Duke Sustainable Cloud Computing

ENVIRON 245 Project Brief, Spring 2021
Group Members: Finn Doherty, Adam Hahn, Madeleine Jones, Vatsal Varma
Clients: John Haws and Liz Milewicz

Research Question: How can Duke maximize the sustainability of its data centers while still meeting cloud computing and storage requirements?

Cloud computing has become a ubiquitous part of today's world, from conveniences like Google Drive, to large scale data storage. Servers, which are large computers equipped for powerful computing and storage, are housed in buildings called data centers (DCs). The global growth of the Information and Communication Technology (ICT) sector has resulted in an increase in energy consumption for DC use[1]. This trend has led to an increase in emissions of greenhouse gases, which are primary contributors to climate change. In response, companies and institutions have begun to explore sustainability as a factor in DC operations.

Duke utilizes on and offsite DCs to meet its computing and storage needs. Duke's onsite data centers were recently consolidated to two main locations: the Telecom building and East Fitzpatrick[2]. These data centers handle online authentication and some faculty research, and hold copies of Duke Health's virtual infrastructure. Duke also contracts much of its cloud needs (including Sakai, the medical center system, and virtual machines) to Microsoft Azure[2].

Part of Duke Library's role on campus is to point clients (consisting primarily of faculty researchers) toward the computing and storage provider best suited for their needs, taking into account cost, efficiency, and accessibility. Recently, Duke OIT and Library Services have become interested in including sustainability as one of these recommendation factors. Analysis of DC operations and sustainability will enable the provision of accurate and responsible recommendations for Duke OIT and other small/medium data center operations.

Objectives:

1. Determine the current environmental impact of DCs in general
2. Identify the levers that make the biggest differences in DC energy efficiency
3. Identify the characteristics of sustainability of Duke's and Microsoft Azure's DC operations.
4. Analyze the factors that influence the recommendations that Duke Library makes to its clients.

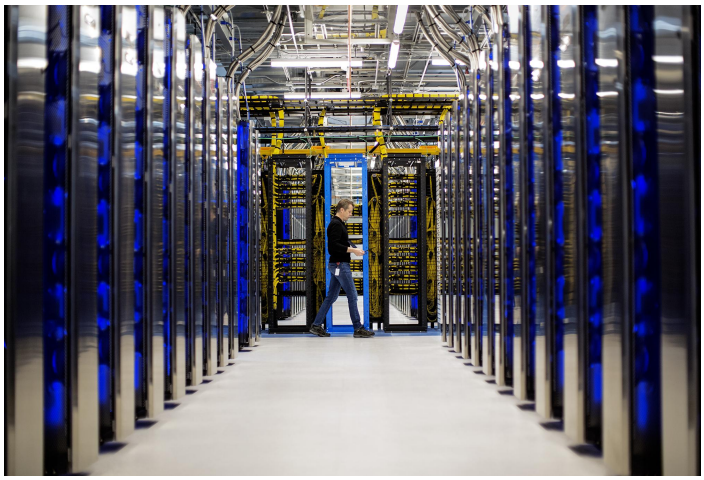


Figure 1. The servers of a Microsoft Azure data center like the ones to which Duke outsources some of its cloud computing.

Methods:

Interviews
Spring 2021

*Purpose: To gain context behind the research questions and the clients' needs.
*John Haws (Director, OIT Service Based Metrics and Reporting)
*Liz Milewicz (Head, Digital Scholarship Services)

*Purpose: To better understand the users, resources, and efficiency of Duke data centers.
*John Haws (Director, OIT Service Based Metrics and Reporting)
*Carl McMillon (Sr Manager, IT)
*Charley Kneifel (Sr Technical Director, OIT)

Web Research

- Research reports regarding DC emissions, energy consumption, and sustainability
- Online web pages about nationwide cloud computing and specifically Azure operations

Social: Improving Duke Library's recommendations will allow their clients to make sustainable decisions that also fulfill their research requirements.

Environmental: Duke's cloud operations and onsite data centers emit climate change-inducing CO_2 gas, the amount of which can be reduced.

Economic: Environmentally-friendly data centers will also be energy efficient and therefore economically efficient. Furthermore, Duke's cloud recommendations factor in cost.

Analysis

Interviews: Interviews were recorded and transcribed, and short summaries including key takeaways were written shortly after the meeting was held. The team analyzed these transcripts by underlining important pieces of information and highlighting according to the decided coding themes.

Web Research: The various reports and online pages that were found were compiled and thematically coded. The sources were then analyzed to determine the current impact of data centers, the factors influencing data center sustainability, and how Duke data centers compare with others. The team used this analysis to create a set of recommendations for Duke Library and OIT.

Results:

What are the environmental implications of data centers?

6.4 **million kWh of electricity are consumed by Duke's Fitzpatrick data center annually** [2]

- A mean of 17,560 kWh consumed per day by Fitzpatrick data center, much greater than an average US home annually (10,649 kWh) [5]
- Fewer servers are required in hyperscale data centers compared to those on-premises due to higher efficiency and closer management

2% **of global CO_2 emissions are from data centers** [1]

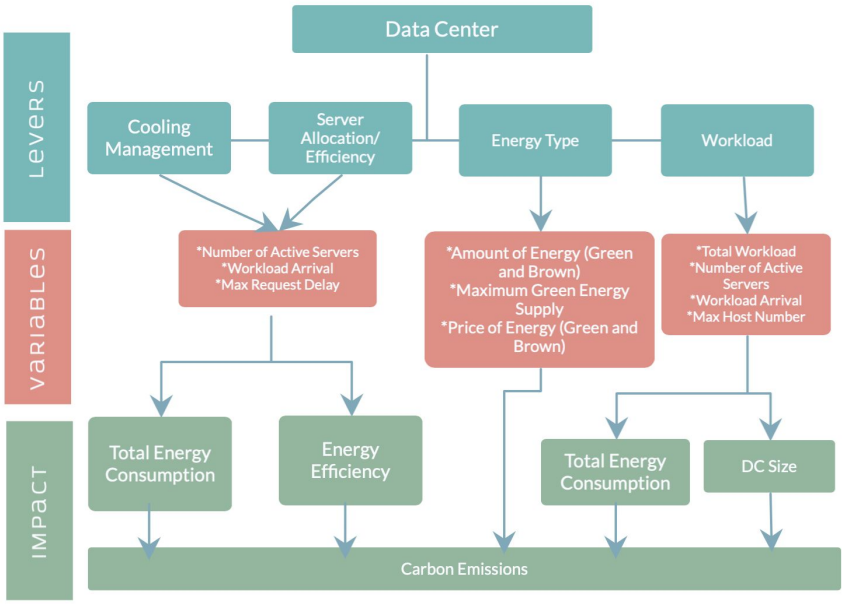
- Data center energy consumption is growing at a rate of 10-12% annually [3]
- This represents massive costs not only for large data centers, but also for companies that use them

70 **billion kWh of electricity are consumed annually by data centers in the US** [4]

- The energy consumption of data centers is primarily attributed to the powering and cooling of servers
- Zombie servers (those no longer serving a purpose but still consuming electricity) represent up to 30% of the energy cost of data centers [1]

What levers can be pulled to reduce the environmental implications? [1,4]

- Primary levers
  - Efficient IT equipment can reduce data center consumption of electricity by over 10%
  - Maximizing operational efficiency and utilization through dynamic server allocation
    - Synchronize server capacity and load
  - Efficient cooling management
    - Proper insulation, streamlined airflow
  - Transference of peak load to the cloud
  - Improving utilization rates
    - Some data centers operate at utilization rates under 15% because they are over provisioned to handle worst case scenarios



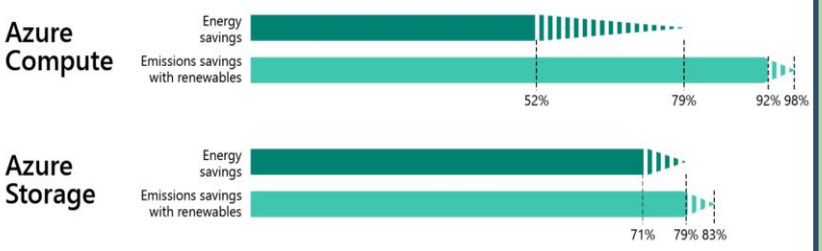


Figure 2. This graph shows the range of energy and emissions (including renewable purchases) savings by companies when they replaced their on-premises computing and storage services with cloud services provided by Microsoft Azure.

Figure retrieved from the 2018 Microsoft Carbon Cloud Study "The Carbon Benefits of Cloud Computing: a Study of the Microsoft Cloud." *Microsoft*, Microsoft, updated 29 July 2020

What are the characteristics of data center sustainability for onsite centers at Duke and data centers managed by Microsoft Azure?

- Microsoft
  - DCs have a Power Usage Effectiveness (PUE) of ~1.125 - low for the industry [4]
  - Microsoft has sustainability goals of carbon neutrality in 2012, 100% renewable energy by 2025, etc. [4]
- Duke University
  - Duke tries to keep 90-95% of servers in their data centers in use at any given time [2]
  - Data centers are cooled by Chiller Plant #2 which centralizes cooling on campus efficiently [2]

What other factors impact Duke Library client recommendations and decisions? [2]

- Grant logistics/Cost
  - Researchers lose access to offsite computing resources (ex: Azure) after grant ends
  - University will charge overhead fees for cloud computing (taken out of grant funding)
- Familiarity
  - Researchers trust and are more comfortable with the onsite method of research computing
  - Cloud computing and third party contracting is relatively new compared to onsite methods

Conclusions and Recommendations for Duke OIT and Library to improve:

Recommendations

- Make computing with Azure easy for researchers, and inform them about the advantages
- Contract out more of Duke's computing/storage to Azure
- Set up cloud computing education for faculty to acclimate to Azure
- Find a way to allow researchers to access their data through Azure even after their grant runs out

Areas of Future Research

- Without the current time and resource limitations, future researchers can:
  - Calculate quantitative data for Duke (PUE) for more direct comparison to Microsoft
  - Use this data to create a cost-benefit analysis of switching completely to Azure, implementing an algorithm to optimize the efficiency, or making no changes at all

Using these recommendations and future research in tandem, Duke OIT could optimize their energy efficiency of their cloud computing. This would:

- Lower costs for the university and researchers
- Reduce energy usage and result in lower emissions
- Form a stronger community dedicated to making positive environmental changes in their work at Duke

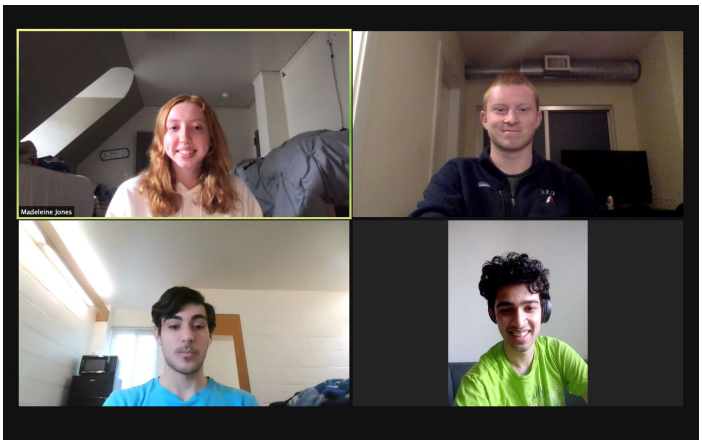


Figure 5. Zoom screenshot of team

Acknowledgements:

Thank you to team clients John Haws and Liz Milewicz for their guidance and support. A special thanks to Charles Kneifel and Carl McMillon for their contribution of information regarding Duke DC operations and sustainability. We would also like to extend our gratitude to Char, Tavey, and Bennett who have supported and taught us throughout this process.

References:

1. Pawlish, Michael & Varde, Aparna & Robila, Stefan. (2012). Cloud computing for environment-friendly data centers. Proceedings. 43-48. 10.1145/2390021.2390030.
2. Kneifel, Charley and McMillon, Carl. Personal interview. 25 Feb. 2021.
3. Renugadevi, T., et al. "Optimized Energy Cost and Carbon Emission-Aware Virtual Machine Allocation in Sustainable Data Centers." *Sustainability*, vol. 12, no. 16, 2020, pp. 6383. *ProQuest*,
4. "The Carbon Benefits of Cloud Computing: a Study of the Microsoft Cloud." Microsoft, Microsoft, 29 July 2020
5. "Electricity Use in Homes - U.S. Energy Information Administration (EIA)." U.S. Energy Information Administration - EIA - Independent Statistics and Analysis, EIA, 9 May 2019.