performance.
- Improve energy efficiency of CRAC units (e.g. more efficient fans, motors, use of variable speed compressors, improved controls, etc.)

Increase the use of "Free Cooling"

As data center HVAC systems typically provide cooling around the clock, every day of the year, they do so during many hours in which the outside air conditions are favorable enough to offset the use of refrigerated cooling. However, anecdotal evidence suggests that few data centers use air-side economizers. As a result, this technology, which has the potential to significantly reduce data center cooling costs, appears to be underutilized.

The reasons for such underutilization are unclear. Perhaps data center designers are concerned that economizers will introduce unwanted humidity or airborne particulates (Shields and Weschler 1998). Perhaps they are concerned that the first and operating costs associated with equipment required to remove humidity and particulates will render such systems uneconomical. This may be further reason to examine the real technological limits of temperature and humidity in data centers. Perhaps they have experienced failure rates similar to those exhibited by economizers in the low-rise commercial buildings industry, and avoid them as a means to maintain high system reliability or lower maintenance costs. These barriers should be explored to determine whether this existing technology could be better utilized. This is one area that may be included for guidance by a newly formed ASHRAE Technical Committee.

Similarly, in many locations, free cooling can be employed to produce chilled water by minimizing the use of chillers. Several different methods can be used to achieve free cooling including direct use of cooling towers, chilled water heat exchangers, and options provided by chiller manufacturers. Although these strategies have been successful in many other chilled water systems, case studies have revealed that free cooling is underutilized in data center applications.

Possible Public Interest Actions:

- Conduct a study to estimate the costs and benefits associated with using economizers in the data center sector in a variety of climate areas.
- Survey manufacturers of data center HVAC equipment to determine the market penetration of air-side economizers in data centers.
- Survey data center operators to identify barriers to more widespread use of air-side economizers.

- Provide research to develop better economizer technologies and operating techniques which have the potential to increase market penetration.

- Provide training and raise awareness of free cooling benefits.

Encourage manufacturers of computer room air-conditioner (CRAC) units to release products with variable-speed compressors

Because of uncertainty regarding power densities of electronic equipment, both now and in the future, CRAC units are frequently oversized. As a result, this equipment typically operates to remove extremely low heating loads.

In some applications, a way to improve the efficiency of CRAC units would be to fit them with variable-speed compressors. Such units would have the potential to be four times as efficient at low loads than at rated load. Indeed, CRAC units with variable-speed compressors were developed by a major manufacturer, but never released. The manufacturer determined that few customers would be willing to pay a premium for such a feature. Given that uncertainty regarding equipment power densities has grown since this manufacturer mothballed this product, perhaps some targeted efforts could stimulate a larger market for this device.

Possible Public Interest Actions:

- Purchase and install CRAC units with variable-speed compressors for one or more showcase projects. Monitor the operation of these units and project their likely costs and benefits. Disseminate the results of these efforts to key stakeholders in the data center industry.

Promote use of highly efficient air delivery systems.

Distribution of air for cooling in a data center can be accomplished in an infinite number of ways, however the paradigm of using raised floor systems for the delivery of air can be problematic in terms of providing cooling where it is needed and with sufficient quantity to cool high heat densities. In addition, energy efficiency can suffer if the lay out of the systems is not optimal, or if unsealed openings are present.

Previous case studies have shown that an HVAC system designed similarly to a traditional building system (similar to those used in large commercial buildings) can be more efficient than the current practice of using raised floor systems with specialized computer room air conditioners.[4] Sound engineering principles, such as providing large, low-pressure drop delivery systems, along with efficient fans and motors that are controlled for varying load conditions, can be utilized in various configurations. Currently these systems are being successfully used with overhead delivery (and overheard wire management) in a small fraction of data centers. Wider acceptance of this practice, and/or using traditional design practices in concert with raised floor

[4] Case studies are available through LBNL's data center website: http://datacenters.lbl.gov

systems could provide large-scale improvements. Old paradigms where computer room air is recirculated using inefficient air movement and cooling devices need to be challenged. Systems utilizing efficiently sized air handlers and cooling coils, perhaps located outside of the data center could be an attractive alternative to current practice.
Possible Public Interest Actions:

- Develop efficient system design concepts based upon benchmarking results and good engineering practice. Hold workshops to present the design concepts to design professionals.

- Demonstrate thermal and energy improvement by optimizing existing raised floor air distribution with an Industry Partner

Determine whether wider ranges of humidity and temperature control can be tolerated by electronics equipment.
Based upon case studies and interviews with designers, it was apparent that many data centers are maintained at approximately 50 percent relative humidity within a tight range, typically plus or minus 5 percent (Nordham, Reiss, and Stein 2001). The need for such tight humidity control is questionable. The reasons given for such control seem to relate to earlier generations of computing equipment where humidified air was needed to prevent static electrical discharges that could damage electronic equipment. In addition, overly moist air could result in water vapor condensing on electronic components, which would also result in equipment damage. These humidity control parameters add energy consumption, both for dehumidifying and/or humidifying. In some cases, dehumidifying is followed by reheat in order to achieve the tight tolerances.

Were humidity levels allowed to fluctuate more, energy consumed to achieve humidity control could be saved. Industry has recognized that this situation is not optimal and is beginning to develop standards and guidelines to enable the design conditions to be relaxed. For example, the manufacturers of electronic equipment used in data centers typically specify relative humidity operating conditions ranging from 20 to 80 percent (Sun Microsystems 2001).

ASHRAE has established a new Technical Committee (TC 9.9) to focus on High Density Electronic Equipment Facility Cooling. This committee is establishing guidelines that should improve the performance of data center facilities. The guidelines focus on several important areas:

- Environmental specifications for various classes of electronic equipment
- Standardized methods of determining temperatures in data centers
- Equipment and facility layout
- Standard reporting of heat load and airflow requirements

These standards, when developed and implemented will help to improve energy efficiency through right-sizing of facility systems, efficient layouts, and broader temperature and humidity limits. Additional research leading to more realistic limits in this area should enable data center designers and operators to control temperature and humidity levels with wider tolerances. Such control will save energy, both for the humidification, and possibly by allowing increased use of outdoor air.

Cooling in a data center provides a workplace environment, and removes heat from critical electronic parts to prolong their life. Looking at these two needs separately may yield some efficiency opportunity. By working closely with IT equipment manufacturers, realistic cooling requirements to protect critical electronic components should be developed. Satisfying true cooling needs may allow relaxation of current practice especially during periods when the data center is not occupied or during periods of peak demand. For example, if even small increases in ambient temperature are acceptable, significant energy savings will result. Often data center operators choose to lower the overall ambient temperature as a solution for "hot spots" in their data center, resulting in overcooling and inefficiency. The ability to tolerate some locally higher temperatures could alleviate this problem.

Possible Public Interest Actions:

- Participate in the development of data center standards through participation in ASHRAE Technical Committee TC 9.9.
- Once guidelines or standards are established, provide training for data center designers and operators in California.
- Use models of data center HVAC systems to estimate the energy savings that could be achieved under a variety of temperature and humidity control scenarios.
- Survey data center designers and equipment manufacturers to research their basis for current data center environmental specifications.
- Contact electronics manufacturers to determine the basis for current temperature and humidity recommendations.

Making Better Use of Existing Energy Efficiency Guidance

Figure 13 Centrifugal Chiller

Selecting more efficient equipment is but one step in optimizing facility system performance and should be included along with more comprehensive system measures. Much of the energy intensive infrastructure equipment in data centers such as chillers, pumps, motors, and transformers are common in other building types. Much information concerning efficiencies of this equipment exists yet case studies and other anecdotal information highlight that data center designers and building owners need to be more exposed to information for designing systems and specifying efficient equipment. Existing guidance for efficient system design, such as that provided by ASHRAE, Cooltools, or DOE's Motor Challenge as well as comparative manufacturer's performance data is under-utilized in the specialized field of data center design.

Possible Public Interest Actions:

- Develop and provide training workshops for design and operations professionals at several locations throughout California concerning use of existing energy efficiency design information.
- Publish a list of available related energy efficiency resources on public websites.
- Develop the basis for incentives for use by Public Utilities to stimulate use of more efficient facility systems and equipment, and specialty components used in data centers.

Improving the Interface Between Building Systems and IT Equipment

Manage computing load and demand

Electrical power for computing equipment and resulting heat loads are relatively constant in data centers as long as the computing equipment set does not change. Unfortunately, computer room equipment is frequently changing due to changes in technology, computing and storage capability, and other business factors. This changes the overall electrical load (and resulting heat load) for the data center and may create the need for redistributing cooling air to relieve local "hot spots". Research into better low cost systems for monitoring and tracking changing electrical loads as well as strategies for allowing systems to respond efficiently to increases or decreases in overall or local loading is needed. One innovative data center professional suggests shifting computing to other less energy intensive areas of the data center - or - even shifting computing to other geographic locations. The ability to sense and respond to local area "hot spots" will enable this concept to succeed.

Possible Public Interest Actions:

- ♦ By working with leading industry partners, determine strategies to optimize infrastructure systems, sensing and controls, operating systems, and IT Equipment to respond efficiently to computing demands.
- ♦ Develop a demonstration project in an operating data center to prove the feasibility of concepts to sense and redirect computing when local area hot spots are detected.

Develop More Thermally Efficient Racks and Enclosures

Efficient air movement inside data centers is critical to energy performance of these facilities. Many racks used to stack IT equipment in data centers are designed to maximize the number of devices that can be accommodated. Many have design features to make them visually attractive or provide security such as locked glass panel fronts. The result is often a closed, densely packed rack system that does not allow for efficient airflow or worse, can contribute to heat transfer problems with the equipment contained in the racks. Racks and equipment enclosures that better facilitate the flow of cooling fluids across hot components would likely improve the efficiency of data center cooling systems. For example, more efficient cooling using air, chilled water, or other mediums may be possible by integrating the design of rack systems with the equipment they contain, as well as with HVAC and electrical infrastructure. The fragmented nature of the market however makes gains in this area difficult. Computer and other IT equipment manufacturers deal strictly with heat transfer issues within their equipment's envelope. Rack manufacturers develop their products independently from the IT equipment suppliers, and the facility systems. An appropriate role for public goods research would be to investigate methods to facilitate the integration of these disparate elements.

Possible Public Interest Actions:

- ♦ Convene workshops bringing the three different areas (computing equipment, rack, building systems) together with a goal of developing an optimized, integrated solution.
- ♦ Develop incentive programs to reward integrated solutions that are more efficient.

Determine Whether Increased Heat Density Makes Economic Sense

Although a guiding paradigm in the IT industry is that customers want greater functional density, the time has come to question whether further increases in density are worth their cost in energy and cooling infrastructure. Open configurations and spatial distribution of IT equipment heat loads (utilizing more real estate) may be better options than development of increasingly complex air cooling systems. There will be break even points where further increases in density are not cost effective compared with spreading out the heat sources. Evaluating options will involve present and future real estate and energy costs, the cost of HVAC systems, and other factors but should also consider the relative energy efficiency of cooling systems for a given solution.

Possible Public Interest Actions:

- ♦ Develop "heat intensity" economic evaluation tools considering expected heat intensity, energy use and electrical power rates, facility and infrastructure cost, etc.

Limits of Cooling mediums

The different cooling mediums (air, water, dielectric fluids) which could be employed in a data center, each have limits for effectively removing heat from computing equipment. Determining limits of the various cooling technologies that could be applied to electronic computing equipment and their scalability would be useful. Some research in this area by CEETHERM and other industry associations, such as iTherm are pursuing this research. The impact on building systems for various cooling mediums is not well understood. An appropriate role for public interest research would be to investigate the efficiency implications to the facility for the various mediums and encourage the use of efficient technologies. This would likely involve a collaboration between industry researchers and developers of new innovative products. Applying the results of studies focused on cooling limits at the processor or computer level, to the impact on building systems would provide useful information in planning and optimizing building systems.

Figure 14 Typical Computer Racks

Possible Public Interest Actions:

- ♦ Collaborate with CEETHERM (http://www.me.gatech.edu/me/publicat/brochures/Mettl/Bro0302.htm, iTherm (www.itherm.org), and other industry associations to assess the limits of current and emerging technology.

Evaluate Emerging IT Rack Systems and Specialty Products

Georgia Tech University and other research organizations are studying heat transfer in IT equipment and computer rack systems, and limited studies are underway. Emerging new products such as new computer rack systems that incorporate plenums and fans to direct cool air to the computing equipment, are being researched or developed by others. California can benefit by collaborating with these organizations to anticipate important trends in IT equipment's increased heat intensity and the related impact on electrical power demands. Increasing heat density will dictate the need to develop more sophisticated cooling solutions. Typically, energy efficiency is not a focus in the development of new products and solutions so long as the devices are cooled adequately. Therefore, an appropriate role for public interest research would be to ensure that energy efficiency is considered while new thermal solutions are being developed. .

Possible Public Interest Actions:

- ♦ Collaborate with CEETHERM (Georgia Tech University) and other researchers. Participate in the iTherm conference – an industry association focused on heat in electronics.
- ♦ Evaluate Emerging Systems and Develop Case Studies to Demonstrate Energy Efficiency

Development of Direct Cooling Systems

Based on prior benchmarking studies, it's evident that HVAC systems can be responsible for half (or more) of the energy consumed by data centers. Much of that energy is used to power fans, which force air through underfloor plenums, and draw air through electronic equipment. Roughly half the energy consumed by data center HVAC systems goes towards powering those fans and cooling the waste heat associated with their operation (Westphalen and Koszalinski 1999). Current HVAC systems exhibit five shortcomings that threaten their continuing dominance in the data center sector:

- The on-board fans incorporated into electronic equipment are especially inefficient and take up room that could be used to house electronic components.
- Air is not an efficient medium for heat transfer. Liquids can move heat much more efficiently.
- The same medium currently used to cool equipment (air cooling) is also used to cool workers. As a result, overcooled workers are endemic in this industry.
- As cool air passes through the IT equipment and rises from the raised floor it is warmed. One study found that air 6 feet off a raised floor was nearly 20°F warmer than air in the plenum. Warm air diminishes the lifetime of the equipment located near the top of racks and cabinets (Schmidt 2001a). The Uptime Institute has found that equipment located at the top of racks is more prone to failure and exhibits shorter life.
- Raised floor forced-air systems are incapable of cooling electronic equipment with a power density that exceeds 150 W/ft2 (Schmidt 2001b).

These shortcomings can be largely mitigated by changing to systems that apply cooling directly to hot electronic components. These systems are typically based on one of two approaches. They either spray refrigerant directly on hot electronic components (Shaw et al 2002, Isothermal Systems Research 2002) or they feature other innovative heat removal features such as micro-refrigeration systems installed directly on those components (Anderson). Using these techniques, a much smaller and simpler forced-air system can then be used to condition the space around the electronic components. While the main driver for such devices is to protect the electronic components, there are also opportunities for energy efficiency gains. An appropriate public interest activity would be to promote energy efficient solutions. The costs and benefits of applying such systems in data centers in a widespread manner are not yet well understood, however.

Possible Public Interest Actions:

- Survey the state of development of direct cooling technologies for electronic equipment and the relative energy efficiency of each.
- Investigate the costs and benefits of these technologies.
- Identify and address barriers to the widespread deployment of these technologies.

Improving Efficiency of IT Equipment

Benchmark Servers for Energy Performance

Not long ago, manufacturers began to release computer servers with low-power microprocessors. These new products allow data centers to pack servers far more densely into a given volume, without appreciably pushing up energy consumption or cooling loads (RLX Technologies 2001). Although, they are now produced by several of the dominant computer manufacturers, such as IBM, Compaq, and Dell, as well as a few start-up manufacturers such as RLX, these low-power servers still constitute a small, but growing, portion of the overall server market (Gartner Dataquest 2002). A typical blade server featuring a low power microprocessor is shown below in figure 14.

Figure 15 Typical Blade server
Source: RLX Technologies

One barrier to the widespread diffusion of low-power servers into the data center sector is that information technology specifiers have no reliable way to compare the *in situ* performance of competing servers. There are widely accepted standards for benchmarking servers for processing speed, but it's not clear how well these standards predict actual performance in a variety of data center environments. These standards only measure processing performance, and there is no standard method to measure input power. The tests on low-power servers have been done only by the manufacturers themselves, which has led to much squabbling in the industry over the methods they used and the reliability of their results. This is one area that the newly formed ASHRAE Technical Committee is working to address to standardize power measurement and reporting. Lastly, because the processors used in low-power servers operate at lower speeds than top-of-the-line servers, there is no consensus on how many low-power servers are required to perform a given task that could be performed by fewer, faster, high-power servers (Compaq 2002). The net result is that specifiers are confused, and few take the trouble to investigate the opportunities to reduce operating costs.

Possible Public Interest Actions:

- Survey benchmarks used for comparing servers to determine their usefulness for comparing energy performance.
- Survey IT specifiers to determine what information, and in what format, would empower them to compare competing servers on the basis of energy performance.
- Identify stakeholders in the server benchmarking field and obtain their cooperation in an effort to develop an energy performance benchmark.
- Develop a new standard benchmarking method for comparing the energy performance of competing servers.
- With the cooperation of industry stakeholders, encourage the development of an independent entity that could test servers to the new energy performance guidelines.

Improve Reliability and Efficiency of Power Supplies

Power supplies convert AC electrical power to DC at various voltages. In data center equipment such as servers they are usually located inside of the electronic product (internal). Power supply efficiency levels of 80 to 90% at full load are readily achievable in most internal and external power supplies at modest incremental cost through improved integrated circuits and better designs [Calwell and Reeder 2002]. With wide variations in efficiency for similar products, careful selection of energy efficient power supplies is important. Comparative information, although not currently available, would be helpful for IT professionals and facility designers to make informed decisions.

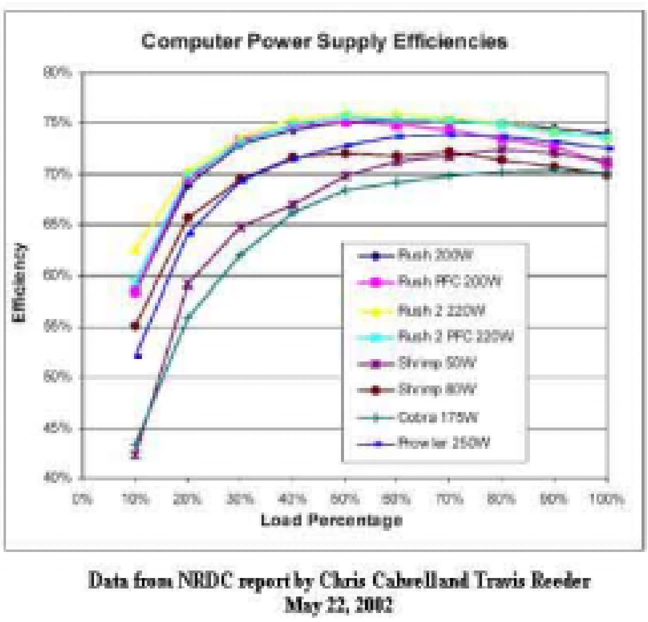

Figure 16 Computer Power Supply Efficiencies

There are often several levels of power conversion occurring in data centers and telecom facilities. Currently, most data centers have significant electrical power losses in their facility systems' supply and distribution including losses in transformers, power line conditioners, UPS, etc. Add to these, the power losses due to inefficient power supplies at the computer, and a significant efficiency opportunity results if an optimal central system can be developed. In addition, current practice introduces significant harmonics in many cases. Alternatives to current practices to deliver appropriate power to the computing device should be investigated by looking at the entire electrical supply chain. This could involve conversion of the main power to DC, elimination of individual small power supply devices, etc.

Possible Public Interest Actions:

- ♦ Benchmark and assess currently available technology for power supplies commonly used in IT equipment. Identify best performers.
- ♦ Benchmark losses in power conversion throughout the entire electrical system including transformers losses, UPS losses, power supply losses etc. with a goal of identifying improvements in efficiency for electrical power delivered to its end use.
- ♦ Hold workshops involving computer/server manufacturers, facility electrical design professionals, and other researchers to investigate system design concepts from the electric meter through to the computing device - whole system approach.
- ♦ Develop standards for reporting power supply performance. Through participation in the ASHRAE committee on High Density Electronics Facility Cooling it may be possible to introduce such a standard. Alternatively, develop a non-proprietary rating system administered by an independent organization.
- ♦ Encourage development of more efficient power supplies:
    - To reduce No load loss.
    - To improve standby performance.
    - Use two different size power supplies to maximize efficiency for a given load

IT Hardware and Software

A number of strategies dealing with IT equipment and software were identified during the roadmap development. Many of the research areas are promising and could yield dramatic energy performance improvement. The specialized and highly technical nature of most of the ideas however, did not seem appropriate for public interest research at the state level. Much research is on going by various industry players in anticipation of higher heat density, yet the motivating force usually is not energy efficiency. It is important to be aware of developments in related research areas to reward or encourage energy saving innovations. Even though development will be performed by the industry itself for issues such as better sleep modes, or passive heat transfer devices, public interest programs can provide a vital role in demonstrations and market transformation activities. Encouraging low-power computing is a complex topic. Makers of highly efficient computer chips, for example, claim their product can cut energy use significantly (e.g., in half), while others point out a limited application of the chips [Anonymous 2001].

Possible Public Interest Actions:

- Foster new and emerging technologies by demonstrating energy saving devices, systems, or strategies. Possible candidates for demonstration projects include:
  - Direct Spray Cooling of Refrigerant onto computer chips
  - Efficient heat sinks
  - Direct Refrigeration of processors
  - Software to redirect computing to eliminate thermal spikes on chips
  - Demonstration of monitoring and control to match cooling to varying heat load in localized areas
  - Demonstration of energy efficient computer cabinets
  - Demonstration of alternate cooling media (N2 cryogenic, spray cooling,) single and 2-phase systems, one fluid from building to chip
  - Demonstration of best practice for idle mode performance

Appendix A

Organizations Associated with Data Centers
Many organizations are stakeholders in the efficient operation of data centers. They include industry associations, individual industrial companies, research organizations, public interest groups, utilities, DOE, EPA, etc. Web sites for selected organizations are listed below:

iTherm
iTherm is an association of computer manufacturers, semiconductor manufacturers, and researchers focusing on heat removal issues at the chip and computer level.
www.itherm.org

CEETHERM
CEETHERM is a collaborative research program carried out at the University of Maryland and Georgia Tech University. Its focus is to research efficiency improvements for data centers from the chip level through to the building systems including distributed generation applications.
http://www.me.gatech.edu/me/publicat/brochures/Mettl/Bro0302.htm

7X24 Exchange
The 7X24 Exchange is an industry association who's goal is to improve end to end reliability by promoting an exchange of information between those that design, build, and maintain data center facilities.
www.7X24exchange.org

Uptime Institute
The Uptime Institute focuses on improving uptime management in Data Center Facilities and Information Technology organizations. Members represent Fortune 500 companies who collectively and interactively learn from each other. They sponsor meetings, tours, benchmarking, best practices, uptime metrics, and provide abnormal incident collection and analysis. They also provide seminars, training in IT, and facilities site infrastructure uptime management, and conduct sponsored research.
www.upsite.com

AFCOM
AFCOM is an association for data center professionals offering services to help support the management of data centers around the world. AFCOM was established in 1981 to offer data center managers the latest information and technology through annual conferences, published magazines, research and hotline services, and industry alliances.
www.afcom.org

Electric Power Research Institute

The Electric Power Research Institute has investigated various aspects of data centers including use of distributed power, power quality, and reliability. They have participated in the Consortium for Electric Infrastructure to Support a Digital Society (CEIDS)

www.epri.com

Silicon Valley Manufacturers Group

The Silicon Valley Manufacturers Group is an industry association in the Silicon Valley area of California. Many data center operators are members of this organization, which generally focuses on improving business competitiveness of the member companies. This focus includes interest in improving reliability and efficiency of operations.

www.svmg.org

United States Telecom Association

The United States Telecom Association is dedicated to serving the needs of the telecommunications community. They provide a forum for the telecommunications industry to define issues, debate policy and set goals. They provide lobbying, and regulatory interface for the industry.

www.usta.org

References

ACEEE, and CECS. 2001. Funding prospectus for "Analysis of Data Centers and their implications for energy demand". Washington, DC, American Council for an Energy Efficient Economy (ACEEE); Center for Energy and Climate Solutions (CECS). July 2001.

The paper includes an overview of data centers; discusses energy use, energy choices, and energy efficiency in data centers; potential impacts of data centers; present and future regulatory issues; and business opportunities in energy services.

Aebischer, B., R. Frischknecht, C. Genoud, A. Huser, and F. Varone. 2002a. Energy- and Eco-Efficiency of Data Centres. A study commissioned by Département de l'intérieur, de l'agriculture et de l'environnement (DIAE) and Service cantonal de l'énergie (ScanE) of the Canton of Geneva, Geneva, November 15.

The study investigates strategies and technical approaches to fostering more energy-efficient and environmentally sound planning, building and operating of data centres. It also formulate recommendations on how to integrate the findings in the legal and regulatory framework in order to handle construction permits for large energy consumers and promote energy efficiency in the economic sectors. Seventeen recommendations grouped in four topics are derived from study conclusions: Transfer of the accord into an institutionalized legal and regulatory framework; Energy-efficiency policies for all large energy consumers; Preconditions, and prerequisites; Operational design of voluntary energy policies.

Aebischer, B., R. Frischknecht, C. Genoud, and F. Varone. 2002b. Energy Efficiency Indicator for High Electric-Load Buildings. The Case of Data Centres. Proceedings of the IEECB 2002. 2nd International Conference on Improving Electricity Efficiency in Commercial Buildings. Nice, France.

Energy per unit of floor area is not an adequate indictor for energy efficiency in high electric-load buildings. For data centres we propose to use a two-stage coefficient of energy efficiency CEE = C1 * c2, where C1 is a measure of the efficiency of the central infrastructure and c2 a measure of the energy efficiency of the equipment.

Anonymous. 2001. Model Data Center Energy Design Meeting. Austin Energy, Austin, TX, Feb 12-13. http://www.austinenergy.com/business/energy_design_meeting.htm

Anonymous. 2002a. 7 x 24 Update: Design & Construction - Issues and trends in mission critical infrastructure design, planning and maintenance. *http://www.facilitiesnet.com/BOM/Jan02/jan02construction.shtml*. July 23, 2002. http://www.7x24exchange.org/.

Anonymous. 2002b. Continuous Availability Review (CAR). The Uptime Institute: Computersite Engineering, Inc. *http://www.upsite.com/csepages/csecar.html*. July 22, 2002.

Anonymous. 2002c. End-to-End Reliability Begins with the User's Definition of Success. The Uptime Institute. *http://www.upsite.com/TUIpages/editorials/endtoend.html*. July 22, 2002.

Anonymous. 2002d. Mechanical Systems Diagnostic Review (MSDR). The Uptime Institute: Computersite Engineering, Inc. *http://www.upsite.com/csepages/csemsdr.html*. July 22, 2002.

Anonymous. 2002e. Site Infrastructure Operations Review (SIOR). The Uptime Institute: Computersite Engineering, Inc. *http://www.upsite.com/csepages/cseior.htm l*. July 22, 2002.

Baer, D. B. Emerging Cooling Requirements & Systems in Telecommunications Spaces, Liebert Corporation.
During the last several years, power density trends, and consequently thermal density trends in telecommunications spaces have become topics of increasing interest. This paper identifies several of the underlying drivers of these trends, project possible outcomes, and assess the impact on cooling system design for these spaces.

Beck, F. 2001. Energy Smart Data Centers: Applying Energy Efficient Design And Technology To The Digital Information Sector. *Renewable Energy Policy Project (REPP): Washington, DC.* (November 2001 REPP).
http://www.repp.org/articles/static/1/1036512059_982708661.html
Both utilities and data center owners face challenges in meeting electricity demand loads with required levels of reliability. However, the bursting of the high-tech stock bubble in 2000 and the 2001 U.S. economic downturn has slowed expansion of data centers. This provides time and an opportunity to examine data center construction and operational practices with an eye toward reducing their energy demands through use of energy efficient technologies and energy smart design practices. As the economy recovers and the next data center rush approaches, best practices can reduce energy use while maintaining or even increasing data center reliability. Energy demands of data centers that support the digital information- and communications-based economy need not be as high as some predict. In fact, data center power demands could be reduced by 20 percent with minimal efficiency efforts, and by 50 percent with more aggressive efficiency measures.

Blount, H. E., H. Naah, and E. S. Johnson. 2001. Data Center and Carrier Hotel Real Estate: Refuting the Overcapacity Myth. Lehman Brothers: TELECOMMUNICATIONS, New York, June 7, 2001. http://www.lehman.com
An exclusive study examining supply and demand trends for data center and carrier hotel real estate in North America. Lehman Brothers and Cushman & Wakefield have completed the first in a regular series of proprietary studies on telecommunications real estate (TRE), including carrier hotels and data centers.

Bors, D. 2000. Data centers pose serious threat to energy supply. *Puget Sound Business Journal (Seattle) - October 9, 2000*.
http://seattle.bizjournals.com/seattle/stories/2000/10/09/focus5.html
To cope with increasing energy demand from data centers, the author discussed feasibilities of two possible approaches: 1) energy industry approach by looking at alternative energy supply 2) construction industry approach by looking at data center energy efficiency. To get there, it is worth investigating five distinct components: (I) Co-generation of power. Presently, standby diesel generators are required to maintain the desired level of reliability at most data center sites, but their exhaust makes most of these generators unacceptable for long-term power generation; (II) Fuel cells offer the promise of very clean emissions and the reasonable possibility for use as standby power; (III) Increased efficiency in data center power distribution systems. There are two separate items that are major contributors to data center power distribution system inefficiencies. The first, power distribution units (PDUs) are available with optional internal transformers that use less energy than the present cadre of K-rated transformers. The second, uninterruptible power systems (UPSs) come in a range of efficiency ratings. If the use of high-efficiency PDUs and UPSs are combined, they offer the potential of a 6 percent saving (IV) Increased efficiency in mechanical cooling systems. In order to ensure data center reliability, mechanical equipment is often selected as a large number of small, self-contained units, which offers opportunities to improve efficiencies; (V) Reductions in energy use by computer, network and storage equipment. Computer manufacturers can do their part by creating computers with greater computational power per watt. They

have been doing this for years as a side effect of hardware improvements, and they can do even better if they make it a goal.

Brown, E., R. N. Elliott, and A. Shipley. 2001. Overview of Data Centers and Their Implications for Energy Demand. Washington, DC, American Council for an Energy Efficient Economy, Center for Energy & climate Solutions (CECS). September 2001. http://www.aceee.org/pdfs/datacenter.pdf.pdf

The white paper discusses data center industry boom and energy efficiency opportunities and incentives in Internet data centers. Emerging in the late 1990's, data centers are locations of concentrated Internet traffic requiring a high-degree of power reliability and a large amount of power relative to their square footage. Typically, power needs range from 10-40MW per building, and buildings are typically built in clusters around nodes in the Internet fiber-optic backbone. During the development boom in 1999 and 2000, projects averaged 6-9 months from site acquisition to operation, and planned operational life was 36 months to refit. Even high energy-prices were dwarfed by net daily profits of 1-2 million dollars per day for these buildings during the boom, creating little incentive for efficient use of energy.

Callsen, T. P. 2000. The Art of Estimating Loads. *Data Center* (Issue 2000.04).

This article discusses the typical Data Center layout. It includes floor plan analysis, HVAC requirements, and the electrical characteristics of the computer hardware typically found in a Data Center.

Calwell, C., and T. Reeder. 2002. Power Supplies: A Hidden Opportunity for Energy Savings (An NRDC Report). Natural Resources Defense Council, San Francisco, CA, May 22, 2002. http://www.nrdc.org

The article discusses the efficiency of power supplies, which perform current conversion and are located inside of the electronic product (internal) or outside of the product (external). The study finds that most external models, often referred to as "wall-packs" or "bricks," use a very energy inefficient design called the linear power supply, with measured energy efficiencies ranging from 20 to 75%; that most internal power supply models use somewhat more efficient designs called switching or switch-mode power supplies; and that internal power supplies have energy efficiencies ranging from 50 to 90%, with wide variations in power use among similar products. Most homes have 5 to 10 devices that use external power supplies, such as cordless phones and answering machines. Internal power supplies are more prevalent in devices that have greater power requirements, typically more than 15 watts. Such devices include computers, televisions, office copiers, and stereo components. The paper points out that power supply efficiency levels of 80 to 90% are readily achievable in most internal and external power supplies at modest incremental cost through improved integrated circuits and better designs.

Compaq. 2002. *Compaq ProLiant BL 10e Delivers Industry Defining Transactions per Watt and Transactions per Square Foot*. White Paper. ftp://ftp.compaq.com/pub/products/servers/benchmarks/BL10e_webbench.pdf

Cratty, W., and W. Allen. 2001. Very High Availability (99.9999%) Combined Heat and Power for Mission Critical Applications. *Cinintel 2001*: 12. http://www.surepowersystem.com

Elliot, N. 2001. Overview of Data Centers and their implications for energy demand. Washington, DC, American Council for an Energy Efficient Economy. Jan 2001, revised June 10, 2001.

Feng, W., M. Warren, and E. Weigle. 2002. The Bladed Beowulf: A Cost-Effective Alternative to Traditional Beowulfs. *Cluster2002 Program*. http://www-unix.mcs.anl.gov/cluster2002/schedule.html; http://public.lanl.gov/feng/Bladed-Beowulf.pdf

Authors present a novel twist to the Beowulf cluster - the Bladed Beowulf. In contrast to traditional Beowulfs, which typically use Intel or AMD processors, the Bladed Beowulf uses Transmeta processors in order to keep thermal power dissipation low and reliability and density high while still achieving comparable performance to Intel- and AMD-based clusters. Given the ever-increasing complexity of traditional super-computers and Beowulf clusters; the issues of size, reliability, power consumption, and ease of administration and use will be "the" issues of this decade for high-performance computing. Bigger and faster machines are simply not good enough anymore. To illustrate, Authors present the results of performance benchmarks on the Bladed Beowulf and introduce two performance metrics that contribute to the total cost of ownership (TCO) of a computing system - performance/power and performance/space.

Frith, C. 2002. Internet Data Centers and the Infrastructure Require Environmental Design, Controls, and Monitoring. *Journal of the IEST* **45**(2002 Annual Edition): 45-52.
Internet Data Centers and the Infrastructure Require Environmental Design, Controls, and Monitoring. The author points out that specifications and standards need to be developed to achieve high performance for mission-critical Internet applications.

Gartner Dataquest. 2002. “Gartner Dataquest Chops Industry's Rapid Growth Expectations for Blade Servers.” Press Release.
http://www4.gartner.com/5_about/press_releases/2002_02/pr20020205d.jsp

Gilleskie, R. J. 2002. The Impact of Power Quality in the Telecommunications Industry. Palm Springs, CA, June 4. http://www.energy2002.ee.doe.gov/Facilities.htm
The workshop addresses the unique issues and special considerations necessary for improving the energy efficiency and reliability of high-tech data centers. This presentation addresses impacts of power quality including voltage sags, harmonics, and high-frequency grounding in telecommunication industry.

Grahame, T., and D. Kathan. 2001. Internet Fuels Shocking Load Requests. *Electrical World* **Vol. 215** (3): 25-27. http://www.platts.com/engineering/ew_back_issues.shtml
This article discusses the implications of the increase for power demand by the Internet's traffic growth on utility planning, operation, and financing.

Greenberg, D. 2001. Addendum to ER-01-15: A Primer on Harmonics. E-SOURCE, Boulder, Colorado, September 2001.
The electrical distribution systems of most commercial and industrial facilities were not designed to operate with an abundance of harmonics-producing loads. In fact, it is only within recent years that such loads have become widespread enough for industry to take notice and to begin to develop strategies to address the problems that harmonics can create. By 1992, concern about the issue had grown sufficiently that the Institute for Electrical and Electronic Engineers (IEEE) developed and published its standard 519, "IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems," which established an approach for setting limits on the harmonic voltage distortion on the utility power system and on the harmonic currents created individual power consumers. Since that time, the electronic loads that give rise to harmonic currents have grown dramatically and are projected to continue growing for the foreseeable future. This being the case, there is and will continue to be a market for technological solutions to the problems that harmonics can cause.

Gross, P. 2002. Needed: New Metrics. *Energy User News*.
http://www.energyusernews.com/eun/cda/articleinformation/features/bnp__features__item/0,2584,82741,00.html

Gruener, J. 2000. Building High-Performance Data Centers. *Dell Magazines - Dell Power Solutions* (Issue 3 "Building Your Internet Data Center").

http://www.dell.com/us/en/esg/topics/power_ps3q00_1_power.htm; http://www.dell.com/us/en/esg/topics/power_ps3q00-giganet.htm
The introduction of Microsoft SQL Server 2000 is a milestone in the race to build the next generation of Internet data centers. These new data centers are made up of tiers of servers, now commonly referred to as server farms, which generally are divided into client services servers (Web servers), application/business logic servers, and data servers supporting multiple instances of databases such as SQL Server 2000.

Hellmann, M. 2002. Consultants Face Difficult New Questions in Evolving Data Center Design. *Energy User News*. http://www.energyusernews.com/CDA/ArticleInformation/features/BNP__Features__Item/0,2584,70610,00.html
While few data center design projects are alike, there are always the twin challenges of "power and fiber." And sometimes, even local politics and human factors. The paper suggested that the consultant should be brought in as soon as a business case is established so criteria can be established and a concept can be developed, priced, and compared to the business case. A planning is necessary before moving on to site selection and refine the concept and again test the business case.

Howe, B., A. Mansoor, and A. Maitra. 2001. Power Quality Guidelines for Energy Efficient Device Application - Guidebook for California Energy Commission (CEC). Final Report to B. Banerjee, California Energy Commission (CEC).
Energy efficiency and conservation are crucial for a balanced energy policy for the Nation in general and the State of California. Widespread adaptation of energy efficient technologies such as energy efficient motors, adjustable speed drives, improved lighting technologies will be the key in achieving self-sufficiency and a balanced energy policy that takes into account both supply side and demand side measures. In order to achieve the full benefit of energy efficient technologies, these must be applied intelligently, and with clear recognition of the impacts some of these technologies may have on power quality and reliability. Any impediment to the application of these energy efficient technologies by the customers is not desirable for the overall benefit to energy users in California. With that in mind EPRI and CEC has worked to develop this guidebook to promote customer adaptation of energy efficient technologies by focusing on three distinct objectives. 1) Minimize any undesirable power quality impacts of energy-saving technologies; 2) Understand the energy savings potential of power quality-related technologies. These include: Surge Protective Devices (SPDs) or Transient Voltage Surge Suppressors (TVSS), Harmonic Filters, Power Factor Correction Capacitors, Electronic Soft Starters for Motors; and 3) How to evaluate "black box" technologies

Intel. 2002. Planning and Building a Data Center - Meeting the e-Business Challenge. Intel Corp. http://www.intel.com/network/idc/doc_library/white_papers/data_center /. Aug 01, 2002.
The paper discusses the keys to success of Internet Service Providers (ISPs) that include 1) Achieve the economies of scale necessary to support a low price business model; 2) Offer added value, typically in the form of specialized services such as applications hosting to justify a premium price. This document provides a high-level overview of the requirements for successfully establishing and operating an Internet data center in today's marketplace. It offers some of the key steps that need to be taken, including project definition, prerequisites and planning. In order to construct a data center that can meet the challenges of the new market, there are three basic areas of data center definition and development: 1) Facilities: including building, security, power, air-conditioning and room for growth; 2) Internet connectivity: performance, availability and scalability; 3) Value-added services and the resources to support their delivery: service levels, technical skills and business processes. The aim is to provide customers with the physical environment, server hardware, network connectivity and technical skills necessary to keep Internet business up and running 24 hours a day, seven days a week. The ability to scale is essential, allowing businesses to upgrade easily by adding bandwidth or server capacity on demand.

Koplin, E. 2000. Finding Holes In The Data Center Envelope. *Engineered Systems* (September 2000). http://www.esmagazine.com/CDA/ArticleInformation/features/BNP__Features__Item/0,2503,8720,00.html

The paper addresses importance of environmental control in data center facilities. Maintaining data center availability requires absolutely reliable infrastructure. A significant amount of this is devoted solely to maintaining stable environmental parameters. And only constant, thorough regulation and testing of these parameters ensures the integrity of the data center “envelope.”

Mandel, S. 2001. Rooms that consume - Internet hotels and other data centers inhale electricity. *Electric Perspectives* **Vol. 26** (No.3). http://www.eei.org/ep/editorial/Apr_01/0401ROOM.htm

The article estimated that the amount of this data center space in the United States nearly doubled in 2000, totaling between 19 million and 25 million square feet by year-end, according to investment analysts. They say they expect another 10 million to 20 million square feet of new space to be added in 2001. Developers are asking electric utilities to supply the buildings with 100-200 watts of electricity per square foot. Since these data centers are new to the economy, there is little historical data on which to base estimates of electricity use for a facility. In addition, the dot.com world makes it difficult for the developer to say confidently how much electricity one of these Internet hotels will use. Source One estimates that tens of billions of dollars worth of electric infrastructure improvements will be needed for data centers over the next few years and that they will consume billions of dollars more worth of electricity. The energy costs are as high or higher than the actual lease costs. Indeed, 50-60 percent of the cost of building a data center is for the power, including batteries, backup generators, and air-conditioning, as well as the cost for utility construction.

Mitchell-Jackson, J. 2001. Energy Needs in an Internet Economy: A Closer Look at Data Centers, July 2001. http://enduse.lbl.gov/projects/infotech.html

This study explains why most estimates of power used by data centers are significantly too high, and gives measured power use data for five such facilities. Total power use for the computer room area of these data centers is no more than 40 W/square foot, including all auxiliary power use and cooling energy. There are two draft journal articles from this work, one focusing on the detailed power use of the data center we've examined in most detail, and the other presenting the aggregate electricity use associated with hosting-type data centers in the U.S.

Mitchell-Jackson, J., J. G. Koomey, B. Nordman, and M. Blazek. 2001. Data Center Power Requirements: Measurements From Silicon Valley. *Energy—the International Journal (Under review)*. http://enduse.lbl.gov/Projects/InfoTech.html

Current estimates of data center power requirements are greatly overstated because they are based on criteria that incorporate oversized, redundant systems, and several safety factors. Furthermore, most estimates assume that data centers are filled to capacity. For the most part, these numbers are unsubstantiated. Although there are many estimates of the amount of electricity consumed by data centers, until this study, there were no publicly available measurements of power use. This paper examines some of the reasons why power requirements at data centers are overstated and adds actual measurements and the analysis of real-world data to the debate over how much energy these facilities use.

Nordham, Reiss, and Stein. 2001. Delivering Energy Services to Internet Hotels and Other High Density Electronic Loads, Part I: Structure of the HiDEL Industry. Platts Research and Consulting, Boulder, CO.

Patel, C. D., C. E. Bash, C. Belady, L. Stahl, and D. Sullivan. 2001. Computational Fluid Dynamics Modeling of High Compute Density Data Centers to Assure System Inlet Air Specifications. Reprinted from the proceedings of the Pacific Rim ASME International

Electronic Packaging Technical Conference and Exhibition (IPACK 2001), © 2001, ASME.
Due to high heat loads, designing the air conditioning system in a data center using simple energy balance is no longer adequate. Data center design cannot rely on intuitive design of air distribution. It is necessary to model the airflow and temperature distribution in a data center. This paper presents a computational fluid dynamics model of a prototype data center to make the case for such modeling.

Patel, C. D., R. Sharma, C. E. Bash, and A. Beitelmal. 2002. Thermal Considerations in Cooling Large Scale High Compute Density Data Centers. 8th ITHERM Conference. San Diego CA.
A high compute density data center of today is characterized as one consisting of thousands of racks each with multiple computing units. The computing units include multiple microprocessors, each dissipating approximately 250 W of power. The heat dissipation from a rack containing such computing units exceeds 10 KW. Today's data center, with 1000 racks, over 30,000 square feet, requires 10 MW of power for the computing infrastructure. A 100,000 square foot data center of tomorrow will require 50 MW of power for the computing infrastructure. Energy required to dissipate this heat will be an additional 20 MW. A hundred thousand square foot planetary scale data center, with five thousand 10 KW racks, would cost ~$44 million per year (@ $100/MWh) just to power the servers & $18 million per year to power the cooling infrastructure for the data center. Cooling design considerations by virtue of proper layout of racks can yield substantial savings in energy. This paper shows an overview of a data center cooling design and presents the results of a case study where layout change was made by virtue of numerical modeling to avail efficient use of air conditioning resources.

PG&E. 2001. Data Center Energy Characterization Study. Pacific Gas and Electric Company (subcontractor: Rumsey Engineers), San Francisco, Feb. 2001.
Rumsey Engineers, Inc. and PG&E have teamed up to conduct an energy study as part of PG&E's Data Center Energy Characterization Study. This study will allow PG&E and designers to make better decisions about the design and construction of data centers in the near future. Three data centers in the PG&E service territory have been analyzed during December 2000 and January 2001, with the particular aim of determining the end-use of electricity. The electricity use at each facility was monitored for a week each. At the end of the report are a set of definitions, which explain the terms used and the components in making each calculation. The three data centers provide co-location service, which is an unmanaged service that provides rack space and network connectivity via a high capacity backbone. About half or more of the electricity goes to powering the data center floor, and 25 to 34 percent of the electricity goes to the heating, air conditioning and ventilation equipment. The HVAC equipment uses a significant amount of power and is where energy efficiency improvements can be made. All three facilities use computer room air conditioning (CRAC) units, which are stand-alone units that create their own refrigeration and circulate air. A central, water-cooled chilled water system with air handlers and economizers can provide similar services with roughly a 50% reduction in cooling energy consumption. Energy density of the three buildings had an average of 35 W/sf. The cooling equipment energy density for the data center floor alone averaged at 17 W/sf for the three facilities. The average designed energy density of the three data centers' server loads was 63 W/sf, while the measured energy density was 34 W/sf. An extrapolated value was also calculated to determine what the server load energy density would be when fully occupied. The average extrapolated energy density was 45 W/sf. Air movement efficiency varies from 23 to 64 percent between the three facilities. Cooling load density varies from 9 to 70 percent between the three facilities.

Planet-TECH. 2002. Technical and Market Assessment for Premium Power in Haverhill. Planet-TECH Associates for The Massachusetts Technology Collaborative, www.mtpc.org , Westborough, MA 01581-3340, Revision: February 20, 2002. http://www.mtpc.org/cluster/Haverhill_Report.pdf; http://www.planet-tech.com/content.htm?cid=2445
This study is pursued under contract to the Massachusetts Technology Collaborative, in response to a request for a "Technical and Market Assessment". It seeks to determine if the provisioning of "premium

power" suitable for data-intensive industries will improve the marketability of a Historic District mill building in Haverhill. It is concluded that such provisioning does improve the marketability, however, not to a degree that is viable at this time. Other avenues for energy innovation are considered and recommendations for next steps are made.

RLX Technologies. 2001. *Redefining Server Economics*. White Paper. http://www.rlx.com/pdf/RLXServerEconWP_v1.0.pdf

RMI, and DR International. 2002. Energy Efficient Data Centers - A Rocky Mountain Institute Design Charrette. Organized, Hosted and Facilitated by Rocky Mountain Institute, with D&R International, Ltd. and Friends. Hayes Mansion Conference Center, San Jose, California. http://www.rmi.org/sitepages/pid626.php
Rapid growth of "mission critical" server-farm and fiber-optic-node data centers has presented energy service providers with urgent issues. Resulting costs have broad financial and societal implications. While recent economic trends have severely curtailed projected growth, the underlying business remains vital. The current slowdown allows us all some breathing room—an excellent opportunity to step back and carefully evaluate designs in preparation for surviving the slowdown and for the resumption of explosive growth. Future data center development will not occur in the first-to-market, damn-the-cost environment of 1999-2000. Rather, the business will be more cost-competitive, and designs that can deliver major savings in both capital cost (correct sizing) and operating cost (high efficiency)—for both new build and retrofit—will provide their owners and operators with an essential competitive advantage.

Robertson, C., and J. Romm. 2002. Data Centers, Power, and Pollution Prevention - Design for Business and Environmental Advantage. The Center for Energy and Climate Solutions; A Division of The Global Environment and Technology Foundation, June 2002. http://www.cool-companies.org; http://www.getf.org
Computers and other electronic equipment will crash at the slightest disruption or fluctuation in their supply of electricity. The power system was not designed for these sensitive electronic loads and is inherently unable to meet the technical requirements of the information economy. For data centers, which play a central role in the information economy, crashing computers cause potentially catastrophic financial losses. The same voltage sag that causes the lights to dim briefly can cause a data center to go off-line, losing large sums of money, for many hours. Data center owners and their power providers must therefore solve several related technical and economic electric power problems. These are: 1) How to assure high-availability (24x7) power supply with a very low probability of failure; 2) How to assure practically perfect power quality; and 3) How to manage risk while minimizing capital and operating expenses

Roth, K. W., Fred Goldstein, and J. Kleinman. 2002. Energy consumption by office and telecommunications equipment in commercial buildings, Volume I: Energy Consumption Baseline. Arthur D. Little (ADL), Inc., 72895-00, Cambridge, MA, January 2002.
ADL carried out a "bottom-up" study to quantify the annual electricity consumption (AEC) of more than thirty (30) types of non-residential office and telecommunications equipment. A preliminary AEC estimate for all equipment types identified eight key equipment categories that received significantly more detailed studied and accounted for almost 90% of the total preliminary AEC. The Key Equipment Categories include: *Computer Monitors and Displays, Personal Computers, Server Computers, Copy Machines, Computer Network Equipment, Telephone Network Equipment, Printers, Uninterruptible Power Supplies (UPSs).* The literature review did not uncover any prior comprehensive studies of telephone network electricity consumption or uninterruptible power supply (UPS) electricity consumption. The AEC analyses found that the office and telecommunications equipment consumed 97-TWh of electricity in 2000. The report concludes that commercial sector office equipment electricity use in the U.S. is about 3% of all electric power use. The ADL work also creates scenarios of future electricity use for office equipment, including the energy used by telecommunications equipment.

Shields, H. and C. Weschler, 1998. Are Indoor Pollutants Threatening the Reliability of Your Electronic Equipment? Heating/Piping/Air Conditioning Magazine. May.

Stein, Jay. 2002. More Efficient Technology Will Ease the Way for Future Data Centers. Proceedings 2002 ACEEE Summer Study on Energy Efficiency in Buildings.

Sullivan, R. F. 2002. Alternating Cold and Hot Aisles Provides More Reliable Cooling for Server Farms. The Uptime Institute. http://www.uptimeinstitute.org/tuiaisles.html
The creation of "server farms" comprising hundreds of individual file servers has become quite commonplace in the new e-commerce economy, while other businesses spawn farms by moving equipment previously in closets or under desktops into a centralized data center environment. However, many of these farms are hastily planned and implemented, as the needed equipment must be quickly installed on a rush schedule. The typical result is a somewhat haphazard layout on the raised floor that can have disastrous consequences due to environmental temperature disparities. Unfortunately, this lack of floor-layout planning is not apparent until after serious reliability problems have already occurred.

The Uptime Institute. 2000. Heat-Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment. The Uptime Institute, Version 1.0. http://www.upsite.com /. http://www.uptimeinstitute.org/heatdensity.html
This white paper provides data and best available insights regarding historical and projected trends in power consumption and the resulting heat dissipation in computer and data processing systems (servers and workstations), storage systems (DASD and tape), and central office-type telecommunications equipment. The topics address the special needs of Information Technology professionals, technology space and data center owners, facilities planners, architects, and engineers.

Thompson, C. S. 2002. Integrated Data Center Design in the New Millennium. *Energy User News*. http://www.energyusernews.com/CDA/ArticleInformation/features/BNP__Features__Item/0,2584,70578,00.html
Data center design requires planning ahead and estimating future electrical needs. Designers must accurately predict space and energy requirements, plus cooling needs for new generations of equipment. Importance of data center reliability is discussed.

Uptime Institute, 2000. Heat Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment. Santa Fe, NM. http://www.upsite.com/TUIpages/whitepapers/tuiheat1.0.html

Wood, L. 2002. Cutting Edge Server Farms - The blade server debate. newarchitectmag.com. *http://www.newarchitectmag.com/documents/s=2412/na0702f/index.html*. July 23, 2002. A blade is the industry term for a server that fits on a single circuit board, including CPU, memory, and perhaps a local hard disk. Multiple blades are plugged into a chassis, where each blade shares a common power supply, cooling system, and communications back plane. Multiple chassis can then be stacked into racks. By comparison, the conventional approach for rack-mounted servers involves only one server per chassis. A chassis cannot be smaller than one vertical rack unit (1U, or about 1.75 inches high). This limits you to 42 to 48 servers in a standard seven-foot rack. A typical blade chassis is much higher than 1U, but several can still be stacked in a rack, allowing upwards of 300 servers per rack, depending on the vendor and configuration. This compact design offers compelling advantages to anyone operating a high-density server farm where space is at a premium.

Indeed, blades are the "next big thing" in servers, and it's probable that any given administrator will have to decide whether to adopt them in the near future.