

DATA CENTER FOR ECOMMERCE BUSINESS CONTINUITY

Prof. Dr. Srisakdi Charmonman

President of Thailand Chapter of the ACM
charm@ksc.au.edu

ABSTRACT

Business continuity is one of the most important factors for business, especially eCommerce, to be successful. If an eCommerce site does not function well when any customer is trying to visit, that customer will most likely go to competing sites. Worse is when a regular and loyal customer is frustrated with the lack of business continuity in his favorite eCommerce site, the site may lost that customer forever. In the old days, business continuity might be attempted by using backup sites. However, in the modern days, data center is more appropriate but there are still problems with electrical outage as well as denial of service attacks and so the combination of data center as well as back up and Internet security precautions may be used. This paper presents ANSI/TIA-942 Telecommunications Infrastructure Standard for Data Center, outsourcing eCommerce data centers, in-house eCommerce data centers, green eCommerce data centers, and sample data center problems such as electrical outage and distributed denial of services.

Index Terms— *eCommerce, data center, ANSI/TIA-942, green data centers, outsourcing data centers, in-house data centers, electricity outage, DDOS.*

1. INTRODUCTION

In the years 1977-1980, the author was the General Manager and Acting Director of the Computer Center at Bangkok Data Center Co., Ltd at the Royal Turf Club. On Saturday's or Sunday's, the computer system was used to support horse racing. If the computer system failed to compute the amounts to pay the bet winners, they would be angry and the mob might be uncontrollable. So, the computer system was designed with two of every part. For examples, there were two central processing units, two sets of tape drives, two sets of disk drives, two printers, two air conditioners, and two groups of operators. Electricity was backed up by a fly wheel system. The large fly wheel had a diameter of a few meters. It was said that had the wheel got out of its position, it would have rolled a few hundred meters. When electricity from the city was available, the wheel would be turning. When there is an outage of the city

electricity, the momentum would keep the wheel turning and producing temporary electricity until the diesel generator started automatically. The fly wheel system and the diesel generator provided business continuity for the Royal Turf Club. However, the system did fail such as when both the air conditioning systems malfunctioned at the same time.

Another electricity backup system was at Bangkok Bank when the author served as full-time Advisor to the President around the years 1980-1983. In that case, hundred of car batteries were used to bridge the gap from the city electricity and diesel generator.

For some important computer systems such as for the US President, three of every parts of the computer systems may be used. Some fault-tolerant computer systems have also been designed.

As of the year 2009, all eCommerce sites cannot continue operating without the Internet. Earlier, some kinds of backup similar to the case of the Royal Turf Club were used but later the concept of “data center” appeared in the 1990s, when more powerful microcomputers of the “server quality” were used to replace mainframes.

According to Wikipedia, “a data center or datacenter or datacentre or server farm is a facility to house computer systems and associated components, such as telecommunications and storage systems” [13]. A data center generally includes redundant or backup power supplies, redundant data communication connections, and environmental controls and security devices. During the 2000s, organizations needed nonstop operation and 24/7 presence on the Internet, and so data centers became more popular. As a matter of fact, they were called “Internet DataCenters or IDC”. A large eCommerce site may have its own in - house data center. A small or medium eCommerce site may use an outsourcing data center.

2. ANSI/TIA-942 TELECOMMUNICATIONS INFRASTRUCTURE STANDARD FOR DATA CENTER

The Telecommunications Industry Association (TIA) was the first organization to release the standards named TIA-942 Telecommunications Infrastructure Standard for Data Center which includes guidelines on a wide range of

subjects [12]. It provides a flexible and manageable structured cabling system using standard media. Later, ANSI (American National Standards Institute) also approved TIA-942 and TIA-942 became “ANSI/TIA-942 Telecommunications Infrastructure Standard for Data Centers” which covers the requirements on site space and layout, cabling infrastructure, tiered reliability, and environmental consideration.

2.1. Site Space and Layout

The data center should be designed with plenty of flexible “White Space”, i.e. empty space that can accommodate future racks or cabinets. For future growth, surrounding empty space should also be available. According to ANSI/TIA-942, a data center should include the following key functional areas such as Entrance Room (ER) which is outside the computer room for better security, Main Distribution Area (MDA) which is location of Main Cross-Connect (MC) such as routers and switches, Horizontal Distribution Area (HDA) which is location of Horizontal Cross-Connect (HC) such as racks for fiber and coaxial cable, Zone Distribution Area (ZDA) which is location of Zone Outlet (ZO) or Consolidation Point (CP), and Equipment Distribution Area (EDA) which is location of equipment cabinets and racks.

2.2. Cabling Infrastructure

Horizontal cabling should be installed with the highest capacity media available to reduce the need for re-cabling in the future. The maximum horizontal distance is 90 meters independent of media type. The maximum channel distance including equipment cords is 100 meters. The maximum backbone fiber optic cabling is 300 meters. The maximum horizontal copper cabling is 100 meters. Computer room and entrance room requirements are such as minimum clear height of 2.6m/8.5ft and minimum door size 1m/3ft wide and 2.13/7ft high.

2.3. Tiered Reliability

The ANSI/TIA-942 includes 4 levels, each specifying architectural, security, electrical, mechanical and telecommunication details. Tier descriptions include information on raised floor heights, watts per square foot, and points of failure

Tier 1 – Basic: 99.671% availability; single path for power and cooling distribution, optional raised floor, UPS, generator; takes 3 months to implement; annual downtime of 28.8 hours, and must be shut down completely for preventive maintenance.

Tier 2 – Redundant Components: 99.741% availability; less susceptible to disruption from both planned and unplanned activity; single path for power and cooling

distribution; includes raised floor, UPS, and generator; takes 3 to 6 months to implement; annual downtime of 22.0 hours, and maintenance of power path and other parts of the infrastructure require a processing shutdown.

Tier 3 – Concurrently Maintainable: 99.982% availability; enables planned activity without disrupting computer hardware operation but unplanned events will still cause disruption; multiple power and cooling distribution paths but with only one path active; takes 15 to 20 months to implement; annual downtime of 16.0 hours, and includes raised floor and sufficient capacity and distribution to carry load on one path while performing maintenance on the other.

Tier 4 – Fault Tolerant: 99.995% availability; planned activity does not disrupt critical load and data center can sustain at least one worst-case unplanned event with no critical load impact; multiple active power and cooling distribution paths includes redundant components, i.e. 2 UPS each with N+1 redundancy; takes 15 to 20 months to implement, and annual downtime of 0.4 hours.

2.4. Environmental Consideration

Several environmental considerations exist within the ANSI/TIA-942 data center standard that are fire suppression, humidity levels, operating temperatures, architectural, electrical, and mechanical system specifications.

3. OUTSOURCING ECOMMERCE DATA CENTER

For small and medium eCommerce sites, it is more desirable to use outsourcing data center or data center service providers than to have in-house data center. Presented in this Section are examples at Fast PC Network, E-Commerce Park, and PACNET.

3.1. Fast PC Network

Fast PC Network [17] is a data center service provider with data centers in Atlanta, Georgia; Boston, Massachusetts; Columbus, Ohio; Houston, Texas; Los Angeles, California; New York City, New York; San Francisco, California; and Seattle, Washington State. The company claims to provide 24x7 power backup system for all equipment with no downtime or disruption because of power outage, state-of-the-art dry-pipe pre-action fire suppression with no critical losses due to fire, high-level onsite security through surveillance cameras and biometric systems to disallow unauthorized entry, and option of adding new systems when required and phasing out older systems as they become outdated.

3.2. E-Commerce Park

The E-Commerce Park Internet Data Center [16] is a data center provider on the island of Curacao, Netherlands, in the government appointed E-Zone, which means that the customers only have to pay 2% annual profit taxes, no import duties, and no sales tax. The data center are completely secured in a fully climate-controlled area that requires biometric identification verification for access.

3.3. PACNET

The headquarters are located in both Hong Kong and Singapore, which are the business centers of Asia. PACNET is a service provider of customer bandwidth requirements. For example, PACNET's products are dedicated and protected bandwidth with a high degree of network transparency, virtually unlimited bandwidth to support the business growth, and unmatched door-to-door connectivity from a single vendor [10].

For the purpose of advanced services, the concept of advanced data center facilities was used by PACNET to provide customers with a secure physical site to host servers, routers, and switches. The company claims to provide PACNET staff to manage equipment 24x7 with conveniently designed rack spaces, reliable heat, air conditioning and ventilation systems, and lighting and power supply. In addition, customers can access to PACNET backbone, a redundant and robust, multi-homed network based on relationship with top tier suppliers and intelligent peering to ensure quality of service [19].

4. IN-HOUSE ECOMMERCE DATA CENTERS

Discussed in his Section are sample data centers from a few of the top eCommerce sites, namely, eBay, Amazon, Target, and Walmart.

4.1. eBay's Data Center

The auction site eBay [5] is the most popular of all eCommerce sites. Founded in 1995, eBay enables eCommerce on a local, national, and international basis with an array of web sites including the eBay marketplaces, PayPal, Skype, Kijiji, Rent.com, and Shopping.com that bring together millions of buyers and sellers every day.

eBay used Sun Microsystems servers, including Sun Fire T2000 servers to reduce Total Cost of Ownership (TCO) for database, and Sun Fire X4100 servers to reduce TCO for search farm.

The eBay's main Teradata data warehouse consists of about 2 petabytes of user data, tens of thousands of users, millions of queries per day, 72 nodes, more than 140 GB/sec of I/O, or 2 GB/node/sec, and hundreds of production databases.

The Greenplum data warehouse consists of six and a half petabytes of user data, 17 trillion records, 150 billion new records/day, which seems to suggest a rate well over 50 terabytes/day, 96 nodes, 200 MB/node/sec of I/O, and 4.5 petabytes of storage.

After having acquired the Awitch-X data center in Phoenix with 135,000 square foot building in 2006, eBay purchased more land at South Jordan and Utah in USA with 250,000 square feet in December 2008.

4.2. Amazon's Data Center

Because Amazon is a technology company, not just an online retailer and data center [8], Amazon uses their technology to support their business. Since 2006, Amazon has provided Amazon Web Service (AWS) a subsidiary of Amazon.com, Inc., for companies of all sizes with an infrastructure web services platform. Amazon with IBM provided Amazon Elastic Compute Cloud (EC2) for database solution. IBM and Amazon Web Services allow developers to use Amazon EC2 to build and run a range of IBM platform technologies. The new "pay-as-you-go" model provides development and production instances of IBM DB2, Informix Dynamic Server, Web Sphere Portal, Lotus Web Content Management, and Novell's SUSE Linux

EC2 supports key development platforms including C/C++, Java, .Net, PHP, Ruby on Rails, Python, as well as Web 2.0, ERP, and others business applications. Using Amazon Web Services, an eCommerce web site can weather unforeseen demand with ease. A pharmaceutical company can rent computing power to execute large scale simulations. A media company can serve unlimited videos, music, and more. An enterprise can deploy bandwidth-consuming services and training to its mobile workforce.

In July 2009, a research from InfiBase stated that 1,422 of the top 500,000 web sites are hosted at Amazon EC2. As Amazon's cloud continues to grow, the company is investing in real-world brick-and-mortar data centers to provide additional capacity. The company recently leased an 110,000 square foot property in northern Virginia to expand its data center footprint. The additional space could help accommodate dramatic growth for Amazon Web Services, the suite of services that allow companies to run their applications on Amazon's infrastructure and pay based on usage. More than 500,000 developers are now using AWS, and Amazon's S3 storage now houses more than 50 billion objects.

4.3. Target's Data Center

Target is an American retailing company that was founded in Minneapolis, Minnesota in 1902 [11]. Net retail sales for the five weeks ended on July 4, 2009 were \$5,690 million

which may be considered a high level success of retail online.

There are three pieces of advice for data center pros considering building a new facility.

Firstly, the building of the Uptime Institute's Tier-3 standard, which means the facility is concurrently maintainable. Portions of the system can be shut down and maintained while still keeping the load available on the raised floor. Maintenance on anything can be done and not affect the operation of the data center.

Secondly, there are facility and security staff on site to respond to instances 24/7. Even if it is nothing more than having someone manning the main security desk checking card access to the data center in case someone needs to get in at 3 a.m. in the morning.

Thirdly, to maintain strict policies and procedures on operations in the data center, Target has four people assigned to raise floor management. No one is allowed to plug in a piece of equipment except those four guys. Also, Target prints labels on every piece of equipment, using an alpha numeric grid system that coincides with its raised floor -- equipment names are based on grid locations, and if it does not have a label, it does not go in the data center.

4.4 Walmart's Data Center

Walmart calls its data center "Walmart's Data Collection SuperCenter" which covers 125,000 square feet. Walmart has been keeping its data center a top secret. It would not even allow any visitor to visit the fortress. It is said that Walmart had enough computer storage to store more than all the information available on the Internet. The exact amount was 460 terabytes, with one terabyte equal to 1000 gigabytes or one trillion bytes or 10^{12} (1000000000000) bytes. Incidentally, Srisakdi Charmonman IT Center has 40 terabytes of memory and that is much more than what Assumption University has to store on the computer in the year 2009. It has also been said that Walmart has created a system to connect video of its staff member each time he check out his file, probably to use facial recognition software to verify the staff member [9].



Figure 1. Walmart's Data Collection SuperCenter

5. GREEN ECOMMERCE DATA CENTER

In the IT space, "Green" means improved energy efficiency with minimal environmental impact. A green data center is a cost-saver in the long run, as older products are replaced with more energy-efficient models, and budget issues are always of primary concern to small and medium-sized businesses (SMBs). With the high cost of electricity and oil, reducing energy use becomes even more important. SMBs have the same green requirements as larger businesses, but generally with lower budgets.

Presented in this Section are sample green data center effort by IBM and Buy.com.

5.1. IBM

From "www.cnet.com" dated 29 May 2009, IBM building green data center at Syracuse [4], by co-operating with Syracuse University and the State of New York to build a new computer data center for the university's campus. Advanced construction and smarter computing technologies were used to make it one of the most energy efficient data centers in the world. It is expected to use 50 percent less energy than a typical data center, and making it one of the greenest computer centers in operation. IBM will provide more than \$5 million in equipment, design service and support, including supplying the electrical co-generation equipment and servers such as IBM BladeCenter, IBM Power 575, and IBM z10 systems. Moreover, the New York State Energy Research and Development Authority (NYSERDA) are contributing \$2 million to the project.

In August 2008, IBM launched "Eco Data Center" in Second Life (SecondLife.com). Visitors who want to check out IBM's latest eco data center could have an avatar in Linden Lab's Second Life, where the facility "exists". The data center is used as a demonstration area for IBM's eco-friendly technologies and services.

5.2. Buy.com

Carnegie Mellon University's Green Design Institute stated that shopping online via Buy.com's eCommerce reduced environmental impact with 35% less energy consumption. In the majority of cases, the Buy.com eCommerce model performed better than shopping at retail stores. Buy.com made available to Carnegie Mellon information about its data center, last mile delivery practices and other sources of energy consumption.

Buy.com is a leading eCommerce company with more than 12 million customers. Buy.com focused on providing a broad selection of high-quality technology and entertainment retail goods at competitive prices. Buy.com offers millions of products in a range of categories, including consumer electronics, computer hardware and software, cell phones, books, music, videos, games, toys, bags, home and outdoor, baby, jewelry, shoes, apparel and sporting goods. Founded in June of 1997, Buy.com is located in Aliso Viejo, California. Buy.com and The Internet Superstore are trademarks of Buy.com Inc.

Buy.com data center electricity used, metered for around 3 years is given as an example. The data covers 932 days from 1/12/2006 to 8/14/2008 with 28 readings. The electricity meter was read on non-uniform intervals but the average interval was around 35 days. Each reading was given in kilowatt-hours (kWh) for the time that had lapsed before the prior reading. This electricity consumption was converted to megajoules (MJ) per shipment. This result in an energy usage of approximately 0.25 MJ/shipment related to the data center. It should be noted that this is a conservative measure of the importance of data centers, as all data center energy usage was allocated to the shipments from their wholesale supplier, which represents only a portion of total buy.com orders, due to missing data [2, 14, 15, 18, 20].

6. SAMPLE DATA CENTER PROBLEMS

Even with the use of data centers, eCommerce sites can still be crippled by a few other problems such as electricity outage and Distributed Denial of Service (DDOS).

6.1. Electricity Outage

There have been five major incidents in the first week of July 2009 in which generators or UPS failures have caused data center power outage that left customers offline [7].

Seattle Data Center outage that disrupted eCommerce started at 11:10 pm Thursday July 2, 2009. An electrical fire occurred in a garage-level electrical room of Fisher Plaza East in Seattle, Washington State, USA [6]. The incident triggered the sprinkler system in an electrical equipment room. Power was restored early Saturday morning, July 4, 2009.

The electrical outage has knocked payment processing provider named "Authorize.net" offline for hours, causing thousands of eCommerce web sites to be unable to take credit cards for online sales. Authorize.net resumed processing from a backup data center.

6.2. Distributed Denial of Service

Major eCommerce sites such as eBay and Amazon which have their own data centers could still be crippled by Distributed Denial of Service (DDOS). A notorious incident occurred in February 2000 [3]. Five popular Internet sites, namely, Amazon.com, Buy.com, eBay.com, CNN.com, and Yahoo.com were attacked by Distributed Denial of Service. It started at Yahoo which was down for about three hours because of a coordinated attack from more than 50 computers, making data requests which overloaded the servers.

In case of eBay.com, the DDOS attack occurred at the end of the business day and lasted until late at night. Several hours after the eBay incident started, Amazon.com was attacked but the service was back to normal within an hour.

From the DDOS attacks of the five popular sites, all the victims tried their best to make sure that similar incidents would not occur again. Similar to all web sites providing services to the public, eCommerce sites are highly vulnerable to DDOS attacks and they need to be online to conduct their business.

6.3 Software Failure at Google App Engine

On 2 July 2009, from 6.45 AM to 12.35 PM, Google App Engine malfunctioned with partial to complete outage [1]. The result of investigation revealed that there was a software bug in a master server which could not take care of an improperly formed filehandle and caused a stack overflow and restarting.

The problem was reported by an engineer at 6.44 AM and the cause could not be identified until 8.00 AM. A similar-looking incident had been seen in a different data center a week earlier and could be solved by an upgrade to a new version of the software. So, the software upgrade was made immediately but it did not work. At 8.35 AM, datastore write failure rose to 100%. At 9.00 AM, another upgrade of the software finished but the system was still failing. At 12.00 PM, the primary App Engine data center needed to be restarted one more time. At 12.35 PM, the system resumed full functionality. It was concluded that even if the system could be repaired, it took too long and all parties concerned must try harder to be able to fix the problem in shorter time if and when it occurs again.

7. CONCLUDING REMARKS

Millions of eCommerce sites all over the world need business continuity to keep their customers. In the early days, redundancy in computing equipment was used. Later, the concept of data center was proposed with 99.671% availability in Tier 1, 99.741% availability in Tier 2, 99.982% availability in Tier 3, and 99.995% availability in Tier 4. However, those figures of availability were meant for data center infrastructure. They did not cover other problems such as electricity outage which may be caused by fire, and Internet security attacks such as Distributed Denial of Services. Therefore, to provide business continuity, all parties concerned still have to continue to find ways and means to try to make the data center running 24 hours a day and 7 days a week to keep their business operating and their customers satisfied.

- searchdatacenter.techtarget.com/news/article/0,289142,sid80_gci1176301,00.html,
- 12. Telecommunication International Association. TIA-942 “Telecommunications Infrastructure Standard for Data Centers”
www.ieee802.org/3/hsg/public/nov06/diminico_01_1106.pdf
- 13. Wikipedia.Datacenter.
en.wikipedia.org/wiki/Data_center
- 14. www.buy.com
- 15. www.buy.com/toc/march-3-2009/66998.html
- 16. www.ecommercepark.com/index.php?page_id=19
- 17. www.fastpcnet.net/datacenter.html
- 18. www.greennm3.com/2009/03/on-line-shopping-13-less-energy-than-brick-and-mortar.html
- 19. www.pacnet.com/enterprise/solution
- 20. www.scribd.com/doc/14664871/Go-Green-Report-2008

REFERENCES.

1. Beckmann, Chris. Google App Engine.
Groups.google.com/group/google-appengine/msg/ba99ded980c8c179
2. Carnegie Mellon University. Carnegie Mellon University Study Finds Shopping Online Results in Less Environmental Impact.
buy.com/toc/march-3-2009/66998.html
3. Dembeck, Chet. One Year Ago: E-Commerce Sites Crippled by Hacker Sabotage.
www.ecommercetimes.com/story/7307.html?wlc=1250391104
4. Haff, Gordon. IBM building 'green' data center at Syracuse.
news.cnet.com/8301-13556_3-10251934-61.html
5. Miller, Rich. eBay Picks Utah for \$334 Million Data Center.
www.datacenterknowledge.com/.../ebay-picks-utah-for-334-million-data-center/
6. Miller, Rich. Major Outage at seattle Data Center
www.datacenterknowledge.com/.../major-outage-at-Seattle-data-center/
7. Miller, Rich. Outage for Rackspace Customers.
www.datacenterknowledge.com/.../outage-for-rackspace-customers/
8. Millier, Rich. Amazon Building Large Data Center in Oregon.
<http://www.datacenterknowledge.com/archives/2008/11/07/amazon-building-large-data-center-in-oregon/>
9. NYnerd. Walmart's Data Center
[\nynerd.com/walmarts-data-center/](http://nynerd.com/walmarts-data-center/)
10. Pacnet. www.pacnet.com
11. Stansberry, Matt. Target builds a better data center.