# Introduction

As the energy and climate crises loom throughout the globe, we are also witnessing an explosive growth in the demand for energy-hungry IT computing infrastructure resources. Most of the global IT data centers that host the IT infrastructure are at the brink of capacity for both energy and floor space.

Building new data centers is an extremely expensive proposition, with costs running anywhere from tens of millions to upwards of a hundred million dollars in capital and operational expenses. Meanwhile, industry analysts estimate that several billion dollars worth of investment in existing server hardware in global data centers is not properly utilized. These under-utilized resources are still consuming expensive energy and floor space in data centers, without doing much value-added work.

The need for more computing power and the inability of organizations to tap into existing, under-utilized hardware resources is driving the demand for more hardware, energy, and floor space. This trend is leading to the building of more energy-hungry data centers. This paradigm, if left unchecked, will no doubt lead to additional unused infrastructure resources, which negatively affect an organization's economic bottom line, stifle its ability to invest in its core business, and contribute to a significant increase in the carbon footprint that is a growing peril for our planet.

## Virtualization for a Smarter Planet

Through smarter use of IT resources, an organization can become more instrumented, interconnected, and intelligent. To address the challenges described above, organizations across the world have already started to leverage smarter technologies, such as *virtualization*. This technology allows us to “slice” a single piece of physical hardware (a server and storage) and the associated physical network resources into pools of many virtual resources. This allows a single physical resource to act as and host many virtual resources, with the promise of reducing IT infrastructure complexity and costs.

Virtualization reduces the energy and floor space footprint required to host the otherwise many physical hardware resources. Mature technologies from several vendors exist in the marketplace today, allowing IT organizations to adopt virtualization and to exploit several infrastructure-related benefits. However, virtualization technologies alone, without the benefit of a smartly engineered transformation effort, have fallen short of providing their promised benefits. These technologies, alone, do not provide the return on investment that resonates with an organization’s lines of business (LOBs). For LOBs, data centers are one of the several enablers of their mission to improve margins, slice down operational costs, reduce their carbon footprint, and attain business efficiencies to fuel their business growth.

Organizations that have adopted virtualization technology have, in general, seen a reduction in their energy, hardware, and floor space spending. However, the successful extraction of overall business value and benefits from virtualization technology has been extremely challenging to realize for the LOB stakeholders. Research and industry experience have revealed that these challenges arise due to an organization’s inability to properly plan, engineer, and manage the transformation effort. These challenges are exacerbated by the complexity inherent in transforming legacy IT infrastructure, applications, business processes, management systems, and the organizational structure, which are all affected by a virtualization transformation. Often, the net results are significant delays in virtualization project schedules, a substantial increase in operations costs to transform the IT infrastructure to a virtualized state, and a higher-than-expected cost to manage them, going forward. The unwanted increase in transformation and ongoing operations costs lead to the erosion of net benefits and cost savings, which would otherwise be compelling business drivers for all key stakeholders of a virtualization transformation.

This book provides an approach based on service patterns to transform IT infrastructures to a virtualized state. It also discusses how to manage IT infrastructures, while optimizing both technical and organizational resources to derive compelling business value for the organization. Our service patterns are derived from our globally acknowledged technical and thought leadership in several areas, including the following:

- Leadership, insights, and experiences from successfully executing virtualization projects and from virtualizing thousands of infrastructure resources (servers, storage, and network) for global organizations across all industry sectors
- Global experience and leadership in applying organizational theory, business process reengineering, and lean methodologies to optimize the organizational constructs and labor models required to support very complex organizational transitions and business environments
- Leadership in the development and management of hardware, software, and systems management technologies for virtualization
- Global leadership in developing and applying "green" technologies for managing global data centers built over several millions of square feet of raised floor space

This book provides an end-to-end approach for virtualization transformations, leading to efficient and cost-optimized virtualized infrastructures and data centers. Our goal is to ensure that virtualized data centers are transformed to provide business value, which transcends the traditional infrastructure cost-optimization business proposition for a CIO.

For organizations embracing virtualization technologies to address the energy and infrastructure resource challenges of the data center, our approach enables their several lines of business to do the following:

- Become more productive in managing their overall costs
- Be agile in provisioning new business services and improving time to market
- Be proactive in catering to rapidly changing business models and client demands
- Become more "charged" as an engine for fueling business growth, with innovation in technology and business process change
- Produce direct value to their clients by optimizing their value chain
- Become more successful at improving their shareholders' equity

Without the benefit of leveraging an approach such as our service-patterns methodology for virtualization transformations, most organizations will continue to stumble as they navigate their way through. They will continue to face extreme challenges in realizing the business value from developing and managing virtualized infrastructures.

## The Organization of this Book

This book is organized as follows:

- Chapter 1 describes various types of virtualization technologies. It also describes how emerging technologies, such as cloud computing, will leverage virtualization and the Web, to play a valuable role in the implementation of IT infrastructures of the future.
- Chapter 2 addresses why virtualization is important for green data centers, which are a critical component of our strategy for building a smarter planet and for reducing the carbon footprint generated by IT resources globally.
- Chapter 3 introduces the concept of a lean transformation. It discusses the eight commonly known types of waste and describes the seven lean levers that are the building blocks of our patterns for virtualization.
- Chapter 4 describes a template for patterns, which we use to document and illustrate our service patterns for virtualization.
- Chapters 5 through 11 present the seven patterns for virtualization in terms of the seven lean levers, namely: Segmenting Complexity, Redistributing Activities, Pooling Resources, Flexible Resource Balancing, Reducing Incoming Hardware Infrastructure and Work, Reducing Non-Value-Added Work, and Standard Operations.
- Chapter 12 describes the essential components of a lean implementation.
- Chapter 13 provides guidance toward building a virtualization business case to demonstrate the financial return that your virtualization project can bring to your organizational stakeholders.

CHAPTER 2

# Why Virtualization Matters for Green Data Centers

Over the past decade, we have all witnessed changes to our lives made possible by IT. Cell phones have been upgraded to smart phones, which are now being used for almost everything except making a phone call. Hybrid cars can parallel-park themselves. Personal video recorders allow you to watch days of prerecorded programming at your convenience. New laws require medical information to be kept for the life of the patient. Just imagine how much data will be kept on someone who is born today and lives to be over 100!

Similar new applications have arisen at the corporate level. However, because IT organizations often spend over 70 percent of their budgets just to maintain previously installed applications, less than 30 percent is left for tactical and strategic improvements.

At the same time, customers have become more dependent on these applications, requiring an increase in requirements for business continuity and security. Many people are familiar with the phrase “mission critical,” whose origins lie in IT applications sent aboard space missions that were not allowed to fail, such as life support systems. Availability was improved with the addition of redundant systems. Today, every industry has mission-critical applications. From neo-natal monitoring in a hospital, to stock-market trading, to monitoring the energy grid, redundant systems are put in place to prevent failure. This redundant equipment requires data center resources, which drive up the cost of IT support.

The dependence on applications has also driven the need to make them more secure. While IT organizations have voluntarily increased security, regulations have been put in place in several industries to minimize the risks associated with a security breach. In the medical profession, for example, the U.S. Health Insurance Portability and Accountability Act of 1996 (HIPAA) contains provisions for the security and privacy of health data. These additional security requirements, and the resources required to support and audit them, continue to drive up the cost of operational IT support.

The demand for new applications, combined with the need to secure them, has also increased the amount of IT equipment required. Data volumes and network bandwidth are doubling every 18 months. Devices accessing data over the Internet are doubling every 2.5 years. This increase in IT equipment has driven a corresponding increase in data center requirements, as there needs to be a place to put all this new equipment.

This equipment also needs to be powered. Studies have shown that IT equipment used about 180 billion kilowatt-hours of electricity in the United States in 2005, roughly two percent of the energy generated in the country. At current growth rates, IT's power requirements are expected to double by 2012. At ten cents per kilowatt-hour, this equates to $36 billion in IT energy costs. Unfortunately, studies have also shown that the cost of energy has increased by 44 percent from 2004 to 2009, with estimates of annual double-digit growth for the next several years. Meanwhile, most IT organizations have flat or decreasing IT budgets, so this increase in energy costs is often offset by decreases in tactical and strategic improvements to the IT environment.

To summarize, the operational challenges surrounding IT organizations are approaching a breaking point because of these trends:

- Electronic data processing demand is increasing for new applications.
- Electronic data processing demand is increasing due to business-continuity redundancy requirements.
- Electronic data processing demand is increasing due to security requirements.
- IT equipment and electricity use is tied to increased demand, raising data center costs.
- Energy costs are also going up, due to per-unit price increases.
- IT organizations must compensate for these increased costs to stay within flat or declining budgets.

The critical lesson to take from this is that IT must reduce its hardware footprint. IT must learn to make more efficient use of hardware, software, and personnel resources.

## Opportunities Abound for the Greening of IT

Fortunately, there are significant opportunities, through virtualization, to address the demand and cost issues in the IT environment. First, let's start by defining green information technology ("green IT"). Green data centers are those that make better use of facilities and information technology integration, resulting in lower energy costs, a reduced carbon footprint, and reduced demands for power, space, and cooling resources. Using this definition, it's clear that green is as much about the economics as it is about the environment.

Opportunities for "going green" exist on both the facilities and IT sides of the data center. Many people were surprised when reports were published indicating that two-thirds of the power that came into a data center never reached a piece of IT equipment. The power that did make it in ended up running servers that used, on average, less than 10 percent of their capacity. In effect, companies were paying for 10 times more than they needed to power equipment and data center resources. There are also power losses associated with the power distribution from the power plant to the data center. Figure 2.1 depicts these inefficiencies. The pie chart on the left shows the ratio between power devoted to computers and power devoted to other elements of the data center. The pie chart on the right shows the ratio of the power used for a server's processor versus other parts of the computer.

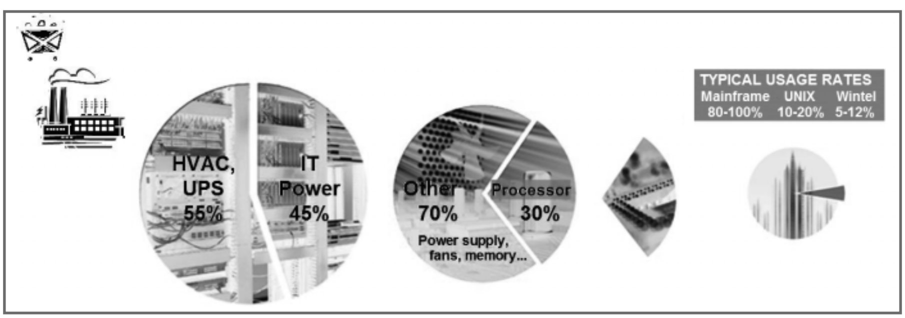


*Figure 2.1: Energy loss from the power plant to the processor.*

The inefficiencies represented in this figure are actually opportunities for the greening of IT. Let's go over them, from the source of the power to its ultimate use:

1. From the power plant to the data center, there is typically a reduction of 100 units of power down to 33 units of power.
2. From the data center to the IT equipment, there is typically a reduction from 33 units of power down to 15 units. This is because about half of the energy is used for cooling the IT equipment and for additional power distribution within the data center.
3. From the IT equipment to the processor, there is typically a reduction from 15 units down to 4.5 units, because the processor requires only about 30 percent of the energy used by the server.
4. From the overall processor to the amount used by the processor to perform productive work, there is typically a reduction from 4.5 units down to 0.45 units, as processors run at 10 percent utilization.

The combined effect of these inefficiencies is that, for every 100 units of energy produced at a power plant, only half a unit is used for productive processing within a server. Put another way, 99.5 percent of the energy generated for IT use is consumed prior to being productively used by the processor. Consider that in relation to the $36 billion in estimated energy costs for IT in 2012. Clearly, companies that can find ways to become more efficient can apply their savings to keep IT budgets flat or even reduced.

Virtualization focuses on improving the fourth inefficiency, by improving processor utilization. In a simple case, doubling the CPU utilization to 20 percent not only cuts the energy requirements in half at the processor level but has a ripple effect across the entire end-to-end energy use. In other words, doubling the CPU utilization for the same amount of IT load can also halve the amount of server hardware needed and, in turn, the amount of data center resources (power, space, and cooling) needed to house the servers. The savings in hardware and data center costs are actually more significant than the associated energy savings.

So, what do we do with the equipment that is no longer needed because of virtualization? It is estimated that one billion computers will become potential scrap by 2010. Older servers were manufactured with a small portion of hazardous substances that pose a toxic risk if not disposed of properly. About half of the companies in the United States already have eco-friendly disposal plans. This figure is likely to increase as government regulations tighten on the proper disposal of IT equipment.

Many companies are starting to offer IT equipment-recycling programs. IBM is one such company. It receives about 40,000 pieces of IT equipment a week. This equipment is either refurbished to be put out for reuse, disassembled to be used as replacement parts for failed components, or sent to landfills after hazardous materials have been removed. Less than one percent of the material received by IBM in this process goes to landfills, with none of the material toxic. This process is a combination of good economic and environmental practices. Reused servers and parts are resold, and the environment is improved by significantly reducing IT equipment in landfills and properly treating hazardous substances.

## IBM Data Center Expertise and "Big Green"

IBM operates hundreds of data centers with over 8 million square feet of data center space on behalf of its customers. Over half of this space is operated by the Information Technology Delivery (ITD) organization, which provides outsourcing services for IBM. ITD's shared data centers span a large range of sizes, from the 300,000-square-foot facility in Boulder, Colorado, to much smaller single-customer data centers. The remaining data center space is operated by IBM's Business Continuity and Recovery Services (BCRS), which provides disaster recovery services, and by its Systems and Technology Group (STG), which provides compute-on-demand services, where customers can lease computing resources on IBM premises. (Cloud computing is not a recent concept for IBM.)

One of the largest accounts that ITD supports is for IBM itself, the IBM Global Account (IGA). Between 1997 and 2007, however, IGA's production workload drastically decreased, as shown in Table 2.1.

*Table 2.1: Decrease in IGA Workload*

| Metric | 1997 | 2007 |
|---|---|---|
| Chief Information Officers | 128 | 1 |
| Traditional Data Centers | 155 | 7 |
| Web Hosting Data Centers | 80 | 5 |
| Networks | 31 | 1 |
| Applications | 15,000 | 4,700 |

The IGA moved from 235 data centers to a dozen, converging into one network in the process and implementing virtualization techniques for server and storage consolidation. By combining organizations, IGA was also able to converge to fewer applications. The reduction in applications improved labor costs; in the previous distributed environment, for instance, each data center required its own IT organization and Chief Information Officer. It also reduced the amount of data center and IT resources required to support the consolidated organization.

Leveraging what it had learned from IGA and other accounts, IBM announced the Project Big Green initiative in May 2007. The announcement stated that IBM would reallocate $1 billion per year to accelerate green technologies. The investment had two components:

- To develop new hardware, software, and services for green IT
- To invest in our internal data centers to inject these newly developed technologies

Based on its own experiences, IBM would offer a roadmap for clients to address IT energy costs. The initiative also provided for the training of a thousand energy specialists within IBM, to help customers address their energy concerns.

The emphasis on energy conservation was not a new one for IBM. From 1990 to 2005, energy conservation efforts at IBM resulted in a 40 percent reduction in carbon dioxide emissions and more than a billion dollars in energy savings. The Project Big Green initiative specified that $100 million of the $1 billion investment would be spent on infrastructure to support best practices in remanufacturing and recycling.

The ability to grow IT business while saving costs through the injection of new technology was a key value proposition for customers. A 25,000-square-foot data center with an average density of 40 watts of energy per square foot consumes $2.6 million in energy costs annually. The use of the new technologies promoted by IBM cuts the energy use of this data center in half, saving $1.3 million per year. Companies with larger data centers save even more. The economic energy savings are matched by the corresponding environmental savings of a reduced carbon footprint.

## New Models for Efficient IT Delivery

With the rapid improvements in IT technology, many IT organizations have established a process for continuous improvement. In fact, the Information Technology Information

Library (ITIL®), the most popular industry guideline for running an IT service, has a whole book devoted to continuous service improvement. The following three models show an evolution of the data center and the vital role that virtualization plays in the continuous improvement paradigm:

- New Enterprise Data Center (NEDC) model
- Dynamic Infrastructure model
- Smarter Planet model

## *The New Enterprise Data Center Model*

In February 2008, IBM launched an updated strategy around the data center that formalized the direction it has taken over the last several years. Based upon previous initiatives, such as Service Oriented Architecture (SOA), Information Systems Management (ISM), Information on Demand (IOD), and IT Optimization (ITO), this strategy brought together major elements to help clients evolve to a more efficient and responsive model for IT service delivery.

The NEDC model has three stages of adoption: simplified, shared, and dynamic. In the simplified stage, IT organizations consolidate data centers and physical infrastructure such as servers, storage, and networks.

As the workload is consolidated, organizations can progress to the shared stage of adoption. To reap economies of scale, virtualization is used to place multiple workloads on the same physical piece of equipment, thereby improving cost, availability, and performance. It is typical in the shared stage to statically map virtual workloads to physical resources.

The dynamic stage of adoption provides policy-based system management that lets virtualized workloads move between physical resources without down time. Based on historical analyses of workload performance, automation functions are written and placed into the system management code to optimize the IT environment by moving the workload dynamically, in response to workload changes. This allows an IT organization to improve the utilization of its physical resources, as fewer excess resources are needed to respond to changes in workload.

## The Dynamic Infrastructure Model

Even with the NEDC vision, additional progress was required as new technology continued to be developed. As of 2007, 85 percent of distributed capacity was still sitting idle. Consumer product and retail industries lost about $40 billion (or 3.5 percent of their sales) due to supply-chain inefficiencies. On average, 70 percent of IT budgets were spent maintaining current infrastructures, instead of adding new capability. Data explosion drove an increase of 54 percent in storage. Better security was also demanded, driven by the fact that 33 percent of consumers notified of a security breach terminate their relationship with the company, perceiving the company to be responsible.

In response to the continuing need to simultaneously improve service, reduce cost, and reduce risk, IBM announced the Dynamic Infrastructure model. The model consists of the elements shown in Figure 2.2.


*Figure 2.2: The Dynamic Infrastructure model for continuous improvement.*

As you can see, virtualization and energy efficiency are two of the elements of the model. Implementing virtualization through the Dynamic Infrastructure model enables organizations to do the following:

- *Reduce operating costs*. Virtualization can consolidate a workload onto fewer servers and accessories. The reduced number of devices, in turn, simplifies management of the infrastructure. The recaptured floor space from workload consolidation can either be used for growth or simply eliminated.
- *Improve service responsiveness*. Properly designed virtualization solutions improve system, network, and application performance. Processing more information in real time enables organizations to make better business decisions. Virtualization techniques also allow organizations to bring new services online quickly, as the wait for physical server installation decreases.
- *Manage availability in a 24-by-7 world*. By designing virtual solutions with high availability characteristics, organizations can increase availability and improve resiliency. They can also manage and secure data without negatively affecting its availability.
- *Dynamically adapt to the peaks of the business*. Virtualization allows organizations to dynamically deliver resources where they are most needed. If included in the virtualization solution, workload balancers can measure resource utilization and be directed through policy-based management to automatically optimize workload on the existing resources.

By implementing energy efficiency through the Dynamic Infrastructure model, organizations can do the following:

- *Reduce energy use*. As part of virtualizing the workload onto fewer physical resources, organizations can implement virtualization on more energy-efficient servers and storage. Improving the facilities infrastructure efficiency reduces the energy needed to supply power and cooling to the IT equipment, further reducing energy requirements.
- *Reduce capital and operating costs*. Improving facilities and IT efficiency extends the life of an existing data center and avoids or delays the capital costs of building new data centers. The reduction of energy usage reduces overall operating costs.
- *Measure and control energy usage*. By understanding energy consumption, organizations can determine the positive effects of such changes as virtualizing servers. This data can then be used to put in place policies to manage and control energy use via techniques such as dynamic image movement to minimize energy consumption.
- *Establish a green strategy*. Many governmental organizations have already unveiled voluntary programs for improving data center efficiency, such as the

> European Union's Code of Conduct for Data Centers and the U.S. Environmental Protection Agency's (EPA's) Energy Star Rating for Data Centers. These voluntary programs are being extended by regulatory requirements for stricter controls over energy consumption, carbon footprints, and the disposal of hazardous substances. In response to these emerging requirements, companies are establishing green strategies to reduce energy usage, which also reduces costs.

Virtualization itself is a key tenet within the Dynamic Infrastructure model to improve service, reduce cost, and manage risk. Virtualization also improves the other tenets of asset management, information infrastructure, business resiliency, and security, as shown here with energy management.

### *The Smarter Planet Model*

The world is getting smaller, flatter, and hotter as we integrate our global economy. The economic downturn requires us to do more with the same amount of resources. The effects of climate change are both a societal and a business concern. Meanwhile, energy use is rising at an unprecedented rate. The same IT techniques being used inside the data center can now be leveraged outside the data center, to reduce costs by eliminating waste in every industry.

This is the premise of the Smarter Planet model. Through the additional use of IT, an organization can become more instrumented, interconnected, and intelligent. Similar autonomic analysis can be performed on the data from each industry, and better decisions can be made not only to improve the quality of products produced but also to reduce their costs.

Smarter water use is produced by applying monitoring and management technologies to help reduce the use of water, as well as related energy and chemicals. Smarter traffic employs dynamic traffic prediction. This provides the basis for dynamic tolling, which leads to reducing congestion. Its byproducts positively improve customer traffic flow and conserve energy. Smarter energy conservation is promoted by analyzing customer usage patterns and providing customized products and services that boost efficiency from the source, through the grid, to the consumer.

Additional analytics produce a significant increase in productivity and reduced costs. These analytics increase the demand for IT resources. Virtualization is a major component in the Smarter Planet model to optimize the amount of IT resources required through higher utilization of physical resources.

## The Five Building Blocks of Green Data Centers

There are five major building blocks for building and maintaining green data centers, summarized in Figure 2.3.

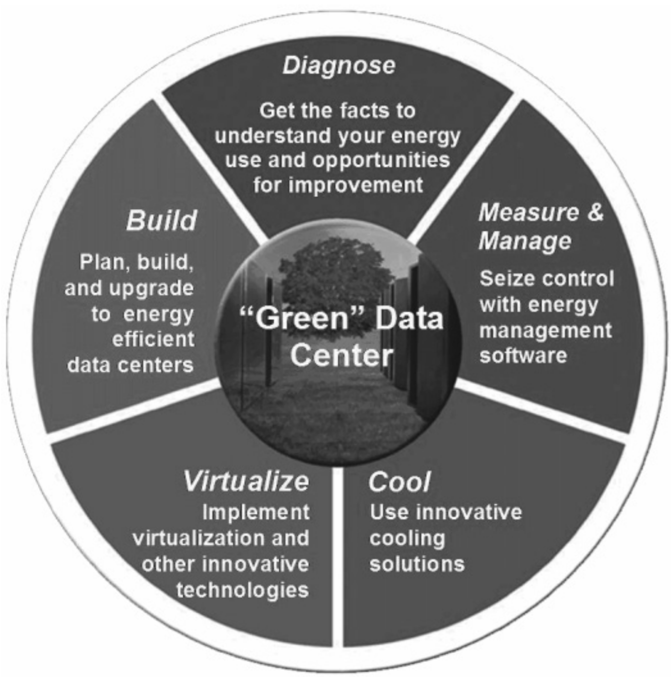


*Figure 2.3: The five building blocks of a green data center.*

### *Building Block 1: Diagnose*

The first building task is to diagnose the data center, getting the facts so that you understand your energy use and opportunities for improvement. There are many metrics to use to diagnose the data center. At the highest level, the measure that has been most widely adopted is the ratio of power coming into a data center divided by the amount of power used by the data center's IT equipment. This metric has been further defined by the Green Grid, an industry organization formed to improve data center efficiency. The metric is called *Power Usage Effectiveness*, or *PUE*. In 2006, a survey showed that data centers typically had a PUE of 3 or higher, indicating that it took three watts of electrical energy coming into the data center for a piece of IT equipment to use one watt. The extra two watts were used to cool the IT equipment, or in power losses within the data center (from such things as charging emergency batteries, transforming to different voltage levels, and converting alternating current to direct current).

In addition to the PUE metric, organizations should also understand the facts around the major subsystems and components of the facilities and IT equipment. On the facilities side, this means keeping track of the utilization metrics for the power (utility, battery, and generator), three-dimensional space (how much space is left for IT equipment), and cooling (how many tons of cooling are required to keep equipment safe from overheating). On the IT side, this means keeping track of the utilization metrics for servers, storage, and network equipment.

At one time, these metrics were obtained manually. It was not uncommon for facilities personnel to walk around the data center periodically and record the metrics. In more recent times, the facilities and IT equipment have increased their instrumentation, interconnection, and intelligence (as indicated in the Smarter Planet model), so these metrics are now collected automatically.

## *Building Block 2: Build*

There are three ways to improve the efficiency of a data center:

- Extend the life of an existing data center.
- Leverage the contributions of multiple data centers.
- Build a new data center.

Each of these ways involves a change of some sort (small or large), which for simplicity is called “build” here. This block includes the planning, building, and deployment of more efficient data centers. The result of the build should improve the PUE and utilization metrics obtained in the diagnose phase.

Extending the life of an existing data center starts by diagnosing which data center component is constrained. It is typical for older data centers to have developed different levels of capacity between major subsystems. For example, they might have 2 megawatts of Uninterruptible Power Supply (UPS, or batteries) and 2.4 megawatts of generators because they purchased UPS in 1-megawatt units and generators in 0.8-megawatt units. As both the UPS and generator capacity are needed for additional IT equipment, the UPS will become constrained before the generator. This mismatch is a source of optimization for data centers. In this example, adding 0.4 megawatts of UPS capacity uses up the excess capacity of the generator and postpones building a new data center.

Leveraging multiple data centers applies the same optimization concepts from a single data center to a collection of them. For example, there might be excess computing power in one data center, and excess storage in another. By understanding the constrained and surplus resources of the various data centers, an organization can move IT workload between them to better optimize its resources. Likewise, it might be possible to consolidate data centers. In many cases, they can be consolidated with a minimal amount of additional resources.

As is best practice for all major IT changes, a detailed business case should be prepared for any data center consolidation, so that you can understand the investments you would need to make and the returns that would be achieved.

## *Building Block 3: Virtualize*

Most of the building blocks for green data centers apply to the physical facilities side of the data center. This block applies to the IT side. Although the primary focus of this book is the virtualization of servers, a number of other types of equipment—storage, network, and desktops—can also be virtualized to improve IT efficiency.

In the simplified adoption stage of the New Enterprise Data Center model, applications are dedicated to specific IT devices. Although this provides isolation and capacity planning at the application level, it is extremely inefficient, as excess capacity from one application cannot be used by another application. Virtualization removes that constraint, as it maintains isolation of server instances in the software, while logically partitioning the hardware in such a way that hardware resources can be shared between partitions when needed. This allows an IT organization to achieve economies of scale by pooling IT resources. It also decreases the amount of excess resources for a particular application, by making sure that sufficient resources are available in the pool.

## Building Block 4: Cool

For the many data centers where the PUE is above two, more energy is being spent on cooling the data center than on IT equipment. Because this is the long pole in the tent, the IT industry has focused on improving cooling technology within data centers, both tactically and longer term. In many data centers, this is "low-hanging fruit" because they did not prioritize the management of airflow. Increases in the cost of cooling are causing them to take another look.

Many organizations can extend the life of their existing data centers by improving airflow. For instance, a best practice in the data center creates hot and cold aisles of IT equipment by facing the exhaust sides of the equipment toward each other, with cold air coming up through the floor via perforated tiles in the cold aisle.

Inefficiencies in airflow are created when hot air mixes with cold. This situation is all too common with cable cutouts in the floor, or racks of equipment with gaps in the rack that allow the air to mix. Low-cost solutions are now available to easily seal cable cutouts and install faceplates in racks of equipment. This improved airflow increases efficiency and reduces the amount of cooling needed. In turn, this efficiency translates to either lower energy costs or the ability to reuse the cooling capacity elsewhere in the data center.

## Block 5: Measure and Manage

As mentioned earlier, initial metrics should be obtained as a baseline for any changes made to the data center. As changes are being made, the effects of these changes should be measured to ensure that they have the anticipated positive result.

As data centers continue to become more dynamic, the need to manage the environment becomes more important. An example is the implementation of the recent American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) guidelines, which recommend that data centers can be operated at increased temperature and relative humidity. It has been suggested that temperatures be increased gradually up to the guideline, in order to understand the effects of the change in different locations in the data center and prevent any dangerous hot spots that might occur.

The latest release of the ITIL (version 3) includes a book devoted to the continuous improvement of IT. With this premise that IT constantly changes and never obtains a final state, you should continue to measure and manage your IT environment, to identify any problem situations or opportunities for improvement.

## Data Center Transformation: A Global Case Study

This section demonstrates how the five building blocks of green data centers can be applied to an older data center. The data center in this example was built in the 1980s for the IT equipment that existed at that time. Its life was extended using the five building blocks.

### *Building Block 1: Diagnose*

The PUE for the Lexington data center was measured at 2.0. This means that for every watt of energy used for IT equipment, an additional watt of energy was needed for cooling and power distribution. On the facilities subsystem level, 43,000 of the data center's 44,000 square feet were in use. The UPS system was at 92 percent utilization (2,227 kilowatts of 2,430 kilowatts), the generator system was at 85 percent utilization (4,229 kilowatts of 4,998 kilowatts installed, noting that additional equipment was installed on the generators that wasn't installed on the batteries), and the chiller system was at 94 percent (1,175 tons of 1,250 tons installed). On the IT level, 60 percent of the single processor servers were operating with a monthly utilization of less than 5 percent. (On average, there was 20 times more equipment on the floor than was needed.)

### *Building Block 2: Build*

The data center was a leased facility, with little room for increases in facilities equipment. We were able to add 170 kilowatts of power to the data center. This created the equivalent of 4,250 square feet at the 40-watts-per-square-foot power density of the existing space. Because there was actually less than 1,000 square feet of physical space available, something had to be done to release some space so that the additional power could be used.

### *Building Block 3: Virtualize*

The majority of the servers had a monthly utilization of less than 5 percent, so the opportunity existed to increase the utilization of each hardware server through virtualization. This would reduce the number of servers handling the current load and therefore create sufficient space to install the new servers enabled by the power upgrade. The net effect of virtualization allowed the data center to expand the IT server capacity by 800 percent, with less than 10 percent additional power capacity and no additional cooling capacity. The number of servers with less than 5 percent utilization was reduced from 60 to 14 percent over a period of two years.

### Building Block 4: Cool

The cooling system at this data center was installed 20 years ago, when there were water-cooled mainframe computers. A data center energy-efficiency survey indicated that about a dozen Computer Room Air Conditioners (CRACs) could be turned off without adversely affecting cooling. By turning these off, the data center was able to reallocate almost 500k kilowatts of energy to other devices.

The survey also turned up several other tactical alterations that could increase the efficiency of the airflow. These included sealing cable cutouts, installing blanking plates in the racks of IT equipment, and rearranging some of the perforated tiles in the data center. The recommendations identified actions that resulted in a 10 percent savings of energy use.

### Building Block 5: Measure and Manage

By implementing the suggestions from the energy study and virtualizing the IT workload, the data center was able to improve its PUE from 2.0 to 1.8, while increasing the IT capacity by a factor of eight.

By measuring the utilization of the newly virtualized servers, the capacity planning organization was able to create a utilization baseline that smoothed out, or "normalized," the workload on the larger servers. In essence, the workload in the virtual images did not all peak at the same time, allowing the overall utilization on a particular server to have less variability than when the workload images ran individually. The IT organization was able to "right size" the virtual servers based on this reduced variability. By rearranging virtual images and updating their size, the organization was able to reallocate six 16-way servers to other projects. This resulted in a savings of $2.4 million in equipment costs.

Metrics on equipment utilization are now reported automatically, and the IT organization looks for optimization projects as part of its continuous improvement program.

## Why Virtualization Is Key for Green IT and Data Center Transformations

There are many reasons why virtualization is key for green IT. We focus here on the top three: efficiency, economics, and transformation.

### *Promoting Efficiency*

It wasn't that long ago when distributed workload was physically mapped to specific hardware. With a multi-tier application, it was common for the Web content to be on one server (or set of servers), the database to be on another, and the application business logic to be on a third. Application developers also needed a place to test fixes against the current production environment and a place to develop new functionality for the application. This often resulted in having nine servers per application: three per environment. As it was difficult to add physical resources to a running environment with minimal technology to share resources, the capacity planners often sized the workload for peak usage instead of average usage. This entire paradigm led to surprisingly low utilization. This was a reasonable approach until the physical data center became constrained and new equipment could no longer be added to it.

Virtualization reduces the waste of IT resources by putting multiple workloads on a single physical resource and sharing the equipment's excess resources. This reduces the amount of physical IT resources needed and its corresponding effect on facility resources. It is not uncommon to combine five older servers into the facility resources (power, space, and cooling) of one new server, increasing average single-digit utilization by a factor of five. This is further improved by the increased IT capacity of a new piece of IT equipment, as processors are faster (can do more work), memory is more compact, and the network bandwidth has improved. A five-to-one consolidation ratio, which is relatively modest, releases 80 percent of the facility's resources for additional growth. Reusing facilities resources makes the facilities more efficient as a byproduct of IT efficiencies.

### *Financial Impacts*

Virtualization has a positive effect on financial capital and expenses. (A detailed analysis of the financial implications of virtualization is presented in Chapter 14.)

On the capital side, eliminating 80 percent of the servers, as indicated in the example, reduces the amount of capital needed for new equipment. Existing equipment can be reused for operations that would have required new investment. In a physically constrained data center, capital requirements for new facilities equipment is reduced, as virtualization provides the method to extend the life of an existing data center and grow IT capacity without the corresponding increase in facilities capacity.

On the expense side, the most obvious virtualization savings for green IT is associated with the reduction in the amount of energy needed to operate the data center. The energy reduction at the server level is multiplied by the PUE of the data center, which provides savings from not having the corresponding energy consumption of the facilities (power and cooling) equipment.

The reduction of IT resources due to virtualization also simplifies the environment, with a corresponding expense savings. An example is fewer network ports (connections) to maintain, with less cost associated with cabling installation and maintenance. The ability to dynamically move workloads between physical resources allows organizations to non-disruptively perform hardware microcode updates, preventing outages and saving the money associated with those outages.

## *Successive Transformations of the Data Center*

Many managers considered virtualization predominantly as a way to reduce IT hardware costs through increased IT utilization. That is just the start. A reduction in servers lowers most other data center costs as well.

The next extension of transformation is the dynamic relocation of workloads without disruption. This is implemented not only to perform preventive maintenance on the physical equipment, but to respond to changes in capacity requirements of the workload. The addition of workload-balancing analysis at the image level provides policy management for dynamic image movement without human intervention. This policy management allows virtual workloads to run at even higher levels of utilization, further reducing IT hardware requirements.

The newest and most extensive transformation enabled by virtualization for green IT is the agility to position IT organizations for the future. An example of this is cloud computing. All of the public cloud computing implementations are based on some level of virtualization. This makes virtualization a prerequisite for those wanting to leverage cloud computing.

## Summary

In this chapter, you learned about the forces driving the transformation of today's data centers. While you could consider these forces pessimistically, they provide the incentive for IT organizations to make the changes necessary to implement future IT requirements.

Virtualization technologies provide an excellent mitigation plan for green IT. In fact, virtualization was identified in the EPA's report to the U.S. Congress as the key IT technology that will improve data center efficiency in the next few years, decreasing the amount of energy consumption needed for the data centers of the future.

In this chapter, you learned the five building blocks for green data centers, examined a case study that applied those building blocks to a specific data center, and saw why virtualization is key for green IT. Although the majority of the building blocks are facilities-related, virtualization on the IT side is considered significant enough to be its own building block.

Green IT is as much about economics as the environment. In this chapter, you've looked at green IT in terms of optimization techniques implemented in the data center to integrate the facilities and IT components. This results in a lower carbon footprint with the same or higher functionality.