Data center defined

The infrastructure behind a digital world

MIETEK GLINKOWSKI – Today's mobile society means that people are consuming and creating data at unprecedented levels – the Internet, search engines, mobile apps, smart phones – all are omnipresent, yet their existence is basically taken for granted. The reality is that all of today's mobile gadgets, and more and more of all business enterprises, depend on the storage, networking and processing of digital data, nearly all of it via or inside a data center. Without question, data centers are the backbone and unsung heroes of the Internet boom, and have become a vital industry for organizations to run mission-critical applications. ABB provides a wide range of products, integrated solutions and expertise that ensure data centers operate safely, reliably and efficiently.

Title picture
In today's world of unprecedented amounts of data use and storage, ABB is helping organizations run mission critical applications.

Current state-of-the-art data centers are highly specialized industrial facilities, full of intricate and interrelated equipment and systems with particular mission-critical needs → 1. Some may be small buildings of 200 m², others the size of 15 soccer fields (about 140,000 m²). Some require 500 kW of power, others 100 MW.

The field is expanding at a tremendous rate. For example, globally, the number of IT racks in 2012 reached 7.7 million – an increase of 15 percent compared to 2011[1]. Estimated growth for data centers this year in the United States was 25 percent with some countries, for instance Turkey, reporting a 60 percent growth. The expansion of the corporate data center industry was well captured in a report by Digital Realty.[2] → 2 shows the most important performance factors and features fueling the expansion of the industry. Energy efficiency and security were viewed as extremely important, whereas consolidation, connectivity and redundancy were rated as very important to somewhat important. ABB provides cost-effective solutions to meet the needs of today's data centers.

1 Segments of the data center industry

There are a variety of distinct industry segments in which data centers are needed.

Colocation/hosting
Many small- and medium-size businesses do not want or cannot afford their own IT infrastructure such as data centers and so they outsource their IT needs to colocation companies. These companies provide IT services, from web hosting, to enterprise IT hosting, to other businesses. This segment of the data center market is clearly focused on revenues from IT; for them the data centers are the primary business offering.

Financials
Banks and other financial institutions such as the New York Stock Exchange (NYSE), NASDAQ, Tokyo Stock Exchange (TSE), etc. need data centers and their high availability to perform financial transactions but data centers per se are not their source of income.

Telecom
From landline digital services to the mobile and smartphone, telecom providers play a major role in the data center industry. Today, virtually all phone services are digital and many of them use VoIP, utilizing the connectivity of the Internet. Major players such as NTT, AT&T, T-Mobile, all own, build and operate data centers.

IT services
Companies such as Google, Amazon, eBay, Facebook and others debuted with the Internet boom approximately 15 years ago. Although these companies rely on data centers as their primary assets, their revenue stream varies from advertising to online shopping. They are innovative in their way of building data centers, providing services and serving customers.

Government
In 1999 the US Federal government operated 432 data centers; in 2013 this number had risen to about 7,000*. This includes everything from the Internal Revenue Service to the Department of Defense and Social Security Administration. For government agencies data centers are a cost.

Healthcare
This segment is expected to grow rapidly with the emerging trend of digitalization of patient records and all medical data from private doctor's visits to hospitalization and major surgeries. For the healthcare industry data centers are a cost.

Corporations, retail, manufacturing, utilities
This includes a large group of private and publicly traded companies in a variety of industries such as oil and gas plastics, retail store chains, and power, gas and water utilities. Although many small and midsize corporations would choose collocation services the larger companies own and operate their dedicated data centers. For example, in Singapore, BP operates its Most of the World (MoW) Mega Data Centre, one of four mega data centers from which BP runs its global IT operations.

Cloud computing is not considered a segment, but rather a service, within the database industry. It is a means of distributing IT applications over a number of physical servers and even physical data centers. There is no longer a direct relationship between an application and a physical device or even physical data center. A good example of this is Apple's iTunes application where data – eg, music, videos, movies – is distributed over a combination of servers and separate Apple data centers. This distribution is dynamic, ie, it depends on resources, availability of IT (as well as power, cooling and several other factors), Internet traffic, etc.

Footnote
* Government Accountability Office of the US Government, 2013, www.gao.gov

Data centers consume large quantities of electrical energy. Current estimates are that up to 2 percent of global energy is consumed by data center enterprises.[3] With the global installed electricity capacity of about 5,000 GW[4] this means data centers consume about 120 GW, almost twice as much as the electricity capacity of Mexico, and more than the countries of Spain or Italy.

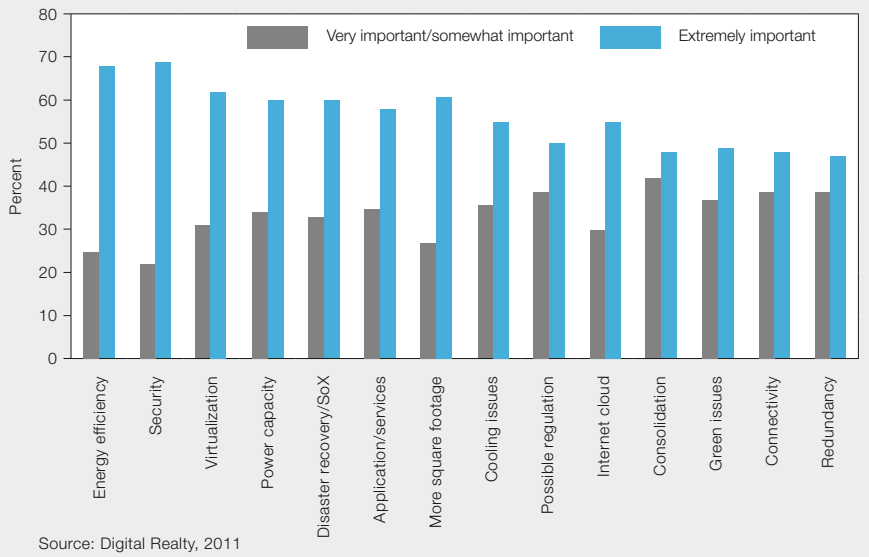
A large variety of software, databases, operating systems and clouds run in data centers.

What is a data center?

Data centers can be defined as three side-by-side infrastructures – IT, power and cooling → 3. The three infrastructures have to be perfectly compatible, matched, and optimized to provide seamless operation of the mission-critical facility → 4.

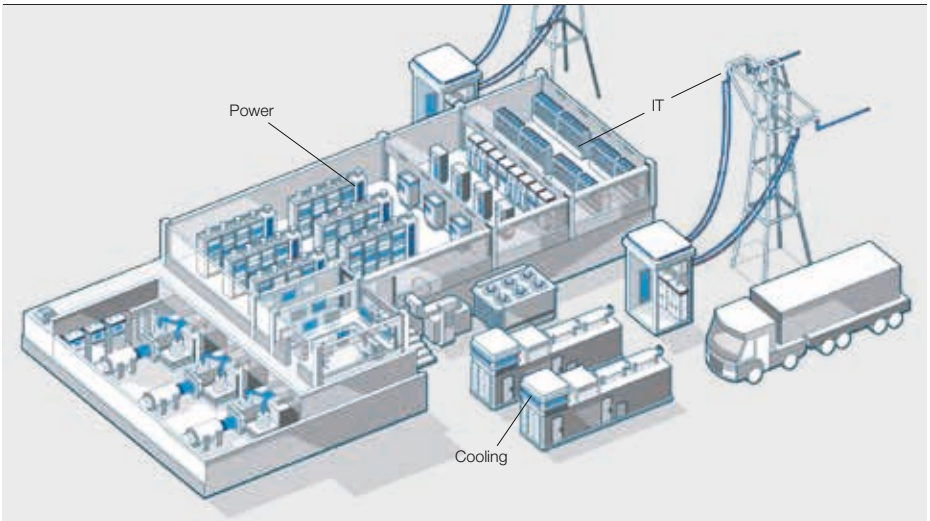
The IT infrastructure contains primarily the IT equipment with its associated software. The equipment is typically grouped into three categories: servers, network switches and storage (memory). Each group has its unique function; however in many cases servers

2 Where do professionals see priorities in data center expansion?

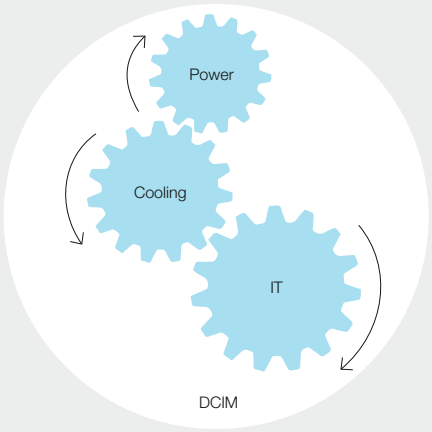


Source: Digital Realty, 2011

3 Physical layout of a generic data center



4 Data center infrastructure



Data center infrastructure includes the interplay of IT, power and cooling with DCIM.

contain storage. This infrastructure is where the main functions of the data centers are implemented and the IT services are delivered. A large variety of software, virtualization, databases, web hosting, operating systems, and clouds run in data centers.

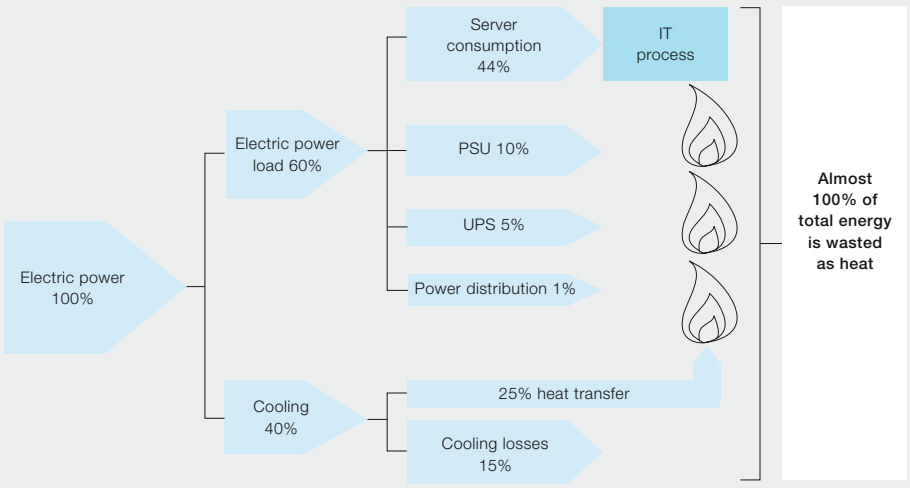
Power and cooling are the two infrastructures necessary to operate the IT equipment. Power is primarily in the form of grid electricity (although there are some exceptions, such as fuel cells). Power is delivered to the IT equipment via complex topologies of transformers, switchgear, gensets (rotating engine generator sets), uninterruptible power supplies (UPSs), busways and automatic transfer switches. The raw power from the utility is transformed, converted, conditioned and distributed to the servers in the IT racks.

IT equipment generates a lot of heat. The power infrastructure accounts for 60 percent and cooling accounts for 40 percent of the energy consumed in a generic data center → 5. The power usage effectiveness (PUE) factor is equal to PUE = 100/55 = 1.82, which is better than the industry average of 1.9. The power infrastructure can be broken down into four components, eventually leading to the IT processes (IT equipment) consuming about 44 percent of the total. Nearly all of the electricity flowing through the power infrastructure and used in cooling is lost as heat.

Footnotes

1 Data Center Dynamics Converged – Media Pack 2012
2 What is Driving the US Market? 2011, Digital Realty Trust
3 The estimates vary from 1.1% to 2.5 %; see multiple sources: http://www.analyticspress.com/data centers.html, www.greenpeace.org, www.forbes.com
4 Data as of 2010 EIA.gov

5 Simplified energy balance for a generic data center



The electric energy is consumed 60 percent by the power infrastructure and 40 percent by cooling.

The power usage effectiviness (PUE) factor would be equal to PUE=100/55=1.82*), which is better than the industry average. UPS losses and all cooling power are counted as overhead power.

The power infrastructure can be broken down into four components, eventually leading to the IT processes (IT equipment) consuming about 44 percent.

Footnote

* The alert reader may be confused that the P_{load} figure is 55 rather than 60. The difference of 5 percent is accounted for by the the UPS losses shown.

This heat has to be removed to assure that the operating temperatures of the equipment stay within the specifications and that the environment around the equipment can be accessed by personnel. Data centers employ very sophisticated and diverse cooling systems to control this environment, including liquid cooling, air cooling, immerse cooling, hot-aisle containment, cold-aisle containment, computer room air conditioners (CRACs) and computer room air handler (CRAH) units. Cooling is the primary component of the energy consumption responsible for the overhead power, ie, PUE factors above 1.0 ➔ 6.

Another component of the infrastructure, data center infrastructure management (DCIM), is becoming increasingly more important. DCIM is a platform to collect, control, integrate, monitor and manage all the systems of the data center. Ensuring that the temperature sensors of the cooling CRAC units are set properly to match the temperature requirements that servers read on their own motherboards is not a trivial task, nor is making sure that the power distributed to the racks of the IT equipment loads the individual feeders in a uniform fashion and does not overload individual cables and circuit breakers. Keeping track of where the IT equipment is located, what purpose it serves, when it needs to be replaced, or who owns it (in the case of a colocation company) is also necessary. All of these functions and more can be handled by a DCIM platform consisting often of both hardware and software to collect the data (eg, temperature, voltage, current, air flow, alarms), process it, display it and enable an operator to make informed decisions. DCIM is referred to as the glue that holds all the components of a data center together – an all-encompassing umbrella for the data center business.

6 Performance indicators

Power usage effectiveness (PUE) is the most common key performance indicator (KPI) for data centers today. It is defined as

$$PUE = \frac{P_{Total}}{P_{IT\ Load}}$$

where P_{Total} is the total power consumed by the data center, $P_{IT\ Load}$ is the power consumed by the IT load. By definition PUE is always greater than 1.0; everything above 1.0 is overhead power consumed by other non-IT loads, such as cooling, lighting, security systems.

The average PUE reported by the environmental protection agency (EPA) in the United States in 2007 was 1.9 (90 percent overhead power consumed). In 2012 Digital Realty reported that the average PUE for non-IT companies was even worse, approximately equal to 2.9.

However there is more to the energy consumption than PUE. For example, if a data center owner improves the energy consumption of the IT load – ie, replacing older servers with the newer technology – and keeps the same cooling system the PUE will actually increase since the denominator of the PUE equation decreases.

In some cases this can be a disincentive to modernize facilities. In other cases in an effort to improve the PUE factor, data centers switch to more water-intensive cooling and therefore consume more water. This is why a new set of KPIs have been introduced, including water usage effectiveness (WUE). Carbon usage effectiveness (CUE) is a data center measurement that takes the total CO_2 and carbon equivalent emissions produced as a result of the data center energy used and divides it by the energy of the IT equipment housed in the data center; the value is in $kgCO_2e/kWh$.

Mietek Glinkowski
ABB Data Centers
Raleigh, NC, United States
mietek.glinkowski@us.abb.com

Further reading

www.abb.com/datacenters