Determining the Sustainability of Current Electronic Information Storage Techniques

Randy Chau **April 4, 2012**

Prepared at the request of Alan Doyle, UBC Records Manager, and Brenda Sawada of UBC SEEDS in partial fulfillment of UBC Geography 419: Research in Environmental Geography, for Professor David Brownstein.

EXECUTIVE SUMMARY

This project is designed to answer a question of sustainability in a field of technology where data is stored electronically. Alan Doyle, a records manager at the University of British Columbia (UBC), has sponsored three questions that will ultimately determine the answer of whether storing data electronically is sustainable fiscally and environmentally. Simultaneously, this project also serves to provide firsthand knowledge of green data center methods for future developments at UBC. The challenges this report has attempted to answer include:

1) What is the current state? (i.e. What are the problems/What are people doing?)
   i) Businesses are running out of room to expand their storage media because they are using too much electricity.
   ii) Organizations and Initiatives are developing to invent standards and influence manufacturers to create greener technologies.
   iii) The amount of carbon emissions put out by data centers are magnified depending on what type of power plant is used to harness this electric energy.
2) What methods are being used to reduce the impact data storage has on the environment?
   i) The invention of Power Usage Effectiveness Tool by Green Grid helps data center users calculate how much of their total power is used by IT equipment themselves—determines efficiency of data center.

ii) Customization of data center hardware marks big improvements to data center efficiency.

iii) Hardware and software upgrades such as Dynamic Voltage Frequency Switching, Solid State Hard Drives, Virtualization, and hybrid UPS systems further increase the efficiency output of data centers and greatly reduces the amount of electricity usage in the data center.

3) What should UBC do? (Recommendations)

i) Currently UBC is trying to integrate scattered small data centers across campus into its newly built pharmaceutical data center so it can utilize sustainable facility specifications.

ii) However, implementation of a structured, centralized system is needed to employ the use of virtualization and further reduce electricity consumption in idling servers.

iii) Finally, a method of records management needs to be initiated to free up cluttered unnecessary information and allow the servers to last as long as possible before a capacity issue comes into play.

INTRODUCTION

When a person creates a document inside a word processing program on a computer, this becomes a piece of electronic information. The user may choose to store this information either on a USB thumb drive or directly on their computer's hard drive. Either way, this information is now a piece of stored electronic data where an environmentalist can start to pose questions of how much energy that piece of equipment is using to allow a user to access it.

Large amounts of electronic data from institutions and businesses require more than just a personal computer to store them. This information is kept inside many servers placed on racks in a facility called a data center. Because these types of facilities contribute to an overuse of electricity, they are negatively impacting the environment.

1.0 Current state in Electronic Information Storage (What are the problems?)

1.1 Business Sector

Whether it is Google, Facebook, or any other corporate business that house data centers to store their information, the space within the facility and the racks that hold servers are expandable in accordance with their need to store information (Ebbers et.al., 2008). However, what happens if a data center reaches its peak power supply but only managed to take up 1/3rd of the space used inside the data center? On one account, the owner would be wasting a large sum of money on this unused space, and on another, it is unsustainable to spend enormous amounts of electrical power on only small amounts of data storage media. A white paper published by IBM indicates that business owners are seeking green options to save energy in order to maximize their data center usage on the same amounts of electricity the data center originally draws. The article also indicates a vice versa scenario where a data center has used up the space within and cannot expand any further (Ebbers et.al., 2008). This addresses another issue whereby the efficiency of equipment and data management comes into play in the role of being sustainable (*see section 2.3-2.4 for more information*).

1.2 Emergence of Data Center Initiatives and Organizations

In today's society, there are standards that set ground rules for the construction of digital storage media and facilities in order to control levels of energy usage and greenhouse gas emissions. Below is a small compiled list of green initiatives and organizations and their contributions toward regulating unsustainable issues within data centers.

Table 1. Environmental regulators that help "green" a data center.

Initiative/Organization	Goal/Objective
Green Grid[1]	An organization that is devoted to invent standards for facility and IT equipment efficiency and has invented a tool called Power Usage Effectiveness (PUE) to measure electrical output from a data center (*see section 2.1 for more information*).
The Climate Savers Computing Initiative[2]	An initiative founded by Google and Intel with a mission to cut carbon dioxide emissions by promoting the development of smart technologies that can improve computer power delivery efficiencies and reduce power consumption.
Facebook's Open Compute Project[3]	By releasing its own data center's green building techniques, they hope to inspire

	other companies to release some of their private sustainable options to the open public. Facebook believes that through opening up they can achieve maximized innovation and a reduction in operational complexity in this field of technology.
Leadership in Energy and Environmental Design(LEED)[4]	Created by the U.S. Green Building Council, this program provides certification through verification of standards from third party independents. In order to be certified, a building, home or community needs to be designed and built using techniques that achieve high performance ratings in key areas of human and environmental health.
American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE)[5]	To invent new technologies in the fields of heating, ventilation, air conditioning, and refrigerating so that both the humans and the environment can benefit in a sustainable fashion.

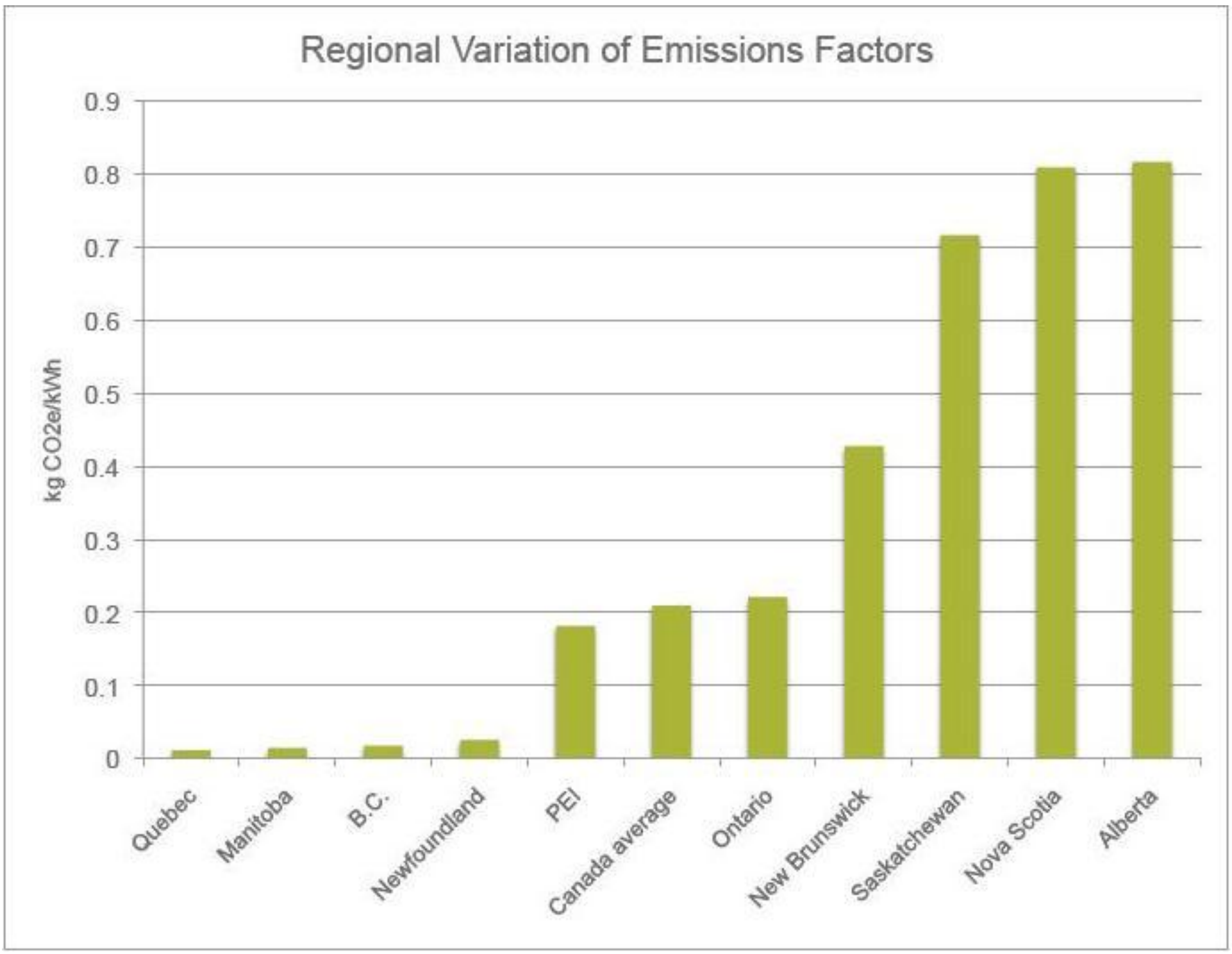
1) Preimesberger, Chris. "Green Grid Gets down to Business." EWeek 24.14 (2007): 11. Academic Search Complete. Web. 31 Jan.2012.
2) “Google Data Centers.” Google. Web. 18 Jan. 2012.<http://www.google.com/about/datacenters/>.

3) Amir, Michael. "Inside the Open Compute Project Server | Facebook." Facebook. 8 Apr. 2011. Web. 3 Apr. 2012.
4) "What LEED Is." *U.S. Green Building Council*. USGS, 2011. Web. 27 Feb. 2012. <http://www.usgbc.org/DisplayPage.aspx?CMSPageID=1988>.
5) ASHRAE. Green Tips for Data Centers. Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2011. Print.

1.3 Role of Energy Source

Through the research of Kristina Welch, ISIS Sauder School of Business- UBC, the importance of where electrical energy comes from can have an enormous impact toward the total amount of greenhouse gas emissions a data center can emit (Welch,2011). If a region is generating electrical energy by coal-fired power plants for example, every year these plants emit about 940 to 1340g/kWh of CO_2 into the atmosphere! If a 4MW data center was powered by one of these plants, you take 4000kW and multiply it by 1340g and multiply again by 8760hours= and you will end up emitting 47million kg of these CO_2 by-products a year.(see Appendix for a summary table of CO2 emissions associated with each type of energy harnessing method.) From this example, the concern of energy saving is greatly emphasized to help reduce our carbon footprint electronic data storage can have on the environment. Figure 1 below shows the emission factors through electricity generation from each province.

Figure 1. Regional Variation of Emissions Factors



Source: http://isis.sauder.ubc.ca/files/2011/08/Assessing-the-Business-Case-for-Data-Centre-Relocations.pdf

2.0 Data Storage Methods That Reduce Environmental Impacts
(What Are the Answers That Can Solve the Previous Problems?)

2.1 Measuring Electricity Usage

The first step to analyzing sustainability in a data center is to measure the energy output from the facility to determine which area consumes the most electricity. Green Grid has devised the Power Usage Effectiveness (PUE) tool, an equation to determine how much of the total energy used in a data center goes to the IT equipment inside it, to help data center owners interpret their obtained data measurements from meters placed at the main facility's components (ASHRAE, 2011).

Equation 2.1:

PUE= Total energy from the data center facility/ IT energy consumption or power

2.2 Rack Size Improvements vs. Inner Hardware Upgrades

Facebook's Open Compute Project has released to the general public their confidential building plans of data centers in order to promote more innovative green technologies to further come out. Through their objective, I have found an answer to how rack size that holds servers is an important factor to making the facility more energy efficient. The standard size of a typical server rack accommodates a server of 1U (1.75 inches) thickness; however, Facebook has customized their racks to be able to hold 1.5U (2.63 inches) thick servers (Amir, 2011). This increase allows the servers to contain larger heat sinks that increase their ability to absorb heat, which ultimately results in less cooling power, thus saving energy (Amir, 2011).

2.3 Hardware and Software Improvements

Once the facility has been upgraded into its suitable green methods, the focus is now turned to the power efficiencies of IT equipment which the data center houses. In my research I have found two hardware and software upgrades that can greatly improve the performance of the equipment while saving energy when the computer is in idle-mode.

First, the server can benefit from a large speed increase with the implementation of the new flash-based solid state hard drives (SSD). This improvement reduces the amount of electricity use because there are no moving mechanical parts within the storage media (Zhang et.al., 2010). A traditional hard drive requires discs that spin at rapid speeds in order to process data stored on it. With the new SSD drives, the second advantageous characteristic is the high literacy rate levels it provides that lets users gain access to data much faster than the older disc drives (Zhang et.al., 2010).

Next, a technology called the dynamic voltage frequency switching (DVFS) is a popular choice for sustainable options. SFU graduate student, Ananth Narayan, explains that this topic is a hardware-based application software that regulates the computer processor once the operating system (OS) has started up (Narayan, 2009). The user has no manual control over this function and DVFS adjusts the voltage and frequency of the processor on the motherboard according to its usage. This means that when a server is idling, the DVFS function will power down the processor so that minimum energy is used in the process.

Thirdly, virtualization is an automated or controllable software, similar in function to DVFS, but regulates the storage media instead of the motherboard features. This means when a server is idling, the virtualization software will power down the SSD or traditional spinning hard discs when it is not in use (Liu et.al., 2009). A popular virtualization software called GreenCloud has the ability to regulate between servers and allow easy access to parts of data located over a span of several servers; however, this Virtual Machine (VM) technology cannot keep the server running when migration processes take place. Therefore, a substitute is mentioned in Liu et.al.'s article called Xen that allows the server to continue running while migration of data is taking place.

Finally, data centers are generally equipped with an uninterruptable supply system (UPS) to provide battery back-up service in the case of a blackout. However, ASHRAE's tenth publication indicates the best type of developed technology is the hybrid-UPS unit/double-conversion on-demand UPS unit. This feature behaves in a similar fashion to the DVFS and virtualization technologies because of its ability to detect certain thresholds allowing the automation of activating or deactivating the component to conserve energy (ASHRAE, 2011).

2.4 Records Management

Zhang et. al.'s Solid State Drive article also mentions in its concluding paragraph the importance of records management. They identified that data management techniques are the most economical options. Therefore, with proper management, one can achieve a small energy consumption value and will not require to add additional hardware due to the servers being full of cluttered, unused information. In another article by Cara Garretson, her interview with Dallas Thornton, IT Director of the San Diego Supercomputing Center, finishes off with a quote stating power consumption awareness is a priority over the issue of equipment lifetime. If taken into account, the amount of money saved from paying less electricity bills will be enough to cover for the cost of server upgrades when the time comes (Garretson, 2007).

2.5 Google Case Study

Google’s release of their data center information on a customized website allows viewers to explore the concepts and terms to describe such a facility and allow researchers to compare certain characteristics among others (Google, 2012). The purpose of adding this case study to the methods section of the paper is to illustrate examples used by Google to show a better application of these green technologies and how it has impacted both the environment and the company itself.

First, by breaking down a general UPS system’s structure and applying a battery on board a server, Google was able to decrease several energy transformations that contributed towards an energy loss through heat.

Next, Google reuses old, slower servers to host programs that require less performance such as G-mail. By doing so Google can not only save money by not having to purchase new servers from time to time, but also reduces greatly the carbon footprint that associates with the manufacturing of new servers.

Thirdly, Google has ties with section 1 of this paper because of its involvement in the creation of the Climate Savers’ Computing Initiative with Intel. This actively shows its involvement with helping society by setting standards to help the creation of more “greener” technology.

This short analysis of Google shows a second example, aside from Facebook’s custom racks, of how customization of their own data center components can help a company reduce their environmental impact as well as boost their PUE to an average of

1.06. The companies not only create an exceptional image of themselves contributing to the environment but also leaves strategies behind so that new technology can hopefully be developed in the near future to help reduce environmental impacts. Of course, one can argue that for such large companies they can afford to boast and open up some of their secrets to gain public attention and in return earn more money, but the fact that these green technologies they have released have been truly beneficial to the environment are something to be praised regardless of what true intentions these companies may have.

3.0 Recommendations for UBC

UBC's current data storage method involves 77 small-sized data centers, reported by UBC Facilities & Capital Planning and Facilities, scattered across campus which is located in non-sustainable facilities with minimal specifications to meet its operating requirements (Welch,2011). Currently, there is a new data center that is being constructed in the new pharmaceutical building and is set to open in September 2012. This new data center incorporates a sustainable cooling strategy which involves using 65ºF (18ºC) water to cool each rack accompanied by energy efficient air conditioning units (Pini, 2010). Through UBC's Information Technology website, the purpose of this data center is to gather the servers housed at the 77 different locations and unite them into the currently constructed energy efficient facility to reduce the institution's consumption of electricity and carbon emissions. However, through an email conversation between Alan Doyle, UBC Records Manager, and Bob Macdonald, UBC IT, I've discovered that the facility is not centralized under a main system that regulates the operations for each server. Along with five other observations these are my recommendations:

Table 2. Future Recommendations for UBC.

Observation	Recommendation
1) Data center not centralized under a main regulatory system. Researchers have complete freedom over their storage media.	Implement a virtualization method within the new data center and apply restrictions as needed for the researchers who are storing their media in this data center. (i.e. scheduled modification times/server downtimes for upgrades, etc.)
2) Data center's media come from 77 different facilities with differing specifications and properties.	Implement one or two mandatory models of servers that have DVFS technology, with SSD hard drives and issue a fee for the migration of data onto this new efficient server and recycle the old servers.
3) Data center's UPS system is standard.	Since it is not likely for UBC to develop their own on-board battery system for each server to replicate Google's strategy, I suggest

	upgrading via ASHRAE's advice on "hybrid UPS unit" to have additional control to reduce unnecessary energy consumption.
4) Data inside these servers are unorganized and take up space.	Enforce data management with temporary retention periods to reduce clutter and free up more space to allow additional storage.

4.0 Conclusion

In this survey report I have mentioned various perspectives from business concerns of saving money to environmental concerns about how the source of energy use can magnify the impacts of carbon dioxide outputs. All of these perspectives address the issue of sustainability as a solution to solving their problems. I have taken these concerns and connected them with stages of solutions required to help "green" a data center in order to solve the question for sustainability. Finally, I have addressed the current situation in UBC and have proposed accordingly certain conditions the facility can lean towards in order to lessen their environmental impact in the near future.

5.0 Further Research

Much more research into this topic is available as this paper only touches on some of the most popular green technologies and problems out there in the world today. Further research in areas of location and positioning of data centers is a must in order to understand the best methodologies one should apply to its facility based on its location. A good place to start off with is an article written by Steve Hamm called “New Green Tech Could Revolutionize Data Centers-Especially in Emerging Markets”. This article talks about implementation of data center facilities around the equator where the use of free cooling is close to impossible due to hot temperatures. However, this article talks about the use of solar energy harnessing directly onto the data center itself which emits close to no carbon emissions.

Also, further research can be done in comparing actual electrical bills cost and the methods they use in similar sized data centers that consume a similar amount of electricity.(i.e. 2MW data centers)

Finally, an interesting field to explore is the Hierarchical Storage Management mechanism that is proven extremely effective in old traditional hard drives but is currently being developed in solid-state hard drives.

6.0 Appendix

Finally, we have a look at the comparisons of ONLY the construction and fuel elements of the two power-generating technologies for a period of 10 years. (I've given a time period of 10 years because at this time period it already is enough to show a significant difference.)

	Coal-Fired Power Plant(650MW)	ITER nuclear fusion plant	Hydroelectric Power Plant (650MW)	Sarnia Solar Farm (60MW expansion)	Wind Power Plant (630MW)
Initial Building Stage	>3 billion dollars	~4.34 billion dollars	~2.2 billion dollars	~300million dollars	~3 billion dollars
Extra Fuel Input Costs/year	$208.9 million dollars	$33,750	N/A (water)	N/A	N/A
Mitigation	$170.8 million dollars	N/A	$3.5 million (downstream bypass facility)	N/A	$13,125 (noise reduction)
10 year runtime total	**~6.80 billion dollars**	**~4.34 billion dollars**	**~2.2 billion**	**~3.25 billion dollars (assume we create enough of these farms to reach peak 650MW)**	**~12billion (assuming actual annual productivity reaches 25% of peak or less)**
CO_2 emissions (g/kWh)	**940-1340**	**20-40**	**4-18* (varies depending on flood area; possibility of exceeding measured outputs from a coal-fired power plant)**	**81-260**	**16-120**

Citation:

Chau, Randy, Alex Lindsay, Martin Kozinsky, Amelia Jung, and Rachel Schott. "Nuclear Fusion: Energizing the Path to Tomorrow." 2012. MS. University of British Columbia, Vancouver. 3 Apr. 2012.

7.0 Acknowledgements

A big thank you to Alan Doyle and Brenda Sawada, my community partners for this project, for their tremendous support in laying out the ground rules for my project and clearing up any troubles I've had along the way. Also, a big thank you to my GEOG312 group members Alex Lindsay, Rachel Schott, Martin Kozinsky, and Amelia Jung as I've used the summary table derived from the nuclear fusion project to show carbon dioxide emissions from each type of energy harnessing source in my appendix!

8.0 Works Cited

1) Amir, Michael. "Inside the Open Compute Project Server | Facebook." *Facebook*. 8 Apr. 2011. Web. 3 Apr. 2012.
2) ASHRAE. *Green Tips for Data Centers.* Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2011. Print.
3) Chau, Randy, Alex Lindsay, Martin Kozinsky, Amelia Jung, and Rachel Schott. "Nuclear Fusion: Energizing the Path to Tomorrow." 2012. MS. University of British Columbia, Vancouver. 3 Apr. 2012.
4) Desposito, Joe, ed. "GreenTouch And Climate Savers Tackle Power Issues." *Electronic Design* (2010): 13. *Academic Search Complete*. Web. 31 Jan. 2012.
5) Ebbers, Mike, Alvin Galea, Michael Schaefer, and Marc Tu Duy Khiem. "The Green Data Center: Steps for the Journey." *IBM*, Aug. 2008. Web. 1 Feb. 2012.
6) Garretson, Cara. "Inside a Green Data Center." *Network World* 24.43 (2007): 49-52 *ABI/INFORM Global.* Web. 1 Feb. 2012.
7) "Google Data Centers." *Google*. Web. 18 Jan. 2012.<http://www.google.com/about/datacenters/>.
8) Hamm, Steve. "New Green Tech Could Revolutionize Data Centers–Especially in Emerging Markets." *A Smarter Planet Blog*. 3 Nov. 2011. Web. 02 Feb. 2012.
9) Liu, Liang, Hao Wang, Xue Liu, Xing Jin, WenBo He, QingBo Wang, and Ying Chen. "GreenCloud: A New Architecture for Green Data Center." *ICAC-INDST* (2009): 29-38. *Compendex*. Web. 31 Jan. 2012.
10) Narayan, Ananth. Power and Energy Management in Data Centers. *Simon Fraser University School of Computing Science*. ACM Surveys, 31 Dec. 2009. Web. 1 Feb. 2012.
11) Pini, Claudio. "UBC IT New University Data Center Operating Conditions." *The University of British Columbia Information Technology.* 22 June 2010. Web. 2012.
12) Preimesberger, Chris. "Green Grid Gets down to Business." *EWeek* 24.14 (2007): 11. *Academic Search Complete*. Web. 31 Jan. 2012.

13) Spafford, George. *Greening the Data Center: Opportunities for Improving Data Center Energy Efficiency*. Ely: IT Governance Pub., 2009. Print.

14) Sukinik, David. "First Green Data Center: Fannie Mae's Urbana Technology Center Confronts Challenges to Earn LEED Certification." *Environmental Design & Construction* 9.6 (2006): 8-12. *Gale Cengage Learning*. Web. 1 Feb. 2012.

15) The University of British Columbia Board of Governors. "Sustainable Development." *Office of the University Counsel*. UBC, June 2005. Web. 1 Feb. 2012.

16) "The University of British Columbia." *New University Data Centre (UDC)*. Web. 3 Apr. 2012.

17) Welch, Kristina. "Assessing the Business Case for Data Centre Relocations." *ISIS Sauder School of Business* (2011). University of British Columbia. Web. 3 Apr. 2012.

18) "What LEED Is." *U.S. Green Building Council*. USGS, 2011. Web. 27 Feb. 2012. <http://www.usgbc.org/DisplayPage.aspx?CMSPageID=1988>.

19) Zhang, Xiao, Xiao-nan Zhao, Yi Lin, and Lei-jie Zeng. "Key Technologies for Green Data Center." *Third International Symposium on Information Processing* (2010): 477-80. *Compendex*. Web. 31 Jan. 2012.