Smarter Data Centers
Achieving Greater Efficiency

Improve energy efficiency and reduce costs

Minimize stranded space, power, and cooling

Monitor, manage, and report across both facilities and IT

Mike Ebbers
Matthew Archibald
Carlos Felipe França da Fonseca
Marc Griffel
Veerendra Para
Michael Searcy

Redpaper

ibm.com/redbooks

International Technical Support Organization

Smarter Data Centers: Achieving Greater Efficiency

October 2011

REDP-4413-01

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

Second Edition (October 2011)

This document created or updated on October 21, 2011.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
BladeCenter®
DPI®
EnergyScale™
FICON®
HiperSockets™
IBM Project Financing™
IBM Systems Director Active Energy Manager™
IBM®
iDataPlex™
Maximo®
Power Systems™
POWER6®
POWER7™
POWER7 Systems™
POWER®
PR/SM™
Processor Resource/Systems Manager™
ProtecTIER®
Redbooks®
Redguide™
Redpaper™
Redbooks (logo) ®
Resource Link™
Smarter Planet™
System i®
System x®
System z®
Tivoli®
X-Architecture®
XIV®
z/OS®
z/VM®
z/VSE™

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

As we move towards becoming a smarter planet and the world becomes more instrumented, interconnected, and intelligent, the demands for data center resources are increasing rapidly. Smaller and more densely packed servers providing greater amounts of computing power can substantially increase power and cooling needs, while growing data volumes necessitate larger storage and network bandwidth capacities. Environmental and regulatory requirements can introduce additional limits on carbon emissions and water consumption.

To satisfy these demands while keeping costs in check, our data centers need to be smarter as well. Comprehensive views of data center inventories, operational and environmental conditions, and consumption across multiple capacity types that span both facilities and IT are required. You can achieve greater efficiency using hardware, software, services, and design both in facilities and IT, but you need a comprehensive data center strategy to tie them together and thus obtain a complete picture of your data center environments.

This IBM® Redpaper™ publication discusses important considerations when creating and implementing your smarter data center strategy. Notable techniques, best practices, and technological advances that can become critical components of success are included, along with methods for bringing them together to gain in-depth knowledge of data center operations. With such insight comes increased resiliency, rapid responsiveness, profitable access to detailed analytics, and reliable planning for the future. Although not all-inclusive, this document provides a guide to getting started, points you to additional sources of information, and suggests ways IBM can partner with you in your pursuit of a smarter data center.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO) Raleigh Center.

Mike Ebbers is a Consulting IT Specialist and ITSO Project Leader who has worked for IBM for 37 years and has published a number of books on various IT solutions. Before joining the ITSO 16 years ago, Mike worked as an Instructor and Developer for IBM U.S. Technical Education.

Matthew Archibald is a Data Center Engineer in IBM Lab Services Data Center Services based out of Research Triangle Park, North Carolina, and Melbourne, Australia. He has been with IBM for eight years and has been working for the last five years in the Lab Services organization. Previous to this, Matt worked as a development engineer in the IBM System x® and IBM BladeCenter® power development lab doing power subsystem design for BladeCenter and System x servers. Matt is also responsible for the development, maintenance, and support of the IBM System x Power Configurator program and holds 15 patents for various data center and hypervisor technologies. Matt has four degrees from Clarkson University in Computer Engineering, Electrical Engineering, Software Engineering, and Computer Science and a Bachelor of Engineering in Electronics Engineering from Auckland University of Technology.

Carlos Felipe França da Fonseca joined IBM Brazil in 2005. In 2010 he moved to IBM Australia to work exclusively with Tivoli® products. He graduated with a degree in Information Technology and has a postgraduate degree in Management of Information Systems. He has over 11 years experience in IT and holds certifications on a number of IBM products and solutions. His areas of expertise also include IBM System x and IBM Systems Director. He co-authored the recent IBM Redbooks® publication *Tivoli Integration Scenarios*, SG24-7878-01.

Marc Griffel is a certified IT Specialist for Virtualization and System Management in Germany and has worked for IBM for almost 20 years. He was actively involved in the transition of the IBM Chip and Design Data Center in Boeblingen, Germany toward becoming a state-of-the-art, energy-efficient data center. He works closely with the IBM Briefing Center, where he is involved in customer briefings providing direction on how to operate more energy-efficient data centers. His areas of expertise cover storage and IBM Power systems virtualization, BladeCenter, SAN and storage management.

Veerendra Para is an IT Specialist for Green IT in IBM India Software Labs (ISL) in Bangalore. He has 10 years experience in IT and has worked for the last three years greening data centers. His primary role is to transform ISL traditional server labs into energy-efficient smarter data centers by introducing innovative green solutions, creating Green DC architectures, and sharing success stories with customers. In the past, he has worked in the systems management field, mainly on IBM System i® and AIX®, and has co-authored three IBM Redbooks publications on System i. He has worked for IBM Global Technology Services and currently works with IBM Software Group India. He holds a diploma in electronics and communications.

Michael Searcy is a Systems Engineer with IBM Global Technology Services in the United States. He is currently assisting with the implementation of the strategy to make IBM data centers smarter with a personal focus on leveraging Tivoli technologies to provide comprehensive data center views that span both facilities and IT. A recent product of this strategy has been the IBM Leadership Data Center located at Research Triangle Park, North Carolina, in the United States. During his 17 years in the IT industry, his experience has spanned all aspects of the delivery process, including leading a team in international test and development projects and in business process development. He holds an honors degree in Electrical Engineering from Auburn University.

Thanks to the following individuals for their invaluable contributions to this project:

Andreas Bieswanger, IBM Germany
Jody Cefola, IBM Poughkeepsie
Steve Cole, IBM Raleigh
Rani Doughty, IBM Australia
Ken Gambon, IBM Boulder
Hendrik Hamann, IBM Yorktown Heights
Mike Hogan, IBM Raleigh
John Miller, IBM Fishkill
Wayne Riley, IBM Beaverton
Jean-Michel Rodriguez, IBM France
Roger Schmidt, IBM Poughkeepsie
Thomas Vogel, IBM Germany
Hans-Dieter Wehle, IBM Germany
Ralf Wohlfarth, IBM Germany

The following authors created the first edition of this Redpaper:

David F Anderson PE, IBM Poughkeepsie
Alvin Galea, Australia Post Corporation
Don Roy, IBM Raleigh
Michael Schaefer, IBM Germany
Marc Tu Duy Khiem, IBM France

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  `https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm`

- Stay current on recent Redbooks publications with RSS Feeds:

  `http://www.redbooks.ibm.com/rss.html`

1

The benefits of a smarter data center

Many data centers are running out of power and space, making it difficult or impossible to add new equipment. Perhaps they have reached the limit of their cooling capacity.

Environmental concerns introduce additional priorities because they can impede a company's ability to grow. Not only must the company ensure it complies with governmental regulations, but customers all over the world also prefer to purchase products and services from companies that have a sustainable attitude.

As sustainability becomes a business imperative, many data center clients are looking at ways to save energy and cut costs so that the company can continue to grow. However, it has become necessary to consider in-depth transformation that extends well beyond simply cutting costs.

Data centers need to provide flexibility to respond quickly to future unknowns in business requirements, technology, and computing models. The data centers must adapt to be more cost-effective, considering both capital and operating costs. Furthermore, data centers require active monitoring and management capabilities to provide operational insights to meet necessary levels of availability, resiliency, capacity planning, and energy efficiency.

In this chapter, we discuss how running a *smarter data center* can help clients move toward a more sustainable business model.

1.1 Overview

The journey toward a smarter data center brings cost and sustainability rewards. Energy consumption and environmental concerns are becoming priorities because they can impede the ability to grow and respond to organizational IT needs.

In September 2007, IBM and the Economist Intelligence Unit issued a report entitled *IT and the environment: A new item on the CIO agenda*[1] *?* This report found that although most organizations said they were *green* organizations, many of them were not actually doing as much as they could. Two-thirds of the two hundred or more executives polled said that their organizations had a board-level executive responsible for energy and the environment; however, only 45% of firms had a program in place to reduce their carbon footprint.[2] In addition, of those that did have a carbon reduction strategy, a majority (52%) had no specific targets for it, although a dedicated few (9%) aimed to be carbon-neutral by 2012.

Almost four years later, in 2011, the Institute for Building and Efficiency issued a report entitled *Energy Efficiency Indicator - 2011 Global Results*[3]. The report found that 76% of the 3,868 decision-makers polled said that their companies have either an energy or carbon reduction goal. However, 33% of respondents said that the lack of funding to pay for improvements was the greatest obstacle in the way to pursuing energy efficiency at their organizations, despite trends charted in Figure 1-1, which shows that, although new server spending has plateaued, management and administration costs, along with power and cooling costs, continue to climb.

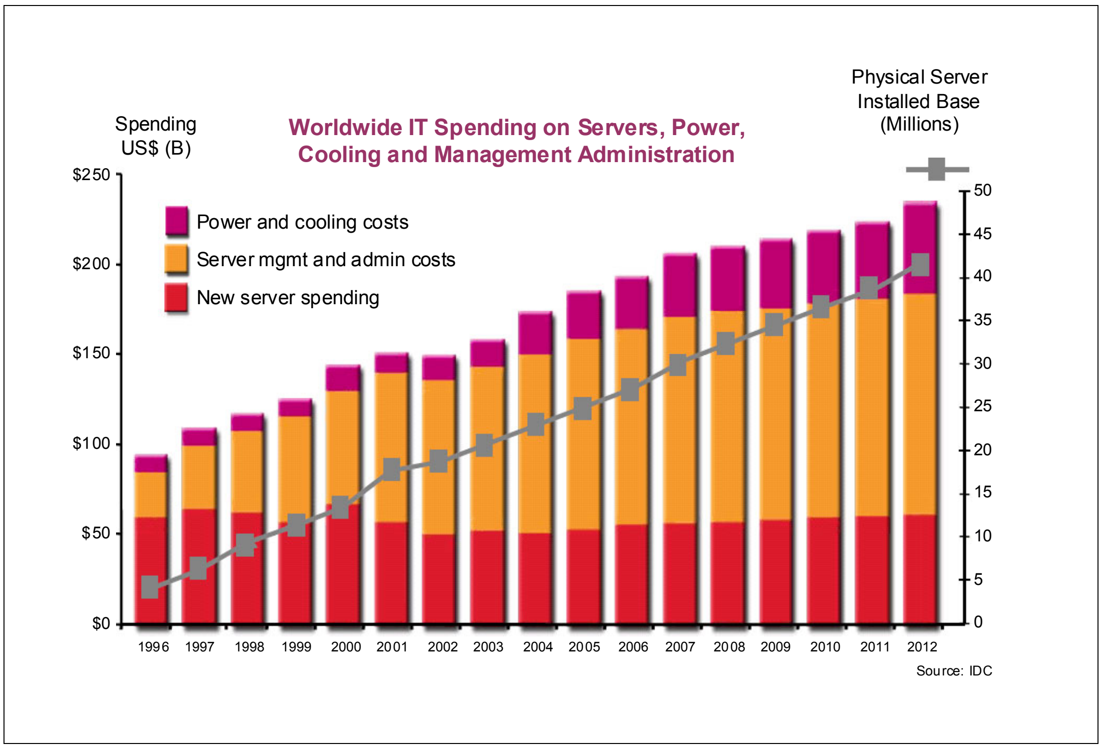


Figure 1-1 Worldwide IT spending on servers, power, cooling and management administration

1 http://www-05.ibm.com/no/ibm/environment/pdf/grennit_oktober2007.pdf
2 The amount of carbon dioxide emitted. See http://www.carbonfootprint.com/
3 "Energy Efficiency Indicator - 2011 Global Results", Institute for Building and Efficiency at: http://www.institutebe.com/Energy-Efficiency-Indicator/2011-global-results.aspx

1.1.1 Managing the increasing cost of energy

The cost of a kilowatt of electricity has risen only slightly in recent years, but the cost of operating servers has increased significantly. The context for this paradox is that server energy consumption is increasing exponentially faster than the associated utility cost. Rising demand has accelerated the adoption of virtualization technologies and increased virtual image densities, driving total server energy consumption higher, while the amortized cost of operation per workload has been decreasing.

1.1.2 Running out of power capacity

Certain companies cannot deploy more servers because additional electrical power is not available. Many utilities, especially those in crowded urban areas, are telling customers that power feeds are at capacity and that they simply have no more power to sell. New server, storage, and networking products give better performance at lower prices but can also be power hungry. The effort to overcome a power supply threshold is a huge investment.

1.1.3 Running out of cooling capacity

Many customer data centers are now 10 to 20 years old, and their cooling facilities are not adapted to present needs. Traditional cooling methods allowed for 2-3 kW of cooling per rack. Today's requirements can reach 20-30 kW per rack, and heat density can easily exceed the design point of the data center.

1.1.4 Running out of physical capacity

Each time a new project or application comes online, new images, servers or storage subsystems are added, resulting in rapid occupation of data center physical space. When images, servers, and storage cannot be added, except by building another data center, this growth can become inordinately expensive.

1.2 Data center energy usage

To understand how to reduce energy consumption, we need to understand where and how energy is used. We can study energy use in a data center by taking three separate views:

- How energy is distributed between IT equipment (servers, storage, network equipment) and supporting facilities (power, cooling, and lighting).
- How energy is distributed between the separate components of the IT equipment (processor, memory, disk, and so forth).
- How the energy allocated to IT resources is really used to produce business results (Are idle resources powered on, using energy without productive results?)

Figure 1-2 shows how energy is used in several components of a typical non-optimized data center. Each component is divided into two portions:

- IT equipment (servers, storage, and network) uses about 45% of a data center's energy.
- The infrastructure that supports this equipment—such as chillers, humidifiers, computer room air conditioners (CRAC), power distribution units (PDU), uninterruptible power supplies (UPS), lights, and power distribution—uses the other 55% of the energy.

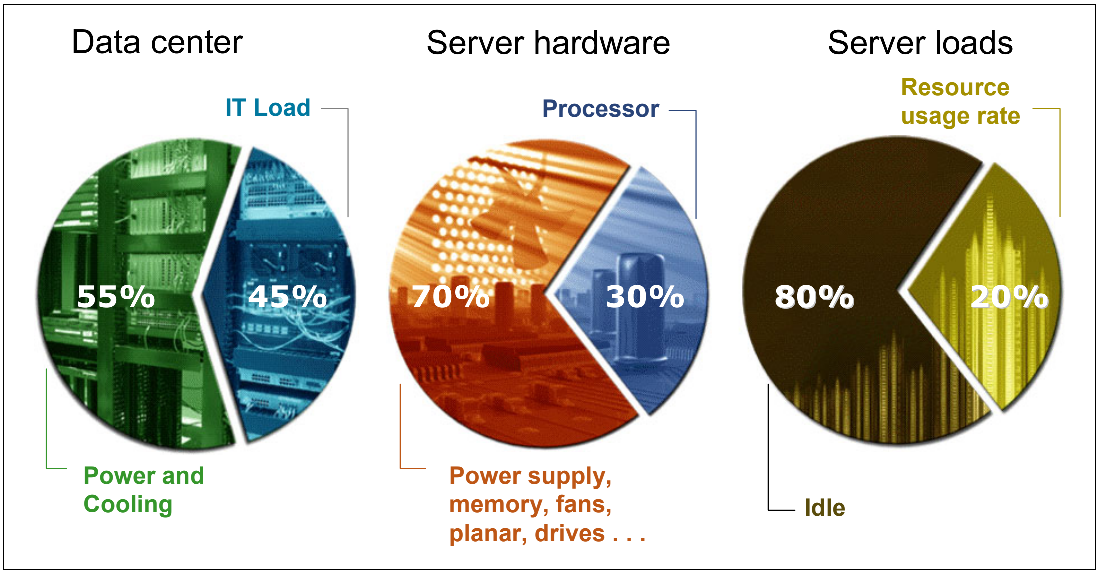


Figure 1-2 How energy is used in a typical data center

Because IT equipment is not using 55% of the energy that is brought into the data center, this portion of the energy is not producing work of value in the form of calculations, data storage, and so on. Fortunately, improving this ratio between infrastructure and IT power consumption is possible. We discuss energy savings and efficiencies for your data center infrastructure in Chapter 4, "Optimizing your site and facilities" on page 57.

Companies must also consider the energy consumption of their IT equipment. In a typical server, the processor uses only 30% of the energy, while the remainder of the system uses the other 70%. Therefore, efficient hardware design is also important. We discuss IT hardware energy savings in Chapter 3, "Optimizing your IT equipment" on page 25.

Although better hardware design can reduce energy consumption, companies must also take into account the utilization of their IT equipment. The utilization rate of certain types of servers can typically be only 20%. Underutilized systems expend a lot of energy on non-business purposes, thus wasting a major investment. Virtualization and consolidation can help utilize the entire capacity of your IT equipment.

IBM POWER® systems and IBM System z® offer integrated virtualization capabilities that are often more optimized than other servers. By using your IT equipment at its full capacity with consolidation and virtualization, you can achieve energy efficiency and maximize the return on investment (ROI). We discuss this in section 3.6, "Consolidation and virtualization" on page 43.

Although large efficiency gains can be made through direct infrastructure and IT improvements, the strategy for a truly smarter data center must span both infrastructure and IT, perceiving the data center as a single, cohesive whole. This perspective can facilitate efficiency gains across the entire facility, enabling assets and workloads to be deployed at their most efficient locations, increasing data center availability and resiliency, and providing a solid foundation for future planning. IBM Tivoli software offers smarter solutions that can help organizations save money by reducing the energy consumed by their data centers and beyond to their general facilities and buildings.

These solutions can increase energy-efficiency on an end-to-end basis and make better use of existing server assets, lower CO_2 emissions and reduce space, power, and cooling costs. We discuss this integrated view in Chapter 5, “Integrating across IT and facilities for the bigger picture” on page 91

1.3 Environmental laws and the company image

Earth’s resources are limited. Oil and coal supplies are finite. Fossil fuel use has been linked to CO_2 increases that can impact climate. The air, water, and mineral resources of the Earth require good stewardship for future generations. Information technology currently accounts for 2% of CO_2 production. However, the global demand for electricity from data centers is expected to triple or quadruple from 2007 levels by 2020.[4]

Today, climate change is a major environmental and political issue. Environmental laws have affected the IT industry. A U.S. a law created in September 2007 authorized the Environmental Protection Agency (EPA) to analyze the growth of energy consumption in data centers. Also, the European Union (EU) has established a directive that will cut its CO_2 emissions by at least 20% of 1990 levels by 2020. The Australian government’s long-term target for carbon pollution reduction is 80% below 2000 levels by 2050.

Beyond respecting environmental laws, being smarter also enables a company to fulfill its objectives of social responsibility as a *green company*. Being green improves the company’s image for its customers and employees as a good company with which to do business and for which to work.[5]

[4] “How dirty is your data? A Look at the Energy Choices That Power Cloud Computing”, Greenpeace International, April 2011

[5] For more information, see the information provided by the U.S. EPA:
`http://www.epa.gov/climatechange`

1.4 The four dimensions of environmental sustainability

There is an opportunity for *smarter organizations* to unlock the value in *green*. Smart systems are emerging with the potential to improve our lives and profoundly impact the environmental challenges facing society and the planet. As illustrated in Figure 1-3, we believe there are four interlocking dimensions of environmental sustainability:

- **Environmental strategies**: A comprehensive "green" strategy reflects many different views of the business and the varied dimensions of resource consumption. It requires a clear understanding of corporate objectives and priorities, plus programs, initiatives and tasks that address environmental and stakeholder concerns, and drive environmental stewardship.
- **Green branding**: Many companies are developing a green strategy to enhance their competitive position, protect and enrich their brand, improve customer communications, and create a "greener" customer experience. In addition to effectively branding products and services to customers, the enterprise must also brand itself as an environmentally sustainable employer.
- **Compliance management**: Government regulations and compliance standards for improving air quality are now top-level business requirements, compelling companies to apply financial and reputation incentives to drive down CO_2 emissions, abide by water and waste management standards, and implement new processes and technologies.
- **Cost-efficient sustainability**: A cost-efficient sustainability strategy balances trade-offs to optimize efficiencies. You should approach and analyze environmental concerns from a holistic viewpoint, while evaluating overall performance goals of cost, service and quality in terms of their interlocking relationships.

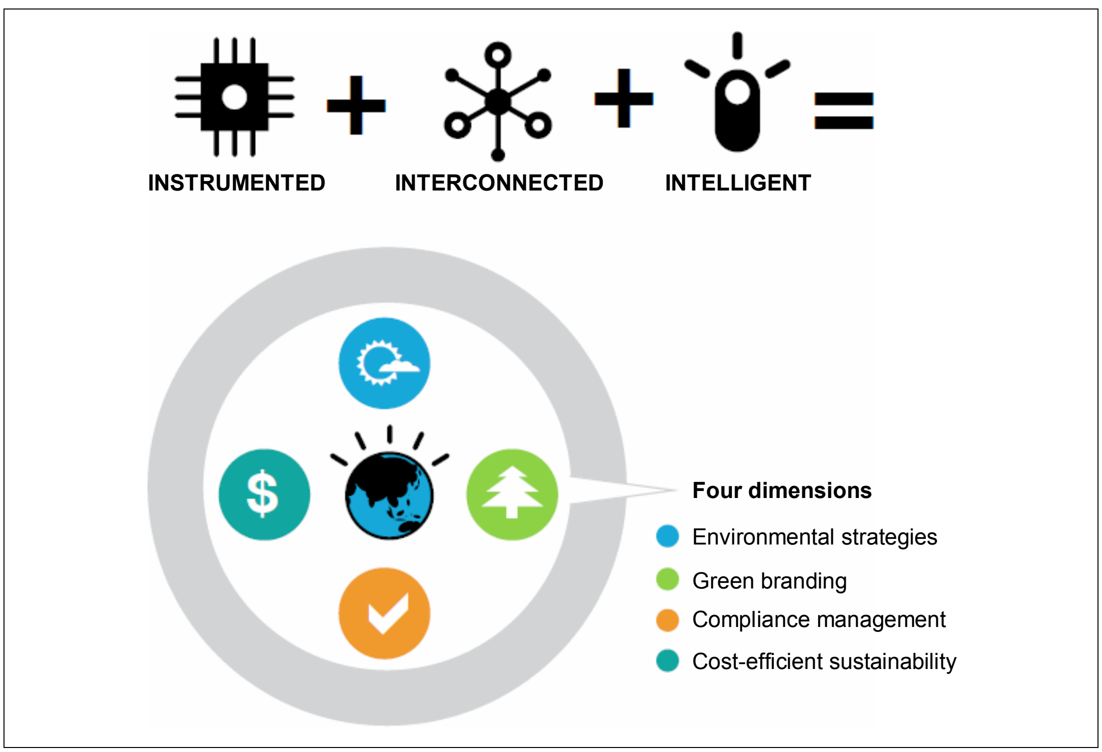


Figure 1-3 An opportunity for smarter organizations to find the value in green

Green awareness can become smart action. Using *smarter technologies* to extend capabilities while reducing costs and risks is more than smart business. Planning for energy efficiency and corporate responsibility with positive results go together. The global economy has made the world more competitive and cooperative. As organizations seek to thrive in a constantly changing world, adapting for sustainability is a business imperative, and smarter data centers can be an integral part of your corporate strategy. Figure 1-4 outlines several of the benefits of a smarter data center.

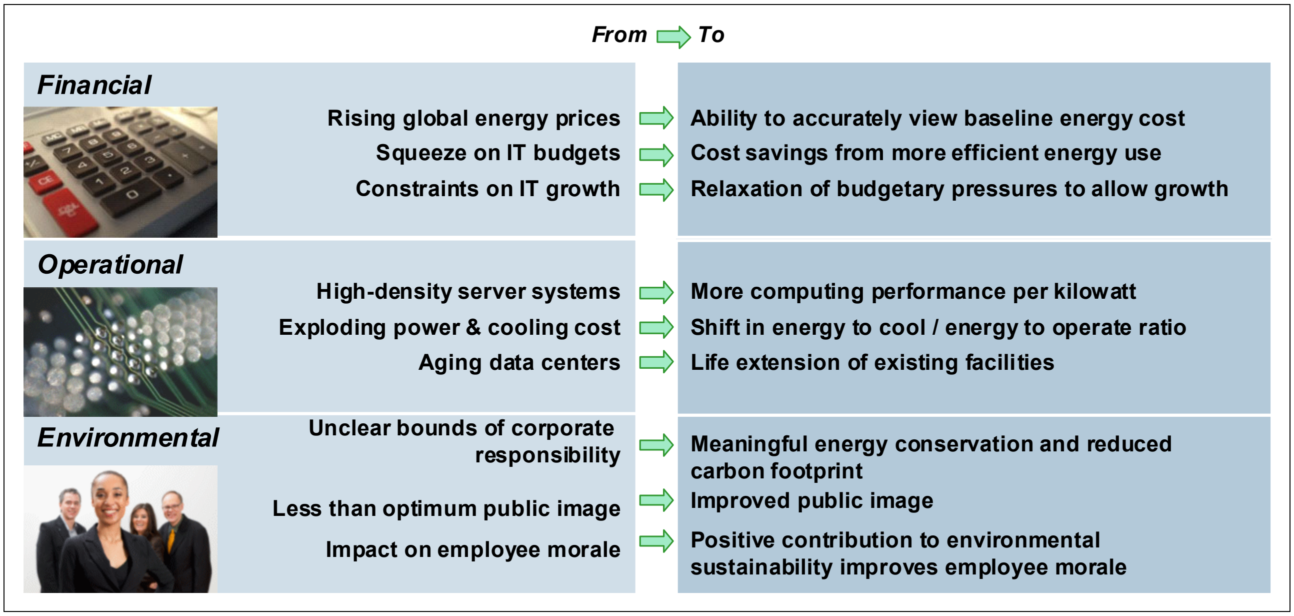


Figure 1-4 The benefits of getting smarter

1.5 Next steps

The remainder of this paper can help you as you develop and implement your smarter data center strategy. In the next chapter, we present several considerations and make several recommendations about developing your strategy.

In Chapter 3, “Optimizing your IT equipment” on page 25 and Chapter 4, “Optimizing your site and facilities” on page 57, we discuss IT systems and data center facilities with recommendations for energy efficiency and optimization.

In Chapter 5, “Integrating across IT and facilities for the bigger picture” on page 91, we discuss how to gather data from across both facilities and IT and take advantage of that data to gain a holistic view of your data center environment

In Chapter 6, “How we can help: IBM services and IBM Business Partners” on page 101, we list ways that IBM and IBM Business Partners can help you. Finally, in Appendix A, “IBM smarter data center: an evolution of energy-efficient operation” on page 109, we list examples of how IBM is committed to *getting smarter.*

> **Note:** Periodically in this document you will see *IBM Redguide™ Case Study* references. For additional information about these and additional case studies detailing how IBM has implemented energy efficiency technologies within our own facilities and for our customers, see the companion Redguide, *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523.

2

Developing your strategy

"As the world becomes more instrumented and interconnected, businesses and institutions need smarter IT infrastructures to handle dramatic increases in data volumes, and to improve data center efficiency."[1]

- Samuel J. Palmisano, IBM Chairman, President and Chief Executive Officer

The data center is shifting.

The trends are unmistakable. As global populations climb, growth markets expand, and we digitally connect all aspects of our world, the demand for data processing continues to skyrocket. The world is increasingly populated by millions of interconnected devices, all generating data that must be gathered, maintained, and analyzed.

Concurrently, technological advances enable more work to be done in a smaller physical footprint. As more servers crowd the same physical environment, capacity is no longer dictated simply by spatial availability. Power, cooling, networking, storage, and other capacity metrics must be managed as well. Environmental concerns and associated governmental regulations introduce a focus on carbon emissions. All of these trends converge in a single location: *the data center*.

Data centers must become immensely more efficient to meet the need while keeping costs in check as the demand for and price of resources continue to rise. However, the realization of this efficiency requires a deep and pervasive transformation in how data centers are designed, managed, operated, populated, and billed, mandating a unified and coordinated effort across organizational and functional boundaries toward a common set of goals.

In this chapter, we discuss several important considerations when planning your smarter data center strategy. Subsequent chapters discuss many of the technological capabilities and best practices that can be applied to both the facility and IT components of your data centers. We then focus on methods for tying views on facility and IT operations together to obtain a holistic view of current conditions while providing effective means to plan for the future.

[1] 2010 IBM Annual Report: http://www.ibm.com/annualreport/2010/

2.1 Recognizing the need and the opportunity

We are at a crossroads of necessity and opportunity. Necessity is driven by rising demand coupled with increased data center management, operational, and construction costs. Opportunities arise in the potential for server consolidation, virtualization, cloud computing, monitoring and management, and even carbon credit trading.

Addressing these needs and exploiting these opportunities will require pervasive change in how our data centers are managed, and insight into improvement opportunities can be found as close as the corner grocery store.

2.1.1 Grocery shopping for data center efficiency

Visiting a well-managed grocery store, you might notice several aspects about the operation of its business.

Each store manages thousands of items on its shelves every day with new stock being delivered and existing stock leaving at the store's points-of-sale. Each store has designated zones hosting products with differing purposes and differing environmental requirements. The store manages stock in three dimensions, knowing what product goes on what aisle, at what location on that aisle, and on what shelf at that location. The store also knows how much room each product occupies. This product placement is often dictated by contractual agreements with suppliers, and prime, eye-level shelf space commands an associated premium.

Shelving units are standardized, based on the nature of the products to be contained on them, and certain units require differing amounts of localized cooling to keep products cool, cold, or frozen.

With frequent sales, each store shifts its stock to end-caps and sales displays. The movement of products within the store or between stores can be driven by time-of-year, weather indicators, excess inventory, or competitive pricing. Each store maintains consistency with product placement, ensuring customers can locate their favorite products easily. At the same time, the store places staples at opposite ends or at the back of the store to encourage spontaneous purchases, balancing consumer and business interests. Also, each store manages thousands of items that expire regularly.

The grocer gathers and consolidates immense amounts of data, performing detailed analytics spanning both facility operations and consumer habits. Management examines store-level operational costs and capacity utilization along with shopping and purchasing patterns, continually seeking improvements in efficiency, profitability, store design, and the consumer experience. Such analytics also provide insight for future planning regarding what inventories to increase, which products to discontinue, and where to build future locations.

Many of the following techniques translate into data center operations as well:

- Three-dimensional inventory tracking that includes physical and logical product details
- Comprehensive operational life cycle processes
- Asset placement based on capacity utilization and asset purpose
- Comprehensive views spanning both the facility and hosted products
- Detailed analytics on product consumption and store operations
- Product shifting in response to operational conditions
- Continual replacement of old stock
- Standardized rack selection
- Localized and variable cooling
- Designated zones based on product type and purpose

Examining the business of an efficient grocery, we can understand the value that comes from applying many of these same techniques to our data center environments, but a successful application will require a pervasive transformation in data center operations, starting with how we define capacity.

2.1.2 Expanding the scope of capacity management

In order to gain large efficiency gains, we must expand our treatment of capacity, both in breadth and in depth. At the macro level, as increasingly powerful servers occupy increasingly smaller physical footprints and generate larger heat signatures, data centers can easily run out of available cooling or power long before physical space is exhausted. Consequently, it becomes imperative to track and manage multiple capacity types concurrently, constantly striving for a balance in consumption across them. Exhausting any one type of capacity leaves other capacities stranded, limiting your growth, increasing the operational costs for the work performed within the environment, and resulting in lost opportunities for additional business without incurring additional construction costs.

Environmental concerns and governmental regulations often necessitate the inclusion of additional capacities, such as water usage and carbon emissions, and larger data transmission demands can require the inclusion of network bandwidth capacity as well.

While the breadth of capacities to manage increases, it deepens as well. Individual servers can be as little as 20% utilized, while data storage demands continue to escalate. Effective utilization of consolidation, virtualization, and cloud computing demands the awareness and management of asset level capacities like CPU and storage.

As a result, a comprehensive view of data center efficiency necessarily spans facilities and IT and identifies both the capacities to manage and the assets that are consuming them (see 2.4.5, “Defining your data standards” on page 15). However, achieving this perspective comes with a set of challenges.

2.2 Acknowledging the challenges

Although the potential for improvement is immense, the challenges to overcome are substantial. These challenges can be both technical and cultural. Acknowledging the reality of these challenges early will enable you to adequately address them in the development and implementation of your strategy.

Fragmented organizational ownership

Separate organizations within the company often assume separate responsibilities for separate aspects of the data center, ranging from the operation of the facility itself to ownership, deployment, and operation of server assets placed on the data center floor.

This fragmented organizational ownership makes it difficult to obtain a comprehensive picture of the data center's contents, conditions, and cost structure. A successful efficiency strategy must bridge all of these organizations, obtaining input and cooperation from each.

Fragmented data availability

Knowledge of a data center has to start with its contents. However, this information is often spread across organizations and lacks a detailed level of granularity that would include asset ownership, operational purpose, physical asset location down to rack slot placement, and server inventories to a level of granularity that includes virtualized servers. Even defined, this information is difficult to maintain with an acceptable degree of accuracy.

Availability of monitoring data can also vary considerably regarding what operational characteristics of assets are being monitored and trended over time using a myriad of varying tools. Data on the environmental conditions of the facility can be fragmented as well.

This fragmentation of both asset inventory and monitoring data makes it impossible to obtain a unified picture of a data center's current contents, asset placement, ownership, and operational conditions that integrates both IT and facilities.

Without this unified picture, effective decisions and efficiency analytics based on actual data become impossible as well. As is often said, you cannot manage what you do not measure.

No comprehensive views on capacity availability and consumption

As described in 2.1.2, "Expanding the scope of capacity management" on page 11, there are an increasing number of capacities to manage at an increasing degree of depth. We often must build comprehensive views across types and facilities.

No integrated strategy

Although separate organizations within a business often have specific strategies to explore and realize efficiency gains, there is often no cohesive, overarching strategy that demonstrates an ability to integrate all of these endeavors together from a comprehensive data center perspective.

Limited financial incentives for energy efficiency

As data center operators must increasingly balance utilization of multiple types of capacity, including power, space, and cooling, effective cost recovery must consider the consumption of all capacities. Stranded capacity of any type is missed opportunity.

However, the ability to perform effective cost recovery is dependent on the definition of capacity types to be managed and accurate measurements of consumption.

Pervasive security, privacy, and legal requirements

With increasing global connectivity and data demands come more pervasive security and privacy responsibilities and the legal obligations associated with them.

Limited cross-organizational communications

In addition to the lack of an overarching strategy that ties disparate organizational improvement efforts together, there is often no shared communication methodology between these efforts for awareness, coordination, exploitation, and avoidance of duplication of effort.

Although each endeavor carries its own business potential, the isolated nature of these initiatives fails to recognize the potential of a planned and communicated combined approach which would yield results where the whole is truly greater than the sum of its parts.

Cultural aversion to change

A successful efficiency strategy requires, to the largest degree possible, coordination across facilities and IT, cost recovery based on consumption across capacity types, and a transition to hosting provider ownership of all aspects of the physical environment. Each of these factors mandates a cultural shift in typical data center operations, which can be difficult to implement, illustrating the need for executive-level participation in your strategy development and application.

2.3 Laying the groundwork

Considering the complexity of a comprehensive smarter data center strategy, it is important to establish realistic expectations and recognize that your strategy will be an evolutionary process that matures over time. Certain improvements can be implemented quickly, but others will take considerable amounts of time and commitment.

While developing your strategy, you should go through each of the following steps, which are discussed at greater length in subsequent sections.

1. Document requirements and assign prioritization
2. Diagnose and benchmark the existing environment
3. Discuss current and desired business, financial, and technical maturity
4. Develop a realistic, incremental plan to achieve stakeholder expectations

The results of this process should establish what you want to achieve, where you are starting from, what destination is realistic, and a plan for getting there.

2.4 Defining your operational, process, and data requirements

"Among our respondents, eight out of ten organizations with process owners have been able to implement standard processes enterprise wide, while only three out of ten did so without process owners. Similarly, seven out of ten with process owners have developed standard data definitions."

- IBM 2010 Global CFO Study[2]

Your strategy will require updates to your business operations, documented processes, and data gathering methods, so it is imperative to start with executive level commitments.

[2] http://www.ibm.com/services/us/cfo/cfostudy2010/

2.4.1 Establishing sponsorship and governance

When constructing a house, you do not immediately start pouring concrete and hammering nails. You need to begin with a client (a sponsor), documented requirements, a blueprint, a plan, and defined teams that will coordinate responsibilities in the construction process. The same is true for your data center strategy.

Obtaining executive sponsorship

Executive-level sponsorship of your strategy and its subsequent implementation is critical for success. Without this level of sponsorship, the definition of a unified set of goals and formalized cooperation between separate functional areas spanning facility and IT operations becomes extremely difficult. The absence of this formalized coordination will significantly impair your odds of success and will inevitably add considerable time and frustration to the endeavor.

Creating an initiative and define responsibilities

After you obtain sponsorship, the effort requires structure. Create a formalized initiative aimed at comprehensive data center efficiency for your company, identifying the critical stakeholders across all operations of your data center environments spanning both facilities and IT. This group of stakeholders will be the primary source for defining and documenting the requirements for your overall strategy. A formalized structure and defined stakeholders ensure that everyone shares a common understanding and is aimed at common goals.

2.4.2 Identifying your business requirements

A successful data center strategy must first identify the business concerns or goals it is intended to address. Defining these business-level requirements early in the process plays an integral role in the selection of the methods and technological means that will be used in implementation. Here are several common data center business requirements.

- Acquire current and accurate knowledge of data center contents and resources
- Manage selection of resource capacity sources and associated costs
- Optimize resource capacity utilization and compute power
- Reduce ratio of facility versus IT consumption
- Measure and improve data center operations and resiliency
- Associate assets, consumption, and conditions with provided IT services
- Exercise process and tools synergy through integration across organizational boundaries

2.4.3 Determining process and data needs

Each item in 2.4.2, “Identifying your business requirements” on page 14 will require its own set of mature processes and a structured data architecture. These process and data requirements must be identified for each business requirement pursued.

2.4.4 Identifying associated processes and ownership

There are multiple activities related to the day-to-day operations of any data center environment and numerous approaches both within facility and IT operations to garner considerable efficiency gains. However, without the proper processes in place, it is difficult to know when to take advantage of the various approaches to maintain efficiency gains on an ongoing basis rather than as a result of a one-time assessment.

You should identify and document the processes involved in the operations of your data center and define the appropriate ownership. You can then create or update these processes to use efficiency techniques as an integral part of your business. See section 2.5, “Assessing your starting point” on page 17 for additional information.

2.4.5 Defining your data standards

To assess progress and communicate effectively, all parties involved in your strategy need to be communicating with a common technical language. This requires the creation of standards, defining what will be measured, how it will be measured, and how this data will be documented, tracked, and maintained.

Defining the capacity types to manage

Much of data center efficiency involves capacity management, but there are many types of capacity, including physical space, power, cooling, storage, CPU operation, and network bandwidth. Governmental regulations or company goals might establish additional capacity limits, such as carbon emissions. As a result, your strategic requirements should define the separate types of capacity you intend to manage and what measurement you will use to gauge the consumption of each. Data center floor space may be measured in square footage, rack capacity in rack units, power in kilowatts, and so on.

Defining your efficiency metrics

A standardized data architecture includes managed capacity providers and consumers. Such an architecture facilitates the creation of related metrics, reports, and role-based dashboards that can be used to assess efficiency of use, determine overall costs, and perform consumption-based billing. A standardized data architecture also enables comparisons across capacity types, different physical sites, and so on.

Differentiating between facility and IT capacity consumers enables the determination and tracking of industry-standard efficiency metrics such as Power Usage Effectiveness (PUE) or its inverse metric, Data Center Infrastructure Efficiency (DCiE), both of which consider the ratio of facility versus IT power consumption.

You might also consider additional efficiency and sustainability metrics introduced by industry consortium *The Green Grid*, such as Carbon Usage Effectiveness (CUE), Water Usage Effectiveness (WUE), Energy Reuse Effectiveness (ERE), and Data Center Compute Efficiency (DCcE).

Associating consumption measurements with logical asset characteristics enables reporting and analysis beyond just efficiency, including cost recovery based on associated services provided by the consuming assets, viewing consumption based on asset ownership, and so on.

Taking advantage of the standard data architecture, facility level measurements can be rolled up to gain regional or business-wide perspectives on capacity provision, consumption, and efficiency.

Defining and prioritizing capacity sources

With a defined set of capacities to manage, the next consideration is capacity sources. For each capacity type to manage, your strategy should define the possible sources for each and how the amount of capacity provided by each potential source is determined and tracked, noting that the available amounts of certain capacities will change over time. Data center physical space is static, but power and cooling capacity might vary based on conditions.

There might also be multiple potential sources for the same capacity type. How will these separate sources of similar capacity be prioritized and differentiated? Power may be provided by the local utility or by on-site power generation from sources such as wind, solar, or geothermal. In addition to computer room air conditioners (CRACs), cooling capacity may be provided by localized liquid cooling or free cooling based on your geographical location and outside conditions.

When prioritizing capacity sources, you might consider related pricing as well. One source of capacity might be exhausted before the data center uses the next source of that same capacity. For example, exhaustion of on-site renewable power might occur before switching to consumption utility power. You might also choose to manage multiple capacities of the same type from the same source, differentiating between them based on unit cost, for instance, tiered pricing on utility power.

Certain capacities might overlap as well. Localized liquid cooling might only be available to the singular rack or racks to which it is attached, but those same racks will also be consuming centralized air conditioning capacity, though obviously at a much reduced level.

Before you can gauge how much of a particularly capacity is being consumed, you must first have a firm handle on how much of that capacity is available.

Defining the assets performing consumption

After you have defined the capacity types you want to manage, the next step is to identify the asset types that will be consuming each of these capacities. The nature of the capacity being consumed will dictate the characteristics of the asset that should be included in your data gathering and associated standard data architecture.

For instance, if you intend to manage data center floor space capacity, you should identify all types of floor-mounted devices as consumers of that capacity. If you measure capacity consumption in square footage, you need to know the depth and width of each floor-mounted device to be placed on the data center floor and include this information in your data center inventory. You should also ensure that your operational processes for asset installation include the gathering and maintenance of this data.

A single identified asset type will likely be a consumer of multiple capacity types. A floor-mounted server will consume floor space capacity, power capacity, cooling capacity, and network bandwidth capacity. If you are managing all of these capacity types, you must ensure that your data architecture and associated processes are gathering the expected consumption of each capacity when placing a floor-mounted server in the data center.

An asset that is identified as a consumer of one or more capacity types may also be configured as a provider of other capacity types. A rack is a consumer of floor space capacity but a provider of rack unit capacity.

Differentiating consumption types

You should also consider the different status states for consumption. The consumption being recorded might be “committed” for an asset that is physically installed, but the actual consumption might not directly measured, “planned” in the case of future asset placements, or “actual” where the consumption is being actively measured on an ongoing basis.

Physical assets often require buffer capacity around their physical dimensions to allow for adequate cabling and air flow. Also. power systems might require failover consumption to ensure continued operations in case of equipment failure.

Your methods for defining capacity and measuring consumption should account for these differing types of consumption.

Although an initial smarter data center implementation might be limited to expected consumption, gathering detailed operational measurements can identify actual consumption, enabling the minimization of operational buffers for increased efficiency.

The more granular the definition of both capacity provision and consumption, the greater the opportunity to identify potential areas for improvement.

The need to identify these varying capacity providers and consumers illustrates clearly the fundamental need for a detailed inventory of your data center's assets, an inventory which contains the physical and operational characteristics of each asset needed to properly manage your data center's use from multiple viewpoints.

2.5 Assessing your starting point

Before you can establish a realistic strategy for reaching your company's smarter data center goals, you must first establish the point from which you are starting. The transformation to smarter data center operations requires a two-pronged evolution, one involving your business processes, and the other involving your technological capabilities.

2.5.1 Assessing business process maturity

In section 2.4.5, "Defining your data standards" on page 15, we discussed the need to identify the business processes associated with the day-to-day operations of your data center, from maintenance to asset placement, to alert and event management. After you have identified these things, you need to assess the maturity of those processes. Much of this Redpaper discusses technical methods for improving the overall efficiency of your data center. These technical methods are simply the tools, but tools alone will not integrate efficiency into your daily business. A mature set of processes is a must.

A comprehensive smarter data center strategy should include an assessment of the existing state of your data center processes with an understanding that these processes will need to evolve along with your technical capabilities. Standards such as Information Technology Infrastructure Library (ITIL) best practices[3] and Capability Maturity Model Integration (CMMI)[4] are effective means for determining your current process maturity and providing direction in its evolution.

2.5.2 Assessing technical capability maturity

There are four functional areas that are integral to obtaining a holistic view of your data center operations.

- **Inventory and life cycle**

  It is imperative to know what assets are deployed within your data center, along with what is coming, what is going, and what is changing. Ideally, this inventory is granular enough to include virtualized servers and detailed enough to include capacity consumption measurements and logical characteristics such as organizational ownership and operational purpose.

- **Instrumented metering**

  It is impossible to know the efficiency status of your data center without continually gathering data about capacity consumption and operational and environmental conditions.

[3] ITIL: http://www.itil-officialsite.com/

[4] CMMI: http://www.sei.cmu.edu/cmmi/

- **Consolidation and condition monitoring**

  A comprehensive view of the data center spans both facility and IT assets. A smarter data center enables collection of data from both aspects of operations and enables that data to be consolidated and analyzed as a cohesive whole. This data collection can provide the foundation for monitoring of abnormal environmental or operational conditions.

- **Integrated analysis and data center optimization**

  After it is consolidated, data gathered from facility and IT operations can be analyzed together, providing the means to identify opportunities for optimization and plan for future needs.

Your company's current ability to satisfy these four functional areas within your existing facilities will determine the length of the journey to making your data centers smarter.

2.5.3 Assessing current data availability

After you have identified the data needed to perform the level of management desired, you can assess the current availability of that data across your data center facilities to determine the gap. The size of this gap might play a significant role in the prioritization of your business requirements fulfillment and the development of the associated implementation plan.

2.5.4 Assessing your current efficiency

Knowing which metrics you intend to use to assess the efficiency of your data centers and the availability of current data, you can assess the current efficiency of those facilities to establish a benchmark against which you can measure progress. If there is insufficient current data to perform this assessment, you will need to gather this data. You will likely need to base your initial assessment on data gathered at a high level and then increase the granularity of that measurement as your capabilities mature. The closer your measurements come to asset level consumption, the more accurate your efficiency metrics become.

2.5.5 Considering important questions

To assess the current efficiency of your data center, ask yourself important questions about the following areas:

- Facilities
- IT equipment
- Utilization rate

The facilities

The following set of questions pertains to your site and facilities:

- How and where is energy being used?
- What is the facility's current DCiE or PUE?
- Are my facilities power and performance-oriented or only performance-oriented?
- Do I invest in a new data center, or can I invest in the evolution of my current data center?
- Is the physical site of my data center adaptable to changes?
- Is my desired level of reliability driving the facility's energy consumption? How much idle capacity for redundancy or resilience exist? Is it enough or too much? Could I eliminate any equipment?

- What support equipment should I choose (uninterruptible power supply, flywheel, generators, power distribution, chillers, CRACs, and so forth)? What are the future trends? Is the facility's infrastructure adaptable to the power and cooling requirements of the next generation hardware? (Note that more IT equipment will be water-cooled in the future.) Does the facility have problems with overheating? WIth humidity?
- Can I use free cooling? (Refer to section 4.4.4, "Economizers to enable free cooling" on page 79.)
- Do power, cooling, or space impact operations today? Which will impact business growth in the future? Can I add future compute capacity within my energy envelope?
- Is my site infrastructure-optimized in the following categories?
  - Airflow and heat removal
  - Power distribution
  - Cooling
  - Lighting
  - Monitoring and management
- Am I able to exploit water for cooling if needed?
- Who can help?

The IT equipment

The following set of questions pertains to IT equipment. It includes design of the hardware but also the options for cooling, powering, and monitoring that exist at the rack level:

- Does the equipment use energy-efficient hardware? Does it use power-reduction features?
- Should I choose power and cooling options at the site, facilities, or rack level?
- Does the hardware provide options for power, thermal, and usage resource monitoring? Do I monitor and control power consumption?
- How is billing of the power usage done?
- Who can help?

Utilization rate

The following set of questions covers the server and storage utilization rate:

- Is utilization of my infrastructure optimized?
- Are there unneeded redundancies?
- Can I consolidate and virtualize?
- How do I go from multiple silos or islands of computing to a *shared model*?
- Do I monitor utilization rates of my resources? What about real-time and trends?
- How is billing done for the services that the infrastructure is providing?
- Who can help?

Additional considerations

The following additional considerations can greatly influence your decision-making:

- With an existing facility, am I in a lease or ownership position? If leasing, what remaining commitment do I have to the facility? If owning, is it in a market which has favorable economic conditions?
- For a new facility, are there additional site selection criteria that you should consider, such as proximity to another strategic site or business partner location, favorable construction and utility costs, presence in a growth market, availability of renewable power or cooling sources, and the potential for extreme weather conditions?

> **Redguide Case Study:** IBM designed the 60,000 square foot, 6 MW power capacity Leadership Data Center in Research Triangle Park, US, to rapidly scale up to 100,000 square feet and 15 MW of IT capacity with a highly scalable modular infrastructure in accordance with demand. See the Redguide *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, for more details.

2.6 Planning the journey

Diligence spent on ensuring your data center strategy is comprehensive and realistic will return immense dividends in its implementation.

2.6.1 Prioritizing for your business

Considering the breadth of scope in making your data centers smarter for your business, it is impractical to attempt to tackle all of the associated challenges at once. Your efforts should be prioritized based on the expected return on your investment (ROI), and the ROI will be unique to your business based on your process maturity, technical maturity, data availability, business specifics, and economic conditions.

2.6.2 Considering your efficiency options

There are a multitude of technical options available to you to improve both the facility and IT operations of your data center. Many of these are noted in Figure 2-1 and are described in more detail in subsequent chapters. The selection of which techniques to use will depend on a number of factors, including your individual business goals and requirements, current process and technical capabilities, cost considerations, and associated time frames.

The ultimate strategy will result in business processes that concurrently take advantage of techniques that enhance both facilities and IT operations, gaining efficiencies in all areas of the business.

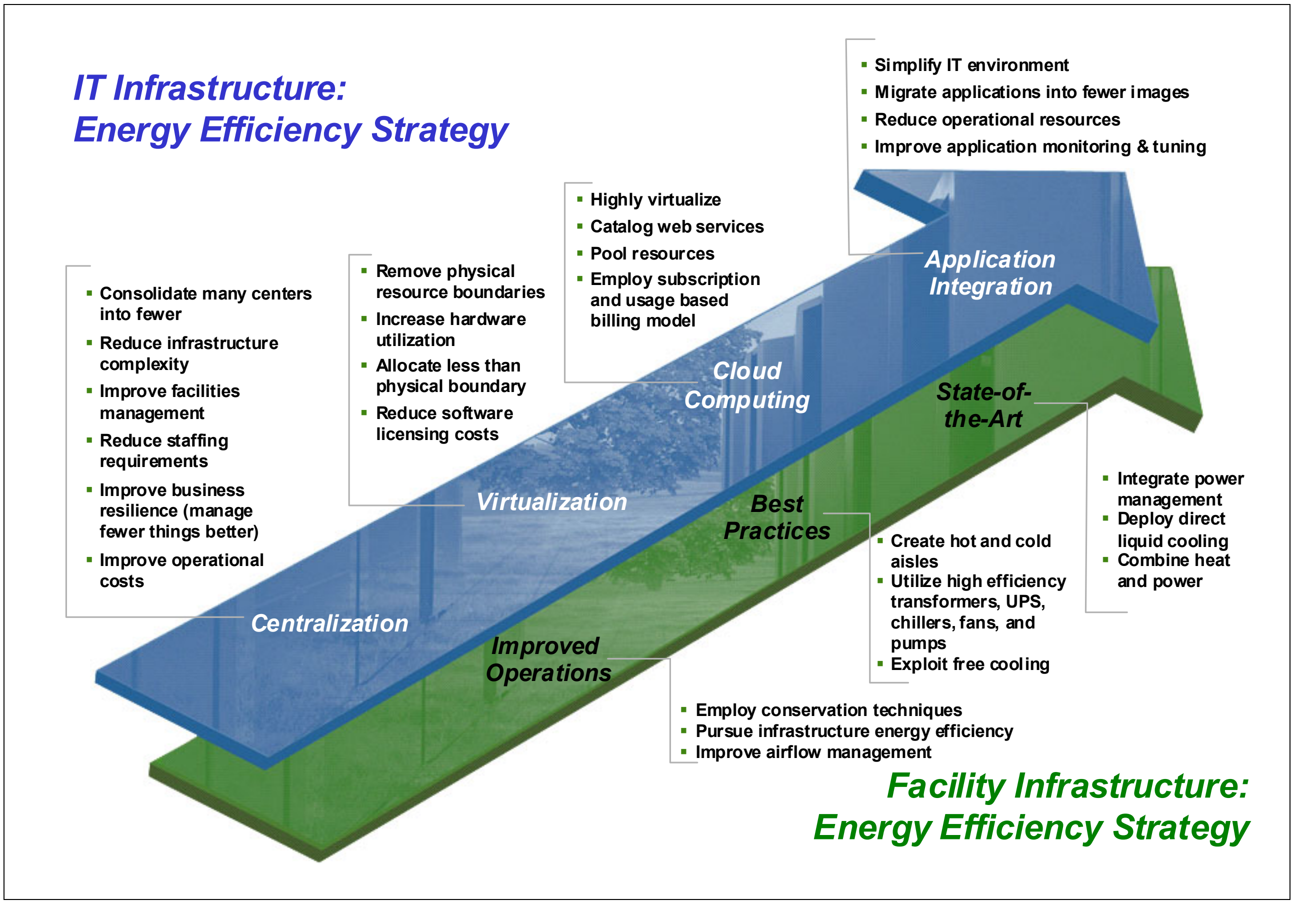


Figure 2-1 A number of efficiency options exist across both facilities and IT

2.6.3 Creating an implementation plan

With your business requirements documented and prioritized, stakeholders identified and committed, the gap between current and desired capabilities defined, and the means and metrics used to gather data and assess progress established, you can create a realistic implementation plan that can be communicated to all involved parties.

2.6.4 Providing a means of ongoing communication

As you develop and implement your smarter data center strategy, it is crucial to have open and ongoing lines of communication, particularly considering the organizational breadth of the undertaking. This can be performed using a series of ongoing, regularly scheduled meetings, the creation of a coordinating board, the designation of an initiative repository, or, more likely, a combination of all of these approaches.

IBM Connections[5] is social software aimed at a business environment. It provides an excellent option for this essential cross-organizational communication and collaboration.

2.6.5 Evolving based on ability

Having determined your starting point according to smarter data center operations and having understood that the integration of these capabilities into your day-to-day business is a continual evolution, you can establish realistic and achievable checkpoints along the way.

2.7 Effective integration into your daily business

Health experts often dislike diets as a means of losing weight because they commonly result in short-term gains only. After the individual stops the diet, the bad habits return, along with the lost weight. Long term success requires a lifestyle change, integrating good eating and exercise habits into your daily routine.

The same can be said for data center efficiency. One-time assessments and the performance of associated one-time actions might yield several measurable benefits, but they will likely be short-term because the same business processes that resulted in the inefficiency continue to be used. The necessary lifestyle changes come when you integrate of a detailed efficiency focus with the processes that govern daily operations.

2.7.1 Assigning efficiency architects

IT solution architects often have little insight into or, for that matter, concern about the operations of the facility providing the hosting for the servers, storage, and networking deployed as part of their deliverables. As long as there is adequate space, power, cooling, and network bandwidth to operate the solution effectively, the intricate details of the provision of that capacity is outside of their scope.

Conversely, facility architects might have limited knowledge of the purpose of the servers being deployed within their hosting environments. As long as their facility can accommodate the spatial, power, cooling, and networking requirements of the equipment to be deployed, what logical solution is operating on those servers is beyond their purview.

However, these singular deployment scopes can miss considerable efficiency gains. Utilization knowledge across capacity types at the facility level, coupled with awareness of server operational purpose at the IT level, enables placement of hardware within the data center at the most efficient possible location, considering both facility capacity and the grouping of like-purposed equipment. See 5.4.3, “Asset placement for efficiency” on page 98 for more information about conscientious architecture design.

---

[5] IBM Connections: http://www.ibm.com/software/lotus/products/connections/

The designation of a particular business architectural role with this cross-organizational awareness can assist in the most efficient use of capacity and ongoing operations of the data center environment.

2.7.2 Recovering costs based on consumption across capacity types

As you start to measure asset-level consumption of capacity across separate capacity types, it is important to integrate this information into your cost recovery process. Including broad-based consumption measurements into your billing process can facilitate greater efficiency efforts.

2.7.3 Monitoring progress with integrated dashboards, metrics, and reports

With access to broad and deep inventory, capacity, consumption, operations, and environmental data, you can take advantage of your efficiency metrics, along with role-based dashboards and longer-term trending reports, to gauge progress and determine next steps.

The same data can be used to monitor the health of your environment and the services provided from it, enabling increases in resiliency and reliability.

2.7.4 Optimizing based on analytics

As you gain a good handle on your data center inventory and the ability to gather data about your facility's operational conditions, you can take advantage of detailed analytics to determine the best options for improving overall efficiency.

2.7.5 Using knowledge for future planning

As you assess how available capacities are consumed and determine operational and environmental pain points, you can better plan for future capacity needs because you will know which types of capacity will be needed where and when.

As detailed in the following chapters, there are many ways to make your data center smarter for your business. Your strategy will help you make the most of these capabilities.

3

Optimizing your IT equipment

Electrical energy, when used for computing, is transformed into heat. Optimizing IT equipment by reducing power and heat at the source has a direct impact on facility efficiencies.

This chapter discusses the IT equipment aspect of a smarter data center. We describe two approaches to improving IT energy efficiency, show how energy efficiency management can fit into an overall systems management environment, and provide tips for moving your IT systems to a more optimized operation.

3.1 Energy flow in a computer system

Understanding how energy flows through IT equipment can help improve its efficiency. Computers transform information from one form to another. However, the energy that we put into the computer does not actually contribute to the information itself. Instead, the energy flows to the devices that hold the information. The physical work—energy consumption—results in a change to the state of the devices over time.

Eventually, all electrical energy that goes into the computer is converted to *heat*. For particular appliances, such as a steam iron, this heat generation is the desired effect when we want to remove wrinkles from clothing. However, in computer information processing, the heat is an unneeded by-product. In fact, we spend additional energy to remove the heat from the data center.

3.1.1 How the electric power is used

After power enters the data center site, it is distributed to the various IT components. The power allocation in a system depends on its architecture and purpose (see Chapter 4, "Optimizing your site and facilities" on page 57). Figure 3-1 illustrates the relative power allocation to the components of several IBM systems. The total amount of power supplied and the distribution of power to the components varies with the workload that is processed. Each system is built with a certain purpose in mind.

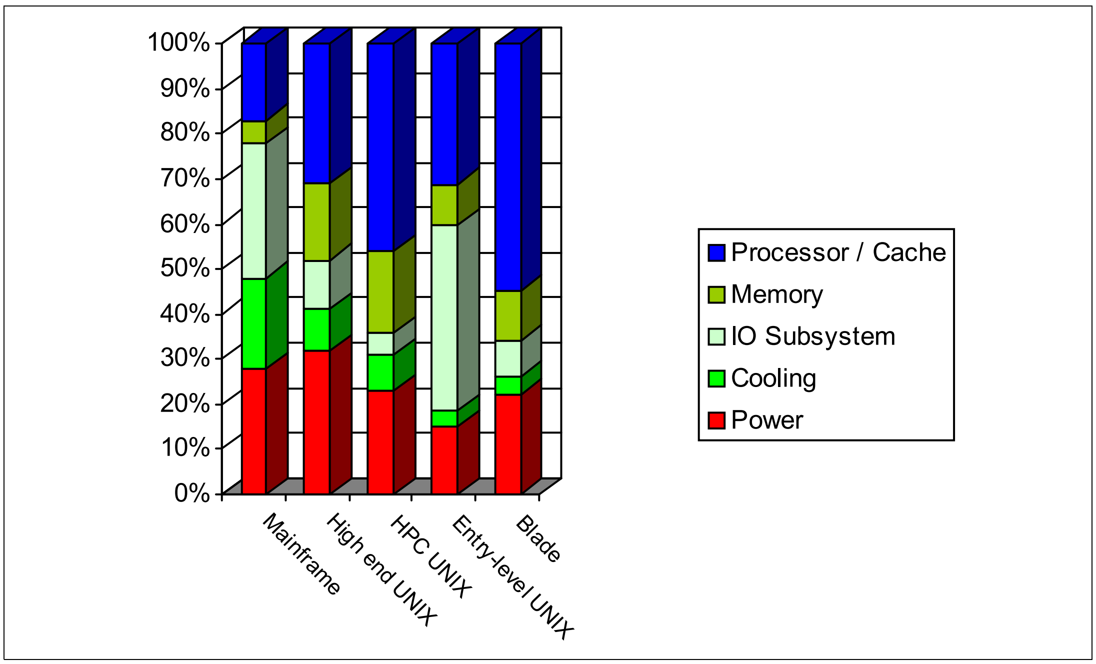


Figure 3-1 Typical relative power consumption by component for typical systems

From left to right, Figure 3-1 shows typical power allocations for mainframe systems, high-end UNIX servers, high performance computing (HPC) servers, entry-level UNIX systems, and blade servers, which represent a power-efficient, typical 1U server replacement. Because the power consumed by the processor makes up approximately 20-30% of the power used for the mainframe, in contrast to the blade's power consumption at over 50%, we use separate strategies for each when optimizing their energy effectiveness.

The red portions at the bottom of the vertical bars show the energy required to provide power to each system. Transforming AC power into DC loses some energy. The efficiency of a transformer (power supply) depends on the load and is non-linear. The most efficient load is 50-75%. Efficiency drops dramatically below a 50% load, and it does not improve significantly with greater loads. The challenge is to balance the system so that each component can operate in the most efficient way.

3.1.2 What to do with the heat

The power that is input is transformed into heat output. To prevent the destruction of the computer chips, the heat must be removed from the chip and then from the system. Figure 3-2 illustrates the removal path from the transistor device to the data center.

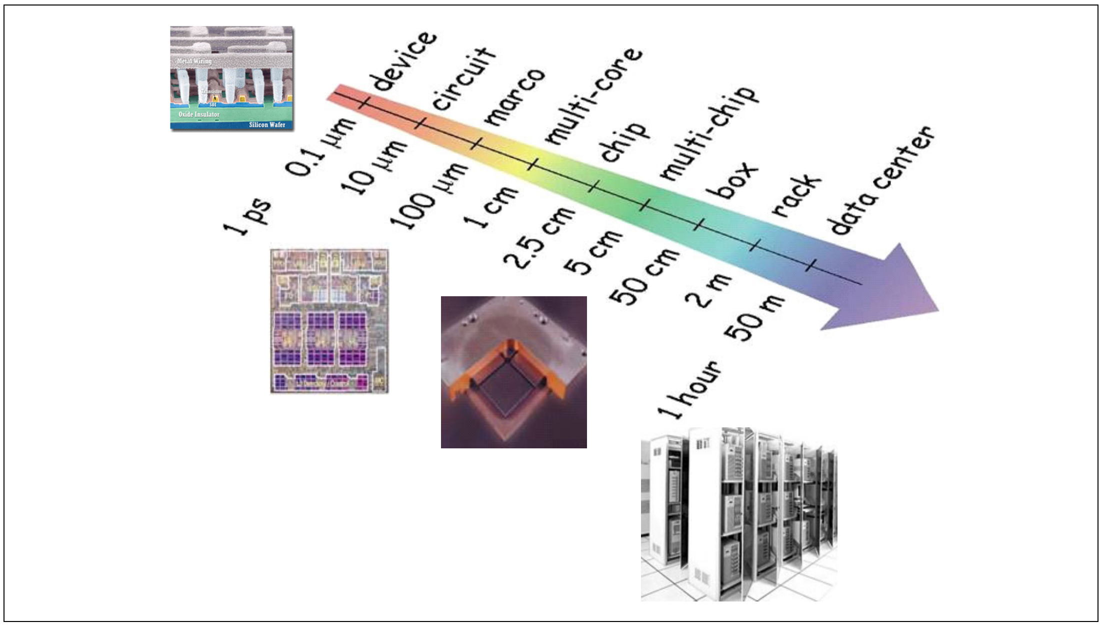


Figure 3-2 Heat flow path from transistor device to data center

Each watt saved in system power results is approximately another watt saved in heat load. These savings also have an effect on the uninterruptible power supply (UPS) and on cooling. Therefore, *reducing system power consumption pays back more than double*, which is a big benefit when moving to a smarter, more energy efficient data center.

Air is an inefficient cooling medium, so liquid cooling is increasing in popularity. Water is currently the most common liquid used for cooling. One liter of water can absorb about 4000 times more heat than the same volume of air. As more heat is generated in a smaller space, water cooled systems are becoming increasingly common. When planning new data centers or revamping current centers, consider that new IT equipment will likely require liquid cooling for efficiency.

3.1.3 IBM hardware solution examples

Let us begin with the system component having the greatest heat density—the *chip*. Usually the heat is removed from the back of the chip by using massive copper plates. IBM actively researches this field and has current solutions that allow for direct water cooling on the chip's backside. The basic principle is to let water stream through microscopic channels in metal plates. After the heat is transferred to water, it can be handled more efficiently.

IBM System zEnterprise

With the new system z196 mainframe, IBM has made tremendous progress in energy efficiency. The z196 system can now be ordered with a *water cooling option*, which can result in large savings on energy costs to cool the system. Figure 3-3 illustrates a sectional view of the redundant Water Conditioning Units, which transfer the cooling energy from the chilled water of the data center to a closed loop internal water system. The Multichip Modules (MCM) are cooled through cold plates, and the rest of the components are cooled through heat exchangers on the back of the system.

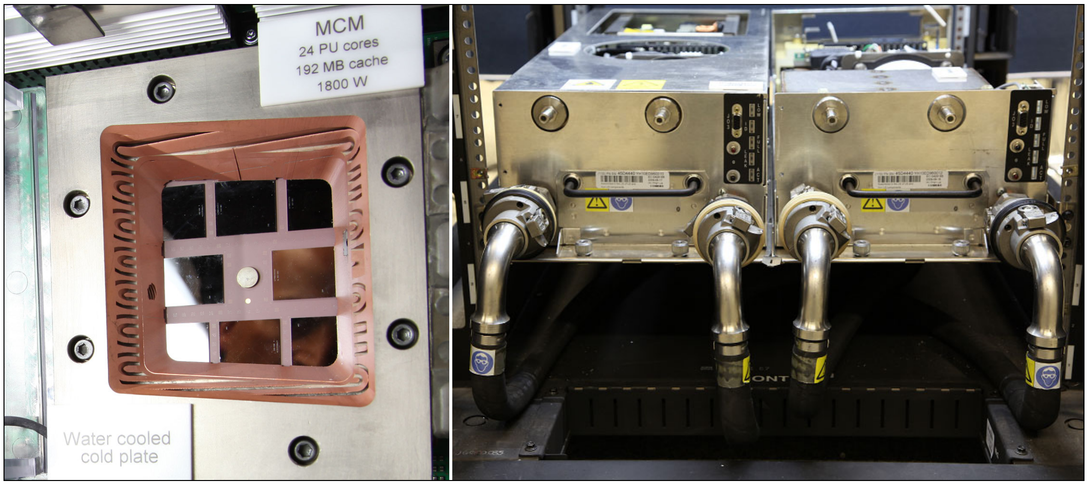


Figure 3-3 IBM z196 MCM Module with cold plate (left) and Water Conditioning Units (right)

Starting with the z114 and z196 each new system is also a hybrid computing solution, which can utilize a zEnterprise BladeCenter Expansion (zBX) unit to support up to 8 BladeCenters. This expansion increases flexibility as each BladeCenter can be equipped with Optimizers to improve performance or with IBM Power7 or System x Blade servers for a heterogeneous workload. The zBX expansion Rack also supports the Rear Door Heat eXchanger which is discussed in more detail in 4.3.1, “Rear Door Heat eXchanger” on page 73

IBM Power 775 supercomputer

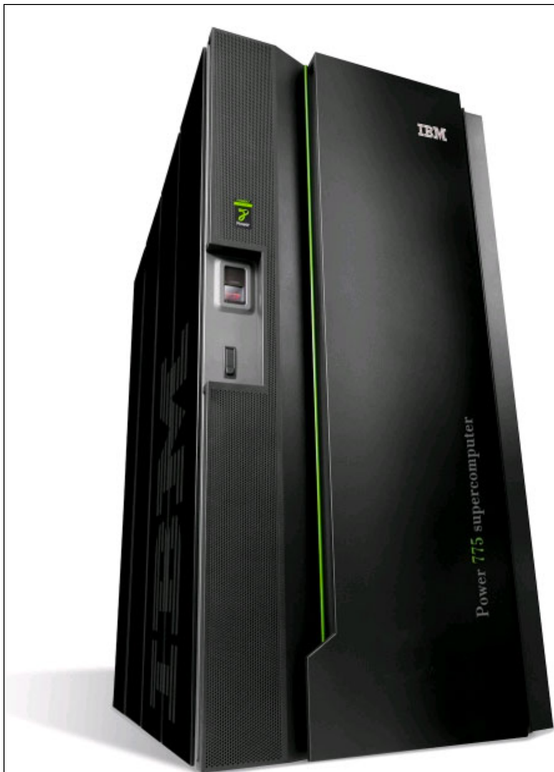
The Power 775 is 100% water cooled and powered from either 3-phase AC or DC to enable maximum data center energy efficiency. It also takes advantage of advanced IBM EnergyScale™ technology. The Power 775 is designed to deliver high performance, scalability, reliability, and manageability for demanding computing applications.

The system racking sequence places the water cooling units in the bottom of the rack, followed by the central electronics complex (CEC) drawers which are also populated from the bottom up. A 2U tall, 30-inch wide, 256-core IBM POWER7™ CEC drawer houses the system planars, redundant 350 V DC power supplies, water-cooled coldplates, and manifolds.

This ultra-dense packing of 256 cores in a 2U drawer coupled with new water cooling technology enables the POWER7 processors of the Power 775 to satisfy extremely demanding High Performance Computing (HPC) applications in a very small physical footprint while effectively removing excess heat.

As shown in Figure 3-4, the IBM Power 775 supercomputer provides several major benefits, several of which are listed as follows:

- Dramatically reduces the development time and cost from the idea stage to the delivery of solutions for the world's toughest engineering and scientific challenges.
- Provides economically viable high-productivity computing.
- Continues operating in the presence of localized hardware failure.
- Minimizes likelihood of operator error.



Operating environment

- Inlet water continuous operating temperature: 2° to 20° Celsius (35° to 68° Fahrenheit)
- Inlet air continuous operating temperature: 5° to 40°Celsius (41° to 104° Fahrenheit)
- Relative humidity: Non-condensing 20% to 80%
- Maximum dew point: 17° Celsius

Power requirements

- AC operating voltages (3-phase V_{ac}) at 50/60Hz: 200 to 240V; 380 to 440V; 480V
- DC operating voltage (5 -wire, V_{dc}); 380 V_{dc}; 520 V_{dc}

Figure 3-4 IBM Power 775 Supercomputer

IBM Rear Door Heat eXchanger

The IBM Rear Door Heat eXchanger, described in more detail in Section 4.3.1, "Rear Door Heat eXchanger" on page 73, can also help reduce the heat load in your data center. It can be used to address single hot spot problems or compensate if the air-based cooling of the data center is at its capacity limit. When a chilled water infrastructure is in place, the IBM Rear Door Heat eXchanger presents an attractive solution because the dissipated heat now bypasses the computer room air conditioner (CRAC) unit and can be more efficiently absorbed by the chillers.

The pictures in Figure 3-5 were taken in a production data center. The displayed rack houses several IBM Blade Centers with IBM Power Systems™ and System x Blades. The IBM Rear Door Heat eXchanger reduces the interior temperature of 46.9° Celsius (114.8° Fahrenheit) to 25.1°C (77°F) on its outside.

Figure 3-5 Heat exchanger in action

iDataPlex

System x presents another approach to solving data center challenges with IBM iDataPlex™. This is a data center solution for Web 2.0, HPC cluster, and corporate batch processing clients who are experiencing limitations of power, cooling, or physical space. iDataPlex servers help pack more processors into the same power and cooling envelope, better use floor space, and *right size* the data center design. With the iDataPlex solution, less power per processor means more processing capacity per kilowatt. The iDataPlex can also run cooler to deliver greater reliability.

iDataPlex offers flexibility at the rack level. It can be cabled either through the bottom, if it is set on a raised floor, or from the ceiling. Front-access cabling and Direct Dock Power allow you to quickly and easily make changes in networking, power connections, and storage. The rack supports multiple networking topologies, including Ethernet, InfiniBand, and Fibre Channel. The following benefits make iDataPlex an excellent option for a smarter data center:

- Efficient power and cooling
- Flexible, integrated configurations
- Increased levels of density
- Cost-optimized solution
- Single-point management
- Data center planning and power and cooling assessment

High-temperature water cooling

As discussed earlier in this chapter, water-cooled systems are more energy-efficient than air-cooled systems, where an equivalent amount of energy can be wasted for cooling as the system uses for IT operation.

Using high-temperature water for cooling can provide even greater benefits. The hot water can be reused directly for additional purposes, such as space heating or desalination. Year-round free-air cooling also becomes a potential option as there is almost no need for refrigerator compressors to cool down the hot water.

Figure 3-6 depicts a hot-water-cooled IBM Blade installed in the Aquasar computer located at the Swiss Federal Institute of Technology in Zurich (ETH), Switzerland. A joint project between ETH Zurich and IBM, Aquasar consumes 40% less energy than a comparable air-cooled computer with an 85% smaller carbon footprint. The IBM System z likewise offers both water-cooling and air-cooling supported by an Active Energy Manager.

Figure 3-6 IBM hot-water-cooled HS22 Blade

3.2 Elements of energy efficiency

So far, we have raised two questions:

- What makes IT equipment energy-efficient?
- How do systems compare to each other?

One approach to answering the two questions is to examine the relationship between the *power* that is used and the *workload* that is processed. This comparison provides a handy concept of an energy-efficient IT system, enabling you to select the most efficient option.

> **Note:** The more energy-efficient a system is, the more workload it can process with a certain power input, or the less power it needs for processing a certain workload. In other words, using energy-efficient equipment means that you can either decrease power and do the same work or increase workload with the same power.

Most efficiency improvements result in a combination of these two solutions. For example, we can assign the workload of computer A to another, more power-efficient computer B and then switch off computer A. This means we can choose from both alternatives at any given time.

> **Note:** For IT equipment energy efficiency, we will primarily look at the power to workload relationship. However, there is another workload-related characteristic we do not want to overlook: the workload-to-time relationship, better known as performance. Using physics, we can see our power-to-workload relationship as an analogy to mechanical work, so we view performance as analogous to mechanical power.

After each change, we want to know that we actually improved our overall energy balance. Obviously, this can only be achieved with proper before-and-after evaluation. The basic starting point and first step in our efforts must therefore be to *enable proper measurement*.

3.3 Using system instrumentation for power measurement

Recalling that it is difficult to manage what you do not measure as you move toward an energy-efficient system, knowing how a system actually behaves is essential. You must understand the system's power consumption along with its heat dissipation.[1] Although there are many systems in the marketplace, we will look at the instrumentation that IBM systems provide.

[1] The instrumentation for workload measurement is beyond the scope of this document. Tracking the operating system's utilization and using accounting measurement tools is (or should be) an established practice for all data centers, whether going green or not.

3.3.1 IBM systems with embedded instrumentation

Virtually all newer IBM systems have built-in thermal sensors and instrumentation for power usage measurement. Although system-specific tools are also available, it is primarily the IBM Systems Director Active Energy Manager™ tool that can read the information provided by these sensors, regardless of the particular system platform (see Figure 3-7). The Active Energy Manager allows us to display current values and to initiate actions based on the overall system status.

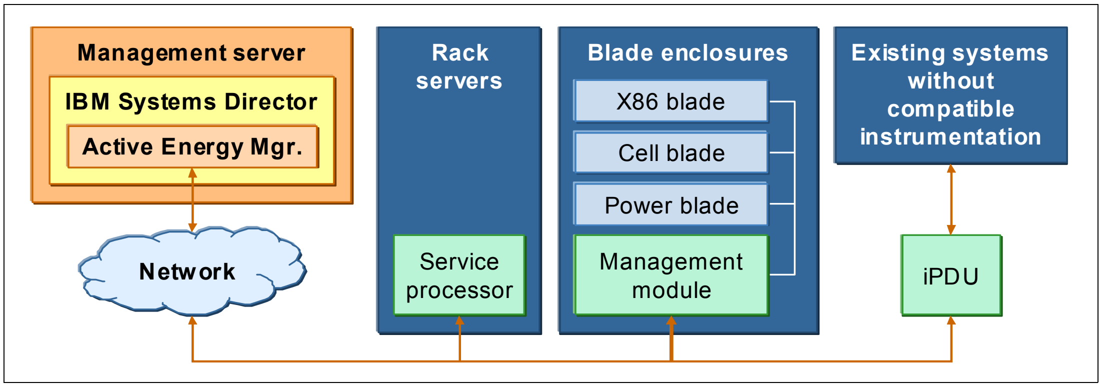


Figure 3-7 How system instrumentation and Active Energy Manager work together

The following overview describes instrumentation in IBM systems:

- The System Activity Display, including the Monitors Dashboard feature of the System z114 and z196, shows the machine's current power and temperature information.
- EnergyScale technology in POWER7 processors is built into POWER blades and IBM Power servers and monitors the actual power and thermal load of the processor.
- System x servers and the BladeCenter chassis provide their power and monitoring information to the Active Energy Manager, described later in more detail.

3.3.2 IBM intelligent power distribution units

For systems that do not have embedded or manageable instruments onboard, intelligent Power Distribution Units (iPDUs) are available[2]. The IBM DPI® PDU+ iPDUs contain versatile sensors that provide power consumption information about the attached devices and, optionally, environmental factors such as temperature and humidity. The iPDU's network interfaces allow for remote monitoring and management through a web browser, any SNMP based Network Management System, Telnet, or HyperTerminal (console). Event notifications can be delivered through SNMP traps or email, and it is possible to send out daily history reports through email. The Active Energy Manager can also be used to manage the iPDUs.

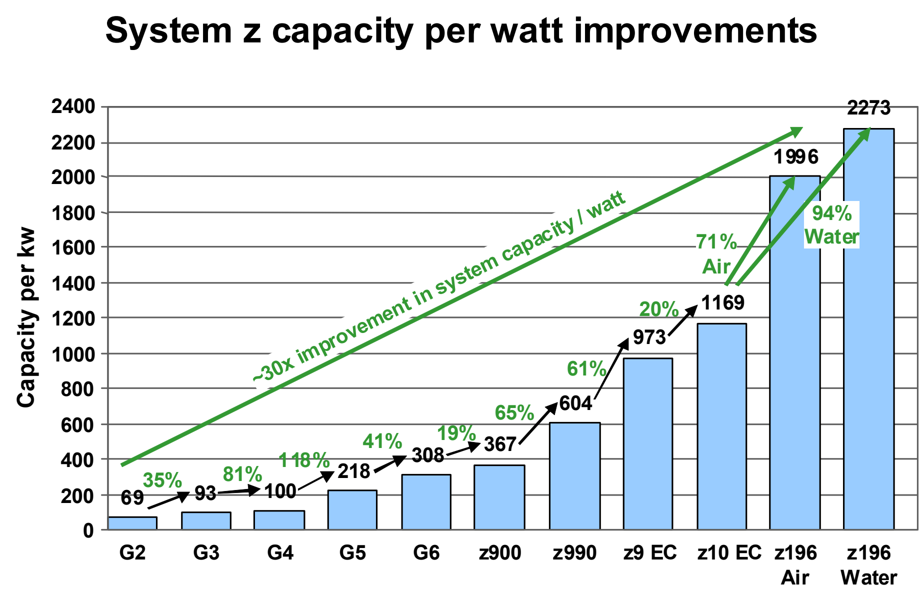
3.4 Managing power on the hardware side

We now concentrate on the possibilities that IBM systems provide for actively managing their power consumption. (In section, 3.5, "Managing power on the software side" on page 40, we show how multiple IBM systems can be managed from a single point.)

[2] http://www.ibm.com/systems/x/options/rackandpower/powerdistributionmonitored.html

3.4.1 IBM System z114 and z196 features

Always an excellent option for server consolidation and virtualization, the latest generation of System z[3] mainframes has continued to improve performance, as shown in Figure 3-8, increasing workload management by 50% over previous versions without increasing power consumption. These improved systems also come equipped with many new hardware capabilities and management functions.



15 years of CMOS: G2 to z196		Net Effect: G2 to z196	
Power increase:	17% per year	Performance increased by:	~300x
Performance increase:	46% per year	Performance / watt increased by:	~30x
Power density increase:	13% per year	Performance / floor area increased by:	~190x

Figure 3-8 System z capacity per watt improvements chart

Below we describe many of the improvements of the new IBM System z family of mainframes.

Hardware capabilities

System z features the following hardware capabilities:

- Power conversion and distribution

  Final voltage and current step-up is performed at point of load and not on the DCA powerboard.

  Benefit: Passive and high efficiency power transformation

- New temperature, humidity and atmospheric pressure sensors

  Temperature and humidity sensors calculate the dew point to optimize the water cooling or refrigeration system, and an atmospheric pressure sensor optimizes cooling fans.

  Benefit: Cooling systems operating at the most efficient level

[3] IBM zEnterprise System: http://www.ibm.com/systems/z/

- Water cooling option

  Chilled building water can feed Water Cooling Units (WCU) in the system, making improvements over an air-cooled solution with refrigerators:

  – Cooling at lower temperatures is unnecessary as extra margins for cooling with refrigerators manage condensation.

    Benefit: Lower temperatures on silicon results in less power consumption

  – You can opt to operate with no refrigerators at all.

    Benefit: Saving of about 500W for each node

  – Most of the heat produced by the system passing directly to building water

    Benefit: Data center air conditioning system load drastically reduced.

  With these improvements, you can save at least 4 kW of energy on a four-book system.

  Note: This function does not apply to z114 systems.

- High Voltage DC

  The system automatically detects if there is a three-phase AC (200V-480V) or DC (330V-550V) power source, enabling support for all industrial voltages. Selecting a high voltage DC source can lead to additional power savings.

  Benefit: Consumption of 2-3% less power using DC due to higher power supply efficiency

  Benefit: No conversion losses on the supply side as uninterruptible power supplies typically convert and reconvert the voltage

- More efficient processor and controller chip design

  The new chip design is 25% more energy efficient than its predecessors.

  Benefit: System running at a 20% higher frequency

  Benefit: Overall System performance raised by 40%

Management functions

The new System z hybrid computing solution also includes the Unified Resource Manager (zManager), a powerful management tool. zManager offers hardware and platform virtualization and management tasks for all diverse systems in the hybrid solution from a single point of control.

From an energy-efficiency point of view, zManager offers energy monitoring and workload and resource management capabilities for the entire infrastructure. Intelligent distribution of various workloads to separate platforms ensures optimal use of system resources. Customers benefit from faster processing and resulting higher energy efficiency.

In addition to zManager, other new management functions help you become more efficient:

- Static power save

  If full system performance is not always required, for instance on weekends, the system can be placed in power save mode once a day through the management interface.

  Benefit: From 3 kW to 4 kW per hour savings using power save mode on a four-node system

  Note: This function does not apply to z114 systems.

- Max potential power

  This new function calculates the maximum power for the current system configuration, considering all system components along with the altitude and temperature where the system is installed.

  Benefit: Ability to rely on actual power consumption instead of working with label power, which normally comes out much higher

  Benefit: Support for data center power sizing

- New monitoring functions
  - Monitors dashboard

    Displays system metrics, including the hybrid part of the system
  - Environment Efficiency Statistics task

    Generates a historical view of the system's power consumption
  - Energy Management Details panels

    Obtains a deeper view of all important power and environmental system conditions
- Power Save and Group Power Save

  These functions enable you to set up three separate power modes. Mode targets can either be zArchitecture resources or Blades.
  - High Performance
  - Low Power
  - Custom

  Benefit: System components logically grouped in a certain manner and groups addressed individually with less effort
- Power Capping and Group Cap

  This feature can help you limit the maximum power on separate system components. BladeCenters support this and have methods to actively reduce power consumption.

  Benefit: Prevention of peak power usage and help facilitating data center power planning

3.4.2 IBM POWER7 EnergyScale

Starting with IBM POWER6®, IBM has offered EnergyScale power management functionality on its Power processor based systems. On IBM POWER7 Systems™[4], EnergyScale has been enhanced further and provides several additional and updated power management functions.

Although many of the features detailed in Table 3-1 are available ready for use, Active Energy Manager can provide access to even greater functionality.

Table 3-1 POWER7 functions

Function	Description	requires AEM
Static Power Save	Reduces the processor frequency and voltage by a fixed amount, reducing the power consumption of the system	No
Dynamic Power Save	Adjusts processor frequency to performance and saves power where applicable	Yes

[4] IBM Power Systems: http://www.ibm.com/systems/power/

Function	Description	requires AEM
Energy-Optimized Fan Control	Adjusts fan speed based on operating temperatures, reducing power consumption	No
Processor Core Nap	Places unused cores in a nap state, operating them in low-power mode	No
Processor Folding	Adjusts the number of processors available for dispatch based on workload demand	No
Energy Scale I/O	Provides power to I/O slots only when needed	No
Server Power Down	Provides information necessary to dynamically move workload and shut down the server	Yes
Partition Power Management	Provides information about power savings settings and the system processor pool	Yes
Power Trending	Reports power consumption of the system	Yes
Thermal Reporting	Reports inlet and exhaust temperatures	Yes
Power Capping	Limits the overall power consumption of the system to a user specified limit	Yes

Many of the power management functions can be configured at the operating system level (AIX) or using the Advanced System Management Interface (ASMI), Hardware Management Console (HMC), Systems Director Management Console (SDMC), or Active Energy Manager. For a deeper understanding of these functions and their availability, see the white paper located at the following site:

http://www.ibm.com/systems/power/hardware/whitepapers/energyscale7.html

3.4.3 IBM BladeCenter features

IBM BladeCenter servers offers an efficient high-density solution for the data center. Smaller and thinner than normal rack mounted servers, up to 14 blade servers can be placed either vertically or horizontally in a single 7U or 9U high BladeCenter chassis depending on the chassis type. Although each blade server provides its own processors, memory, storage, and networking controllers, the BladeCenter chassis provides a shared infrastructure for power, cooling, I/O network switching, and management, minimizing the use of physical space and energy.

The BladeCenter family of products[5] includes five chassis: BladeCenter E, BladeCenter H, BladeCenter HT, BladeCenter S and BladeCenter T. Each has its own unique set of characteristics to meet a wide range of customer needs.

All BladeCenter chassis provide immense flexibility and lower the total cost of ownership (TCO) through diverse I/O support, eliminating the need to purchase additional switch gear while lowering cabling complexity. All chassis models support Ethernet, Fibre Channel, and SAS modules and industry standard switching from Cisco, Brocade, and others. The BladeCenter H chassis also provides support for high performance protocols such as 10 GB Ethernet, Virtual Fabric, Convergence Enhanced Ethernet and Fibre Channel over Ethernet (CEE/FCoE), and 4X InfiniBand.

[5] IBM BladeCenter: http://www.ibm.com/systems/bladecenter/

All five chassis support a common set of blades that are available on separate processor platforms, including Intel, AMD, POWER and Cell. For Intel applications, IBM now offers four-socket blade servers with the ability to dynamically reprovision between two-socket and four-socket blades remotely and MAX5 memory scaling, providing unmatched memory density per socket in the blade space. With the MAX5 expansion, you can significantly reduce the number of deployed CPUs for a desired memory footprint and lower power consumption per workload.

With the available power configurator[6] tool supporting both BladeCenters and System x, you can estimate the power requirements for a specific configuration. The results, which are usually below the machine's label rating, can be used for better power management and electrical infrastructure load estimates, reducing stranded power in the data center.

Supplementing their efficient redundant power supplies, BladeCenters also offer many power management options depending on the platform, including power capping on Intel blades and EnergyScale on POWER7 blades. Also, like many other IBM server offerings, BladeCenters can be managed with IBM Systems Director Active Energy Manager.

Note: As shown in Figure 3-9, the BladeCenter S chassis can also move certain workloads out of the data center altogether. Designed for office and branch environments, the all-in-one BladeCenter S can hold up to six blade servers, offers redundant storage area network (SAN), I/O, and disk capabilities, and can be equipped with an optional Office Enablement Kit which includes an acoustical module, front locking door, air filters, and more.

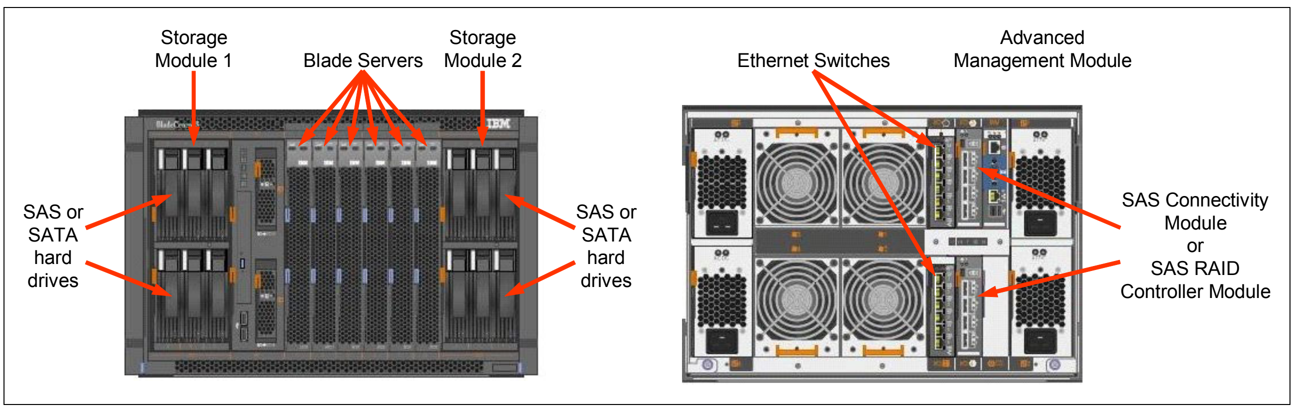


Figure 3-9 IBM BladeCenter S

3.4.4 IBM System x features

The latest System x family[7] has many of the same key features as the IBM BladeCenter, such as:

- Reduced power consumption with the use of low power 40 W processors, low voltage 1.35V DDR 3 memory, 2.5" Serial Advance Technology Attachment (SATA) hard drives, and high IOP solid-state drives.
- Efficient cooling across all System x and BladeCenter systems using an efficient cooling design called Calibrated Vector Cooling and industry leading efficient power supplies.

[6] The IBM System x and BladeCenter Power Configurator can be downloaded from this site: http://www.ibm.com/systems/bladecenter/powerconfig/
[7] IBM System x: http://www.ibm.com/systems/x/

- Immense affordability, flexibility, and choice with the IBM fifth-generation X-Architecture®, eX5, featuring innovations that deliver unparalleled capabilities and continuing to take advantage of open, standards-based components. The external MAX5 memory expansion option, which decouples server memory from the system processors, enables unprecedented memory expansion, server performance optimization, and support for large database deployments or superior virtualization density.
- Pre-installation planning with the IBM System x and BladeCenter Power Configurator supporting proper sizing of the installation environment and electrical infrastructure.
- Power and thermal monitoring of servers and infrastructure with IBM Systems Director Active Energy Manager.
- Power capping with Active Energy Manager (see section 3.5, “Managing power on the software side” on page 40).
- Extensive support of virtualization environments to enable server consolidation (Red Hat KVM, VMware, Xen, and Microsoft Hyper-V) and power consumption optimization of watts per workload.

> **Note:** Developed by Alinean Inc., the IBM Systems Consolidation Evaluation Tool can provide a preliminary, high-level analysis of TCO and ROI savings when consolidating using a System x or BladeCenter solution.
>
> http://www.ibm.com/systems/x/resources/tools/sconevalltool_intel/

3.4.5 ACPI and friendly power-saving features

The Advanced Configuration and Power Interface (ACPI) allows the operating system to take power-saving actions, such as putting the processor to sleep, spinning the disks down, and others. Although primarily used for mobile and desktop computers, certain servers, like the IBM System x, offer these features too. Note, however, that your system’s mainboard, chipset, timer, BIOS, UEFI, CPU, and the operating system itself have to be ACPI-enabled. ACPI is broadly available in the marketplace and is well-supported.

Processor-specific features like Intel Enhanced SpeedStep, AMD PowerNow!, or Cool’n’Quiet™ can also help reduce overall system power consumption. By reducing the processor clock frequency and core voltage, power requirements decrease while providing thermal relief as well. Figure 3-1 on page 26 shows the candidates for this type of power management, which are those systems whose processor plays a dominant role in power consumption.

3.4.6 High-input voltages

IBM has designed all System z and certain Power7 systems (Power 775 & 795) with a 200-480 volts of alternating current (VAC) three-phase power input so that they can be installed in any location in the world. In certain instances, the customer can chose to run a system from a high voltage (380-480 VAC) or a lower voltage (200-240 VAC) power source. When this choice is available, the higher voltage group provides the following advantages:

- Smaller line cords, line cord connectors, and building wiring all reduce installation cost and complexity.
- Step-down transformers on the computer floor (such as 480V-to-208V) are not required, which can reduce the installation cost significantly.
- System z and Power7 systems and typically the building infrastructure that feeds power to them offer higher energy efficiency, reducing energy costs.

- As energy prices increase, the savings become more significant.
- Using the UPS system in front of the system at the building level, eliminating the building transformer, results in even greater savings.

Many of the latest IBM products also support high voltage DC, which can generate even greater energy savings. For further information see 4.7.5, “DC versus AC” on page 83.

3.5 Managing power on the software side

From a hardware perspective, there are many power management options which function out-of-the-box and provide great energy savings. However, using appropriate software and management tools can help you further tune your hardware and attain even greater benefits.

3.5.1 IBM Systems Director Active Energy Manager

The primary tool for power management of IBM systems is Active Energy Manager (AEM)[8], which is an extension to the IBM Systems Director product. It provides a common systems management environment for all of the IBM platforms. Active Energy Manager offers the following basic features:

- Measures and displays current power and temperature data from the managed systems.
- Provides trend data over selected periods of time.
- Sets power capping for IBM systems if their firmware supports it, and manages the Power Saver mode for Power Systems.

[8] IBM Systems Director Active Energy Manager: http://www.ibm.com/systems/software/director/aem/

Figure 3-10 shows a scenario that demonstrates how AEM can help in optimizing your rack or BladeCenter layout, as follows:

1. Assume that power trending indicates that the server's actual power consumption remains below its labeled rating. Find a certain power usage that the server never exceeds.
2. Employ AEM for power capping. Cap the power for the systems to the maximum observed level. Remember, power capping sets the limit to something that should not happen under normal circumstances because of the significant performance impact it implies. However, obviously, the allocated power for the rack as a whole and the power actually consumed are now under control. The power that was previously over-allocated for each single system can now be managed at a rack-level.
3. Use the previously over-allocated power effectively by adding more systems to the rack. Ideally, install new systems with a better energy footprint, replacing less efficient machines. See 3.6, "Consolidation and virtualization" on page 43 for more information.

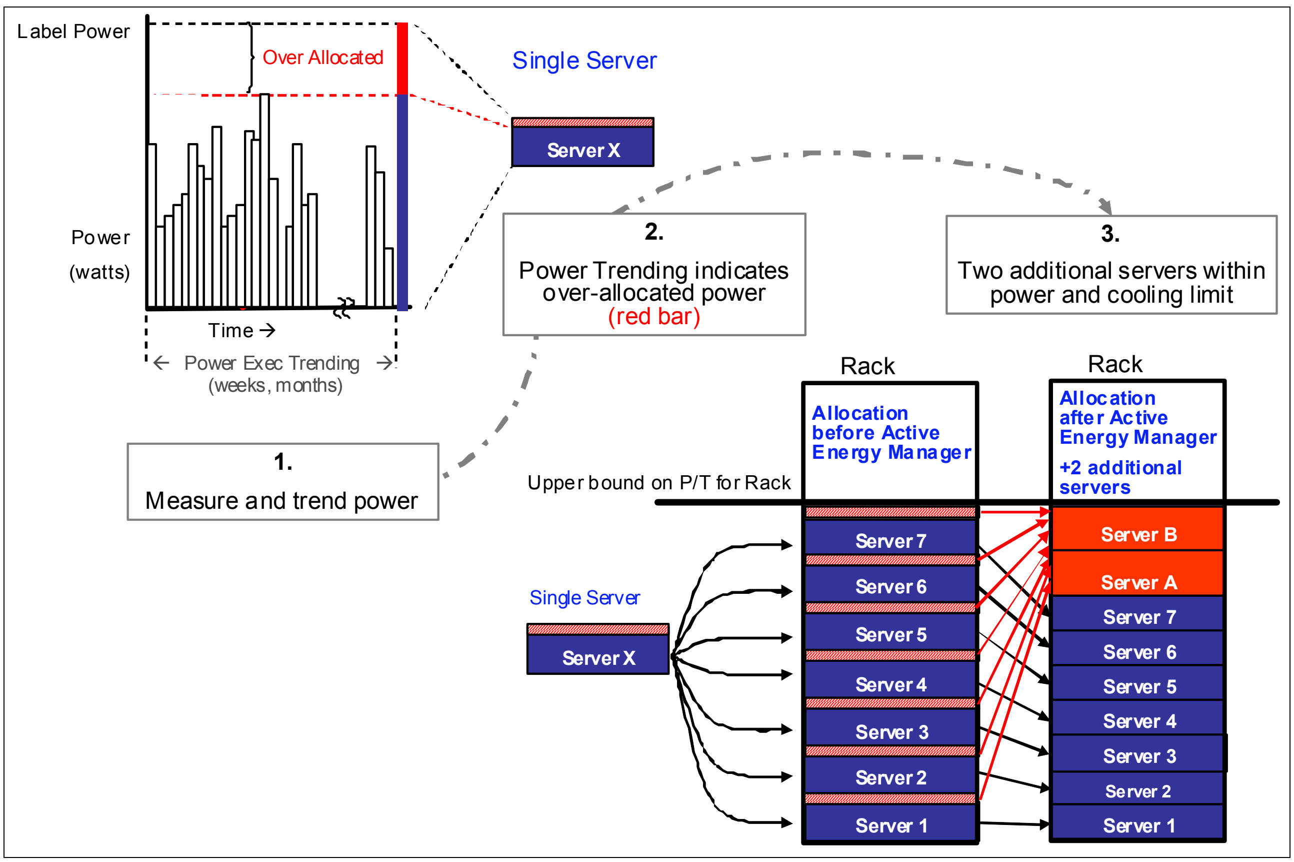


Figure 3-10 Optimize rack layout with the help of Active Energy Manager

When you have a power management infrastructure in place, you can take additional steps. For example, you can locate hot spots in your data center, perhaps supported by physical location information about your servers. Alternatively, after you know the hot spots, you can prevent them by provisioning your applications to *cool* servers. If your utility provider offers load dependent rates, or day and night rates, you can optimize your power consumption accordingly. Many options are available. However the first step is to build the infrastructure.

3.5.2 IBM Systems Energy Estimator

Although many customers use equipment label power for power sizing in their data centers, the IBM Systems Energy Estimator[9] can help you identify the actual energy required for an IBM Power system along with its associated heat load.

To do so, the tool will ask you to provide the actual system configuration, including the type, model, number of drawers (if applicable), processors, memory, disks and I/O. Out of this data the tool is able to provide you an estimate on how much energy the system will use and how much heat it produces. You can also import sysplan files generated with the IBM System Planning Tool or e-config files generated with the IBM e-config tool.

Figure 3-11 shows a typical energy and heat estimation for a one-drawer IBM 9179-MHB system.

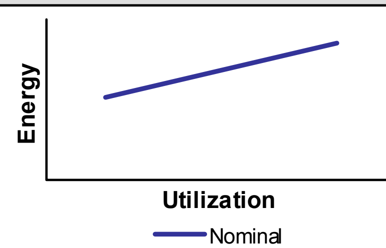
System information

Model:	780-9179-MHB
Cores:	16 cores of 4-64
Processor:	IBM®POWER7
Clock speed:	3860 MHz
Configured Memory:	128 GB

Energy Estimate

Nominal Mode

The following estimation was made without enabling any power saver modes. Thus the cores are running at the published frequency and the system capacity is as published. In Active Energy Manager, this is described as "no power savings mode". Using Active Energy Manager to specify power saver modes may significantly reduce energy consumption.



Estimated energy at 75% CPU utilization:	1.059 Watts 3.615 BTU/hr
Estimated energy at full utilization:	1.109 Watts 3.786 BTU/hr

Figure 3-11 IBM Systems Energy Estimator

Note: This estimated power requirement is typical and does not represent worst-case power conditions.

[9] IBM Systems Energy Estimator: http://www-912.ibm.com/see/EnergyEstimator

Although the IBM Systems Energy Estimator is available for IBM Power Systems only, IBM also provides the System x and BladeCenter Power Configurator[10] and the Power Estimation Tool for IBM System z196.[11]

3.6 Consolidation and virtualization

Having discussed the concept of power management earlier, we now look at the second key energy efficiency technique for workload management: *consolidation*. We concentrate our workload together so it will be processed by as few machines as possible and have these machines work at their most energy-efficient level. A powerful tool to help us do so is *virtualization*.

> **Note:** Workload is not limited to the applications a computer has to run. Workload can also represent storage usage or network usage.

3.6.1 Consolidation: A key in energy efficiency

Figure 3-12 on page 44 illustrates the idea of *consolidation*. Let us assume we have four systems, each running two applications (APP). Also, each machine consumes 2 kW power, 8 kW in total. However, as can be the case for small x86 servers, they are utilized at only 10%.

[10] http://www.ibm.com/systems/bladecenter/powerconfig/

[11] For the IBM System z Power Estimation Tool, navigate to the Power Estimation Tool under the Planning and Tools section of IBM Resource Link™ at http://www.ibm.com/servers/resourcelink

If we are able to consolidate these eight applications to a single, more powerful server and run this server at a utilization of 70% with a power usage of 4 kW, this single server can operate more energy-efficiently, according to our definition in 3.2, “Elements of energy efficiency” on page 32. In addition, if we perform a simple power management technique of switching off the previous four systems, the result is total power consumption of 4 kW and a 70% utilized system.

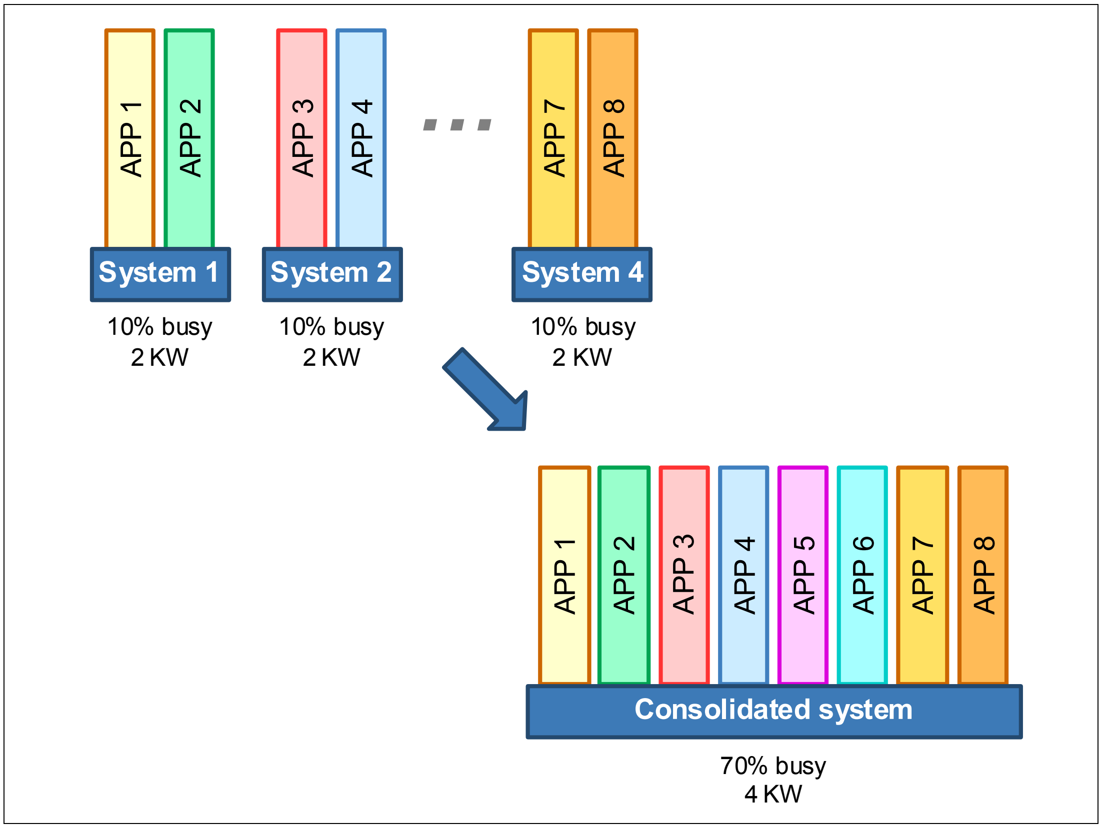


Figure 3-12 Consolidation of applications from under-utilized servers to a single, more efficient server

As we have indicated earlier, a decrease in overall power consumption is not the only factor in reducing energy usage. Hand-in-hand with the power reduction goes the same amount of heat load reduction and another add-on for the infrastructure. This double reduction is the reason why consolidation is an enormous lever to moving to a smarter, more energy-efficient data center.

However, a particular drawback of consolidation is that none of systems 1 through 4 is allowed to be down during the time that the respective applications are moving to the consolidated system. Thus during that migration time, higher demands on resources might occur *temporarily*.

Redguide Case Study: St. Lawrence College in Ontario, Canada, took advantage of a Scalable Modular Data Center and extensive server and storage consolidation and virtualization to consolidate two data centers into one and reduce the number of managed servers from 70 to 12. See the Redguide *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, for more details.

3.6.2 Virtualization: The greenest of technologies

In addition to consolidation is *virtualization*, the concept of dealing with abstract systems. It can dramatically reduce the amount of IT equipment needed in a data center. Virtualization eliminates the physical bonds that applications have to servers, storage, or networking equipment. A dedicated server for each application is inefficient and results in low utilization. Virtualization enables “car pooling” of applications on servers. The physical car (server) can be fixed, but the riders (applications) can change, be diverse (size and type), and come and go as needed.

The example in Figure 3-12 on page 44 shows how specific applications were moved to another system with a more conservative energy footprint. In our simple case, we assume all systems are running at the same operating system level. However, what if the applications require separate operating system levels or even completely separate operating systems? This is where virtualization comes into play.

> **Important:** Virtualization enables us to take the old system *as a whole* and reimplement it on the target machine.

The term virtualization is widely used and has several definitions, as follows:

- Creating logical instances of a computer system consisting of CPU, memory, and I/O capabilities.
- Assembling other virtual components.
- Consisting of a virtual CPU or virtual memory and disk.
- Serving as a virtual network between a virtual computer and the outside world.

For a virtual system to do real work, it must run on a real system. Obviously, additional intelligence is required to do this. There are pure software solutions, a system’s firmware can offer virtualization features, or such features cant be hardwired into the system. Many current processor architectures, such as those included with IBM System z and IBM Power machines, have virtualization features integrated, which software solutions can take advantage of. IBM also offers kernel-based virtual machines (KVM), which add hypervisor functionality based on the built-in support of Intel and AMD processors by the operating system kernel, enabling KVM to support various operating systems directly on the hardware layer. In the field, various other solutions are available, such as VMware vSphere, Microsoft Hyper-V, and Xen.

To continue with our example, using virtualization gives a slightly different picture, as shown in Figure 3-13. Instead of moving the applications to the consolidated server, we now virtualize the existing systems 1 through 4 on our consolidation target. The effect is clear: not only is the application moving, but its complete operating environment has moved with it.

Taking a closer look, we find other attractive opportunities:

- Consider the three separate systems. To communicate, they require a network infrastructure such as NICs, cables, and switches. If our virtualization system supports network virtualization, this infrastructure is no longer needed. The virtualized systems can communicate using the virtualization system's capabilities, often transferring in-memory data at enormous speed. Performance and energy efficiency increase because the network components are dropped. Once again, this reduces the need for site and facilities resources.
- Each of the separate systems has its own storage system, namely disks. The virtualized systems can now share the disks available to the virtualization system. By virtualizing its storage, the virtualization system can provide optimal disk capacity—in terms of energy efficiency—to the virtualized systems.

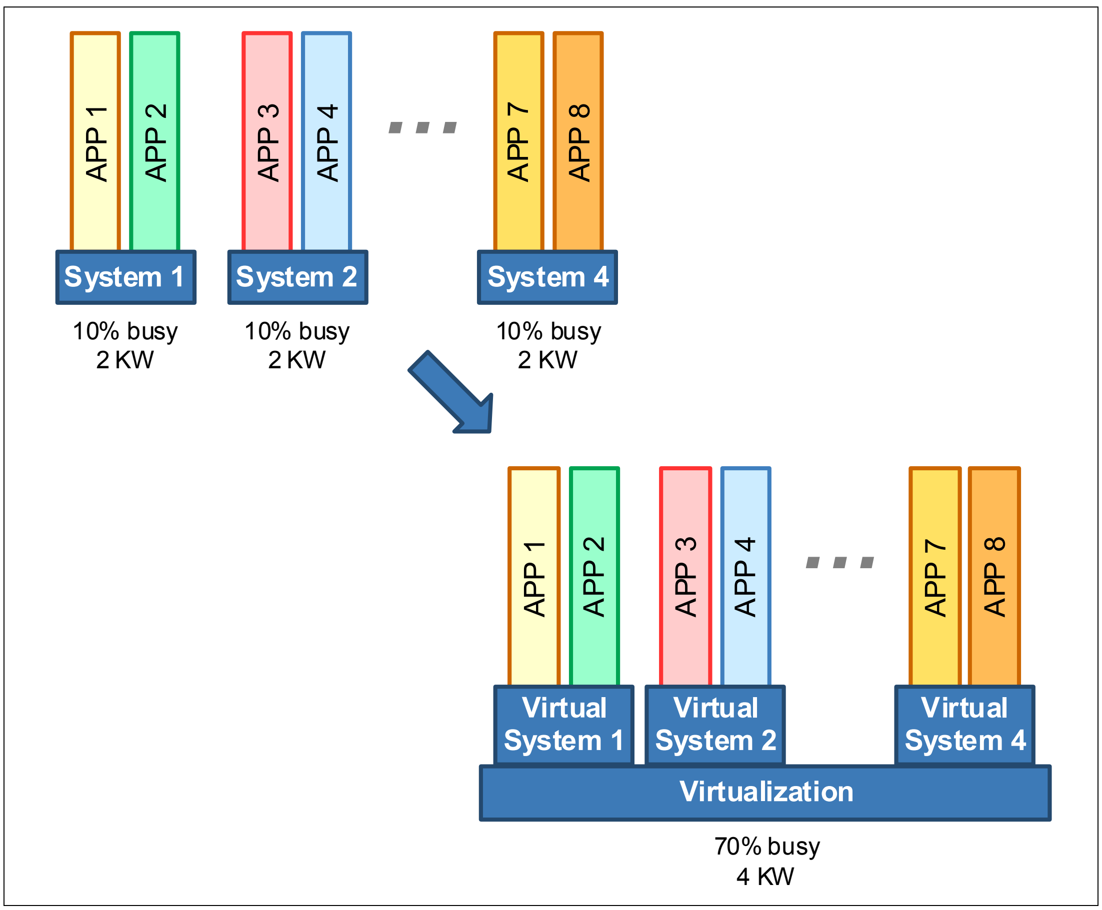


Figure 3-13 Virtualization allows us to consolidate systems the way they are

3.7 Virtualizing servers

This section discusses the IBM techniques that are available for server virtualization, the most attractive approach to consolidation. In many cases, it is the easiest and most effective way to transfer workload from inefficient, underutilized systems to efficient, well-utilized equipment.

It is worth noting that since 1964, the mainframe, with its shared architecture and built-in virtualization, has been the gold standard for server virtualization. System z is highly regarded as a consolidation platform in terms of workload diversity and number of virtualized instances while maintaining high service quality.

3.7.1 Partitioning

Partitioning is sometimes confused with virtualization, but the partitioning feature is rather a tool that supports virtualization. Partitioning is the ability of a computer system to connect its pool of resources (CPU, memory, and I/O) together to form a single instance of a working computer or *logical partition* (LPAR). Many of these LPARs can be defined on a single machine as long as resources are available. Of course, other restrictions apply, such as the total number of LPARs a machine can support. The power supplied to the existing physical computer system is now used for all of these logical systems, yet these logical systems operate completely independently from each other.

Important: LPARs each work independently at the maximum performance of the underlying system. All partitions share the energy provided to the overall system.

LPARs have been available on the IBM System z since the late 1980s and on IBM Power Systems for the last decade. Although the System z and IBM Power System partitioning features differ in their technical implementations, they both provide a way to divide up a physical system into several independent logical systems.

3.7.2 Special virtualization features of IBM systems

Besides providing logical computers, each running an independent operating system instance, IBM systems offer further virtualization features, such as these:

- Micropartitioning: Assignment of fractions of CPUs to virtual machines. In uncapped mode, this assignment can become dynamic.
- IBM HiperSockets™ on System z: A network virtualization technique that allows LPARs to communicate through memory instead of using dedicated networking peripherals. This allows for a very large communication bandwidth compared to normal networks.
- Live Partition Mobility on IBM Power Systems: Moving active LPARs from one physical machine to another without disrupting the respective applications.
- Virtual Ethernet on IBM Power Systems: Similar technique to HiperSockets. LPARs can communicate using the Hypervisor. Instead of using traditional networking peripherals, the communication takes place within the machine.
- Virtual I/O on IBM Power Systems: Use of a virtual I/O server (VIO server) on an IBM Power System for further I/O virtualization. The VIO server can, for example, divide up physical disks in the machine and offer them as single virtual disks to the LPARs. It is also possible to set up a Shared Ethernet Adapter (SEA) on the VIO server which enables Network Connectivity for the LPARs through one interface.

- Multiple image facility (MIF) on System z: Using channel sharing among IBM Processor Resource/Systems Manager™ (IBM PR/SM™) logical partitions. With MIF, multiple logical partitions can directly share channels and optionally share any of the control units and associated I/O devices configured to these shared channels. MIF also provides a way to limit the logical partitions that can access a reconfigurable channel or a shared channel to enhance security. Dynamic Channel Management takes MIF a step further to dynamically adjust the channel configuration in response to shifting workload patterns. It is a function of the Intelligent Resource Director and works together with Workload Manager (WLM) to virtualize I/O paths out to the controllers.
- IBM z/VM® virtualization technology on System z: Using the Processor Resource/System Manager (PR/SM) to divide your System z into as many as 60 LPARs. z/VM virtualization technology allows you to create almost any number of virtual machines on these LPARs. Virtual machines can also host several types of workloads on operating systems such as IBM z/OS®, Linux, IBM z/VSE™, z/TPF or z/VM.

The partitioning technique is based on a *hypervisor*. The hypervisor or *virtual machine monitor* (VMM) establishes a virtual machine to the operating system. In theory, the operating system is unaware of this intermediate layer between itself and the hardware (full virtualization). That is the case for System z. However, for IBM Power Systems, their operating system must be aware of the Hypervisor (paravirtualization). Therefore, AIX and Linux for POWER architecture include the necessary support.

3.7.3 Other virtualization techniques

Other virtualization techniques are available, in addition to partitioning. Popular in the market are KVM, VMware vSphere, Xen and Microsoft Hyper-V. Also, hardware manufacturers extend their products to support virtualization.

Another technique is operating system level virtualization. One operating system on a machine is capable of making virtual instances of itself available as a virtual system. IBM System Workload Partitions (WPARS) on an AIX level exemplify this technique. In contrast to the other techniques, all virtualized systems run on the same operating system level, which is the only operating system the machine provides. This dependency on a single operating system level places obvious limitations on this technique, for instance if we want to consolidate servers running separate operating systems such Windows and Linux.

Attention: In addition to the network virtualization products mentioned in this Redpaper, another popular virtualization technique is to combine related applications on one central server or complex. This allows the networking between them to be done internally at computer speeds rather than network speeds, and it saves the cost of networking hardware and software.

3.8 Virtualizing and optimizing storage

Computer systems are not the only candidates for virtualizing; storage can be virtualized too. This section describes several storage optimization options, including the IBM SAN Volume Controller, which provides a virtual pool of storage consisting of SAN-attached physical storage devices. In addition, the IBM System z platform includes storage virtualization capabilities in its hardware and its Data Facility Storage Management Subsystem (DFSMS™) software under the z/OS operating system.

These devices can be part of an overall Information Lifecycle Management (ILM) solution being offered by IBM. This approach starts from a business perspective and aligns the cost for storing information with the information's value. Its foundation is tiered or virtualized storage. Although ILM is not within the scope of this paper, it should be noted that optimizing your storage landscape by adapting it to your actual needs can be a significant green strategy.

3.8.1 Data deduplication

Storage demands are rising rapidly. At the same time, identical data can often be stored multiple times in the storage infrastructure, resulting in wasted storage capacity.

Data deduplication is a method to reduce the amount of data by consolidating copies or fractions of files. New and existing data is split up into *chunks*. These chunks are then checked for redundancy. If any data duplicates are found, pointers are shifted to reference a single copy. To do this, IBM offers separate products for separate types of implementations:

- IBM Tivoli Storage Manager: Offers client and server-based deduplication processing
- IBM ProtecTIER® and N Series: Offers storage system-based deduplication processing

Establishing data deduplication can be efficient in several ways:

- Reducing storage footprint
- Saving energy as you need less disk capacity to store your data
- Freeing up disk space on the whole storage infrastructure for new data
- Reducing network traffic as mirrored or backup data is not sent multiple times

For further information about data deduplication, refer to the following IBM Redbooks publication:

Implementing IBM Storage Data Deduplication Solutions, SG24-7888

http://www.redbooks.ibm.com/abstracts/sg247888.html

3.8.2 Tiered storage

Depending on the associated technology, storage hardware options can vary considerably regarding efficiency and cost. You can spend a large amount for a fast state-of-the-art Solid State Disk (SSD) or less money for a large, slow-spinning SATA disk.

In a tiered storage environment, data is classified and assigned dynamically to separate tiers. For example, you can use expensive but fast storage components to store often-accessed, mission-critical files while utilizing cheaper but slower storage for less used, non-critical data. Deploying storage media based on the nature of the data to be stored can improve efficiency and reduce costs.

3.8.3 IBM SAN Volume Controller

The SAN Volume Controller (SVC) is a hardware device that brings storage devices in a SAN together into a virtual pool. This makes your storage appear as one logical device to manage. To connected computers, the SVC offers virtual disks as ordinary SCSI devices. On the SAN side, SVC integrates various storage subsystems, even multivendor, and takes care of the correct block mapping between the SAN devices and the virtual disks for the computers. Figure 3-14 illustrates how it works.

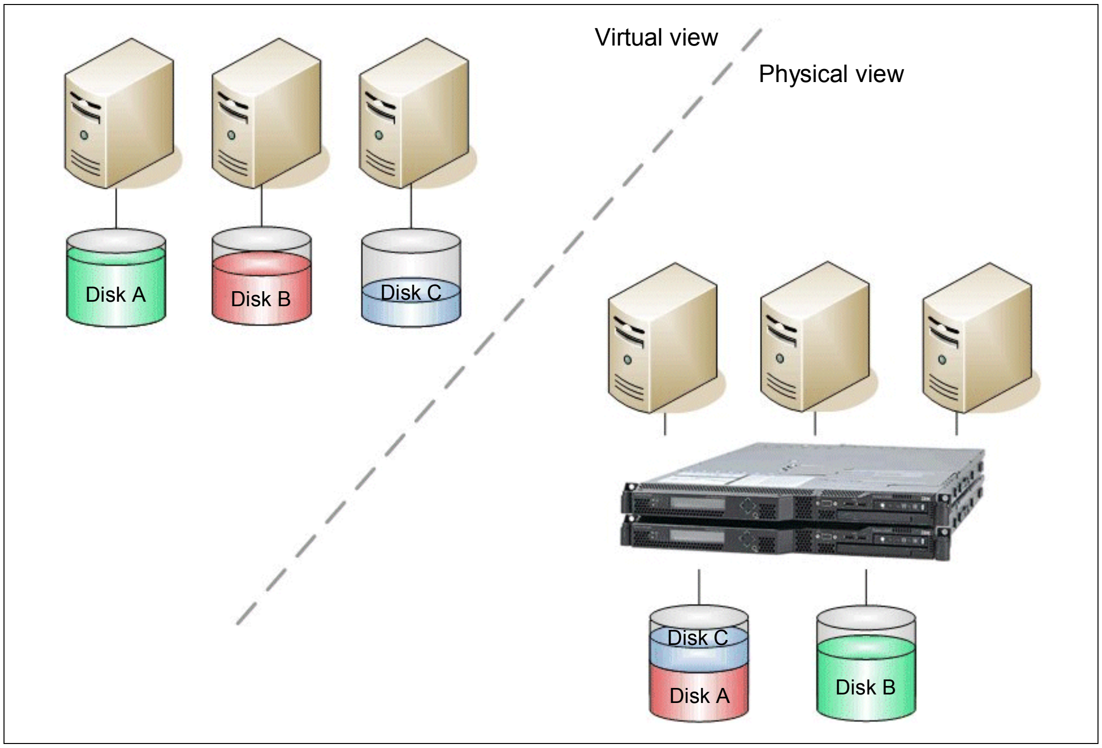


Figure 3-14 Storage virtualization: virtual view and physical view

Obviously, storage virtualization is another tool for consolidation. If underutilized, disks can be virtualized and consolidated in a SAN, and the data can reside in more efficient storage pools.

The SVC supports migration of data among the connected device and to remote sites for redundancy or backup purposes. Also, it can help manage storage hierarchies, where low-activity or inactive data can be migrated to cheaper storage. The integrated cache, on the other hand, is able to improve the performance of lower-tier storage. The data migrations are managed transparently, so they do not interrupt the applications.

The following points make the SVC an attractive tool for an energy efficient storage strategy.

- Data migration from older to newer, more efficient systems can happen transparently
- Tiered storage enables you to use media with a smaller energy footprint, while the SVC cache improves its performance.
- Consolidation of the system individual storage devices to virtual storage has the same effect of increasing storage utilization, as is shown for server virtualization.
- Thin Provisioning - SVC supports space-efficient virtual disks. Logical Unit Number (LUN) capacity will only be used if there is new data written to the device.

Storage virtualization requires more effort than server virtualization, often requiring a reexamination of the existing storage landscape. During consolidation, large amounts of data must be moved from the old systems to the consolidated storage system. This can become a long task that requires detailed planning. However, when done, the effect can be enormous because now storage can be assigned to systems in the most flexible way.

3.8.4 IBM Storwize V7000

With built-in solid state drive (SSD) optimization and a modular design, the IBM Storwize V7000[12] greatly simplifies the addition of virtual storage to your environment. The IBM Storwize V7000 can significantly increase storage utilization and reduce energy costs with high efficiency capabilities such as these:

- Virtualization features increase disk utilization by up to 30%
- Thin provisioning dramatically reduces disk storage needs
- Easy Tier technology coupled with solid-state drives provides high performance at lower costs

3.8.5 IBM XIV Storage System

IBM XIV® is high-end disk storage designed for simplified expansion and integration with multiple business applications including virtualization, database deployments, and detailed analytics. Although the base XIV model provides excellent price-to-performance support for most applications, the XIV Gen3 takes advantage of comprehensive hardware upgrades to deliver up to four times the throughput, giving applications a significant performance boost.

The grid architecture of the IBM XIV Storage System[13] optimizes performance in virtualized environments and integrates seamlessly with cloud technologies, supporting rapid growth. Self-tuning, IBM XIV automates balanced data placement across all key system resources, eliminating hot spots.

3.8.6 Virtual tapes

Looking at the cost of storage, tapes are the cheapest media on which to store data. They offer the largest storage volume at the lowest cost, which is the reason that they are the optimal backup media. Tapes have a long latency compared to disks. This is not always a benefit. Currently, data centers are faced with a time limitation for backing up and probably restoring their data because the time frames for backups shrink while the amount of data to back up expands. For this reason, many sites prefer large disk-based backup systems instead of tapes.

Tape virtualization might be a solution to this problem. A virtual tape server behaves just like a tape library, but a very fast one. This is made possible with internal disk arrays and a migration strategy to export to and import from real tape libraries.

[12] IBM Storwize V7000: http://www.ibm.com/systems/storage/disk/storwize_v7000/
[13] IBM XIV Storage System: http://www.ibm.com/systems/storage/disk/xiv/

The IBM Virtualization Engine TS7700 is a mainframe virtual-tape solution that optimizes tape processing. It has a fully integrated tiered storage hierarchy of disks and tape and can take advantage of the benefits of both technologies. Deploying this innovative subsystem can help reduce batch processing time, total cost of ownership, and management overhead. For example, if the batch process is constrained, the TS7700 Virtualization Engine's processing power, cache capacity, and support for 4 GBps IBM FICON® attachments can help alleviate bottlenecks and reduce the batch window. Another system, the TS7520 Virtualization Engine, is a general-purpose system for non-mainframes.

3.9 Integrating energy and systems management

In 3.5, "Managing power on the software side" on page 40, we introduced Active Energy Manager as the tool of choice for all IBM platforms. However, in an overall systems management environment, power management is only one aspect of many. This section discusses how to integrate AEM into the IBM Tivoli systems management platform.

IBM Tivoli Monitoring is the tool to monitor the status of your IT environment. It allows you to monitor your operating systems, databases, and servers throughout distributed environments with a flexible, customizable portal. A monitoring agent sits in place to tap the desired data from the monitored system.

From the Open Process Automation Library (OPAL), the IBM Tivoli Monitoring Power Management Agent is available. This agent constitutes the interface between AEM and IBM Tivoli Monitoring. With this agent in place, IBM Tivoli Monitoring can monitor the power, temperature, and CPU usage of the respective systems. Figure 3-15 shows how the components interact.

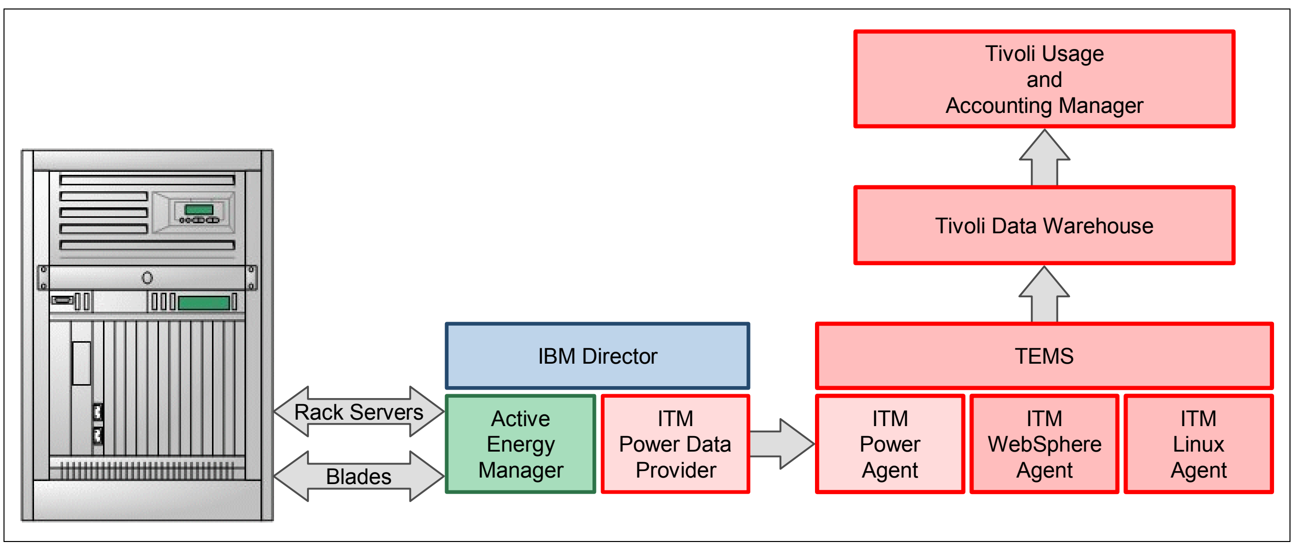


Figure 3-15 Integration of AEM with Tivoli systems management tools

Having this entry point into the Tivoli environment enables you to employ all the well-known features of IBM Tivoli Monitoring and other tools with which it interacts. You can also add the performance aspect discussed in section 3.2, "Elements of energy efficiency" on page 32.

Optimizing for power and performance might include the following scenarios:

- Reprovisioning a server based on the machine's environmental temperature or overall rack power consumption to another rack in a cooler area of your data center. For a temperature alert in IBM Tivoli Monitoring, trigger the reprovisioning in IBM Tivoli Provisioning Manager.
- Power capping a single server having a temperature problem, perhaps because of an obstructed airflow, until the problem is solved on-site.
- Feeding power, temperature, and CPU usage data into the IBM Tivoli Data Warehouse. The IBM Tivoli Usage and Accounting Manager, can correlate with accounting data. Charge the IT users according to their CPU and correlated power usage.

Multiple opportunities emerge after the AEM data becomes available to the Tivoli environment. For additional details about how Tivoli monitoring can be used to integrate views across both your data center facility and IT operations, see Chapter 5, "Integrating across IT and facilities for the bigger picture" on page 91. Due to the flexible nature of the Tivoli toolset, user setup might be complex. IBM services can help you find the best-fit solution.

3.10 Looking to the cloud

IBM defines *cloud computing* as a new consumption and delivery model of IT resources inspired by consumer Internet services. Enabled by virtualization, (service) automation, and standardization, cloud computing enables self-service, sourcing options, and economies-of-scale.

In essence, cloud computing is an architectural framework by which one or many organizations can deploy, manage, and retract any workload, public or private. Cloud computing addresses business needs from a self-service, automated workload perspective. The concept collectively addresses all the aspects of modern computing, from components (SAN, network, servers, or software) to implementations (virtual desktops, hosted applications, email, and so forth) in a comprehensive, cohesive solution.

Cloud computing provides flexibility for both the customer and the provider. The cloud customer can both acquire and release resources when needed to quickly adapt to their changing business needs. Multiple delivery models, including Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), empowers customers to select the solution that best satisfies their requirements.

As you move up the ladder from IaaS to PaaS to SaaS, the level of deployment standardization increases. Standardization of the underlying delivery architecture enables the cloud customer to focus on the requirements that are unique to their business while expediting deployment time and greatly reducing operational costs. A solution that necessitates the deployment of a unique, physical infrastructure might require weeks or even months to deploy on its own. The same solution can be deployable in hours using a standardized environment provided by the cloud. The greater the level of allowable standardization, the greater the associated efficiency.

The cloud model also facilitates greater efficiency within the data center. Separating the physical hosting environment from the delivered service empowers the data center operator to select standardized, cost-effective, and energy-efficient hardware. Standardized logical environments in the cloud are highly virtualized, resulting in high utilization rates, reduction in energy costs, and support requirements.

A standardized hosting environment also enables the rapid movement of workloads within a data center and even between data center sites in response to any number of events and criteria, such as extreme weather conditions, utility power disruptions, equipment failures, lower utility costs, availability of renewable energy sources, the geographical location of end users, and so on. This ability to rapidly adapt to changing conditions and requirements translates to increased resiliency and reliability.

The ability to rapidly shift workloads based on operational and environmental conditions is a perfect example of a benefit that can be gained by obtaining integrated views of your data center that span both facility and IT operations. Changes in facility conditions can prompt a coordinated shift in IT operations. For more information about this integration, see Chapter 5, "Integrating across IT and facilities for the bigger picture" on page 91.

IBM provides all types of cloud solutions, from IaaS to SaaS, and operates and manages both public and private cloud environments, providing immense solution flexibility. IBM also provides comprehensive services to help you develop your cloud computing strategy, determine which workloads are good candidates for a cloud implementation, and transition to the most efficient cloud operation.

Although not all workloads are ideal candidates for the cloud computing model, the use of cloud computing services where appropriate can result in substantial improvements in your business operations, and IBM can help you achieve them.

For detailed information about all of the IBM cloud computing services, visit this site:

http://www.ibm.com/cloud/.

3.11 Where to start

First, measure energy, cooling, and inlet temperatures where possible. Trending analysis can help identify efficiency needs and opportunities. Find out what tools your systems provide. The objective is to make your energy consumption measurable. This is the most important step for all future decisions.

Aggressively drive consolidation of servers and storage using virtualization. When you are prepared for a virtualized infrastructure, consolidation comes easy. Ask the following questions:

- Which servers should be replaced in the near future? Which of the new, energy-efficient systems would be able to take on this workload because they are underutilized?
- Which systems, if any, are complementary loaded? Consider the following points:
  - Low CPU and high memory usage can nicely coexist with high CPU and low memory usage systems in a partitioned environment.
  - Look for non-overlapping workloads such as day versus night jobs.
- Is any system completely filled with non-business-critical work (lowering performance or response time)? If yes, virtualize it. Pair it with another virtual system on which you run occasional peak loads, such as weekly or monthly balancing. By setting appropriate

entitlements, the peak load can displace the non-critical load while the overall system usage stays close to 100%.

- Are there individual *infrastructure servers*, such as Lightweight Directory Access Protocol (LDAP) servers, domain name servers (DNSs), and licensing servers? These servers are often built on a single system that has a low utilization. If so, virtualize and consolidate to get rid of these inefficiencies. Virtualization in this case also helps in reducing the cost of redundancy. Move the virtualized image to another machine instead of providing a hot standby

  Use virtualized storage for new and virtualized servers. Although this can require start-up costs, it pays back quickly. Instead of having terabytes of unused disk space being turned around in individual machines, consolidate and *right-size* your storage. Moving a virtual operating system from one machine to another is as simple as connecting the target machine with the virtualized disks

4

Optimizing your site and facilities

This chapter describes the various components of the site and facility infrastructure, the types of infrastructures that are available, and why they are an integral part of your data center.

To be energy-efficient, the data center must utilize a high-efficiency infrastructure and exploit best practices. The life span of the infrastructure is three to five times longer than most IT hardware and requires a much larger investment. Thus it is important to understand the timing of updating your infrastructure. This chapter also offers tips for utilizing best practices to increase the efficiency of your infrastructure.

The site and facility infrastructure can be divided into the following five sections, each of which is dependent on the other:

- Data center cooling
- Heating, ventilation, and air conditioning (HVAC)
- Uninterruptible power supply (UPS)
- Power
- Standby generator or alternative power sources

IBM Business Partners, such as APC Schneider, Emerson Liebert, Eaton, Coolcentric, Rittal, Panduit, and Anixter, also have a variety of energy-efficient solutions that we discuss in this chapter.

4.1 Increasing the efficiency of your data center infrastructure

Making your data center smarter is all about efficiency and saving energy. What infrastructure can you change? What best practices should you employ to become more energy-efficient and reduce your energy consumption? How should you integrate your IT and facilities to gain and sustain efficiency?

4.1.1 Reducing power consumption with innovative technologies

Applying innovative technologies within the data center can yield more computing power per kilowatt. IT equipment continues to become more energy-efficient. With technology evolution and innovation outpacing the life expectancy of data center equipment, many companies are finding that replacing older IT equipment with newer models can significantly reduce server power and cooling requirements and free up valuable floor space. For example, IBM studies have demonstrated that blade servers can reduce server power and cooling requirements by 25-40% over 1U technologies.

Although replacing equipment before it is fully depreciated might seem unwise, the advantages that new models can offer (lower energy consumption and two to three times more computing power than older models), combined with potential space, power, and cooling recoveries, are usually enough to offset any lost asset value.

IBM Business Partners such as APC[1],Eaton[2] and Emerson Liebert[3] all can provide innovative green solutions focusing on power, UPS systems, and cooling, along with facilities monitoring and management. IBM technology partners, such as Panduit[4] in conjunction with Anixter[5], provide data cabling solutions. These partners also develop data center solutions that help customers optimize their physical infrastructure through simplification, increased agility, and operational efficiency. For instance, companies such as GE[6] design lighting solutions to save energy.

1 http://www.apc.com
2 http://www.eaton.com/EatonCom/ProductsServices/index.htm
3 http://www.emersonnetworkpower.com/en-US/Brands/Liebert/Pages/default.aspx
4 http://www.panduit.com/Solutions/DataCenter
5 http://www.anixter.com/
6 http://www.ge.com/products_services/lighting.html

4.1.2 Addressing environmental condition variations

Thermal measurement of the data center is a major step when looking to improve energy efficiency. IBM Measurement and Management Technologies (MMT), provides a powerful solution which can generate dramatic savings. Figure 4-1 shows you the MMT life cycle, starting with real-time and high resolution measurement gathering, followed by measurement-based modeling to provide the foundation for detailed environmental management and control for data center cooling optimization.

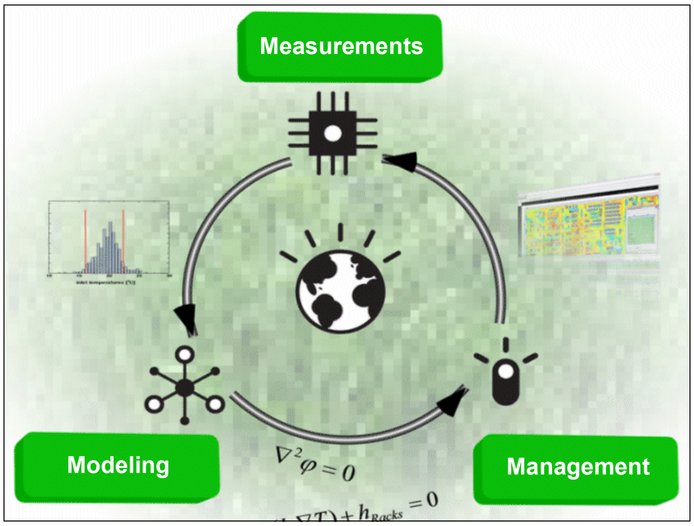


Figure 4-1 MMT life cycle

How it works

You can be initially capture a data center thermal snapshot with a mobile measurement cart, which collects thousands of temperature and humidity data points at multiple elevations as it moves through the data center. Based on this data, a physics-based modeling technique generates a 3-D thermal map. Using this map, you can easily identify environmental inefficiencies and optimize your data center floor space.

MMT 1.5 brings additional functionality and value. Reading from a network of wired or wireless temperature, humidity, flow, and pressure sensors installed throughout the data center, environmental conditions are fed continuously into the data center MMT model, generating a near-real-time dynamic environmental map. Connectivity to Active Energy Manager and the data center building management system can provide additional data about IT and facility hardware, and support for additional sensor types can provide insight into air particulate concentrations and more.

How you can profit

MMT delivers a consolidated and integrated view of your data center. Hence you can start managing it from an overall perspective:

- Optimize floor space (racks and systems)
- Reduce hotspots
- Improve air conditioning flow
- Optimize air conditioning use
- Increase chiller set point
- Remove plenum temperature variations
- Optimize server workload
- Dynamically monitor data center environmental conditions
- Activate alerting services

Figure 4-2 shows a before-and-after thermal map. After powering off one chiller, the surrounding zone warms but still maintains a temperature well within operational tolerances. Also depicted is a version of the MMT cart, which can be used for obtaining initial measurements.

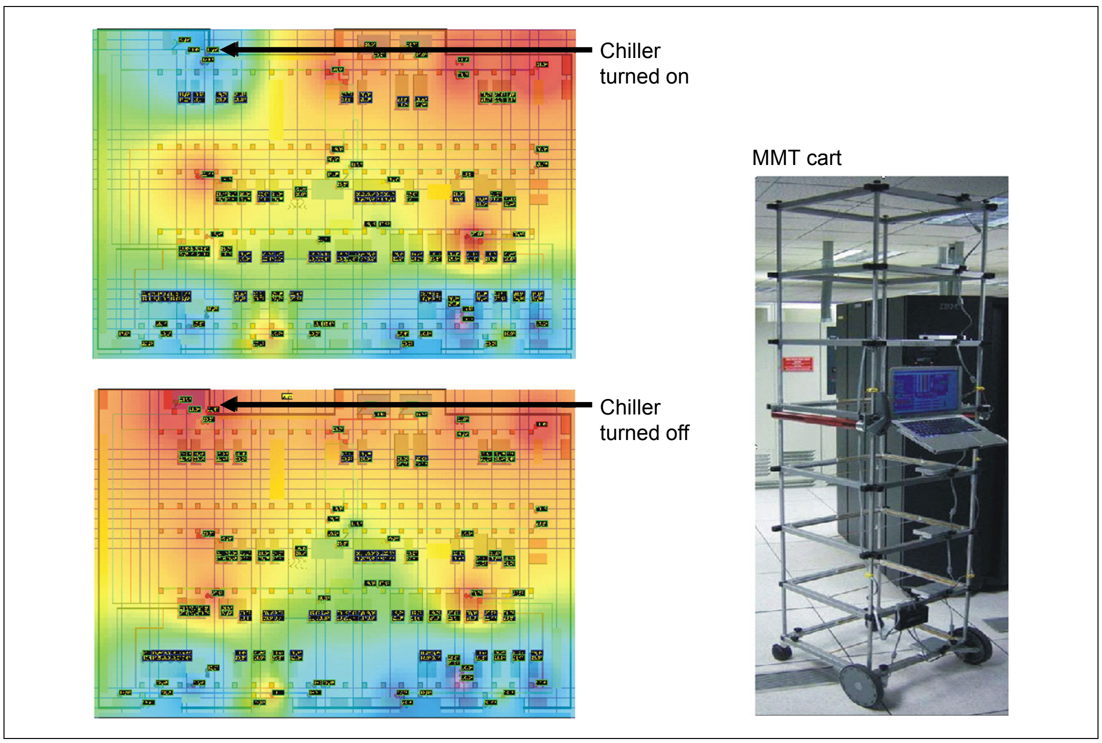


Figure 4-2 MMT thermal maps and measurement cart

MMT is available as a service from the IBM Systems and Technology Group (STG) Lab Services team. It is also integrated into the comprehensive Tivoli product line with IBM Maximo® Asset Management for Energy Optimization.

Additional information about MMT and related services provided by IBM STG Lab Services can be found at the following website. For more information about Maximo Asset Management for Energy Optimization and its integration with IBM Tivoli Monitoring for Energy Management, see Chapter 5, “Integrating across IT and facilities for the bigger picture” on page 91.

Measurement and Management Technologies for Data Centers

https://researcher.ibm.com/researcher/view_project.php?id=2377

4.1.3 Reducing cooling requirements

You must consider a number of factors when you develop a plan for improving power and cooling efficiency by reducing the heat generated in the data center. Improvements in rack and room layout can increase energy efficiency with a relatively low initial investment. Consider the following improvement opportunities:

- Organizing IT equipment into a hot aisle and cold aisle configuration.
- Right positioning of the IT equipment so you can control the airflow between the hot and cold aisles and prevent hot air from recirculating back to the IT equipment cooling intakes.
- Taking advantage of supplemental cooling options, such as water or refrigerant heat exchangers.
- Improving rack cooling efficiency by employing a rear door heat exchanger or an enclosed racking system to dissipate heat from high-density computer systems before it enters the room. Similarly, relatively simple airflow management improvements can boost energy efficiency:
  - Take advantage of the current capacity by clearing under-floor blockages and implementing effective cable management.
  - Ensure that floor openings match the equipment thermal load by adding or removing vented tiles at the equipment air intakes.
  - Consider adding ducted returns. Ultimately, companies must consider organizing their data centers into thermal zones, assigning a defined set of IT equipment and floor space to specific HVAC or CRAC units. This type of space and thermal planning will eliminate hotter areas of the room (hot spots) that challenge cooling systems and enhance system reliability by helping to avoid heat-related hardware failures. Also you can avoid cool spots.
  - Visualize and understand the thermal profile of your data center with IBM services such as MMT. Also, recommendations for diagnosed inefficiencies will result in further energy and cost savings.

Companies such as APC and Emerson Network Power have worked as Business Partners with IBM to devise cooling solutions for energy efficient data centers.

4.1.4 Improving physical infrastructure

Energy efficiency for infrastructure equipment has significantly improved in recent years. Replacing chiller or UPS systems that have been in service for 15 years or more can result in substantial savings. New best-in-class UPS systems can operate with as much as 70% less loss than existing UPS equipment. New chiller systems can improve efficiency by up to 50%. New chiller plants also can be installed with variable-speed drives, reducing pumping system energy usage and allowing better integration of the liquid cooling system into the chilled water infrastructure. Water-side economizers, which utilize outside air to directly cool the chilled water, can further reduce the energy required to cool the data center.

The capacity and efficiency of chilled water systems can be augmented with thermal storage systems that store energy generated at night, when chillers typically operate more efficiently, and then release this energy during the day, when energy costs are higher.

Cold air delivery to the data center also can be made more efficient, either through central HVAC systems or through CRAC units with variable speed drives. Central HVAC tends to be more efficient, because the systems are larger and more amenable to taking advantage of no-cost cooling when outside air temperatures are sufficiently low to provide some or all of the cooling requirements. CRAC units, alternatively, provide greater flexibility in managing the data center.

Even without upgrading facilities equipment, companies can save energy and gain cooling capacity by relaxing stringent relative humidity and temperature requirements for their data centers. As these specifications are usually driven by the presence of hot spots, understanding the thermal profile of your data center and removing those hot spots (and cold spots) will allow temperature and relative humidity requirements to be relaxed, helping to reduce the energy required to operate the data center.

In addition to cutting back on power usage inside its data center, a company can also reduce its carbon footprint[7] by taking advantage of options for more eco-friendly sources of power. Integrating renewable energy into the power supply, including solar, wind, hydro, and bio-mass generated energy, is an efficient way to reduce dependency on fossil fuels. Companies in certain geographic areas are eligible for government rebates for deploying alternative energy resources, and companies with the flexibility to relocate or plans to open new data centers are choosing locations that are rich in renewable energy sources as part of their corporate environmental strategy.

A well-organized cabling infrastructure can simplify data center management while facilitating energy efficiency. IBM, along with its Business Partners such as Anixter, are offering structured cabling solutions for new or reorganized data centers. As the cabling infrastructure is the core of every data, voice, and multimedia network, it is also a key to data center availability and manageability.

Also, increased efficiency can be as close as the light switch. Occupancy-based lighting controls combine simple implementations with quick efficiency returns.

[7] The amount of carbon dioxide emitted. See http://www.carbonfootprint.com/

4.2 Optimizing cooling across the data center

By addressing several of the most common cooling issues first, you will regain cooling capacity with a relatively low up-front investment. There are a number of opportunities for improving cooling efficiency within the data center. Improvements to airflow management, rack and room layout, and localized cooling all can increase energy efficiency with relatively low up-front investments.

> **Note:** Certain parts of this content are taken from a publication by IBM Fellow, Dr. Roger Schmidt, presenting a CIO's guide to *The green data center.* This publication can be found at the following address:
>
> http://www-07.ibm.com/systems/includes/content/optimiseit/pdf/CIO_Guide_to_Green_Data_Center.pdf

4.2.1 Managing airflow

Where possible, avoid mixing hot and cold air. To increase airflow efficiency, you must have a clear path for the cool air to travel under the raised floor to get to the loaded areas. Above the raised floor, allow a path for the hot air to return back to the CRAC units. The following tips can help improve the airflow management:

- Add or remove vented tiles.
  - Remove vented tiles from hot aisles and open areas.
  - Add tiles to heavy heat load areas.
  - Lower the vented tiles with dampers to deliver smaller amounts of air in low heat areas, and open them fully for high heat areas.
  - Replace any tiles that have unused cutouts with full tiles.
- Improve airflow through racks.
  - Where possible, prevent the mixing of cold and hot air. Install blanking plates in racks that have empty space.
  - Provide large openings in racks for the cool air to bypass the heat load generated by the servers and for hot air to cycle back to the CRAC unit. Air always takes the path of least resistance. This leaves cold air in the rack to be drawn through the server again.
  - Populate unused module bays in the rear of BladeCenter chassis with appropriate blanking plates/fillers.
- Seal cable cutouts and penetrations.

  Cutouts in raised floors affect the air distribution and reduce the static air pressure under the floor. Block the openings by using brushes, foam, or pillows. This will allow more air to get to where it is needed.
- Clear under-floor obstructions.

  Excessive under-floor obstructions can lead to an increase in static pressure. High static pressure can have a reverse impact on the airflow under and above the raised floor.

  Remove under-floor obstructions such as:
  - Unused cables and wiring
  - Unused under-floor equipment or communication boxes

Figure 4-3 shows an open cable cutout. In this situation, air is being released, reducing under-floor static pressure and efficiency.

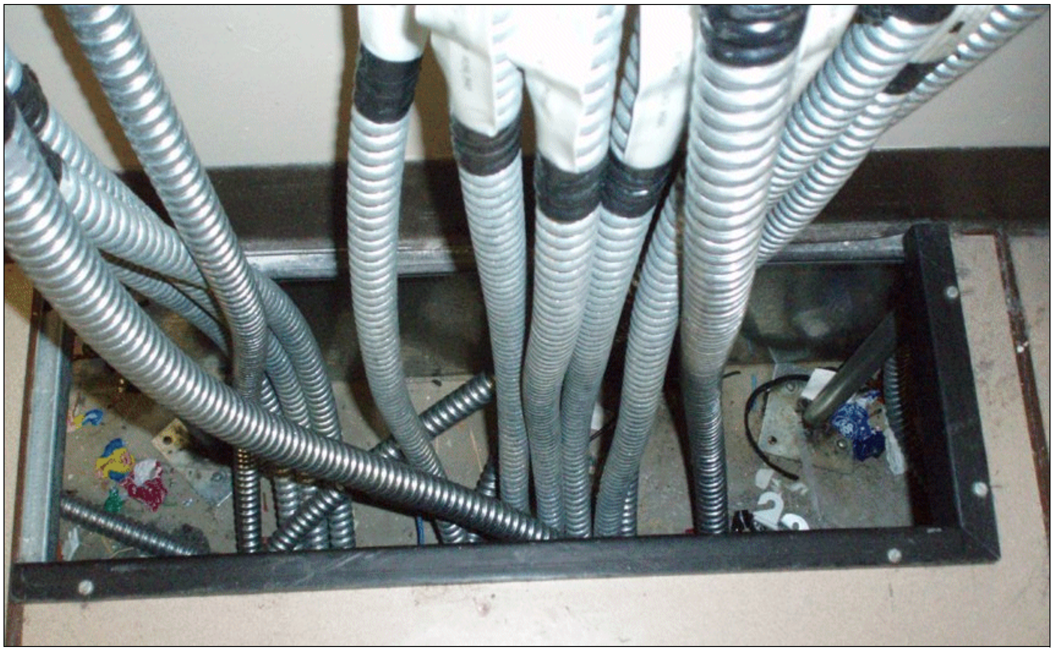


Figure 4-3 Open tile cutout

Figure 4-4 shows pillows used to seal the cable cutout. This solution dramatically reduces cool air escaping into unnecessary areas.

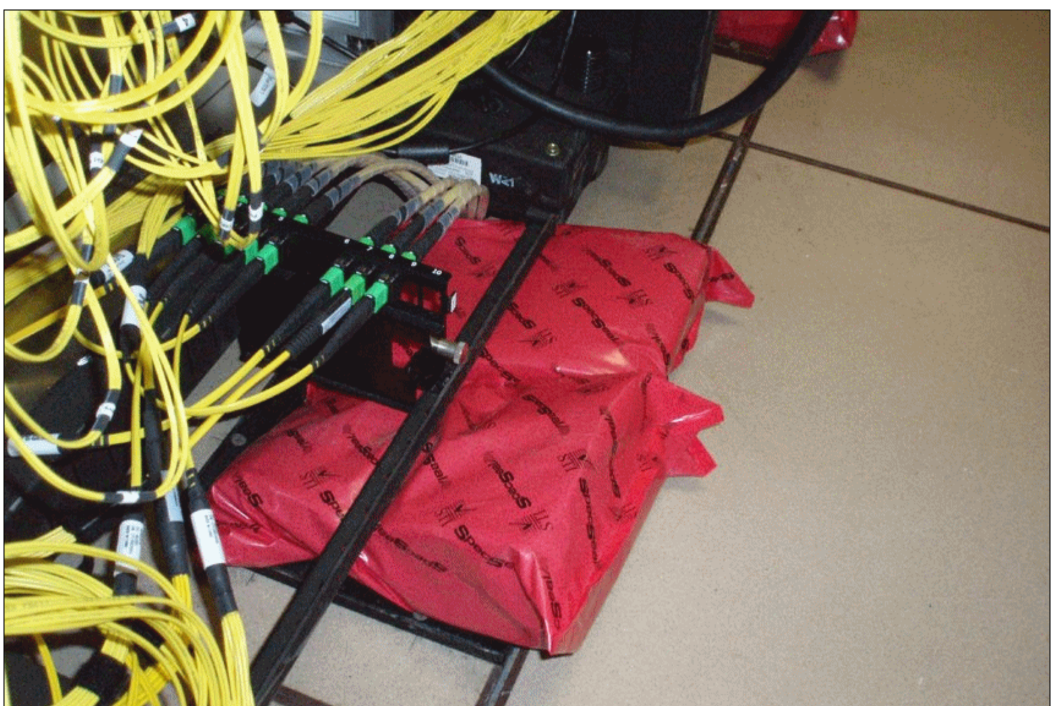


Figure 4-4 Sealed tile with pillows

Figure 4-5 shows excessive cabling obstructing airflow.

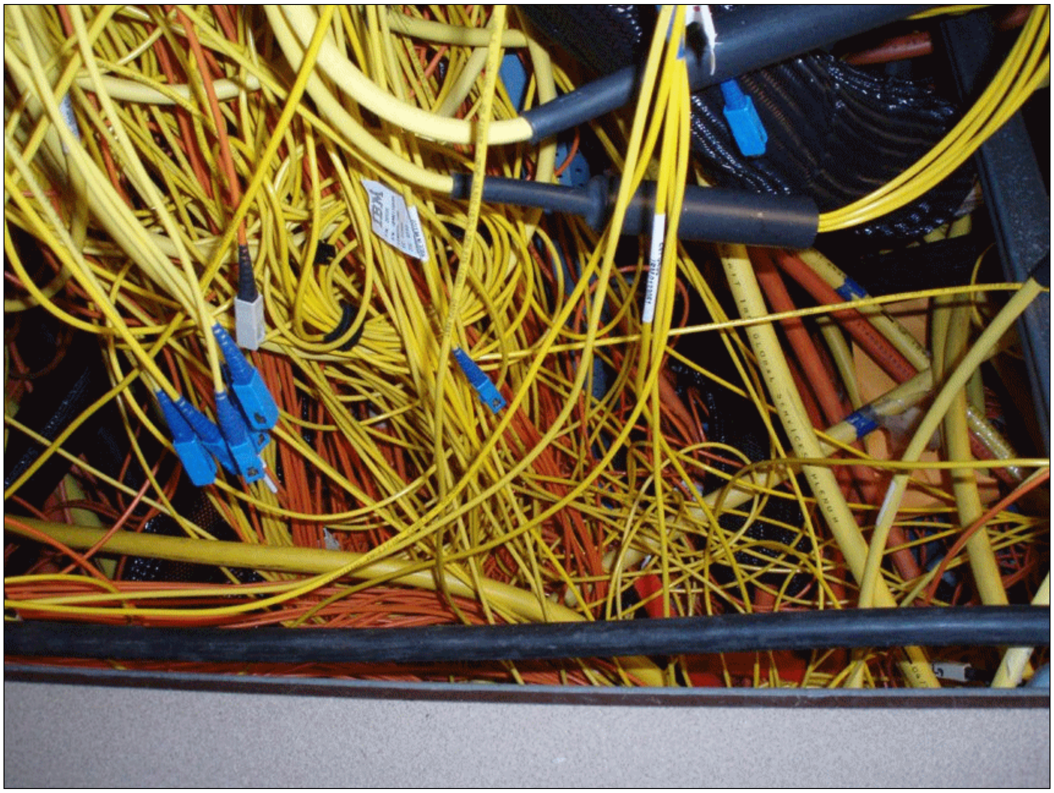


Figure 4-5 Under-floor blockages

4.2.2 Implementing structured cable management

Using overhead and vertical cable management trays helps reduce the number of cables under the raised floor. As new cables are installed, they should be placed in overhead trays, and obsolete cables should be removed. Structured cable management, as shown in the before and after images included in Figure 4-6, not only facilitates better airflow, but it also simplifies operational management and problem determination.

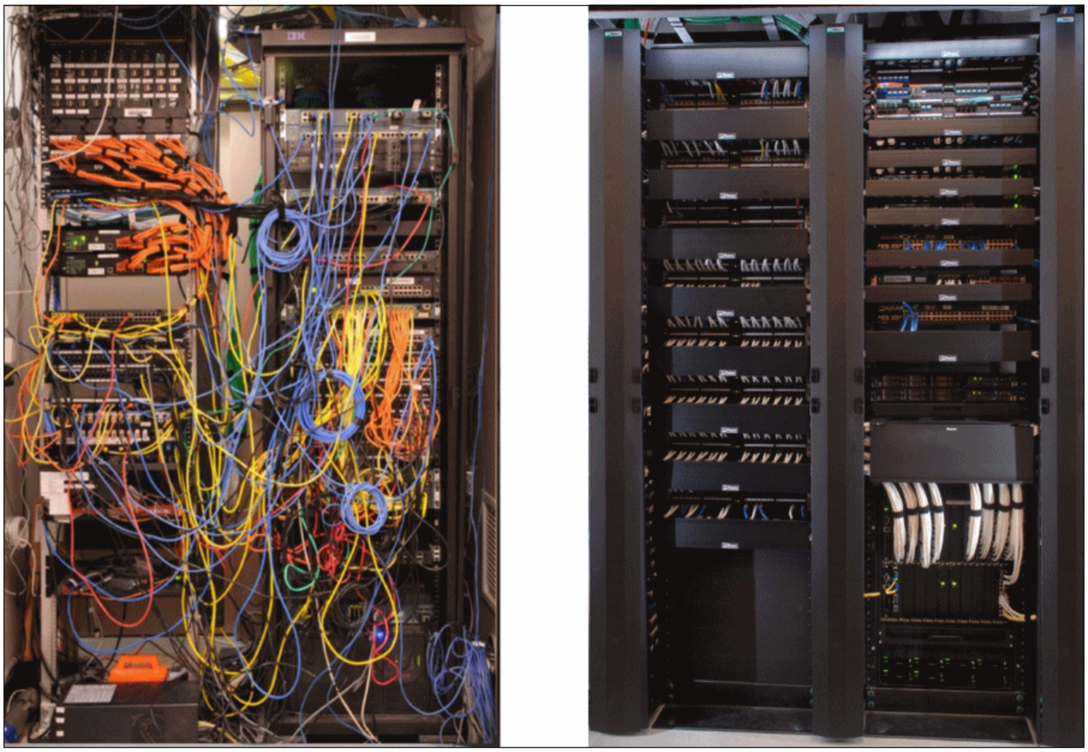


Figure 4-6 Before and after images of structured cable management

Note: All IBM systems facilitate efficient cable management, including the IBM System z, which supports overhead cabling for greater flexibility.

4.2.3 Considering raised floor height recommendations

Current raised floor modeling suggests a height to support at least 600 mm (24 in) of unobstructed space, enabling a clear path for cool air to travel. Some new raised floors are 900 mm (36 in) high, which allows additional quantities of air to circulate for extreme cooling requirements. Figure 4-7 shows such a higher under-floor arrangement.

For low-raised floors, such as 150-300 mm (6-12 in), do not place equipment close to CRAC units because low airflow or reverse flow can occur from the perforated tiles.

Under-floor partitions can also be placed under the raised floor to direct air into the desired areas.

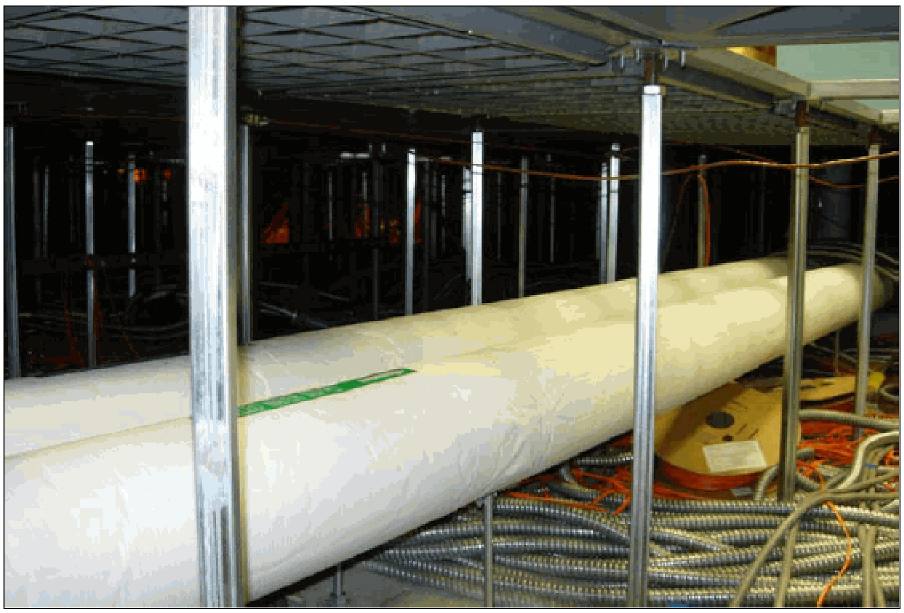


Figure 4-7 Under-floor view with chilled water pipes

4.2.4 Exploring insulation options

Consider the following measures you can implement for data center insulation:

- Insulate the data center walls and ceiling.
- Seal penetrations at the perimeter of the data center.
- Install double glazed windows.
- Install weather strips on doors.

These simple steps can assist in maintaining the temperature and humidity set-points of the data center and increase thermal efficiencies.

4.2.5 Configuring hot aisles and cold aisles

A hot aisle and cold aisle configuration enables much better airflow management on the raised floor for both hot and cold air. This type of configuration offers an opportunity to route hot and cold airflow using independent air corridors that minimize air mixing and improve efficiency.

Note: It is not always practical to move existing equipment; however, great efficiencies can be gained if doing so is possible.

Figure 4-8 shows a hot and cold aisle configuration, as recommended by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE).

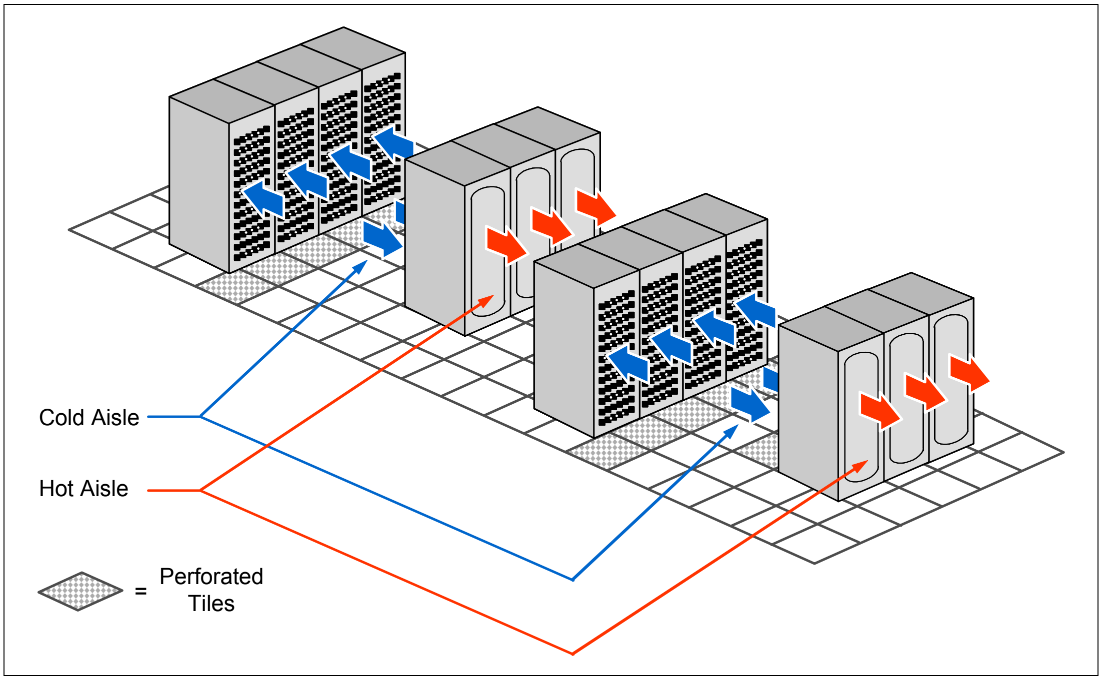


Figure 4-8 Hot and cold aisle configuration

Figure 4-9 shows the thermal flow benefits achieved by hot and cold aisles. The hotter the air returning to the CRAC unit, the greater heat transfer that is achieved, increasing the efficiency of the CRAC unit.

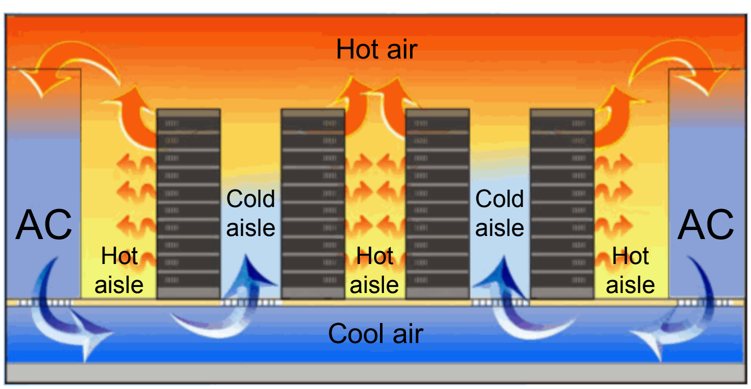


Figure 4-9 Thermal flow of hot and cold air

If flexibility exists, locate CRAC units facing the hot aisle rather than cold aisles. The under-floor velocity pressure will be maximized in the cold aisles for this layout versus locating the CRAC at the end of the cold aisle.

Hot aisle and cold aisle configurations can be supplemented with additional airflow management options, which can help prevent recirculation of the hot exhaust air into the inlets of the servers.

Stationary: In this solution, hardened panels are assembled from the ceiling onto IT racks to create a door-accessible enclosure, thus separating cold zones from hot zones. In this type of solution, rearranging IT racks becomes difficult.

Hot air chimneys: Typically in this solution, chimneys are tightly fixed onto IT racks to channel hot air into return air ducts. This avoids mixing of hot air with cold air. Refreshing your IT racks can become a challenging task when deploying this solution.

PVC curtains: In this solution, typically PVC curtains are used to separate hot and cold aisles. This is a low-cost solution that can be quickly and easily deployed. It is also easily configurable, making rack replacements and rearrangement simpler. However, you must verify the fire ratings of the curtains before implementing this kind of solution.

Figure 4-10 depicts an IBM data center with a hanging partition solution for managing the hot and cold aisles.

Figure 4-10 IBM SWG India data center

> **Note:** It is important to be aware of local fire codes and to contact your insurance providers before implementing airflow containment options within your data center.

IBM Measurement and Management Technologies (MMT), discussed in 4.1.2, “Addressing environmental condition variations” on page 59, can be a vital tool in establishing your initial aisle configuration, in the decision to add supplemental airflow management options, and in the ongoing operation of your cooling infrastructure.

Figure 4-11 depicts the MMT thermal profile of a data center before deployment of hot and cold airflow management. Without containments, the cold air is being circulated all over data center, reducing CRAC efficiency.

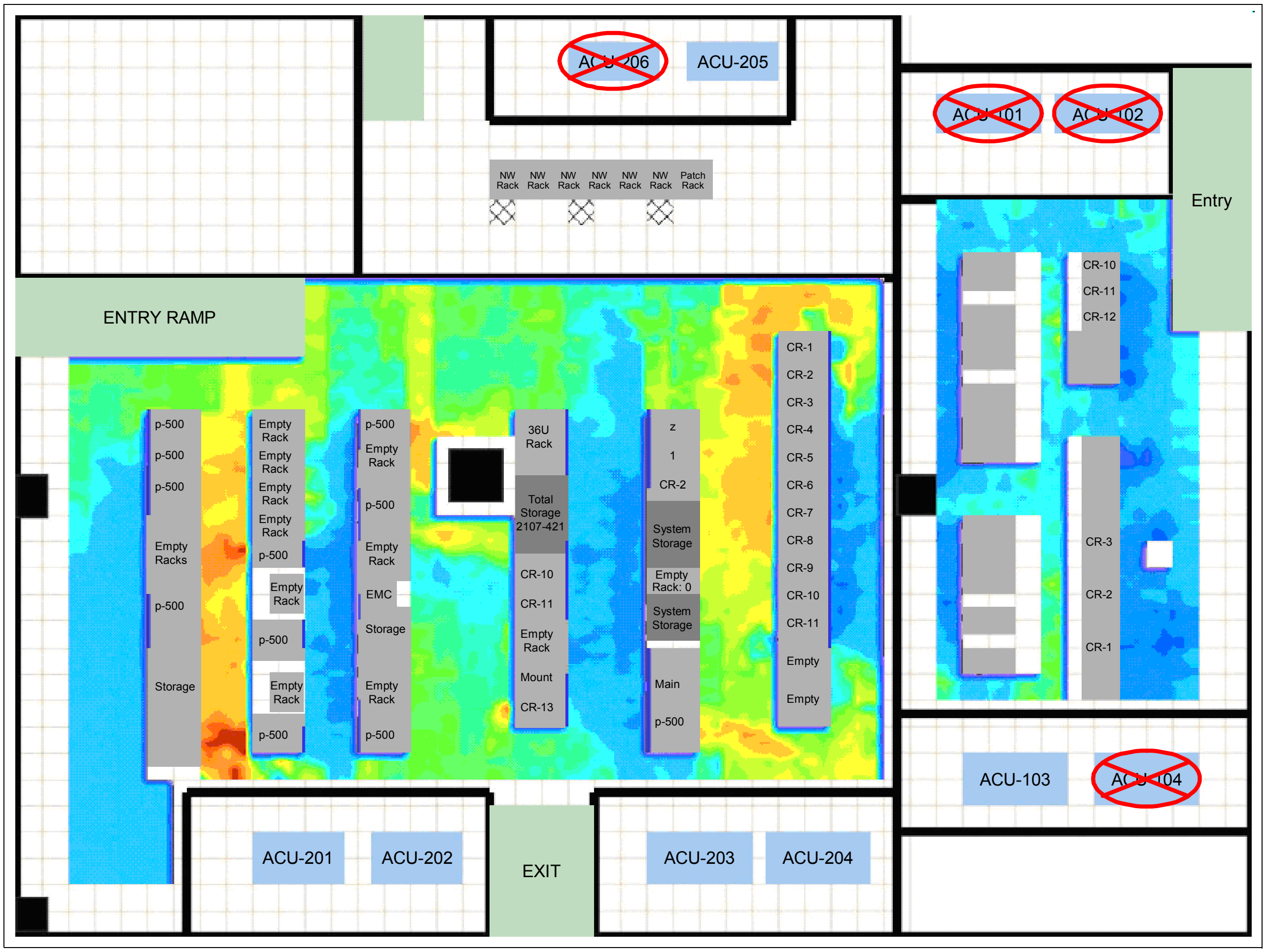


Figure 4-11 MMT scanning results before hot and cold airflow management deployment

Figure 4-12 shows the thermal conditions after the deployment of additional airflow management. With cold aisle containment in place, the cold air is only circulating where it is required. This reduces stress on the CRAC units and increases their running efficiency.

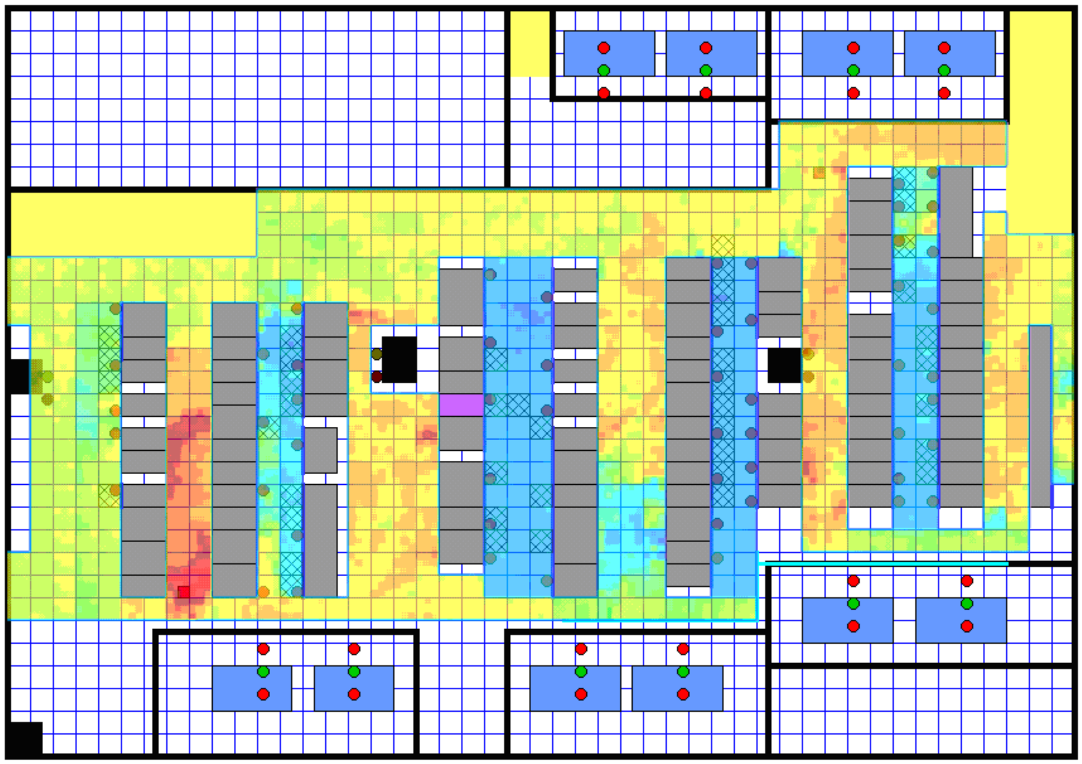


Figure 4-12 MMT scanning results after hot and cold airflow management deployment

4.2.6 Controlling direct rack air flow

Another option for cooling optimization is to manage the flow of chilled air directly.

The classic approach is to distribute chilled air through cooling vents on the raised floor in front of every rack. This can result in an unfavorable air, humidity, and temperature profile as cold air is intermixed with hot air. This typical configuration can require large amounts of cold air to be blown toward the racks to ensure adequate cooling for systems in the upper levels, which is very energy-inefficient.

One solution to this problem is to use an air duct to directly control the flow of cold air to the rack. The air inlet duct provides a controllable and balanced air flow at a constant temperature and humidity to the computer equipment installed in the server rack.

> **Important:** Because of the constant temperature all over the rack, it is possible to raise chilled air temperature supplied by the air conditioning. This translates to tremendous savings in cooling costs.

Figure 4-13 shows a drawing of a rack air inlet duct and the installation at the IBM Research and Development Data Center in Boeblingen, Germany.

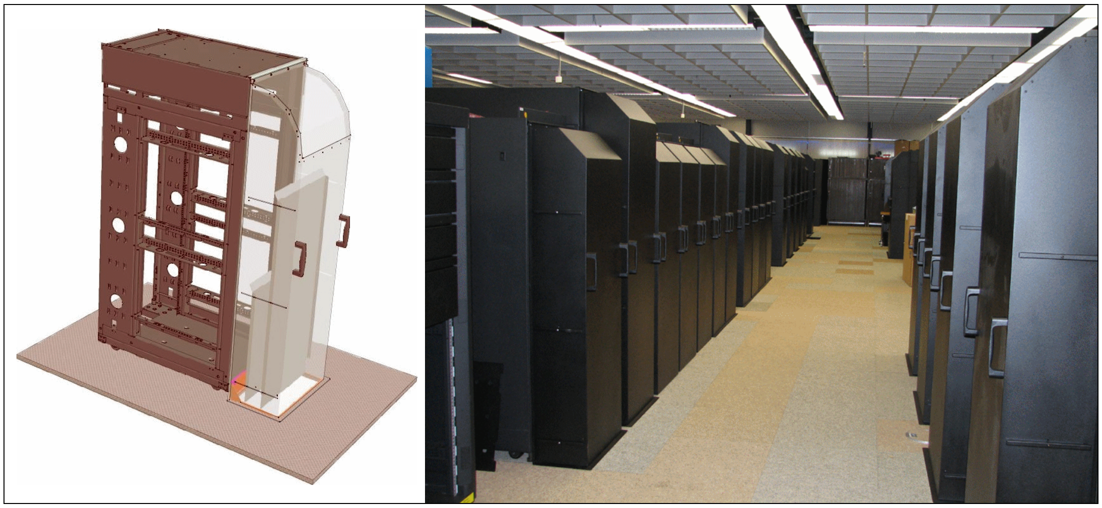


Figure 4-13 Air duct installation at IBM Research & Development center at Boeblingen

The improvements in the environment include:

- Consistent air, temperature, and humidity spread on the whole rack
- Extended hardware life expectancy due to lower operating temperatures
- Increased energy efficiency due to reduced chilled air leakage and reduced intermixing of cool and hot air
- No active components
- Noise reduction
- Easy to install
- Adaptive air flow control

4.3 Localizing cooling

Traditional IT facility cooling designs utilize multiple CRAC units that are located around the perimeter of the data center. The cold supply air from the CRAC units must travel the distance from the unit to the server racks and then return back to the CRAC units.

To resolve hot spot issues raised by high density racks in data centers, cooling equipment manufacturers offer alternative cooling solutions to remove the heat load in the data center. The objective is to provide localized liquid cooling by locating the heat exchangers closer to the problem. Locating heat exchangers directly at the heat source lessens the need for the CRAC unit. This increases the efficiency of the remaining CRAC units and the capacity of the available cooling within the data center. These heat exchangers are all scalable.

The Liebert XD system[8] from Emerson Network Power offers a number of components for localized cooling. APC's InfraStruXure solution[9] provides integrated rack-optimized packages.

[8] http://www.emersonnetworkpower.com/en-US/Products/PrecisionCooling/HighDensityModularCooling/
[9] http://www.apc.com/products/infrastruxure/

To take advantage of localized cooling, there is a requirement for chilled water in the data center. Chilled water in data centers is not new. It was used in the 1980s for mainframes. The need has returned as air cooling is unable to keep up with the thermal demand produced by high density servers. Several viable options based on this strategy have been developed for racks:

- Front or rear mounted fin and tube heat exchangers
- Internal fin and tube heat exchangers either at the bottom or mounted on the side of a rack
- Overhead fin and tube heat exchangers
- Internal server cooling

Localized cooling options are not limited to water based cooling. Many IBM Business Partners such as APC, Liebert and Coolcentric can also provide refrigerant-based (DX) cooling at rack and row level.

> **Redguide Case Study:** Xcel Energy included hot/cold aisle configurations, in row cooling, and localized water cooling in a data center consolidation effort that freed 2,000 square feet of floor space. See the redguide, *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, for more details.

4.3.1 Rear Door Heat eXchanger

The IBM Rear Door Heat eXchanger is a water-cooled passive device that serves as the rear door of an IBM enterprise rack. The Rear Door Hear eXchanger is designed to cool up to 35kW of heat from a high density rack. The hot air is exhausted from the rear of the server. The hot air passes through the Rear Door Hear eXchanger and is cooled. The heat is then transferred back to the chilled water loop, bypassing and reducing stress on data center CRAC units. The recommended chilled water temperatures are necessary for the successful operation of this device.

Figure 4-14 shows how Rear Door Hear eXchangers operate in conjunction with a Coolant Distribution Unit (CDU), and Figure 4-15 on page 74 shows the immediate impact on a rack's temperature profile of water flow through the heat exchanger.

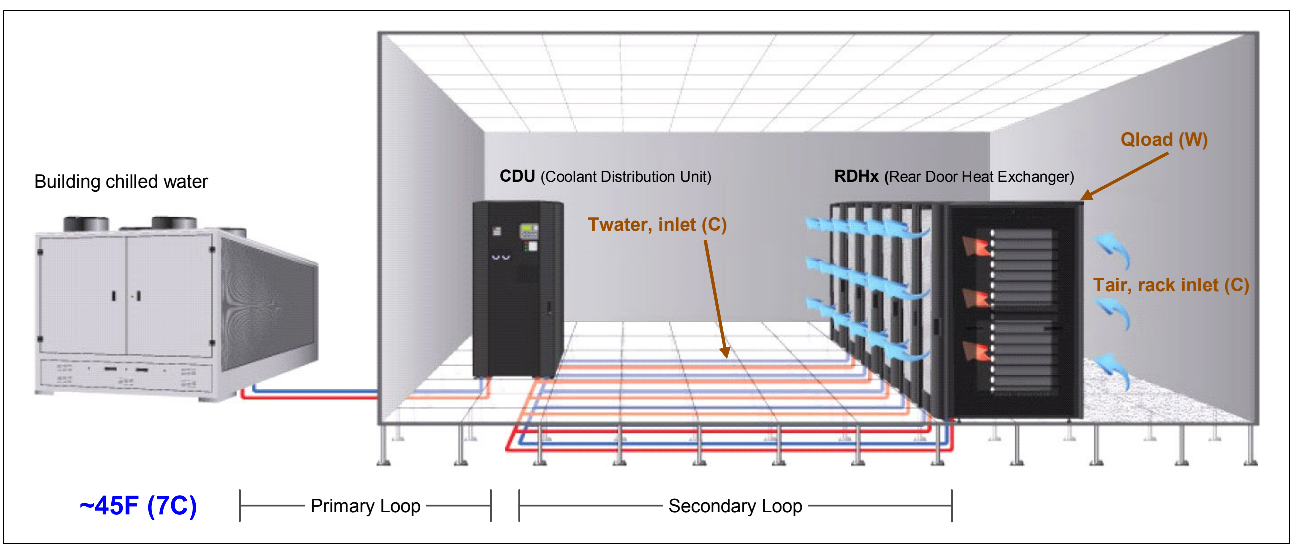


Figure 4-14 Rear Door Heat eXchanger operation

Here are a few benefits of a Rear Door Hear eXchanger installation:

- Completely neutralizes heat loads of up to 35kW
- Has no moving parts to fail
- Requires no energy at the Rear Door Hear eXchanger
- Quickly deployable

4.3.2 Rear Door Heat eXchanger case study

In a case study of rack level testing with high power density servers, IBM BladeCenter servers were used in the rack. Functional 1U servers were also used to determine heat extraction. These tests were conducted in a room where the inlet temperature could be maintained between 23-25 degrees Celsius. The water flow rate was maintained at 31 LPM in the test cases, and the water inlet temperature was kept at 17 degrees Celsius.

The results for the BladeCenter and 1U rack servers showed a 50-60% heat extraction of the input heat load. The increase in airflow impedance to the servers resulted in a negligible 1 degree Celsius increase in the CPU temperatures due to a slight airflow loss to the servers. The exit air temperature from the Rear Door Heat eXchanger was reduced by an average of 25 degrees Celsius from the discharge end of the servers.

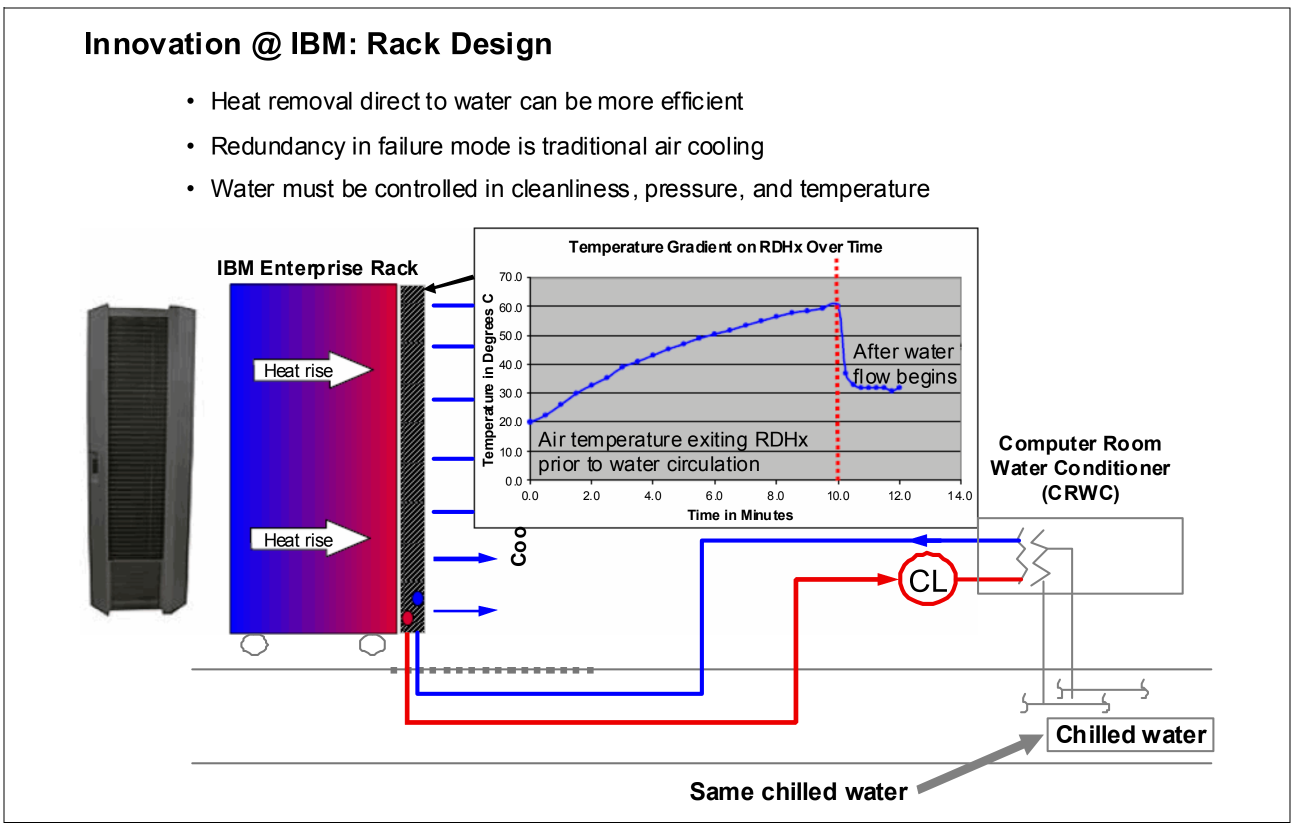


Figure 4-15 IBM Rear Door Heat eXchanger

A graph of results from using the Rear Door Heat eXchanger is illustrated in Figure 4-16.

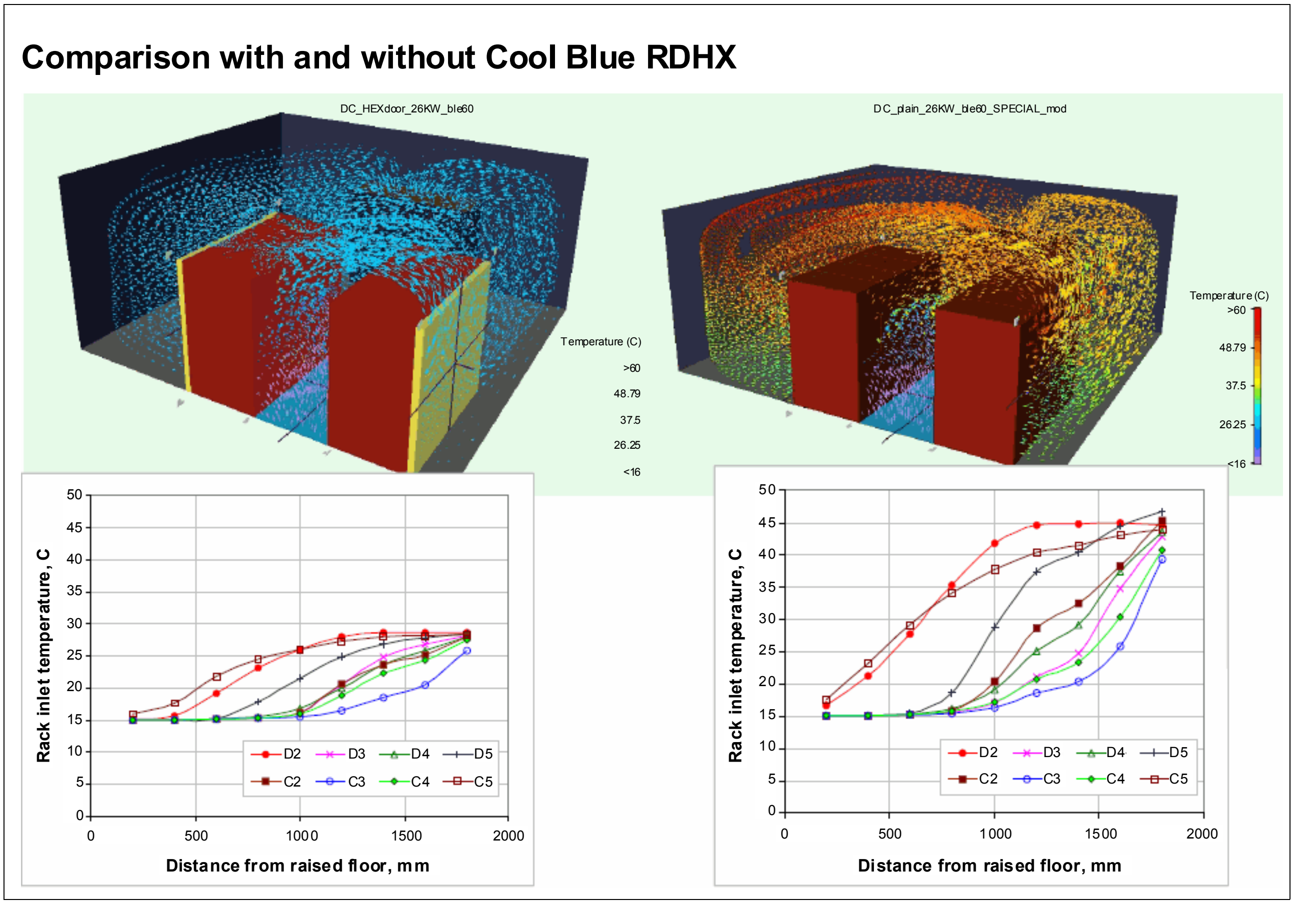


Figure 4-16 IBM cool blue rear door heat exchanger comparison

Note: Rear Door Hear eXchanger doors can be customized to suit your IT racks. IBM Global Technology Services can help you determine and satisfy your heat management requirements. For more information see http://www.ibm.com/services/

4.3.3 Modular water unit

Figure 4-17 shows a modular water unit (MWU) located in a POWER 575 rack. The MWU supplies water internally to the server. The MWU acts as a heat exchanger transferring heat collected from the servers through a closed loop to the facility water loop. This increases efficiency by transferring heat back to the chilled water loop directly and bypassing the CRAC units.

Note: The size is significantly smaller than the technology that was available 15 years ago.

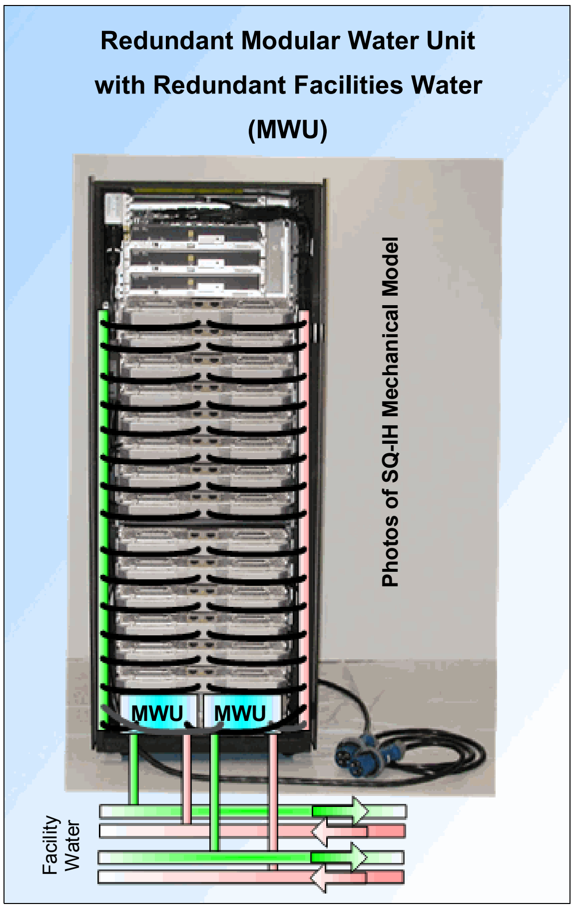


Figure 4-17 POWER 575 with MWU

4.3.4 Enclosed rack cooling

Enclosed rack cooling or self contained-rack cooling provides increased capacity for high density racks. These environments can be located on both slab and raised floor. The self-contained environment, using APC technology, allows the hot air from the rear of the servers to be conducted through the in-row cooler and then out to the front of the rack to start the cycle again. Increased efficiencies are gained by directly conducting the hot air over the condenser coil within the enclosed rack. The hot air cannot mix with the cold air, allowing maximum heat transfer to take place. The amount of heat rejection is dependent on the size of enclosure. Correct sizing is important for this feature to work.

Figure 4-18 shows an enclosed rack cooling example that is provided by Rittal.[10] The solution consists of redundant power distribution, one service rack, and a Liquid Cooling Package (LCP) enclosed by two system racks.

Figure 4-18 Rittal Rimatrix

The LCP unit takes care of cooling the enclosed racks with a redundant heat exchanger connected to chilled water facilities. One major advantage of this solution is that chilled water inlet temperature can be around 21^{o}C (70^{o}F) for cooling redundant 15 kW loads. As this solution works independently of any external air conditioning, it can even be placed outside of the air conditioned data center.

4.3.5 Sidecar heat exchanger

A sidecar heat exchanger is a water cooled heat exchanger located on the side of a rack to remove 100% of the heat generated by a rack of servers.

4.4 Enhancing HVAC

There is no doubt that HVAC is the highest consumer of energy in the whole data center. Yet energy efficiency for chillers has significantly improved in recent years. Replacing chillers and pumps that have been in service for 10-15 years or more can result in substantial savings. New chillers can reduce energy consumption by up to 50%. Each type of chiller operates differently depending on environmental conditions. It is import to understand what is the correct type of chiller for your environment. A number of components make up HVAC:

Chiller Removes heat from a data center. Sometimes used to describe the device that chills water that is used as the transfer medium for heat removal.

Pump Circulates water through the data center using CRAC units, in-row cooling units, and rear door heat exchangers.

Air handling unit (AHU) Cools or heats outside air depending on the ambient temperature and the required internal air temperature.

[10] Rittal RimatriX5: http://www.rimatrix5.de

4.4.1 Types of chillers

Air-cooled and water-cooled chillers each can use various types of compressors such as reciprocating, centrifugal, and screw driven.

Air cooled chillers reject the heat with a condenser coil. Air is blown over the condenser coils with a fan so heat is transferred into the atmosphere. These units use more power than water-cooled systems because the heat transfer works more effectively on a wet surface than a dry surface.

To further increase the efficiency of an air-cooled chiller, a pre-cooler evaporator can be placed on the air intake side. This pre-cools the air up to 10 degrees Celsius (17 degrees Fahrenheit) before the air passes over the condenser coil, which allows a greater transfer of heat.

Water-cooled chillers are generally used for very large capacity applications. They reject heat into a separate water source, typically a cooling tower, that uses the evaporation process to reject the heat into the atmosphere. However, if you are close to a river or lake, this water source can be used as a cheap and energy efficient option to reject this heat through immersion cooling if permitted by local environmental regulations.

The cooling tower requires more maintenance because condenser water tends to build up mineral deposits and requires chemical dosage to eliminate the chance of Legionnaires' disease. A water-cooled chiller requires much less power than an air-cooled chiller. The heated water by-product can also be used for heating buildings or swimming pools. The efficiency differences between the three compressors are as follows:

- *Reciprocating* compressors are the least efficient. They are usually set to run at three stages, which are 33%, 66%, and 100%. When the load requirement is only 25%, the additional 8% is generated and then discarded.
- *Centrifugal* compressors are more efficient because they have fewer moving parts.
- *Screw-driven* compressors are the most efficient. They are able to stage to the exact load that is required, from as little as 10% to 100%.

Note: Chillers that continually operate above 75% lose efficiency, so if you are able to reduce the heat load within the data center, this will benefit the chiller load. For every watt you save in the data center, you will save 1.25 watts at the chiller.

A new technology chiller system can improve efficiency by up to 50%. New chiller plants also can be installed with variable-speed drives, reducing pumping system energy usage and allowing better integration of the liquid cooling system into the chilled water infrastructure. Water-side economizers, which use outside air to directly cool the chilled water, can further reduce the energy required to run the chillers.

4.4.2 Variable-speed drive pumps

Variable-speed drive (VSD) pumps self-adjust the output pressure depending on the demand. A standard pump operates at 100% load regardless of the demand, making it very inefficient. To increase your pump's efficiency, VSD control units can be added to your existing pumps, increasing their efficiency and reducing the electrical load.

4.4.3 Air handling units

Air Handling Units (AHU) cool or heat outside air, depending on the outside ambient temperature and the required internal air temperature of the data center. Energy efficient AHUs can reuse a percentage of conditioned air, hot or cold, for energy savings and increased efficiency. They can also be used to centrally control the humidity levels in the data center, relieving workload on the data center CRACs, if the allowable humidity is in line with AHSRAE guidelines. By using AHUs on makeup air to control humidity in the DC, significant energy savings can be realized versus trying to control humidity at the individual CRAC level.

4.4.4 Economizers to enable free cooling

Two types of economizers are air-side and water-side.

Air-side economizers

Air-side economizers can be used as a free cooling unit. However, these units work best where there is a constant supply of cool and clean air. The consistency can be maintained by running these units overnight. The outside air economizer directly draws in outside air for use by the data center.

Attention: Be aware of the following issues with air-side economizers:

- Temperature and humidity control is pivotal
- Gaseous contamination can enter your data center and impact equipment
- Particle contamination can enter your data center and impact equipment

Water-side economizers

Water-side economizers use cool outdoor air to generate chilled condenser water that is used to partially or fully meet the cooling demands of the facility. When the outside air is cool enough, the water-side economizer will take over part or full load of the chiller. This can result in a number of free cooling hours per day.

4.5 Cool less, save more

The American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Technical Committee 9.9 (TC 9.9) has developed new guidelines, establishing four classes of data center operation with each class defined by an operating envelope of acceptable environmental conditions of temperature and humidity. The guidelines are intended to provide data center operators with the information necessary to operate in the most energy-efficient mode while still achieving the reliability required by their business.

Figure 4-19 shows a psychrometric chart, *ASHRAE Environmental Classes for Data Centers*, that depicts the data center classes and their associated acceptable operating envelopes.

Increasing the data center temperature set points by just one degree can save energy costs dramatically by reducing the load on CRACs, enabling better utilization of outside air through more free cooling, and reducing the load on chillers by increasing chilled water temperature.

IBM has increased the temperature in its own data centers from 20 degrees Celsius (68 degrees Fahrenheit) to 22 degrees Celsius (72 degrees Fahrenheit).

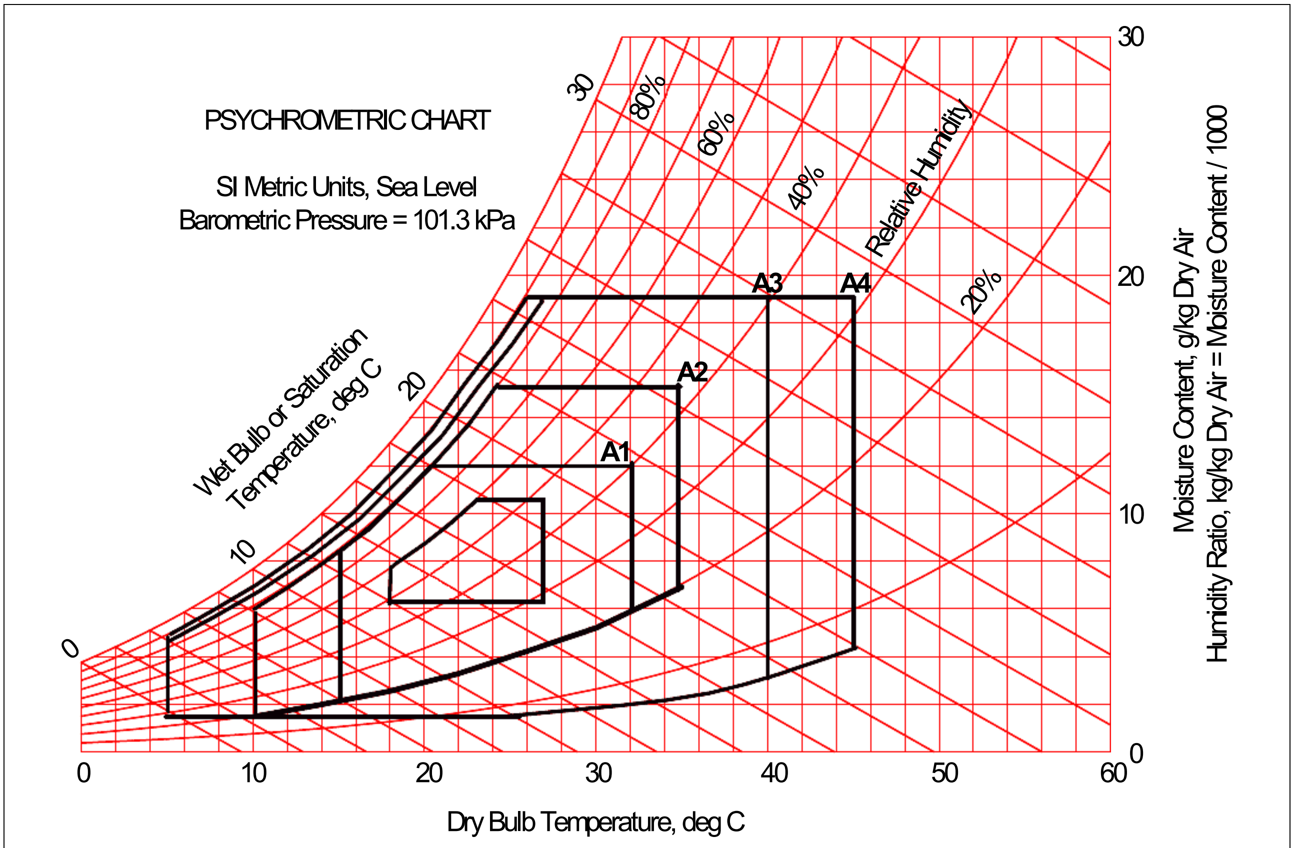


Figure 4-19 ASHRAE Environmental Classes for Data Centers

Note: Refer to the ASHRAE white paper *2011 Thermal Guidelines for Data Processing Environments* to read more on ASHRAE guidelines.

In addition, as described in 3.5.1, “IBM Systems Director Active Energy Manager” on page 40, IBM Systems Director Active Energy Manager (AEM) can monitor, manage, and provide trend analysis for power and heat for various IBM series of systems. AEM can provide substantial cost savings by controlling energy consumption and improving energy use through its power capping, standby, and power save recommendations.

Apart from these techniques and solutions, consider adopting industry best practices such as rearranging perforated tiles as needed based on evolving conditions, implementing remote consoles, using remote power ON/OFF solutions for servers, restricting physical access to your data centers, and sealing even small air leaks inside the data center. These practices help in sustaining the energy efficiency of your data center.

> **Note:** A periodic air quality check for data centers can be a good best practice to avoid corrosion (gaseous contamination) inside the data center and reduce component failure rates. As noted in 4.1.2, "Addressing environmental condition variations" on page 59, IBM Measurement and Management Technologies (MMT) can be a valuable tool in monitoring environmental and air quality conditions in your data center.

4.6 UPS options

An uninterruptible power supply (UPS) is a device that continues to supply electricity during a utility power failure. UPS units are often used to support critical computing environments.

Energy efficiency for UPS systems has significantly improved in recent years. Replacing a UPS system that has been in service for 10 years or more can result in substantial savings. New best-in-class UPS systems are approximately 70% more efficient than older ones.

4.6.1 Components and types

Various types of UPS units are available, and all have varying design and performance characteristics. The major components of a UPS unit are as follows:

Inverter The inverter converts direct current (DC) to alternating current (AC). If the power fails, the inverter provides 115 or 240 volts AC from a 12 volt DC power store.

Rectifier The rectifier converts AC to DC. It provides DC power to ensure batteries remain charged. In larger UPS units, when the rectifier is used in conjunction with an inverter, it acts as a premium power filter. This is achieved by the rectifier converting AC to DC and then the inverter converting DC back to AC. During this process, power spikes and dips are removed and clean power is delivered to the IT equipment.

Power Store Although batteries are the most common power store, fuel cells and fly wheels are also widely used. The DC electricity is stored to be used by the inverter in case of a power failure.

Each UPS type caters to separate applications. Understanding your equipment is important because one size does not fit all. The following types are the most common UPS units:

Standby UPS
This type of UPS is commonly used for personal computers or single servers. The inverter starts only when the power fails. This is very efficient and cheap.

Line Interactive UPS
This type of UPS is commonly used for 5-15 servers. When the power fails, power flows from the battery through the UPS to the servers. The inverter is always connected to the output. This provides filtering compared with the standby UPS.

Double Conversion UPS
This type of UPS is used for larger data centers. It recreates the AC power by using a rectifier to change from AC to DC and then using an inverter to change back from DC to AC. This provides the best quality power and protection and is usually the most expensive option.

Although UPS units are typically rated in kilovolt-amperes (kVA), for example, 330 kVA, the usable power is measured in watts. The efficiency of the UPS unit depends on the power factor (PF) of the unit, as shown in the following examples:

- A 330 kVA UPS unit operating at 0.8 = 264 kW of usable energy
- A 330 kVA UPS unit operating at 0.9 = 297 kW of usable energy

As you can see, the unit with the higher PF rating provides more usable energy.

4.6.2 Flywheel technology

As new flywheel technology emerges, there is always the option to replace old UPS batteries with a new flywheel. UPS batteries have a limited life span. The quality of the battery and the number of times it is discharged can limit its life to 10 years or even less in certain cases. The flywheel can operate in a greater temperature range than batteries, has a much smaller footprint than batteries, and does not have to convert AC to DC and back to AC.

The advantages of the flywheel are its efficiency and size. The disadvantages are its limited hold-up time, and it does not condition power. However, having at least one flywheel in your UPS bank is not a bad strategy.

4.7 Examining your facility power

This section discusses additional considerations when planning for more efficient power.

4.7.1 Utility power supply

Utility companies sell power at a rate of X cents per kWh. How much of this do you get to use? Inefficiencies from the utility pole to your equipment can be as great as 20%. For example, 1 kW of power into your facility can equate to 0.8 kW usable power for your equipment.

4.7.2 Power factor correction

Power factor correction (PFC) on your site's electricity supply can regain some of that lost power. With PFC, 1 kW of power-in equates to 0.95 kW usable power for your equipment. For sites that use 2500 kW to 3000 kW, the pay-back period is three to four years.

4.7.3 Power distribution and resistance

Undersized wiring and lengthy cable runs from the power distribution point to the rack can lead to energy losses due to resistance. Ensuring the correct sizing of cables and limiting line distance can improve the efficiency of the power delivered to the server.

In-row UPS units and power distribution boards save on both wiring and power loss due to resistance.

4.7.4 Intelligent power distribution unit

Intelligent power distribution units (iPDU) can be used to measure and manage power consumption at the individual outlet level in data center racks. Interfaces between iPDUs to the IBM Systems Director Active Energy Manager (AEM) can provide a complete view of individual asset energy consumption within the associated racks, enabling the power management capabilities of AEM to be used for greater energy efficiency.

For additional information about the AEM, see 3.5.1, "IBM Systems Director Active Energy Manager" on page 40. For additional details about how iPDU data gathered by the AEM can be integrated with IBM Tivoli Monitoring for Energy Management. See 5.3.1, "Multi-system data consolidation" on page 95.

4.7.5 DC versus AC

High voltage DC holds promise for reduced energy consumption and increased efficiency by reducing the number of necessary power conversions, eliminating conversion losses in the process. However, adoption of high voltage DC within the data center has been tentative. Safety concerns and a lack of comprehensive standards, coupled with the widespread availability and lower costs of AC components and continued AC power management improvements, have hindered transition from AC to DC.

When deciding whether or not to take advantage of high voltage DC, consider the following questions:

- Should we change our infrastructure for this improvement?
- Will AC computer power supplies continue to improve?

> **Note:** Many IBM systems, including the IBM System z196 mainframe and the IBM Power 775 Supercomputer, support both DC and AC power sources.

4.7.6 On-site power generation

For information about on-site power generation when examining your facility power, see Section 4.8.2, "On-site power generation" on page 84.

4.8 Reviewing generator options

The last infrastructure component that we discuss is generators.

4.8.1 Standby generator

Today's standby generators are designed to have greater fuel efficiency than older style generators and to be much greener when releasing CO_2 emissions into the atmosphere. They also have a much faster start and transfer time—less than 30 seconds—which supports the use of flywheel and fuel cell technologies to replace UPS batteries.

The standby generator is a key piece of the site infrastructure for high availability data centers. UPS systems can energize the data center for a few minutes or even hours. However, without a standby generator, the HVAC systems that supply the cooling to the data center will not operate, and the likelihood of overheating the data center is very great. High-power availability is a key in achieving a high-availability data center.

A standby generator supplies power when the utility power supply is not available. Both the standby generator and the utility power supply are connected to an automatic transfer switch (ATS). When the utility power supply is not available, the generator automatically starts, taking approximately 40 seconds to assume the load. During this time, the UPS supports the data center load. Without the UPS, the data facility loses power and IT systems power down.

Standby generators typically run on diesel fuel or natural gas. These units have a life span of 15 to 20 years. In areas that have a good utility power supply, these units might not get much run time, so it is important to maintain and test them.

When sizing these units, be sure to account for all the infrastructure that maintains the data center, including chillers, pumps, CRACs, UPSs, AHUs, and other site infrastructure.

4.8.2 On-site power generation

On-site power generation can augment or replace utility power as follows:

- Sites that have limited utility power can use on-site generation during peak periods.
- Sites that have access to cheap fuels, where the cost of electrical generation is cheaper than the cost of utility power, can use on-site generation.
- Utility power might be unreliable, so on-site power generation is useful.

An on-site power source can also be more efficient and reliable than a remote source in the reduction of power conversions, reduction of distribution losses, and reduction of susceptibility to distribution failures.

Co-generation

One option for on-site power generation that is a good fit for data center environments is cogeneration, or combined heat and power (CHP). As illustrated in Figure 4-20, the heat by-product from electrical generators can be leveraged to assist with ancillary heating and cooling needs.

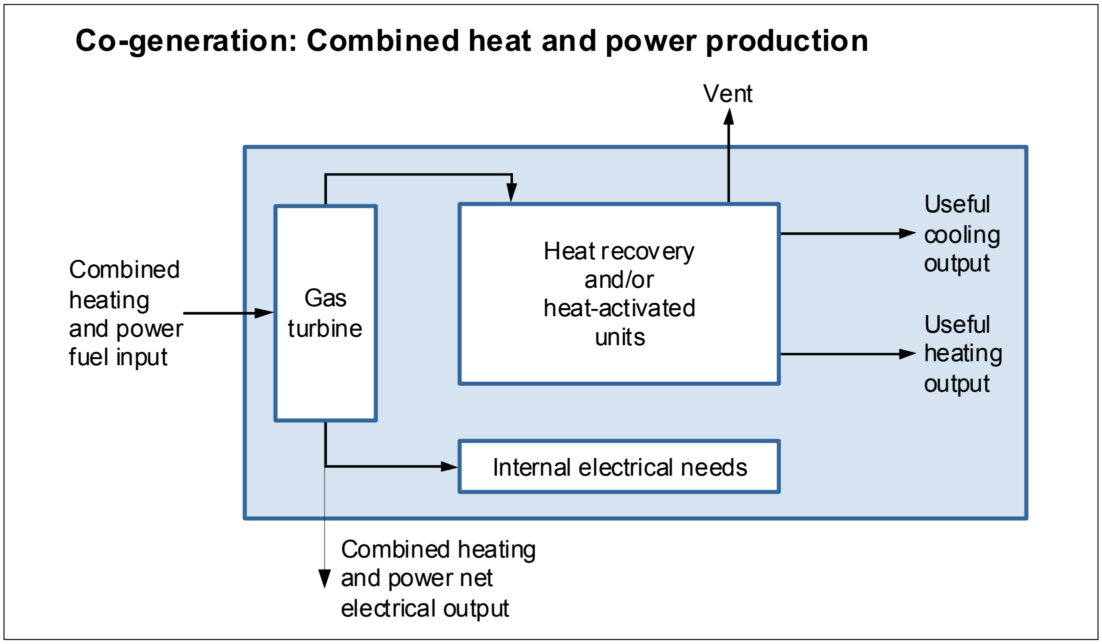


Figure 4-20 Cogeneration schematic

Data centers have much higher energy utilization intensities (20 to 100 watts per square foot) than typical commercial buildings, with all of this energy being converted to heat as it is used within the facility. Considering the level of heat generated, an effective CHP configuration for data centers is to use this heat to power absorption chillers. Recovered heat in the form of steam or hot water can be used to power a chiller which, in turn, can be used for facility air conditioning or, less commonly, to feed chilled water to water cooled racks. Figure 4-21 shows the efficiency and emissions benefits of CHP for data centers.

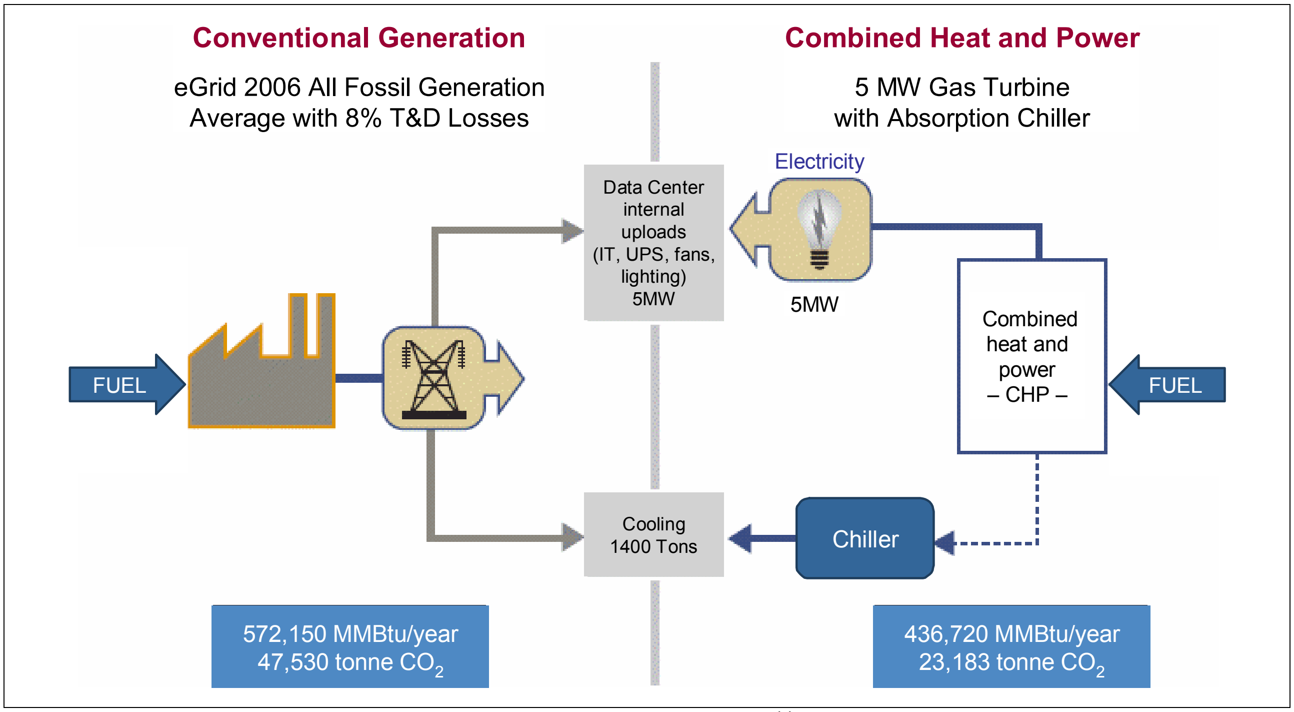


Figure 4-21 Efficiency and Emissions Benefits of CHP for Data Centers[11]

> **Note:** For more information about combined heat and power and the U.S. EPA CHP Partnership program, visit the EPA site at `http://www.epa.gov/chp`

Solar power as a lean energy resource for your data center

With rising energy costs and interests in environmental sustainability, clean and renewable sources of power can be an option for on-site generation as well. Though a bit more expensive to implement, the costs can often be recovered considering the long lifespan of most data center facilities.

[11] U.S. EPA Combined Heat and Power (CHP): `http://www.epa.gov/chp`

Depending on your geographic location and regional climate patterns, solar power can be a good supplemental energy source for your data center. Figure 4-22 shows a possible reference configuration and potential energy savings of a solar power generation design for data centers and DC powered servers.

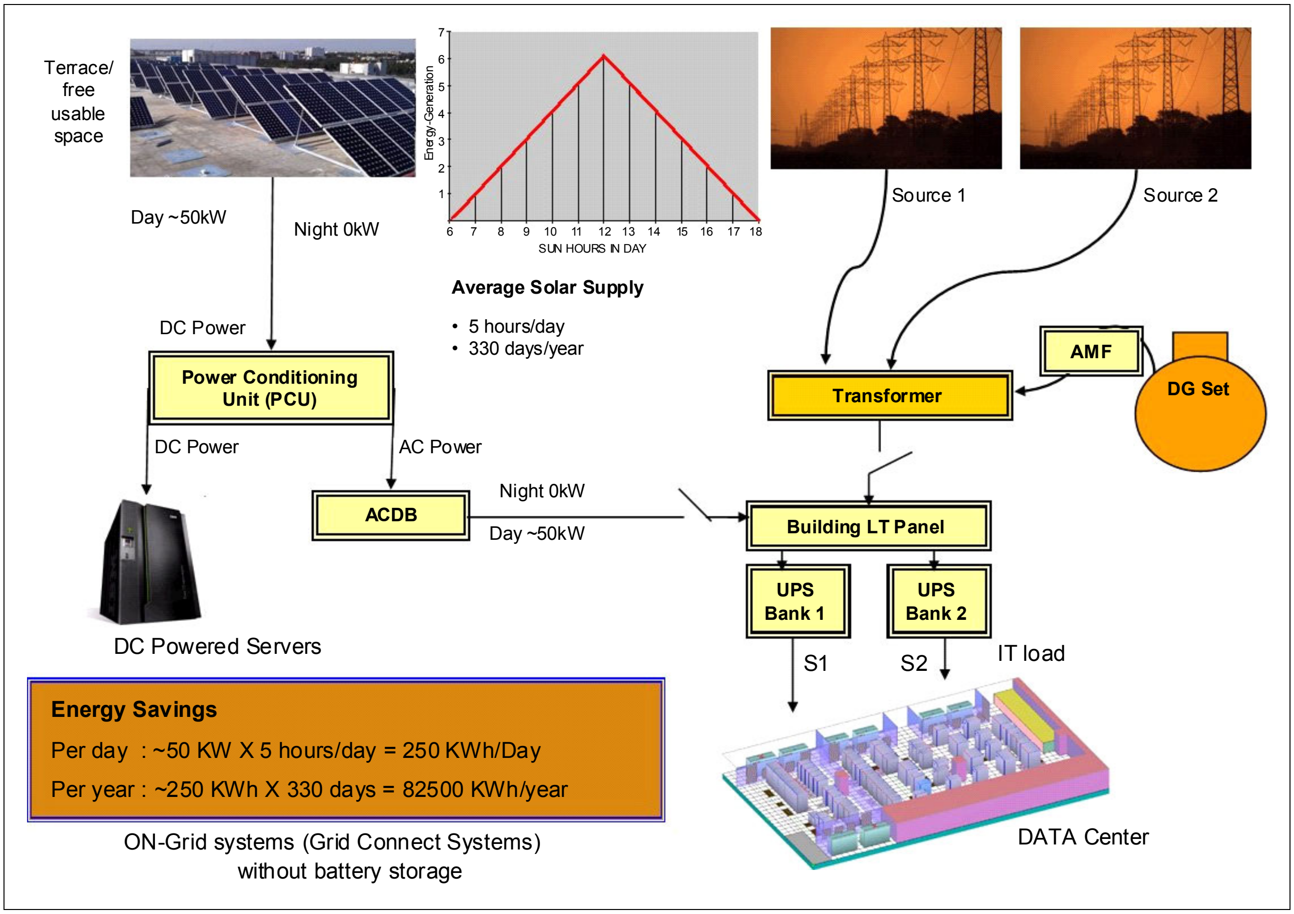


Figure 4-22 Sample solar power design for data centers and DC-powered servers

The following website sponsored by NASA's Earth Science Enterprise Program can be a useful data source for your clean energy project:

http://eosweb.larc.nasa.gov/sse/RETScreen/

Other technologies evolving for on-site generation

The following additional technologies are evolving as additional options for on-site power generation:

- Fuel cells, which can eliminate the need for UPS because reliability is 99.99% or better
- Nuclear power, which can eliminate the need for UPS because reliability is 99.99% or better
- Wind power, an evolving solution for energy efficient buildings in areas where the wind is continuous, although not as developed as other technologies for data centers

Redguide Case Study: The IBM India Software Lab in Bangalore supplements its utility power supply with an alternative clean energy source to lower energy costs. See the Redguide, *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, for more details.

4.9 Revising your existing data centers

You will not be able to manage or control costs if you are not measuring your data center energy consumption and efficiency. Ideally, awareness of both consumption and efficiency spans from the individual servers in your data center all the way to your utility provider. As you recognize where the inefficiencies are, you can start addressing them by employing more energy efficient solutions, technologies, and best practices.

As discussed in this chapter, there are many approaches that can be taken to extend the life of your existing data center facilities, reclaiming space, power, and cooling capacity for efficiency while concurrently improving data center reliability and resiliency.

4.9.1 Investigation, consolidation, and replacement

Consolidating and replacing old IT equipment can have a direct impact on the site and facility infrastructure. For every 1 kW of energy saved in the data center, an additional 1.35 kW is saved on the site and facility infrastructure, which results in a 2.35 kW saving from the utility provider.

As parts of your infrastructure progressively reach the end of serviceable life, investigate all the options available and their impact on other components.

Adding more energy-improving components to your existing infrastructure can increase the energy efficiency and capacity of your infrastructure.

The following steps will make your the data center more efficient:

1. Assess the data center environment.
2. Implement no-cost and low-cost initiatives.
3. Replace the oldest systems first; they are more likely to fail and are less energy efficient.
4. Invest in systems that save the most energy.

The following steps will make your site and facilities more efficient:

1. Assess the site and facility infrastructure.
2. Replace the oldest systems first; they are more likely to fail and are less energy-efficient.
3. Invest in infrastructure that saves the most energy.

IBM services can assist in all of these areas. See Chapter 6, “How we can help: IBM services and IBM Business Partners” on page 101 for more details.

4.9.2 Investigating industry initiatives

Also investigate the following industry activities and guidelines:

- Green Grid, which is a consortium of IT and data center equipment companies to help you measure your data center efficiency:
  - Establish metrics for data centers.
  - Develop standards and practices to improve energy efficiency in the data center.

  See the Green Grid website:

  http://www.thegreengrid.org/
- The American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE) offers the following ways to help:
  - Maintains various subcommittees to provide information and guidance on operations.
  - Develops standards for the design of efficient data center cooling.

  See the ASHRAE website:

  http://www.ashrae.org/
- The Uptime Institute offers information about the following areas, along with many others:
  - Data center availability and reliability information.
  - Operating criteria for adding energy efficiency to a data center.

  See the Uptime Institute website:

  http://www.uptimeinstitute.org/

5

Integrating across IT and facilities for the bigger picture

To achieve the greatest degree of efficiency from your data center environment, you must have a comprehensive perspective that spans both the facilities and IT equipment that are present on the data center floor. Such a perspective starts with your data center inventory and the operational lifecycle that keeps it up-to-date. That inventory is supplemented with instrumented metering that gathers data about the ongoing operational conditions of your data center.

That data can then be consolidated, enabling monitoring across all aspects of data center operations for potential problem conditions. That consolidated data set can also provide the foundation for detailed analytics, role-based dashboards, historical trend reporting, and future planning.

The development of this integrated, comprehensive perspective is the subject of this chapter.

5.1 Tracking your inventory comprehensively

To effectively manage the capacity of your data center, it is important to know what equipment is installed on the data center floor, including what is coming, what is leaving, and what is changing. The utility of this inventory can be greatly enhanced by establishing standards regarding the asset types to be included in that inventory, a unique identification methodology (both physical and logical), and the physical and operational characteristics to be documented for each asset. While an initial focus on physical assets is expected, the ideal inventory should contain both physical and logical assets, including virtualized servers.

The level of granularity of your inventory, coupled with the granularity of your measurements, will dictate the degree to which you can comprehensively manage capacity. For instance, you cannot manage disk utilization if your inventory does not include disk capacity and your measurements do not include disk consumption.

If physical asset data can be expanded with logical characteristics, such as organizational ownership and operational purpose, metrics and reports across asset types or assets fulfilling a similar purpose can be gathered. Servers performing similar work can be co-located within the data center, enabling increasingly efficient cooling designs and power-saving controls. See 5.4.3, “Asset placement for efficiency” on page 98.

Without a detailed and well-maintained inventory, in-depth analyses of data center trends and identification of improvement opportunities become considerably more difficult.

You can take advantage of tools such as Maximo Asset Management (MAM)[1] and Maximo Data Center Infrastructure Manager (DCIM)[2] to track asset characteristics, process associated work orders, and place assets on the data center floor.

5.1.1 Three-dimensional inventory management

The ideal inventory will track asset placement within the data center in three dimensions, down to the slot placement of servers within hosting racks. This level of detail enables visibility into rack-level capacity to minimize the underutilization of rack space.

5.1.2 Centralized inventory control

Within the Internet, the end user of a website has little or no knowledge of or concern about the type of physical server providing the hosting of the site they are visiting. If the providers of that site are using a hosting service, they may also have limited knowledge of the actual equipment used by the host. They simply have requirements related to functionality, availability, and performance they expect to be met.

This separation of functionality from physicality empowers the hosting provider to have much greater flexibility when it comes to hardware selection, physical placement, and ongoing operation for greatest efficiency.

The closer your organization can get to having the facilities team exercise complete ownership of the selection and placement process of hardware on the data center floor, the greater the opportunity for high levels of operational efficiency.

---

1 Maximo Asset Management: http://www.ibm.com/software/tivoli/products/maximo-asset-mgmt/

2 Maximo DCIM: http://www.ibm.com/software/tivoli/products/maximo-data-center-infrastructure/

This arrangement is one ultimate promise of cloud computing. As the service requestor relinquishes ownership of the physical environment, it empowers the service provider to become much more efficient in capacity management.

5.1.3 Power and cooling traceability

When an asset is placed on the data center floor, it is important to trace the circuit connectivity of that asset back through to the power source and to document that path. Done diligently, this can improve power utilization while simultaneously preventing the inadvertent overloading of a circuit, power panel, or Power Distribution Unit (PDU). For dual-corded devices, tracing this path can also prevent the connectivity of both power cords to the same power source infrastructure, preventing the failure of one PDU, generator, or UPS casing asset, which can result in service downtime and potential lost revenue. Mistaken connectivity of assets to singular power supplies is one of the most frequent causes of service outages.

While more fluid, the association of an asset with the related infrastructure providing cooling for that asset can provide much greater levels of control over cooling utilization. As shown in Figure 5-1, tools like Maximo Asset Management for Energy Optimization[3] (MAMEO) can actively display the zone of influence for each computer room air conditioner (CRAC) within the data center, along with the usage of its cooling capacity. As CRACs are adjusted, the layout of these zones adjusts automatically.

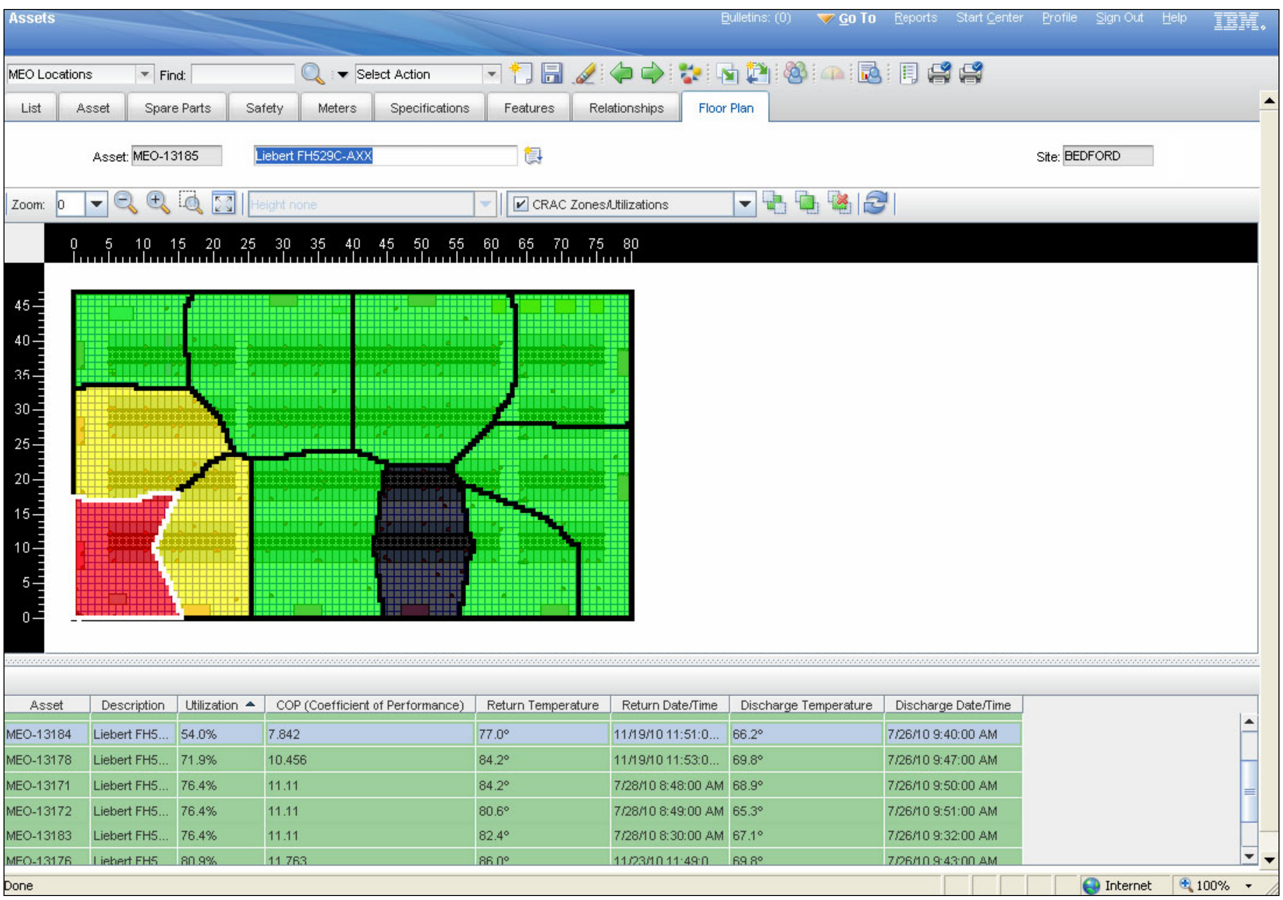


Figure 5-1 Maximo Asset Management for Energy Optimization

[3] Maximo Asset Management for Energy Optimization:
http://www.ibm.com/software/tivoli/products/maximo-asset-mgmt-energy-optimize/

5.1.4 Autonomic inventory reconciliation

A detailed inventory is only as good as its accuracy. As assets are moved into, out of, and around in a data center, the results of a one-time inventory can quickly become obsolete. Although an established process for change management using tools like Maximo Asset Management can go a long way in providing ongoing inventory maintenance, there is always the possibility of someone bypassing the process or following it incorrectly, resulting in inventory errors.

However, the closer an organization can get to establishing standards for rack selection, installation orientation, and configuration (such as door openings), the more effective an autonomic method for performing inventory reconciliation using tools like Radio Frequency Identification (RFID), can become.

5.2 Gathering data from across your environment

You need a comprehensive perspective of data center operational conditions to facilitate efficiency and increase flexibility, resiliency, and dependability.

5.2.1 Measuring across IT and facilities

A comprehensive perspective should include data from facilities and IT assets and ongoing operational and environmental conditions within the data center.

Operational data from facilities equipment such as PDUs, generators, uninterruptible power supplies (UPS), CRACs, and air handling units (AHUs) can be obtained by connecting to existing building and power management systems.

Detailed power consumption data can be gathered by performing branch circuit monitoring using intelligent PDUs (iPDUs) or by connecting to asset-specific monitoring tools like IBM Active Energy Manager (AEM).

Environmental conditions can be obtained through static sensors gathering data about temperature, pressure, humidity, and air flow. Such sensors can also be used to assess particulate concentrations in the air, which can be a concern for facilities looking to make greater use of local free cooling options.

Asset-level utilization metrics can be gathered with monitoring tools like IBM Tivoli Monitoring, providing visibility into the utilization of CPU, disk, storage, network bandwidth, and the like. AEM can also be used to gather data about asset-level thermal conditions.

Only when a comprehensive data set is gathered can it be analyzed in detail to determine consumption trends and identify opportunities to improve efficiency. The more granular the level of detail gathered, the more readily improvement opportunities can be identified. The same data set can also be used to substantially improve service availability and resiliency.

However, to be effective, common data standards and an overall data architecture are a must.

5.2.2 Relationship management

In 5.1.3, “Power and cooling traceability” on page 93, we noted the value of tracing IT assets to the related infrastructure components providing power and cooling capacity. Logical relationship management can provide additional benefits. If your asset inventory includes details about logical characteristics like associated organizational ownership and provided business services, this information can be associated with operational conditions to proactively identify and respond to potential problems.

5.3 Consolidating data and monitor conditions

Examining data at the individual system manager level, for instance information on the building or power management system or the IT asset manager, can be invaluable for performing detailed and fine-tuned analytics. Integration and examination of a comprehensive data set obtained across management systems offers a different perspective and additional value.

5.3.1 Multi-system data consolidation

Tools like IBM Tivoli Monitoring for Energy Management[4] can be used to consolidate data from multiple sources within your data center, as illustrated in Figure 5-2 on page 96. Native connectivity support for building management and automation systems like Siemens APOGEE can be used to gather facilities data, or standards like OPC can be used to connect to and gather data from a multitude of facility and power management systems.

Tivoli Monitoring for Energy Management can connect directly to AEM to integrate thermal and power data from IT assets, and support for multiple iPDUs can enable data gathering at the individual power outlet level.

Connectivity support for vendor sensor solutions and IBM Measurement and Management Technologies (MMT) enables the gathering of environmental conditions data such as temperature, pressure, and humidity.

[4] Tivoli Energy Management: http://www.ibm.com/software/tivoli/solutions/green/

Merging with asset utilization data from IBM Tivoli Monitoring and inventory data from Maximo Asset Management (MAM), you can obtain a nearly total picture of the data center's contents and operational and environmental conditions.

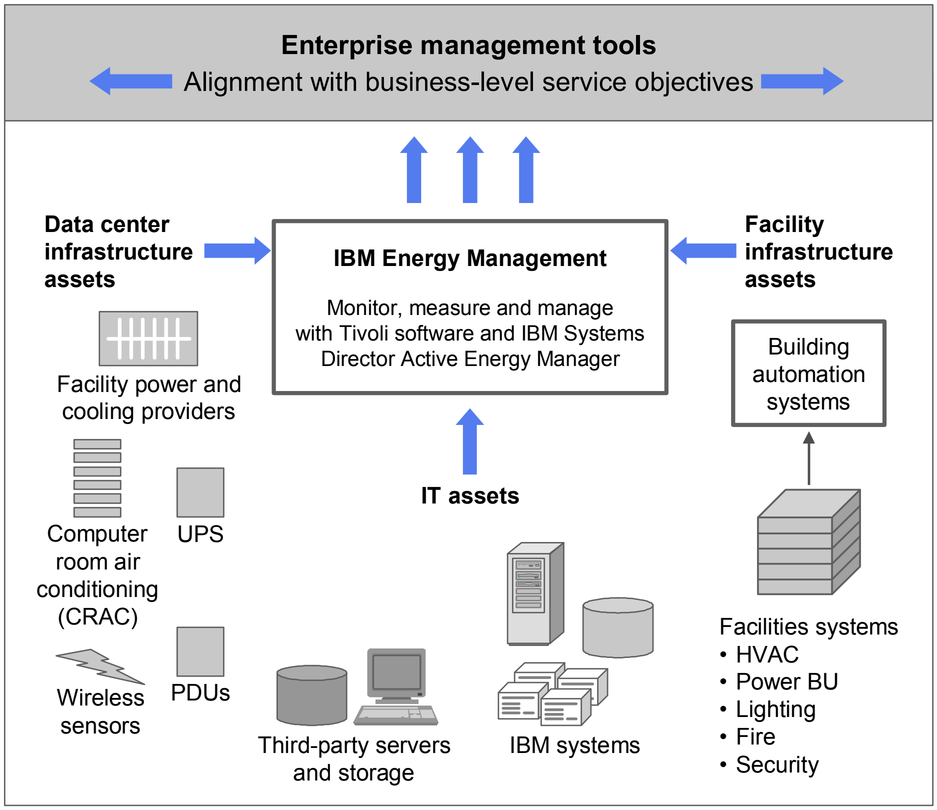


Figure 5-2 Tivoli Monitoring for Energy Management

5.3.2 Integrated alert and event management

With all of these data feeds available, you can configure conditional scenarios under which events should be generated. Such events can be gathered, correlated, and managed using Tivoli Netcool/OMNIbus with the option of triggering alerts and associated work orders.

The detection of a thermal hot spot from a failing CRAC unit can be associated with the IT assets in the impacted data center cooling zone and the services provided by those assets, enabling those workloads to be transitioned to backup systems before a potential server failure, avoiding service downtime and lost revenue.

5.3.3 Recovering costs based on actual consumption

Effective capacity management becomes more realistic in the presence of associated incentives, and few incentives are more capable than a bill for consumption. As data center operators face a need to place a much greater focus on capacity consumption beyond simply physical space, the scope of cost recovery must expand as well. Tivoli Usage and Accounting Manager can be a valuable tool for this purpose. However, billing based on consumption only becomes possible when there is detailed and accurate consumption data.

5.4 Performing integrated analyses

As the data from multiple data center sources is consolidated, it becomes important to perform integrated analyses of that data.

5.4.1 Role-based dashboards, trend reporting, and environmental mapping

With a consolidated and comprehensive data set in place, detailed role-based dashboards, historical trend reporting, and mapping of environmental conditions become possible.

As shown in Figure 5-3, tools like Tivoli Business Service Manager (TBSM) and IBM Mashup Center can be used to generate role-based dashboards. Executive-level presentations can display near real-time data center information such as efficiency metrics like PUE and DCiE, energy consumption by asset type, energy costs, and carbon emissions.

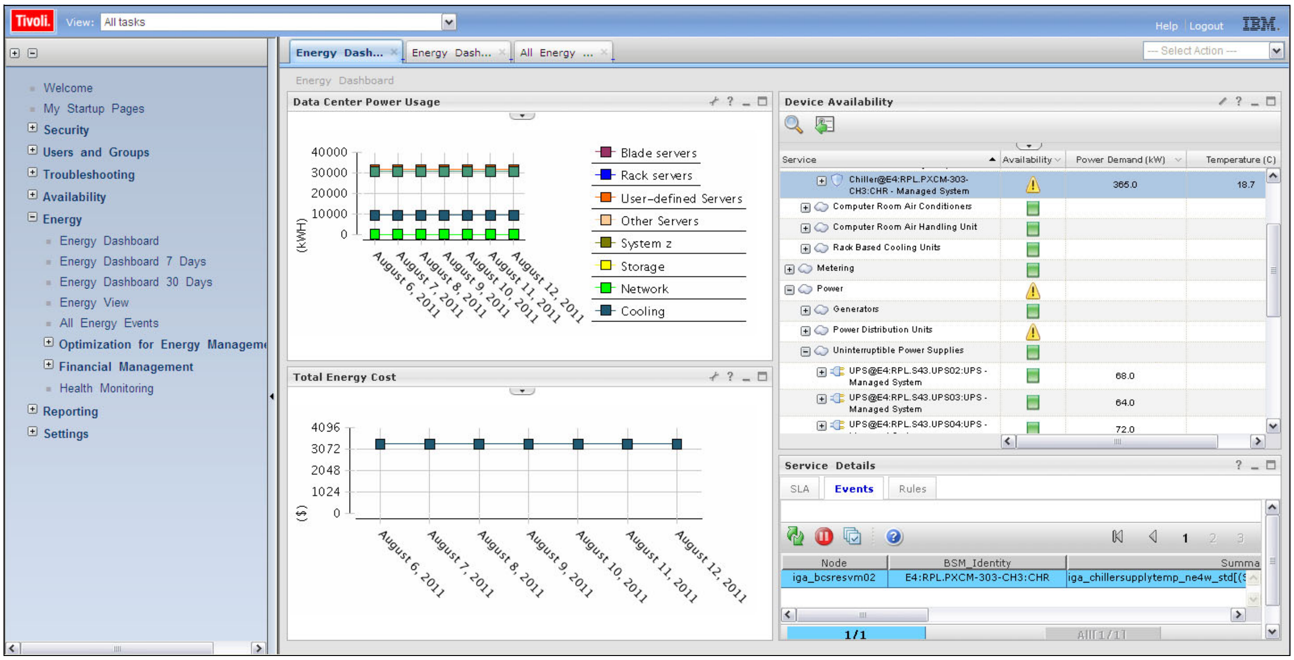


Figure 5-3 View data center operational and efficiency conditions in role-based dashboards

Facilities data from building and power management systems can be displayed alongside IT asset data from AEM and ITM, and historical trends can be displayed alongside near real-time events from OMNIbus. Detailed historical trending energy reports can be generated using Tivoli Common Reporting, and such reports can even be configured to be generated and distributed automatically.

You can take advantage of Maximo Asset Management for Energy Optimization environmental and facilities data gathered with Tivoli Monitoring for Energy Management to display interactive environmental maps of the data center floor, portraying thermal and humidity conditions, hot and cold spots, and CRAC cooling zones.

Redguide Case Study: The IBM Green Innovation Data Center in Southbury, Connecticut uses a broad Tivoli tool set to monitor, manage, and report on its data center environment. See the Redguide *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, for more details.

5.4.2 Comprehensive capacity management

Detailed knowledge of a data center's inventory and the ongoing consumption of physical space in three dimensions, power, cooling, and server-level capacities presents powerful management options across all capacity types. Underutilized capacity from total physical rack space down to individual server CPUs, can be identified and remedy options considered.

5.4.3 Asset placement for efficiency

When placing new IT assets on the data center floor, you want to place them in the most energy-efficient location possible, ideally taking into account available space, power, and cooling capacity and the asset's logical purpose to the business and availability requirements.

With insight into spatial consumption in three dimensions, you can readily identify open rack space. Coupled with knowledge of current consumption of PDU, panel, and circuit power capacities, and CRAC-level cooling capacity from maps of data center cooling zones, you can effectively place a new asset in the environment where it makes the best use of available capacity across all three.

Knowledge of server purpose can take the provision of high and low-density sections of the data center to the next level of granularity. For instance, high-density servers that provide high-availability services expected to operate at high performance on a continual basis, generating considerable heat footprints, can be increasingly co-located on designated racks with focused cooling technologies such as rear-door heat exchangers and in specified areas of the data centers taking advantage of liquid cooling or high-efficiency air conditioning units.

Inversely, servers fulfilling lower performance requirements can be co-located as well and targeted for power-saving modes and reduced cooling requirements.

5.4.4 Planning for the future

As you obtain consumption data across multiple capacity types within your data center environment, over time you can start discerning overall trends. Such knowledge becomes invaluable when performing the business planning for future capacity needs, empowering you to effectively determine when, where, and what type of capacity your business will need as you progress.

Note: Any data set is only as good as the processes in place to utilize it. Although not covered in this document, the existence and maturity of associated processes is imperative for a successful implementation of a smarter data center.

5.5 Looking beyond the data center: smarter buildings

While the focus here has been on data centers, many of the same methodologies can be applied to any smarter building, whether it be a data center, office building, hospital, or airport. The same architecture put in place to make your data center more intelligent can be expanded to include the total building or business campus where the data center resides. IBM Intelligent Building Management (IIBM)[5] provides an integrated suite of tools that can enable you to deploy a smarter building initiative quickly and effectively. For more information, see *Smarter Cities Series: Understanding the IBM Approach to Efficient Buildings*, REDP-4735 at the following address:

http://www.redbooks.ibm.com/abstracts/redp4735.html/

[5] IBM Intelligent Building Management: http://ibm.com/software/tivoli/products/intelligent-building-mgmt/

6

How we can help: IBM services and IBM Business Partners

IBM and IBM Business Partners have solutions that can help make your data center facility and IT infrastructures energy-efficient. INBM can help you meet these challenges:

- Continued rise of overall energy costs and demand
- Rising cost of air conditioning to cool data centers or computer rooms
- Data centers at or near their limits of space, power, or cooling capacity
- Concern over power surges and shortages affecting business
- Government-imposed levies on carbon production and regulations on power consumption and waste
- Increasing drive toward centralized data centers
- Too many servers that are partially used or not used to capacity
- Need for an environmentally safe way to dispose of old software or hardware

Our energy-efficiency solutions are based on proven technologies utilized by IBM and its customers to realize immediate, tangible returns on investment through the implementation of more efficient technical, operational, and management practices. By focusing on energy efficiency, you can lower your overall energy expenses and optimize your existing server capacities to free up budget for business-critical solutions while becoming more environmentally and operationally sustainable.

This chapter discusses services and assistance that IBM and IBM Business Partners can deliver.

6.1 Services IBM can deliver

IBM professionals can help at the following levels by providing information for hardware and software solutions and recommendations for designing and building highly efficient data centers:

- Systems and technologies
- Management systems
- Data centers

6.1.1 IBM Global Technology Services

IBM Global Technology Services (GTS) can help with the following categories in your move toward having a smarter data center:

- Diagnosis
- Building
- Cooling
- Virtualizing and simplifying
- Management, monitoring and enhancement

Diagnosis

IBM can help assess energy efficiency of your data center by using power management analysis and thermal analytics. IBM also studies options of server and storage virtualization and consolidation. IBM provides a fact-based business case and can detect major opportunities for improvement with reduction of energy costs. Actions at the thermal level can help eliminate hot spots (regions of high power density) and undesired intermixing of hot and cold air. Payback on the investment can be achieved in as little as two years, thereby covering the cost of the assessment in the first year.

Building

IBM can provide expertise to customers, based on the experience of building millions of square feet of data centers for clients worldwide. Planning, building, or upgrading a new data center provides the perfect opportunity to rationalize the data center strategy as a way for the customer to gain substantial savings on capital, operations, and energy efficiency.

Cooling

We can help you implement IBM cooling technologies designed to increase computer capacity in your data center while maintaining or even lowering energy consumption. We can also help you maximize the usable space in your data center and operate it reliably and more efficiently.

Virtualizing and simplifying

Virtualizing servers and storage devices can increase processing loads and boost individual utilization rates. Virtualization enables organizations to increase utilization rates of their computing resources while managing them in ways that offer more resource flexibility. We can help you implement IBM hardware and software solutions for virtualization that help you move toward a smarter data center requiring fewer physical servers, thereby reducing energy demands and expanding IT capacity.

Management, monitoring and enhancement

Tracking energy usage helps customers control their billing but also aids in allocating energy where it is needed. IBM can assist you implement technologies that help you effectively monitor and manage space, power, and cooling usage of your data center environment. Our technologies can provide IT and data center managers with a single view of actual energy usage across systems and facilities within the data center. These technologies also allow you to use near-real-time and historic data to initiate actions to improve resource utilization, availability, and service, while reducing energy consumption and cost.

> **Note:** IBM has four types of services to address facilities problems you might have:
>
> - Data Center and Facilities Strategy Services
> - IT Facilities Consolidation and Relocation Services
> - IT Facilities Assessment, Design and Construction Services
> - Specialized Facilities Services

Table 6-1 shows examples of IBM GTS offerings.

Table 6-1 Examples of IBM GTS offerings that help clients move toward a smarter data center

Data center energy efficiency assessment	Provides a comprehensive, fact-based analysis that prioritizes tactical plans across the data center to help improve efficiency and reduce costs. Provides from 15% to 40% annual energy savings and includes an industry standard comparison (MPG for your data center), identifies low performance areas and actions for improvement and provides a business case to prioritize investment.
Data center infrastructure energy efficiency optimization	Allows clients to work one-on-one with IBM Data Center Services Power / Thermal Development Engineers to formulate a balanced plan to improve efficiency, reduce total cost of ownership, and maximize the aggregate IT equipment supported by an existing data center infrastructure.
IT facilities assessment, design and construction services	Helps you create stable, security-rich, energy-efficient, future-ready data centers and enterprise command center facilities. The end-to-end services include a review of your existing data center's reliability, points of failure, growth, floor space, and power and cooling needs. IBM uses analytics to improve data center investment governance by understanding vendor and technology trade-offs, helping to reduce total data center costs by 15% to 30% over the data center's useful life.
Installation of Rear Door Heat eXchanger	Provides a simple, cost-effective, energy-efficient solution to solve hot spot problems within the data center. The overall approach is to provide and oversee a simple step-by-step process for implementing of the Rear Door Heat eXchanger units.
Server consolidation efficiency study	As an attractively priced streamlined assessment, provides you with the facts required to justify initiatives for optimizing your server environments. Identifies solutions that can help to reduce energy consumption in transitioning to a smarter data center and quantifies savings to reduce TCO and energy costs.
Server optimization and integration services	Helps you create a cost-effective, scalable, flexible, and resilient server infrastructure to support business applications using industry-leading practices based on IBM experience and intellectual capital.
Storage optimization and integration services	Helps you reduce complexity, optimize performance, and manage growth by creating cost-effective, highly utilized, scalable, and resilient storage infrastructures. Helps maximize energy efficiency through consolidation, virtualization, and storage optimization leading to a smarter infrastructure.

Server and storage power / cooling trends and data center best practices	Helps you understand the current and future power, and cooling and I/O demands that IT equipment places on your existing or planned data center infrastructure.
Data migration services	As a suite of technologies and services, enables the movement of stored data—typically included in energy assessment recommendations—in a nondestructive manner regardless of server platform or storage array vendor.

Note: The following links are to information about site and facilities, servers, and storage:

- `http://ibm.com/services/siteandfacilities`
- `http://ibm.com/services/server`
- `http://ibm.com/services/storage`

6.1.2 IBM Systems and Technology Group

The IBM Systems and Technology Group (STG) can help you optimize the utilization of your data center and system solutions by focusing on new technologies emerging from IBM product development labs. IBM STG also provides support to the IBM Innovation Centers through their network of briefing, benchmark, design centers, and services support.

Cross Platform Consulting Services and Data Center Services

IBM STG Cross Platform Consulting Services provides many services aimed at the data center to help you find ways to improve data center energy productivity and effectiveness. As a highly trained extension of IBM development laboratories with the most up-to-date knowledge of industry trends, IT technologies, and best practices, STG can compare the energy efficiency and productivity of IBM offerings against other available products and how they can integrate with other systems in your data center. You will be able to understand the IBM approach for reducing carbon emissions and how workload-optimized systems can help extend the life of your data center.

Table 6-2 shows examples of IBM STG offerings that can help you move toward smarter data centers.

Table 6-2 Several of IBM STG offerings for smarter data centers

Measurement and Management Technologies and Data Center Thermal Analysis	Enables you to use Measurement and Management Technologies (MMT) and IBM experience from worldwide data center analysis engagements. Creates an environmental profile of your data center to identify possibilities for cooling efficiencies and smarter energy use. Identifies, diagnoses, and solves thermal hot spots and over-cooled areas.
Power and Cooling Trends and Data Center Best Practices	Helps you understand the current and future power, cooling, and energy demands that IT equipment will place on your existing or planned data center infrastructure. Also helps you make smarter power and cooling decisions for the future of your data center, your business, and the planet, contributing to energy efficiency for an IBM Smarter Planet™.
IT Systems Energy Efficiency Assessment	Analyzes your existing IT systems energy efficiency and provides a full view of any negative consequences or wasteful areas. Strategically assesses and sizes the data center benefits and potential energy savings of IT virtualization, consolidation, or rationalization. With this type of feedback, you can grow within the capacity constraints of an existing data center and avoid or delay the expense, effort and interruptions of expansion or relocation.

Data Center Thermal Analysis and Optimization	Looks at the thermal profile of your data center to identify, diagnose and remedy trouble spots or energy inefficiencies. Assists you in visualizing and better understanding the potential impact of new or proposed changes to cooling capacities and loads. Also helps you more efficiently deploy IT equipment by evaluating current and potential data center thermal capacity and installed equipment.
Active Energy Manager Implementation Jumpstart	Provides the expertise and experience of IBM Systems Lab Services to speed up the analysis and deployment of IBM Systems Director Active Energy Manager for IT power management and transfers key skills to your IT and facility management staffs.
Data Center Readiness Assessment for Rear Door Heat Exchanger	Assesses the feasibility and readiness of your data center before you deploy Rear Door Heat Exchanger, which is a simple, cost-effective and energy-efficient solution to cooling problems and thermal capacity management of your data center.
Data Center Baseline Cooling Assessment	Offers a baseline and high-level assessment of your data center cooling capacity. Verifies the availability and adequacy of your current cooling capacity, reports any areas of concern, and documents recommendations to increase capacity, utilization, and availability.
Data Center Power and Cooling Planning for iDataplex	Provides pre-installation, planning, and guidance for the power, cooling, physical layout, and mechanical systems required for highly dense iDataPlex systems.

Note: Learn more about IBM STG Cross Platform Consulting Services and Data Center Services at this website:

http://www.ibm.com/systems/services/labservices/solutions/

IBM offers over 500 services of all types. We highlight several of the relevant services in this section. They are also described, along with other services, on the IBM website, located at this address:

http://www.ibm.com/services/us/index.wss/allservices/

To find a particular services, enter relevant keywords in the search field of the website.

6.1.3 IBM Asset Recovery Solutions

The disposition of obsolete information technology has become a major issue for corporations. International Data Corporation estimates that most corporations store old computer equipment for up to three years at a cost of U.S. $360 per machine and then pays an additional U.S. $218 for its eventual disposal. In the face of these statistics, it is clear that customers faces real challenges every time a decision is made to upgrade IT equipment. Customers want to maximize recovery costs at every opportunity. In addition, disposal methods must comply with applicable environmental regulations. Ideally, customers want one single-source provider instead of using a patchwork of separate services.

IBM Global Financing (IGF) can meet these challenges and more. Asset Recovery Solutions, an offering from IBM Global Asset Recovery Services (ARS) division, offers a suite of highly competitive solutions to dispose of IT equipment and equipment related to IT from IBM and other sources. The equipment can include hardware, fax machines, and printers. IBM has worldwide product remarketing and logistics capabilities complemented by state-of-the-art systems and brings these best-of-breed capabilities to the customer, with the reach, resources, and knowledge to serve its customers anywhere in the world, providing real peace of mind.

Figure 6-1 provides an overview of Asset Recovery Solutions.

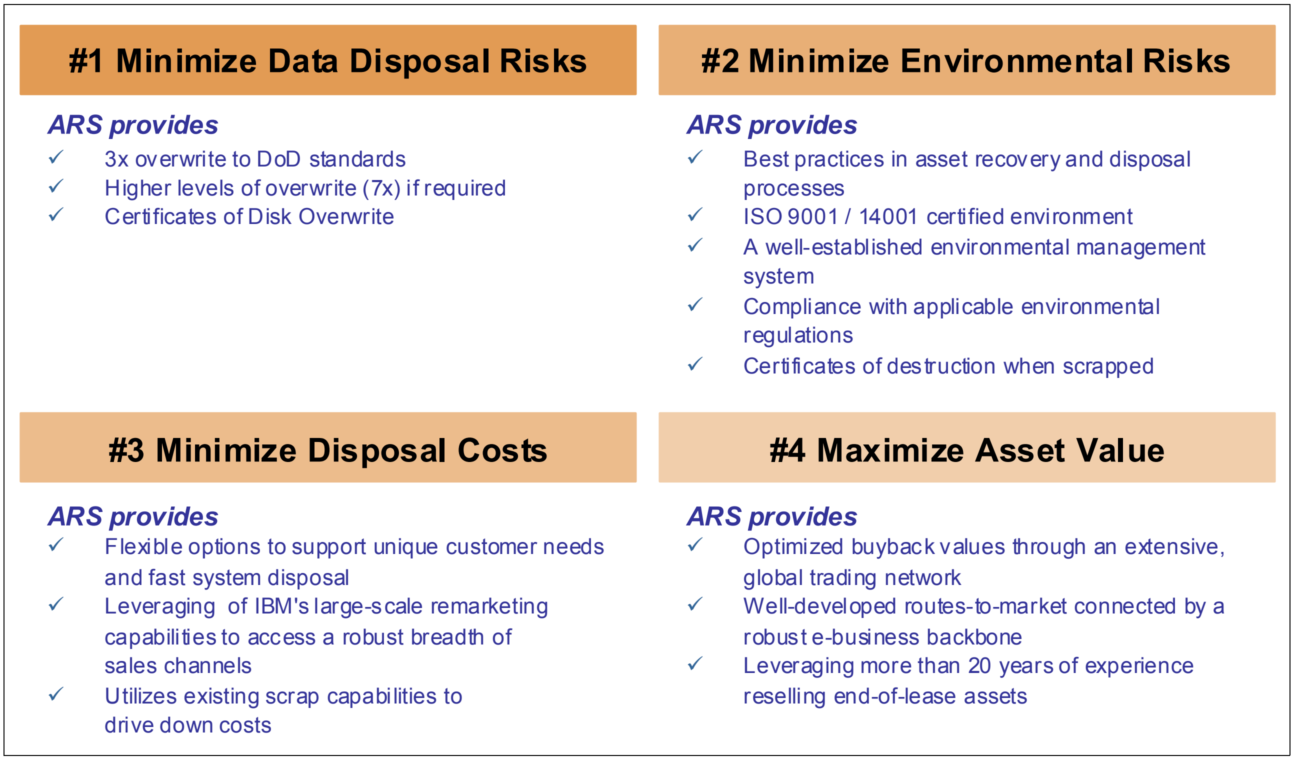


Figure 6-1 Asset Recovery Solutions

IBM methods provide a proper processing of used equipment in a hassle-free, secure, and environmentally-friendly manner. You can also minimize data disposal risk with IBM disk overwrite services and realize the maximum residual cash value of your retired assets through our buyback services.

For more information, see the following website:

`http://www.ibm.com/financing/us/recovery`

6.1.4 IBM financing solutions

IBM Project Financing™ provides financial flexibility to design and build energy-friendly data centers. Offerings includes financing for select facility design and construction, building and structural upgrades, infrastructure equipment, IT hardware, software, and consulting from IBM and other vendors. We provide industry knowledge to ensure payment structures in line with the useful life of various asset classes.

IBM Project Financing offers the possibility to control cash flow and let the customer decide when to draw funds and how much to draw. Easy one-time negotiations cover the entire project and let you finance IT and data center equipment and all related services for design, construction, and relocation. Project Financing is a one-stop shop for a single, integrated IBM solution. IBM can provide consistent terms globally for all leases and loans, enabling worldwide competitive pricing. You can be prepared for current and future geographical expansion. Financing is available for planned expenditures, and the flexibility to address unforeseen challenges as they arise. For more information, se the following website:

http://www.ibm.com/financing

> **Important:** IBM can help you plan, implement, and claim incentive rewards for energy efficiency projects that are available for your geography. See the following IBM Redbooks publication for more information: *Energy Efficiency Incentives for the Data Center: Assessment and Measurement*, REDP-4670. Get more information at this site:
>
> http://www.redbooks.ibm.com/abstracts/redp4670.html

6.1.5 IBM smarter data centers

IBM is continuing to evolve its own data centers with efficiency, environmental greenness, and sustainability in mind. You can find more information about smarter data centers at these sites:

http://www.ibm.com/services/smarterdatacenter
http://www.ibm.com/cio/smarterdc/
http://www.ibm.com/ibm/green/index3.shtml

6.2 IBM Business Partners

IBM has a number of Business Partners who can assist in offering the right energy-efficient solution for your data center. With strategic alliances and offering best innovative technologies with best practice solutions, these partners can help you achieve a state-of-the-art data center.

The Green Solutions Center in Research Triangle Park (RTP) showcases IBM and partner innovations for energy efficient data centers. It was designed and built with the following IBM Business Partners:

- APC provides power and cooling solutions. The company often partners with IBM Global Technology Services (GTS) to support scalable modular data centers.[1]
- Emerson Network Power has been providing cooling products since the 1960s, when the IBM System/360 mainframe was introduced.[2]
- Eaton is known for delivering power quality solutions for data centers.[3]
- Panduit develops and provides data center solutions that help customers optimize their physical infrastructure through simplification, increasing agility and operational efficiency.[4]
- Crestron provides automated room control systems with power management to save on energy.[5]

1 http://www.apc.com/products/infrastruxure/
2 http://www.emersonnetworkpower.com/en-US/Products/PrecisionCooling/
3 http://www.eaton.com/EatonCom/Market/Datacenters/index.htm
4 http://www.panduit.com/Solutions/DataCenter/
5 http://www.crestron.com

IBM has also joined forces with Neuwing Energy Ventures to enable organizations to receive Energy Efficiency Certificates (EEC) for projects that reduce energy.[6]

Other organizations and programs to watch include these:

- Leadership in Energy and Environmental Design (LEED)[7], a voluntary consensus-based national rating system for developing high-performance, sustainable buildings. Developed by U.S. Green Building Council (USGBC), LEED addresses all building types and emphasizes state-of-the-art strategies for sustainable site development, water savings, energy efficiency, materials and resources selection, and indoor environmental quality. LEED is a practical rating tool for green building design and construction that provides immediate and measurable results for building owners and occupants.
- Energy Star[8] is a joint program of the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Energy (DOE). Its purpose is to encourage efficient use of energy, thereby protecting our environment and saving us money. With the help of Energy Star, Americans have saved considerable amounts of energy to reduce greenhouse gas emissions and to save significantly on their energy bills.

  Energy Star offers help with energy management for businesses, which can produce a considerable savings for a company's bottom line and for the environment. Partnering with Energy Star helps a company set goals, track savings, and reward improvements. Top performing buildings can earn the ENERGY STAR, which is a mark of excellence for energy efficiency that is recognized nationally.

[6] http://www.neuwingenergy.com
[7] http://www.usgbc.org/LEED
[8] http://www.energystar.gov

A

IBM smarter data center: an evolution of energy-efficient operation

The IBM smarter data center strategy takes advantage of mature tools and techniques to perform rigorous analysis to determine the facts that enable informed business decisions. The smarter data center model also includes the use of new consumption models. The end result is a data center that operates efficiently and manages resources effectively while providing the desired service level.

This Appendix describes the actions IBM has taken to assess, plan, and implement leadership data center strategies to help the business operate smarter.

A.1 A history of leadership in helping the environment

For many years, IBM has been committed to energy efficiency and to the protection of the environment.

To establish an exhaustive list of accomplishments attributed to IBM for the protection of the environment would be difficult. However, here are several significant accomplishments:

- In 2006, IBM becomes a founding member of *The Green Grid.*
- In 2007, IBM receives the U.S. EPA's SmartWay Excellence Award.
- In 2008, IBM Global Asset Recovery Services becomes one of the first recipients of IDC's Green Recycling and Asset Disposal for the Enterprise (G.R.A.D.E.) certification.
- In 2010, the IBM Leadership Data Center in Research Triangle Park becomes the first IBM data center to be certified LEED Gold.
- In 2011, IBM announces it has exceeded its 2007 pledge to double compute capacity in three years without increasing energy consumption by increasing compute performance per megaWatt-hour by a factor of 2.1.[1]

A.2 Smarter technology for a smarter planet

IBM continues to invest in developing industry-leading technologies that bring value and innovation to clients. From advanced chip design with the POWER7 and Cell processors to the early development of time sharing, the predecessor to modern virtualization, and from data center modelling and optimization to the elastic cloud environment, IBM has been a pioneer in technologies that have changed the way companies think, act, and do business.

Syracuse University, in partnership with IBM, operates one of the world's greenest and smartest data centers with a focus on coupling new technology with new uses for existing technology.[2] This new way of thinking about data center design and operation leads to more efficient, cost-effective computing for clients. Technologies like tri-generation, high voltage DC power distribution, all liquid cooling, and real-time data center environmental monitoring, all combine to increase the deployment and management of commodity system hardware.

Leadership in technology development extends outside the systems with the IBM-designed Aquasar super computer for the Swiss Federal Institute of Technology Zurich. IBM delivered a first-of-a-kind hot water-cooled supercomputer that consumed up to 40 percent less energy than a comparable air-cooled machine. Through the direct use of waste heat to provide warmth to university buildings, Aquasar's carbon footprint is reduced by up to 85 percent.[3]

IBM continues its focus on smarter technologies with the following achievements:

- In 2009, IBM announces Canada's most powerful supercomputer for analyzing climate change at the University of Toronto. This supercomputer uses innovative xCAT and Moab software for innovative power management.
- On February 16, 2011, the IBM Watson supercomputer wins the IBM Jeopardy challenge
- On June 09, 2011, IBM introduces Intelligent Building Management software with installations at Tulane University, The Cloisters of the Metropolitan Museum of Art, and IBM Rochester that can reduce energy consumption by up to 40% and reduce facility management costs 10 to 30%.

[1] IBM Environmental News: http://www.ibm.com/ibm/environment/news/datacenterperformance_2011.shtml
[2] IBM Press Release: http://www-03.ibm.com/press/us/en/pressrelease/28946.wss
[3] IBM Press Release http://www-03.ibm.com/press/us/en/pressrelease/32049.wss

A.3 IBM Research Triangle Park Leadership Data Center

Recognizing the need to support new compute models, IBM made the $362 million investment in the IBM Research Triangle Park Leadership Data Center (LDC) to help clients operate smarter businesses, organizations, and cities.

Announced on February 4, 2010, the new data center demonstrates technologies from across IBM. The LDC uses advanced software virtualization technologies that enable access to information and services from any device with extremely high levels of availability and quality of experience. The facility aggressively conserves energy resources, saving costs and speeding services deployment through a smart management approach that links equipment, building systems, and data center operations. The LDC is a showcase of how IBM technologies and practices can help organizations operate smarter.

Figure A-1 The 60,000 square foot IBM Leadership Data Center in Research Triangle Park, North Carolina

> **Note:** For more on the IBM Leadership Data Center see the following website:
>
> http://www.ibm.com/cio/smarterdc

A.4 Future directions

The IBM strategy going forward focuses on the high level of interconnectivity between workloads, IT systems, and facility infrastructure and takes advantage of the evolution of IT resource consumption and delivery. Smarter data centers will be designed to deliver desired service levels in a flexible environment that reduces complexity and increases operational efficiency through automated processes, giving businesses the agility needed to adapt to changing business environments. The data center strategy of the future is automated and extensible.

7

Conclusion: The journey to smarter

Today's CIOs face many concerns in their data centers to achieve business sustainability. The cost of energy is growing, as is public awareness about data centers impacting our environment. Data center power, cooling, and space capacities are at or near their limits. With new environmental laws coming and company images at stake, energy efficiency is a necessity, from cost and sustainability perspectives and also as a way to reduce organizational risk.

7.1 The smarter data center

With the continually increasing demands being placed on the world's data centers, coupled with the rising demand for and costs of resources, it is imperative that our data centers become smarter. Smarter data centers make the most of existing resources, are flexible to respond to changing demands, are monitored and managed to quickly and even proactively respond to changing conditions, and provide the foundational data for detailed analytics to effectively plan for future growth.

A smarter data center must monitor and manage the provision and consumption of multiple types of capacity, including physical space, power, and cooling, to balance utilization and minimize stranded capacity. Equally important is obtaining the greatest business benefit from consumed capacity by maximizing efficiency across the entire data center environment, spanning both facilities and IT.

7.1.1 Maximizing your efficiency and extending the lifespan

In this Redpaper, we have described multiple technologies, methods, and best practices that you can use to extend the life of your existing facilities and equipment by recouping lost capacity and maximizing efficiency.

To maximize the efficiency of your physical space, we mentioned the importance of three-dimensional inventory management, server consolidation and virtualization, storage virtualization and management, site consolidation, and the deployment of the latest server technologies, which are providing ever-increasing amounts of computing power in diminishing physical footprints.

To maximize the efficiency of your cooling capacity, we described best practices like hot and cold aisle configurations, direct airflow management, localized liquid cooling, free cooling options, and the importance of environmental conditions analyses to identify and eliminate thermal hot spots.

To maximize the efficiency of your power capacity, we included details on embedded hardware power management options, software-based power management capabilities, power estimation utilities, power monitoring down to circuit and outlet-level consumption, and the importance of power connectivity relationship management for both effective planning and avoidance of downtime.

While it is critically important to extend the life of your existing facilities, the need to develop new facilities is inevitable. Included were considerations to make and questions to ask when planning new facilities, from selecting the site to exploiting scalable modular data centers that can quickly grow along with your business.

7.1.2 Integrating across the data center and beyond

Many of these methods have focused on the facility operations of the data center, others on IT operations, but the greatest efficiency is obtained when these perspectives are integrated, allowing you to perceive all parts of the data center in its entirety. Thus we also described the benefits of detailed inventory maintenance, operational condition measurements and data consolidation, role-based integrated dashboards, historical data trending, and detailed analytics for future planning. The capabilities of cloud computing demonstrate the potential associated with this integration.

Many of these methodologies can be extended for effective management of additional capacity types and beyond the data center to your business facilities, creating smarter buildings.

7.1.3 Your smarter data center strategy

As mentioned early on, there is no one-size-fits-all solution. As you can see, there are innumerable options from which to choose. Which methods and techniques to employ to make your data centers smarter depends on the unique specifics of your data center environments. Determining which approach to take and then successfully implementing it requires a comprehensive smarter data center strategy.

When defining your strategy, executive level sponsorship and cross-organizational cooperation are imperative. The strategy should define the business requirements to satisfy and detail the capacities that will be managed. It should define the standards, assess existing conditions, and establish realistic goals. Also, it should integrate efficiency methods into the day-to-day operations of the business.

7.1.4 IBM: A valuable partner

With so many options to consider and decisions to make, it can be vital to have a business partner in making your data centers smarter, particularly a partner with over 100 years in business, experience in implementing efficiency methods across 8 million square feet of data center space owned or managed around the world, and expertise to determine the best strategy for the success of your business.

A companion to this publication, the Redguide *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523, provides several case studies of how IBM has deployed many of the methodologies described in this Redpaper within our own facilities and with our customers for large efficiency gains.

The next step is to contact your IBM representative and allow us to help you develop and implement your own smarter data center strategy for a healthier bottom line and a foundation for the future.

B

Additional material

This paper refers to additional material that can be downloaded from the Internet as described in the following sections.

Locating the web material

The web material associated with this paper is available in softcopy on the Internet from the IBM Redbooks web server at this address:

ibm.com/redbooks

Select **Additional materials** and open the directory that corresponds with the IBM Redpaper form number, REDP4413.

Using the web material

The additional web material that accompanies this paper includes the following files:

File name	*Description*
REDP4413.zip	Zipped presentation

System requirements for downloading the web material

The web material requires the following system configuration:

Hard disk space: 10 MB minimum

Downloading and extracting the web material

Create a subdirectory (folder) on your workstation, and extract the contents of the web material .zip file into this folder.

Related publications

We consider the publications listed in this section particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks publications

For information about ordering these publications, see “How to get IBM Redbooks publications” on page 122. Note that several of the documents referenced here might be available in softcopy only.

- *Smarter Data Centers: Accelerating the Move to a Smarter Planet*, REDP-4523
- *Energy Efficiency Incentives for the Data Center: Assessment and Measurement*, REDP-4670
- *Deploying a Cloud on IBM System z*, REDP-4711
- *IBM zEnterprise System Technical Introduction*, SG24-7832
- *IBM zEnterprise System Technical Guide*, SG24-7833
- *Tivoli Integration Scenarios*, SG24-7878
- *Smarter Cities Series: Understanding the IBM Approach to Efficient Buildings*, REDP-4735
- *Performance and Capacity Implications for a Smarter Planet*, REDP-4762
- *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940
- *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590
- *IBM System Storage Solutions for Smarter Systems*, SG24-7945
- *Implementing IBM Storage Data Deduplication Solutions*, SG24-7888
- *IBM System x Private Cloud Offering, Advanced Configuration: Architecture and Implementation Guide*, REDP-4771
- *IBM eX5 Implementation Guide,* SG24-7909, available here:

  http://www.ibm.com/redbooks

Online resources

These websites provide further information about the focus of this paper:

- IBM Integrated Service Management

  http://www.ibm.com/ibm/servicemanagement/
  http://www.youtube.com/user/ibmservicemanagement#p/c/0/N68T1H8XGAA
- IBM Asset Recovery Solutions

  http://www.ibm.com/financing/us/recovery
- IBM Global Financing

  http://www.ibm.com/financing

- IBM smarter data centers

  http://www.ibm.com/services/smarterdatacenter

- IBM services A-Z

  http://www.ibm.com/services/us/index.wss/allservices/

- IBM IT services

  http://www.ibm.com/services/

- IBM System x and BladeCenter Power Configurator

  http://www-03.ibm.com/systems/bladecenter/resources/powerconfig.html

- IBM Global Technology Services: *Defining a blueprint for a smarter data center for flexibility and cost-effectiveness*, April 2011

  http://public.dhe.ibm.com/common/ssi/ecm/en/gtw03031usen/GTW03031USEN.PDF

- IBM Scalable Modular Data Center: *The art of the possible: Rapidly deploying cost-effective, energy-efficient data centers*, February 2008

  http://www.ibm.com/services/us/its/pdf/smdc-eb-sfe03001-usen-00-022708.pdf

- IBM CIO: Sustainability

  http://www-935.ibm.com/services/c-suite/cio/sustainability.html

- IBM Government industry

  http://www-935.ibm.com/services/us/gbs/bus/html/bcs_government.html?g_type=pspo

- APC

  http://www.apc.com

- TE Connectivity: A Green Data Center

  http://www.adc.com/dm/us/en/truenet/greendatacenter

- American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)

  http://www.ashrae.com

- Anixter

  http://www.anixter.com

- Carbon Footprint

  http://www.carbonfootprint.com

- Eaton

  http://www.eaton.com/EatonCom/ProductsServices/index.htm

- Emerson Network Power

  http://www.emersondatacenter.com

- Energy Star

  http://www.energystar.gov

- Environmental Protection Agency

  http://www.epa.gov/climatechange

- GE Ecomagination

  http://www.ecomagination.com

- SearchDataCenter.com from TechTarget, Inc.

  http://searchdatacenter.com

- The Green Grid

  http://www.thegreengrid.org
- U.S. Green Building Council (USGBC)

  http://www.usgbc.org
- Uptime Institute

  http://www.uptimeinstitute.org
- VMware VMotion technology

  http://www.vmware.com/products/vi/vc/vmotion.html

Other publications

These publications are also relevant to the topic of this paper:

- *Energy Efficiency Indicator - 2011 Global Results*, Institute for Building and Efficiency, 2011
- *How dirty is your data? A Look at the Energy Choices That Power Cloud Computing*, Greenpeace International, April 2011

ASHRAE publications

The American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) underwrites Technical Committee 9.9 (TC 9.9), which is concerned with mission-critical facilities, technology spaces, and electronic equipment.

TC 9.9 offers a number of books to help with data energy efficiency and best practices, as follows:

- Green Tips for Data Centers (2010),
  ISBN/ISSN: 978-1-933742-95-3
- Real-Time Energy Consumption Measurements in Data Centers (2010),
  ISBN/ISSN: 978-1-933742-73-1
- Thermal Guidelines for Data Processing Environments, Second Edition (2009),
  ISBN/ISSN: 978-1-933742-46-5
- Design Considerations for Datacom Equipment Centers, Second Edition (2009),
  ISBN/ISSN: 1-931862-94-X
- Best Practices for Datacom Facility Energy Efficiency, Second Edition (2009),
  ISBN/ISSN: 978-1-933742-27-4
- Particulate and Gaseous Contamination in Datacom Environments (2009),
  ISBN/ISSN: 978-1-933742-60-1
- High Density Data Centers—Case Studies and Best Practices (2008),
  ISBN/ISSN: 6331-2-403-000000-43

How to get IBM Redbooks publications

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications, and Additional materials and order hardcopy Redbooks at this website:

ibm.com/redbooks

Smarter Data Centers

Achieving Greater Efficiency

Redpaper™

Improve energy efficiency and reduce costs

Minimize stranded space, power, and cooling

Monitor, manage, and report across both facilities and IT

As we move towards becoming a smarter planet and the world becomes more instrumented, interconnected, and intelligent, the demands for data center resources are increasing rapidly. Smaller and more densely packed servers providing greater amounts of computing power can substantially increase power and cooling needs, while growing data volumes necessitate larger storage and network bandwidth capacities. Environmental and regulatory requirements can introduce additional limits on carbon emissions and water consumption.

To satisfy these demands while keeping costs in check, our data centers need to be smarter as well. Comprehensive views of data center inventories, operational and environmental conditions, and consumption across multiple capacity types that span both facilities and IT are required. You can achieve greater efficiency using hardware, software, services, and design both in facilities and IT, but you need a comprehensive data center strategy to tie them together and thus obtain a complete picture of your data center environments.

This IBM Redpaper publication discusses important considerations when creating and implementing your smarter data center strategy. Notable techniques, best practices, and technological advances that can become critical components of success are included, along with methods for bringing them together to gain in-depth knowledge of data center operations. With such insight comes increased resiliency, rapid responsiveness, profitable access to detailed analytics, and reliable planning for the future. Although not all-inclusive, this document provides a guide to getting started, points you to additional sources of information, and suggests ways IBM can partner with you in your pursuit of a smarter data center.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

REDP-4413-01