FA19 Proceedings Cloud Computing Engineering
e516

Gregor von Laszewski

Editor

laszewski@gmail.com

https://cloudmesh-community.github.io/pub/vonlaszewski-e516-proceedings.epub

November 23, 2019 - 05:20 AM

Created by Cloudmesh & Cyberaide Bookmanager, https://github.com/cyberaide/bookmanager

FA19 Proceedings Cloud Computing Engineering

Gregor von Laszewski

FA19 Proceedings Cloud Computing Engineering

4 REFERENCES

1 PREFACE

Sat Nov 23 05:20:06 EST 2019

1.1 Disclaimer

This book has been generated with Cyberaide Bookmanager.

Bookmanager is a tool to create a publication from a number of sources on the internet. It is especially useful to create customized books, lecture notes, or handouts. Content is best integrated in markdown format as it is very fast to produce the output.

Bookmanager has been developed based on our experience over the last 3 years with a more sophisticated approach. Bookmanager takes the lessons from this approach and distributes a tool that can easily be used by others.

The following shields provide some information about it. Feel free to click on them.

pypi v0.2.28 | License Apache 2.0 | python 3.7 | format wheel | status stable | build unknown

1.1.1 Acknowledgment

If you use bookmanager to produce a document you must include the following acknowledgement.

> *"This document was produced with Cyberaide Bookmanager developed by Gregor von Laszewski available at https://pypi.python.org/pypi/cyberaide-bookmanager. It is in the responsibility of the user to make sure an author acknowledgement section is included in your document. Copyright verification of content included in a book is responsibility of the book editor."*

The bibtex entry is

```
@Misc{www-cyberaide-bookmanager,
  author =    {Gregor von Laszewski},
  title =     {{Cyberaide Book Manager}},
  howpublished = {pypi},
  month =     apr,
  year =      2019,
  url={https://pypi.org/project/cyberaide-bookmanager/}
}
```

1.1.2 Extensions

We are happy to discuss with you bugs, issues and ideas for enhancements. Please use the convenient github issues at

- https://github.com/cyberaide/bookmanager/issues

Please do not file with us issues that relate to an editors book. They will provide you with their own mechanism on how to correct their content.

2 REPORTS

2.1 Report List

Here we summarize the Projects and teams. Place the document in your github folder under project/report.md. Please make sure to use proper markdown and not githubs common mark. Once the reports are created they will at one point also show up in the proceedings.

2.1.1 CLOUD STORAGE-STORAGE-LOCAL PROJECTS

- see cloudmesh-storage to start.
- all need to contribute to the local provider
- see the commandline interface.
- develop Python API.
- develop OpenAPI REST services for it.
- if this works technically we should be able to copy between all clouds as we have all services for each cloud and each speaks the same protocol. This means your benchmark must include all clouds

HID	Status	Names	Project
160	approved	Shreyans	Azure to Google Cloudmesh Storage Provider for Virtual Directories
152	approved	Pratibha	Google to/form AWS: Cloudmesh Storage Provider for Virtual Directories
162	approved	Shivani	Oracle to/from Local Storage Service (Residential) 534
155	approved	Ketan	Azure to/form AWS blob Storage Service

2.1.2 CLOUDMESH COMPUTE PROJECTS

HID	Status	Names	Project
162	approved	Shivani	516: Oracle Compute Service Residential
165	approved	Zhi	Google Compute Service
170	approved	Chenxu	Azure Cloudmesh Compute
169	approved	Harsh	Developing Cloudmesh interfaces for Google Cloud Platform

2.1.3 PI PROJECTS

HID	Status	Names	Project
148	approved	Sub	Federated Kubernetes Cluster with Raspberry PI
158 150	approved	Daivik Akshay	Raspberry Pi Cloud Cluster for Spark, Raspberry Pi Cloud Cluster for Hadoop

2.1.4 DATABASE ABSTRACTION PROJECTS

HID	Status	Names	Project

HID	Status	Names	Project
147	approved	Harsha	Abstract database management on Multicloud environments for the NIST Big Data Reference Architecture Azure, Google, SQL
141	approved	Bala	Abstract database management on Multicloud environments for the NIST Big Dara Refernce Architecture AWS, Oracle, SQL NIST BigDataInterface reference implementation for data base absractions on 2 clouds with 2 differnt technologies, one is a SQL based the othe is a NoSQPL based (mongo)
156	approved	Manikandan Nagarajan	Abstract database management on Multicloud environments for the NIST Big Dara Refernce Architecture AWS, Azure, Mongo, NIST BigDataInterface reference implementation for data base absractions on 2 clouds with 2 differnt technologies, one is a SQL based the othe is a NoSQPL based (mongo)

2.1.5 OTHER PROJECTS

HID	Status	Names	Project
167	approved	Bill Screen	Cloudmesh Virtual Directory Life Cycle Service, This needs to run on 2 clouds + locally. TTL = time to live is one, attribute for lifecycle management. Bi9ll explained this is more than ttl.
166	approved	Brian Funk	Cloudmesh Frugal with AWS, Azure, and Google Cloud.
168	approved	Deepak Deopura	AWS to/from Azure data transfer using APIs. -Extended version can be push data in SQL base warehouse (for example snowflake warehouse. It may be out of scope for now for this project purpose.
144	approved	Andrew Holland	Introduce Encryption Functionality for Cloudmesh Configuration File
146	approved	Kenneth Jones	Serverless API Performance Analysis
159	approved	Austin Zebrowski	Apache Airflow data pipelining between HDFS and Amazon Redshift and Cloudmesh
171	TBD	Jagadesh	Abstract Streaming Interfaces for the NIST Big data Architecture
172	approved	Nayeem	Cloudmesh key group management
174	approved	Sahithi R	title missing
151	approved, participants unclear	Qiwei Liu, Yanting Wan	Cloudmesh Cloud AI Services

HID	Status	Names	Project
140	approved but some details are missing	Mohamed Elfateh	Cognitive Response Simulation using Cloudmesh Cloud AI Services
153 164	approved	Anish, Siddhesh	Hadoop/Spark Cluster abstraction layer - dynamically changing available machines/storage available to hadoop/spark instances using orchestration tool and command line
161	approved, too limited with focus on R, needs to do more than R	Jim Nelson	Batch Processing using Kubernetes and Cloudmesh batch processing virtual cluster for batch processing of arbitrary programs including R. Possibly cloudmeh-flow would help, but we do the jobs on cubernetes, while the bactch jobs are formulated as workflows.

2.2 Cognitive Response Simulation *using* Cloudmesh Cloud AI Services

Mohamed Elfateh Abdelgader, fa19-516-140

2.2.1 Introduction

The main objective of this project is to provide AI capabilities on cloud. The developed functionalities will be implemented on two different cloud platforms. The scope of work for this deployemnt is to build cababilities on the cloud allows recieving feeds in a data form (either structured/unstructured) from end users and the service shouls interpret these feeds an responds back an automated cognitve responce, this reponse could be advices or actions based on the data nature. As the design and requirements are evolving during this project, scope of work shall be expanded as well and this will be reflected and documented here rapidly.

2.2.2 Technologies Used

- Python.
- Flask.
- OpenAPI.
- Development manchine is in MacOS

2.2.3 Design

[x] Database design completed.

2.2.4 Implementation

TBD

2.2.5 Challenges

- Finding enough resources of a similar approach.

2.2.6 Limitations

TBD

2.2.7 Conclusion

TBD

2.2.8 References

1. TBD
2. TBD

2.2.9 Progress

- [x] learned how to create a cloudmesh command
- [x] leraned how to use cloudmesh Config()
- [x] learned how to start a vm on openstack with cloudmesh

- [x] drafted report

- Week Sep 29 - Oct 5

  - [x] Collecting info about how to build prober APIs using OpenAPI.
  - [x] learning about different DBs Mongodbg and SQLlite in order to decide which is the most prober fit.

- Week Oct 6 - Oct 12

  - [x] Installing MongoDB and start drafting DB design.
  - [x] loading datafiles in Dev environment.

2.3 Deployement Of Database in Multiple Cloud ☁

Balakrishna Katuru(Bala) - **fa19-516-141**.

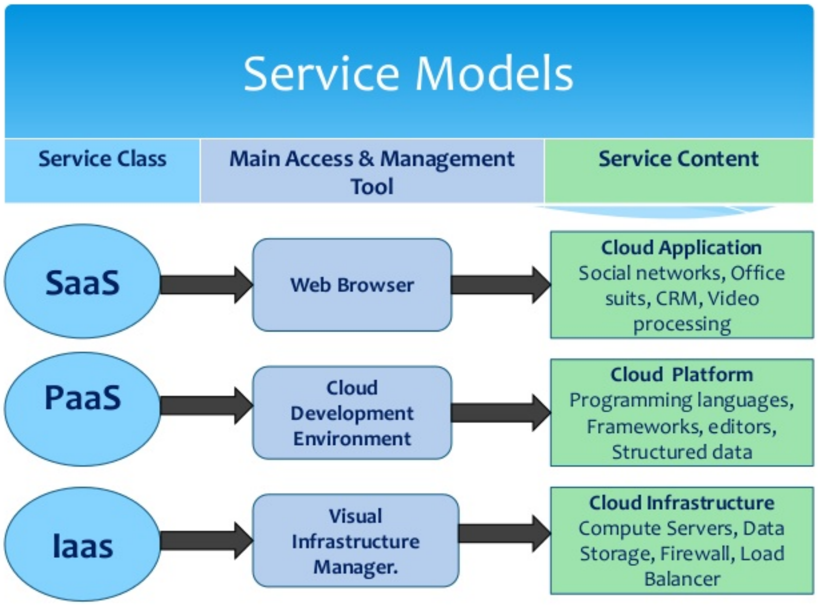
2.3.1 Objective

Deploying the Database in multiple clouds.

2.3.2 Facts and comparisions about top could service providers

This section discuss about the details of various cloud service models, public cloud service market share of major service providers. And also what are the challenges the customers expereinced while using cloud services.

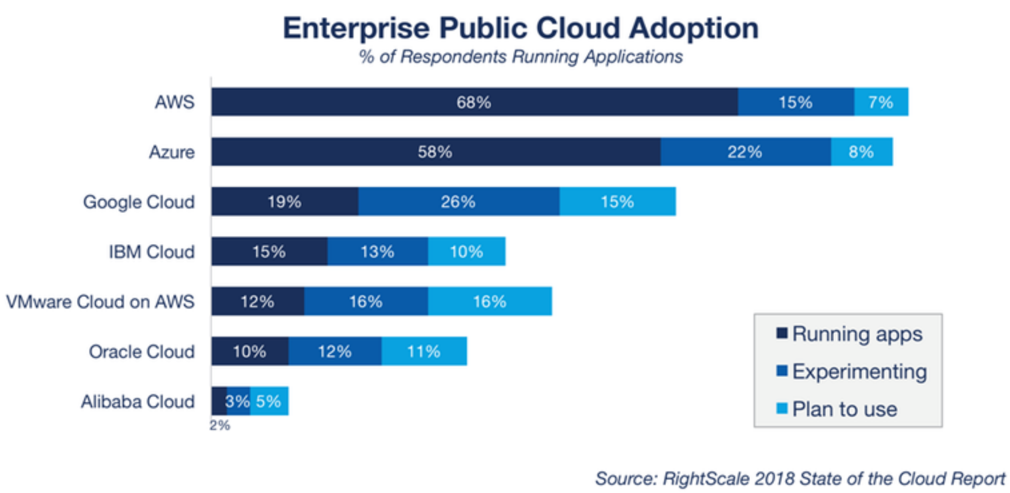
2.3.3 Types of Cloud – Services Models



Image

2.3.4 Enterprice Public Cloud Adaption

Look at the major public cloud providers stake.



Top four Cloud service providers Scorecard in 2018.

AWS vs. Azure vs. Google vs. IBM Enterprise Scorecard

Area	AWS	Azure	Google	IBM
% Adoption	68%	58%	19%	15%
YoY Growth in Adoption	15%	35%	26%	50%
% Adoption in Beginners	47%	49%	18%	14%
% with Footprint >50 VMs	58%	44%	17%	14%
YoY Growth in Footprint> 50 VMs	14%	38%	42%	56%

AWS leads

Other vendors lead AWS

Source: RightScale 2018 State of the Cloud Report

RIGHTSCALE

The challenges while using the Cloud services are given below

Job Title	#1 Challenge	#2 Challenge
C-level Executive / Owner	36% – Cost management	36% – Security
VP / SVP	50% – Cost management	Evenly split between other options
Director	24% – Security	22% – Cost management
Manager	26% – Control	24% – Cost management
Engineer	24% – Cost management	20% – Security
Architect	40% – Cost management	27% – Security
Developer	30% – Cost management	23% – Performance

Chart 3

2.3.5 Project Synapsis

- Can be considered cloud storage(Example: Create a Bucket in case of Google) to deploy the Database.
- Adapt a cloud based storage service using REST service & APIs.
- The Database will be deployed in multiple clouds using the appropriate storage service.
- Use an abstract API to deploy the Database independent of the underlying infrastructure.

2.3.6 Technologies

- Python
- Cloudmesh
- AWS
- API / Rest

2.3.7 Overview

This application will be developed using AWS SQL databaseas back end and Python Flask to create front end UI.

Backend: My SQL server database will be created on AWS. And then a table to be created to store the data Ex: Employee Name, Employee ID, Department, Salary, Phone, Email.

Frontend: Frontend will have basic controlsto input teh data like employee details. With user interaction, CRUD operations will be performed. Display all records, add new record, update existing record and delete the existing record.

Modules: Table, flask, mysql.

2.3.8 Implementation and Deployement steps

1. Creation of cloud accounts
2. Creating directories:
3. Creating Flask Module: Create app.py script to import the flask module.
4. Creating Database Configuration: Create db_config.py Python script to setup the MySQL database configurations to connect the database.
5. Creating Main Script: Create main.py script, that will define all URIs or Action paths to perform CRUD operations with user interation.
6. Implementing REST End Points:

2.3.9 Test Scenarios

1. When execute the python main.py, the server will start on default port 5000.
2. Success massage after insert data in to table. And also teh message after successful Update/delete operation.
3.
    4.

2.3.10 Participants

Bala

2.3.11 References

1. https://blog.rapidapi.com/how-to-use-an-api/
2. https://www.devteam.space/blog/top-10-cloud-computing-services-providers/
3. https://www.zdnet.com/article/top-cloud-providers-2018-how-aws-microsoft-google-ibm-oracle-alibaba-stack-up/
4.

2.4 Encryption of Cloud Secrets ☁

- Andrew Holland
    - repo: fa19-516-144
    - email: hollanaa@iu.edu

- Contributors
- Forked Branch

2.4.1 Introduction

The Cloudmesh project lacks certain security capabilities and secure coding practices. General improvements are required to address three primary concerns.

First, the code base requires a general audit on all files using openssl. Currently two files are responsible for the security protocols and usage of openssl. Namely the cloudmesh-cloud/cloudmesh/security/encrypt.py and cloudmesh-cloud/cms/management/configuration/security/encryption.py files. These files require a full audit and update to use well-defined modules.

Second, security related configuration files require obscuring secrets. Presently the cloudmesh.yaml file responsible for configuring Cloudmesh has all data presented in clear-text. The `cms config cat less` command obscures the passwords within the config file, but only for the screen. Any malicious user with read access to the file would be able to extract all passwords set within Cloudmesh. The config file should encrypt the password bytes by default to decrease the attack surface on Cloudmesh.

After addressing the primary concerns Cloudmesh will be capable of being further extended to integrate password managers directly into secrets management. This could include software such as Keepass. If successful, Cloudmesh can automate the generation, storage, and access of the keys and config.

2.4.2 Implementation

2.4.2.1 Encrypting Cloudmesh.yaml

2.4.2.1.1 Automating Key Management

2.4.2.1.1.1 SSH-Agent

Original plans included integrating ssh-agent to automatically retrieve passwords for key operations (such as encryption). This goes against the functionality of the SSH-Agent. As referenced in the IETF informational documentation for ssh-agent found here details that the agent should only be used for signing data.

The ssh-agent is used to prove possession of the private key without exposing the private bytes of the key. This is done by generating a public-private key pair and sending the public key to the server. When you attempt to ssh to the server it will request a signature of some random data. When you retrieve the data you request the ssh-agent signs it and this signed data is sent back to the server. Since the signature can be validated by the public key you prove the private key is within your possession. Notice this is different than using the actual key bytes.

To give a practical example of the agent being unable to provide private key bytes we can reference the ssh_agent demo directory.
In short, we will use a public-private key pair to encrypt some data.
Even if the private key is added to the ssh-agent a password will be prompted.
Please read the README within the directory further explanation.

2.4.3 Proposed Software to Integrate into Project

- Keepass(2)

- python module: python cryptography
- ssh-keygen

2.4.4 Related Concepts

- Elliptic Curve Cryptography
- Advanced Encryption Standard Galois Counter Mode (AES-GCM)

2.4.5 References

2.4.6 Tasks

2.4.6.1 Openssl Security Audit

Task Lead: Andrew
Status: In Progress Forked Branch

Last Update: Added symmetric and asymmetric encryption to CmsEncryptor class

2.4.6.2 Encrypting Cloudmesh.yaml Secrets

Task Lead: Andrew
Status: Pending
Forked Branch

Last Update: Add first test script for decryption and encryption

2.4.7 Progress

2.4.7.1 Week of Monday Nov. 11th

2.4.7.1.1 Andrew

1. add get_path() function to Config() class to return yaml dot path(s) to key
2. implemented means to get value from config given a dot path to key the get()
   This was necessary since the Config.get() function couldn't handle dot paths.
3. Wrote script to begin testing encryption of config values
4. Began integration into cloudmesh by testing Config.set with encrypted data

2.4.7.2 Week of Monday Nov. 4th

2.4.7.2.1 Andrew

1. Implemented asymmetric encryption using rsa
2. wrote script to demonstrate ssh-agent cannot be used with encryption
3. added README for running the scripts
4. added CmsHasher class to hash data (can be used for unique file name generation)

2.4.7.3 Week of Monday Oct. 28th

2.4.7.3.1 Andrew

1. Investigated using ssh-agent modules paramiko, and ssh2-python

2. Investigated ssh-agent documentation. Discovered use case is for signing only.
3. Implemented symmetric encryption using AES-GCM

2.4.7.4 Week of Monday Oct. 21st

2.4.7.4.1 Andrew

1. Finished Analysis of openssl related files found here
2. Established weekly meeting time with partner.
3. Discussed cms key and cms keygroup commands with partner.
4. Updated project.md to address concerns related to partner's project.
5. Added initial writings for book chapters within the /project/chapters dir
6. Added initial encryptor and pem_handler to replace old encrypt.py

2.4.7.5 Week of Monday Oct. 14th

2.4.7.5.1 Andrew

1. Created the KeyGroup.py file to handle key groups

    1. Need to investigate the purpose of SecGroup.output

2. Designed queries

2.4.7.6 Week of Monday Oct. 07th

2.4.7.6.1 Andrew

1. Forked cloudmesh-cloud to local repo

    1. fork located here

2. Edited cms key –source=FILEPATH to now parse filepath argument

    1. Within key-group branch
3. Investigating how to add new local-keygroup collection to database

4. Installed robo3t to observe changes to local mongodb

2.4.7.7 Week of Monday Sep. 30th

2.4.7.7.1 Andrew

1. Researched password managers for future integration discovered

    1. kpcli
    2. gopass
    3. kedpm
    4. keepass2 cli
2. Submitted PR for debian 9 installation of mongo
3. Installed docker on local system to ease testing
4. Began audit of cms-cloud/cms/security/encrypt.py-bug check /project/audit.md

5. Took second pass look through the encrypt.py-bug. Wrote questions for Gregor

2.5 Abstract Database Management On Multicloud Environments (with focus on Azure and AWS3) ☁

Harsha Upadhyay, fa19-516-147

Contributors

2.5.1 Objective

Abstract database management on Multicloud environments for the NIST Big Data Reference Architecture AWS, Azure.

2.5.2 Introduction

We will be providing database abstractions to host arbitrary databases in arbitrary cloud environments. In order to verify that the database provisioning multi cloud environment works, we will be providing a detailed test to manipulate data in database. This will include standard database functionality. The implementation is being conducted as part of API REST services and we will be using following clouds:

1. Amazon
2. Azure
3. and Local DB

We are providing pytests to deploy and execute the verification of the correctness of this services.

2.5.3 Motivation

Clouding computing is a market emerging trend. It provides on-demand availability of computer system resources, databases, storage etc. without direct maintenance of the platform by the user of cloud services.

A cloud database is a database which runs on a cloud computing platform. This platform can be private, public or hybrid. When we talk about database, there are two models,

1. Traditional cloud model
2. Database as a service (DBaaS)

Cloud database as a service is becoming more and more populer these days because of the following main reasons:

- no physical infrastructure needed
- can be scaled quickly and efficiently
- mostly self-managed database with less administrative overhead

In this cloud computing project, cloud database as a service feature will be used with the objective of creating a functionality to deploy a database in multiple cloud environment.

Here is a quick reference table giving the listing of database services available from different cloud services provider market leaders,

List of Database Products from Popular Cloud Service Providers

AWS

Auora*
RDS*
Redshift*
DynamoDB
ElastiCache for Memcached
ElastiCache for Redis
DocumentDB
Neptune
Timestream
QLDB

Azure

SQL Database*
Cosmos DB
Database for MySQL*
Database for PostgreSQL**
Database for MariaDB **
SQL Server on Virtual Machines
Database Migration Service
Cache for Redis
Table Storage
Data Explorer

Google

Cloud SQL*
Cloud Spanner*
Cloud Bigtable*
Cloud Firestone
Firebase realtime Database
Cloud Memory Store

MongoDB

MongoDB Atlas**

Oracle

NoSQL Database
Database Exadata Cloud Service
Database Exadata Cloud at Customer
Autonomous Data Warehouse
Database Cloud Service :Bare Metal
Autonomous Transaction Processing
Database Cloud Service :Virtual Machine

IBM

IBM Db2 on Cloud *
IBM Db2 Warehouse on Cloud
IBN Cloudant
IBM Cloud Databases for MongoDB
IBM Cloud Databases for Elasticsearch
IBM Cloud Databases for etcd
IBM Cloud Databases for PostgreSQL
IBM Cloud Messages for RabbitMQ
IBM Compose for JanusGraph
IBM Compose for ScyllaDB
Hyper Protect DBaaS for MongoDB
Hyper Protect DBaaS for PostgreSQL

SAP

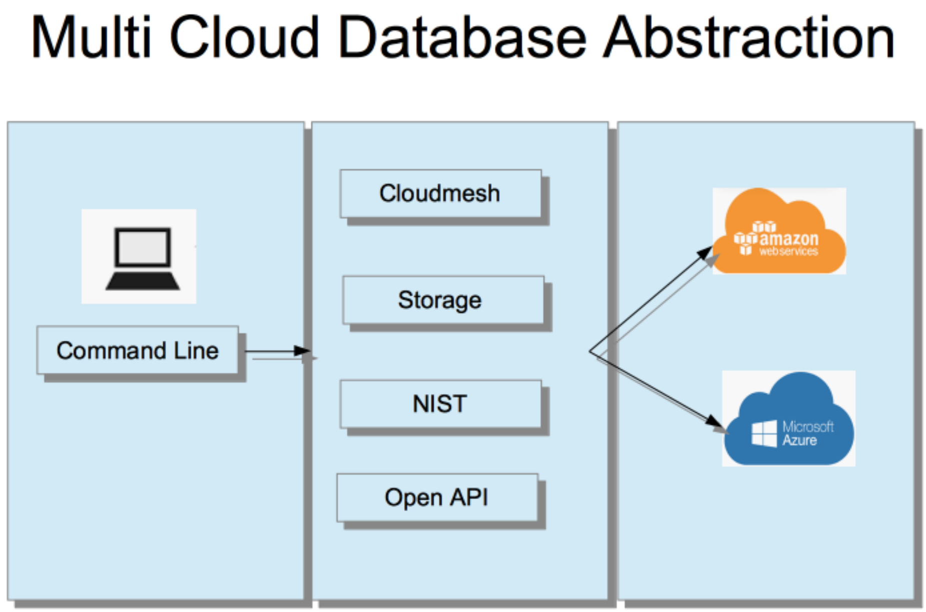
SAP Cloud Platform Big Data Services
SAP Cloud Platform SAP HANA service*
SAP Cloud Platform ASE service
SAP Cloud Platform Document service

* Relational
** Open Source

Comparison of Cloud Database

Type	AWS	Azure	Google
Relational	RDS, Aurora	Azure Database Azure SQL Database Database for MySQL Database for PostgreSQL Database for MariaDB	Cloud SQL Cloud Spanner
NoSQL- Key/Value	DynamoDB SimpleDB	CosmosDB, Table Storage	Cloud Bigtable Cloud FireStore Firebase Realtime Database
N0SQL- Document	DocumentDB	CosmosDB	-
NoSQL-Graph	AWS Redshift	Azure SQL Data Warehouse	Big Query
In-Memory	ElastiCache for Redis	Cache for Redis	Cloud Memorystore
TimeSeries	Timestream	Time Series Insights	Google Cloud Bigtable
NoSQL-Graph	Neptune	CosmosDB	Cloud Spanner*
Ledger	Quantum ledger Database QLDB	-	-

2.5.4 Architecture Diagram



2.5.5 Technology Detail

- Cloudmesh

  Cloudmesh is a multicloud architecture system which offers single architecture for using multiple cloud provides at the same time. Adbvantage of using cloudmesh is that it not only provides a REST based API but also commandline shell which makes easier to switch between clouds using single variable.

- REST / Open API

  APIs are sets of requirements that govern how one application can communicate and interact with another. connexion Open API 3 will be used in this project

- Python Scripting

  Python is a most popupar programming language which provides vast variety of libraries. Python can be used for developing web, desktop, scientific or any other application. Python will be used as scripting language in this project.

- AWS Redshift/ RDS & Azure SQL Database

  AWS and Azure are two market leading cloud services provider from Amazon and Microsoft respectively. AWS and Azure both offer number of database services. In project, relation SQL database from these cloud providers will be used.

2.5.6 Implementation Plan

2.5.6.1 Step 1: Database Object Creation

Create Database objects like tables, views on a cloud DB (Azure SQL Database) Python script to create objects

2.5.6.2 Step 2: Open API .yaml file

Create a yaml file * to get database , schema and DDL listing from one cloud environment (e.g Azure or AWS) * to create/copy database schema and DDL in other cloude environment

database.yaml

2.5.6.3 Step3: Setup Config file

Create a config file with environment detail for all environment used in the project

cloudmesh.yaml

2.5.6.4 Step4 : Integration with cloudmesh

Set up command line

2.5.7 Progress

- Azure account created
- A database created on Azure SQL Database
- Docker setup on local
- Python script to test connection to databse and deploy table
- AWS account creation
- Code diretory & file strucure set up complete
- Database.py created for get ,set Db and Schema
- Database.py tested for get ,set Db and Schema

2.5.8 Refernces

1. Cloud Computing by von Laszewski https://github.com/cloudmesh-community/book/tree/master/books
2. APIs and Python libraries https://cloud.google.com/python/docs/reference/
3. Google Cloud APIs https://github.com/googleapis/google-cloud-python#google-cloud-python-client
4. Cloudmesh Storage Open API https://github.com/cloudmesh/cloudmesh-storage/blob/master/cloudmesh/storage/spec/openapi_storage.yaml
5. NIST https://github.com/cloudmesh/cloudmesh-nist/blob/master/spec/database.yaml
6. Azure SQL Database https://azure.microsoft.com/en-us/services/sql-database/
7. Azure Cloud Database Services https://azure.microsoft.com/en-us/product-categories/databases/
8. AWS Cloud Database Services https://aws.amazon.com/products/databases/
9. Google Cloud Database Servcies https://cloud.google.com/products/databases/?hl=pl
10. Oracle Cloud Database Services https://www.oracle.com/database/cloud-services.html
11. IBM Cloud Database Services https://www.ibm.com/cloud/databases
12. MongoDB Cloud Databse Servcies https://www.mongodb.com/cloud

☁

2.6 Federated Kubernetes Clusters With Raspberry Pi

Sub and Gregor von Laszewski, fa19-516-148

laszewski@gmail.com

Insights

2.6.1 Abstract

2.6.2 Introduction

2.6.3 Related Work

2.6.4 Architecture

We leverage or improve https://github.com/cloudmesh/cm-burn

Our goal is to craete and contrast the creation of Federated kubernetes clusters. We have different models

1. The clusters are owned by a single user
2. The clusters are owned by multiple users

Goal is to create a federation of them. In each case the fedaration can be achieved in one of two ways.

1. consider a big kubernetis cluster that integrates all resources
2. consider a cluster of kubernetes clusters

This requires some investigation into kubernetes

Some images are shown in cm-burn that we may want tocopy here.

2.6.5 Technologies used

- cloudmesh cm-burn
- cloudmesh-inventory
- Kubernetes
- Docker

2.6.6 Benchmark and Evaluation

2.6.7 Conclusion

2.6.8 Other documentation files

- `sdcard-setup.md` contains documentation on cm-pi-burn and how the Pi images are modified/initialized. Also refer to the `cm-pi-burn.md` file in the `cm-burn` repo.

- `pi-setup.md` contains documentation on post-burn configuration that must be done after a Pi is booted up for the first time.

- `docker.md` contains documentation on creating the Docker image used for testing.

2.7 Cloudmesh Cloud AI Service ☁

The cloudmesh Cloud Ai service will provide AI capabilities that are running on the different cloud, e.g. chamelon, asure. For the example functions, linear regression, principle components analysis and so on will be provided so that users can utilize the computing power of the clouds to train their models. The cloudmesh cloud AI service will administrate multiple clouds and determine which cloud to use for scheduled tasks.

Link to the project:

- https://github.com/cloudmesh/cloudmesh-analytics

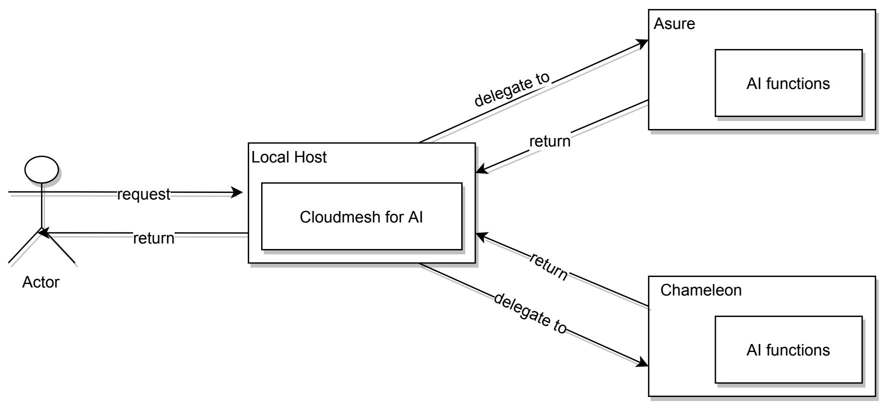
Insights:

- https://github.com/cloudmesh-community/fa19-516-151/graphs/contributors
- https://github.com/cloudmesh/cloudmesh-analytics/graphs/contributors

2.7.1 Group Members

- Qiwei Liu
  - hid: 151
  - Link to the hid repo
    - https://github.com/cloudmesh-community/fa19-516-151
- Yanting Wan
  - hid: 170
  - Link to the hid repo
    - https://github.com/cloudmesh-community/fa19-516-170

2.7.2 Architecture Design



architecture

The architecture primarily contains four objects:

- User which is the actor
- The application running on local host using cloudmesh will manipulate multiple cloud instance, decides delegate computational tasks to which cloud
- The AI services will be running on the Asure, Chameleon, and Chameleon cloud, exposing APIs to incoming requests, and return the return the result to the local host

2.7.3 Technical Analysis

2.7.3.1 Operating System

- Mac OS

2.7.3.2 Databae

- Mongodb

2.7.3.3 REST

- OpenAPI: The REST API will be defined by using OpenAPI specification
- Swagger editor: The swagger editor is used to write API documentation based on the OpenAPI standard
- Flask: The web application framework that handles incoming requests

- Connexion: Connexion is an application on the top of Flask that will map the REST API documentation to python functions on Flask
- Pytest will be the testing framework

2.7.4 Benchmark

2.7.5 Reference

2.8 Development

2.8.1 Progress

2.8.1.1 Week 6

Qiwei Liu

1. Set up flask web application framework
2. Set up the test framework and testing data based using sqlite3
3. Done file upload, list file
4. Set up chameleon instance

Yanting Wan

1. Set up connexion, and uses it to map Opean API(yaml) file.
2. Done uploading file locally, testing it on Swagger-ui
3. Done testing a ai function with locally stored dictionary as parameter, testing it on Swagger-ui

2.8.1.2 Week 7

Qiwei Liu

1. Update folder structure
2. Gregor update folder structure, refactor file routes, refacter to analytics route
3. Change folder structure
4. Change folder structure by putting functions under cloudmesh directory
5. update project report
6. Update report.md
7. Update linear regression tests and the exception handling
8. Change cloudmesh package structure
9. Merge remote-tracking branch 'refs/remotes/origin/master'
10. Add more tests to file operations
11. Done prototyping linear regression
12. Merge remote-tracking branch 'refs/remotes/origin/master'

2.8.1.3 Week 8

1. Migration to cloudmesh-analytics https://github.com/cloudmesh/cloudmesh-analytics

Yanting Wan

1. Download a virtual box to run Ubuntu 19.04 system.
2. Reinstall cloudmesh-cloud, mongoDB in Ubuntu 19.04.
3. Start a VM in Chameleon.

4. Create venv and install requirements in migrated project folder.

- [ ] Problem1: cannot ssh into VM

2.8.1.4 Week 9

Qiwei Liu

1. Add cms command to start and stop the server

Yanting Wan 1. Install Docker in ubuntu VM. 2. edit Dockerfile and run a simple example in Docker.

2.8.1.5 Comments on Files

2.8.2 Work Breakdown

2.8.3 Example Usages

1. Upload a file to the server that will be further processed

```
curl -X POST "http://localhost:8000/cloudmesh-ai-services/upload" -H "accept: application/json" -H "Content-Type: multipart/form-data" -F "file=@learn.rkt"
```

2. Checking the uploaded files

```
curl -X GET "http://localhost:8000/cloudmesh-ai-services/list-files" -H "accept: application/json"
```

3. Contracting a json file which contains the file name, and the parameters for the linear regression to the REST API. The output will be save on the server that could be downloaded.

```
curl -X POST "http://localhost:8000/cloudmesh-ai-services/linear-regression/linear" -H "accept: */*" -H "Content-Type: application/json" -d "{\"file_name\":\"string\
```

The next version will encapsulate the server request command and user can only provide the content body, For example，

```
cloudmesh ai upload "linear_regression.csv"
```

2.9 Cloudmesh Storage Provider for Virtual Directories - AWS to/from Google

Pratibha Madharapakkam Pagadala

fa19-516-152

https://github.com/cloudmesh-community/fa19-516-152/blob/master/project/report.md

2.9.1 Introduction

This project is to develop API and rest services to manage and transfer files between different cloud service providers. A cloudmesh based command will be implemented to transfer files present in a storage queue from a source to target cloud provider. In this instance, the functionality will be implemented for AWS and Google Cloud. For performance evaluation py tests will be created.

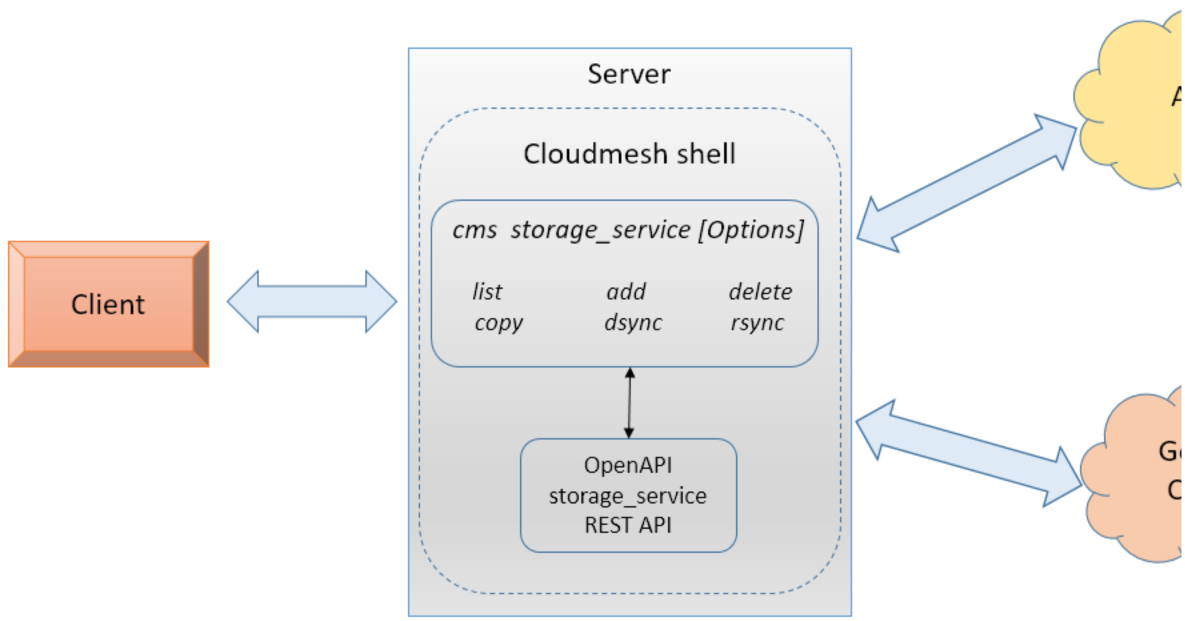
2.9.2 Motivation

Multiple cloud providers offer storage solutions to manage data in the form of files. The intention here is to build a command which can provide functionality to read the files from a queue and move them from a source to a target cloud provider's storage. In this method, users will be able to split or move the data across different cloud providers that provider cheaper solutions.

2.9.3 Requirements

- AWS, Google cloud accounts
- cloudmesh-storage API
- REST API
- cloudmesh-storage queue API
- Python API TBD

2.9.4 Architecture Diagram



Architecture

Description

- Client intiates a cms storage_switch command with options such as

    1. Recursive file copy from source to target
    2. List the files
    3. Add the files
    4. delete the files
- Cloudmesh storage_switch command will run on the local server. According to the options and arguments, this would delegate the functions between AWS and Google Cloud.

- Storage and Utility APIs on AWS and Google cloud.

2.9.5 Technology Used

- cloudmesh-storage
- Python
- REST
- AWS S3 Storage

- Google Storage
- OpenAPI

2.9.6 Progress

- Update report.md with architecture and command details.
- Generated command using cloudmesh shell - cms storage_service command
- Create AWS account and Google Cloud accounts
- Access AWS account using cloudmesh commands

Next Steps * Update provider classes for AWS and Google Cloud. * Explore the Google's gsutil api and read on how to use it in project. * Copy file from aws to google cloud * Copy file from google cloud to aws.

2.9.7 References

- https://github.com/googleapis/google-cloud-python#google-cloud-python-client
- https://aws.amazon.com/s3/
- https://boto3.amazonaws.com/v1/documentation/api/latest/guide/resources.html
- https://github.com/cloudmesh/cloudmesh-storage/tree/master/cloudmesh/storage

2.10 Spark Cluster Management Abstraction Layer ☁

Anish Mirjankar fa19-516-153
Siddhesh Mirjankar fa19-516-164

Insights: https://github.com/cloudmesh-community/fa19-516-153/graphs/contributors

2.10.1 Problem

In various enterprise data pipelines, there is a lack of multi-cloud architecture, often due to services like Spark being natively integrated into clusters such as AWS Elastic MapReduce. These data pipelines can benefit from a provider-agnostic solution that will encompass all their available options, rather than forcing them to choose a cloud platform over another. This can be especially beneficial to data teams that require dynamic storage solutions and want the flexibility to move between cloud platforms with ease.

2.10.2 Proposal

We will be exploring options for an implementation of Apache Spark that can be managed remotely from a multi-cloud orchestration service. We will abstract the storage and compute initalization within Spark to run parameterized jobs from this service. This will allow the performance bottlenecks of high-performance data transfer to be contained within the cluster itself, rather than a data source.

2.10.3 Action

In order to solve this problem, we will be implementing a Nomad and Kubernetes cluster, and generating a standalone Spark image that will run parameterized jobs, utilizing all of the available multi-cloud options available to the orchestator as well as all of the compute instances. We will also be implementing a testing service that will provide the cluster with the access to compute resources and storage that the jobs will need to run.

2.10.4 Solution

The solution is composed of 3 main parts, creating a cluster, interacting with the cluster, and deploying jobs to the cluster.

We are developing a command, `cms cluster`, that will perform all of these actions efficiently.

The following commands will be integrated into the cloudmesh service:

```
cluster create -n NAME -p PROVIDER [HOSTNAMES]
cluster add -n NAME HOSTNAME
cluster remove -n NAME HOSTNAME
cluster kill -n NAME # only cloudmesh - bring every machine involved in server down
cluster info # find all clusters
cluster info -n NAME # find info about given cluster (query the address for either kubernetes or nomad)
cluster submit -n NAME JOB
cluster list
```

Source

2.10.4.1 Interaction

We are interacting with the nomad and kubernetes REST APIs to dynamically modify and interact with the cluster/agent configurations while jobs are running. For each interaction, cloudmesh queries the appropriate provider's API to perform the action to avoid managing a local state.

2.10.4.2 Initialization

Using this mechanism, cloudmesh will be able to simultaneously initialize and prepare machines in a cluster while building and deploying the images.
The initialization and preparation steps will submit the requested shell script to each machine added to the cluster: - Kubernetes - Nomad

2.10.4.3 Deployment

When submitting a job to each of these providers, cloudmesh will first build the requested image: - Hadoop - Spark - **TODO** - Cloudmesh - if a remote instance is needed

And submit the jobfile to the cluster using the provider's REST API.

Sources: - Kubernetes - Nomad - Hadoop

2.11 Cloudmesh Data Transfer Service for AWS S3 and Azure Blob

Ketan Pimparkar, fa19-516-155

2.11.1 Abstract

TBD

2.11.2 Objective

Provide cloudmesh users an API and REST service to transfer files, directories from data storage of one cloud service provider to other cloud service provider. This packge will consider AWS S3 and Azure Blob storage for current implementation.

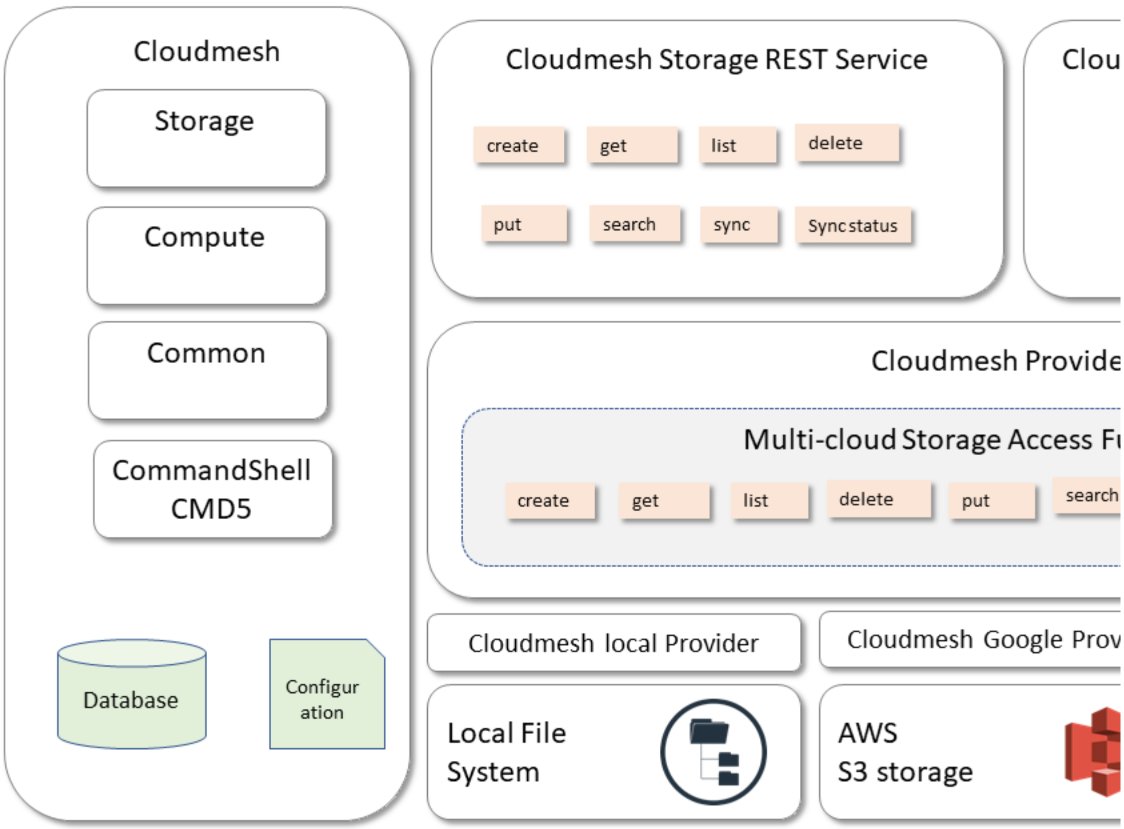
2.11.3 Motivation

Cloud technology evolves at a very fast rate. Due to which, policies and facilities provided by cloud service providers change as well. There could be various practical scenarios in which users want to transfer the data currently stored in AWS S3 to Azure Blob. Such scenarios could be change in pricing policy or storage capacity rules of AWS S3 or Azure Blob.

Cloudmesh is a multicloud platform. With inclusion of data transfer service, a highly optimized and simple to use methos will be made available to cloudmesh users.

2.11.4 Architecture

- Architecture diagram:



CM Transfer Architecture Diagram

Diagram credit: Prof. Gregor

2.11.5 Technology

- AWS S3
- Azure blob storage
- Python
- cloudmesh storage
- OpenAPI 3.0.2
- REST

2.11.6 Usage

- API:
- The code is available here

```
Usage:
    transfer config [--file=ip_file]
    transfer --id=<transfer_id> --data=<file_name> [--copy=True|False]
    transfer status --id=<transfer_id>
    transfer statistic

This command is part of CloudMesh's multicloud storage service.
Command allows users to transfer files/directories from storage of one
Cloud Service Provider (CSP) to storage of other CSP. Current
implementation is to transfer data between Azure blob storage and AWS
S3 bucket.

Arguments:
  transfer_id   A unique id/name assigned by user to each transfer
                instance
  file_name     Name of the file/directory to be transferred
  Boolean       True/False argument for --copy option. When False,
                data will be removed from source location
  ip_file       Input file used to configure 'transfer' command

Options:
  --id=transfer_id        Specify a unique i/name to the transfer
                          instance
  --data=file_name        Specify the file/directory name to be
                          transferred
  --copy=True|False       Specify is the data should be kept in
                          source location after the transfer
  --file=ip_file          Specify the file to be used for
                          configuration of the transfer instance
  -h                      Help function

Description:
  transfer config [options..]
        Configures source/destination and authentication details
        to be used by transfer command

  transfer [options..]
        Transfers file/directory from storage of one CSP to
        storage of another CSP

  transfer status [options..]
        Returns status of given transfer instance

  transfer statistic
        Returns statistics of all transfer processes

Examples:
  transfer --id="Dummy transfer" --data=dummy_file.txt --copy=True
```

- REST service:

TBD

2.11.7 References

- AzCopy: https://docs.microsoft.com/en-us/azure/storage/common/storage-use-azcopy-s3?toc=%2fazure%2fstorage%2fblobs%2ftoc.json
- AWS Boto3 API: https://boto3.amazonaws.com/v1/documentation/api/latest/index.html?id=docs_gateway
- Cloudmesh manual: https://cloudmesh.github.io/cloudmesh-manual/preface/about.html
- Install Azure python SDK: https://docs.microsoft.com/en-us/azure/python/python-sdk-azure-install
- Azure python API usage: https://github.com/Azure-Samples/storage-blobs-python-quickstart/blob/master/example.py
- Cloud computing book by Prof. Grogor: https://laszewski.github.io/book/cloud/

2.11.8 Benchmarks

TBD

- Benchmark report to be created

2.11.9 Testing

TBD

- PyTest report to be created

2.11.10 Progress

- done. Installation of Cloudmesh and mongoDB in Windows10 Pro system

- done. Creation of AWS EC2 instance and S3 buckets
- done. Access AWS account using Cloudmesh CLI
- done. Define architecture of the transfer service
- done. Define docopt of the transfer service
- Define test cases
- done. Creation of Azure account
- Copy files from local storage to AWS S3 and Azure Blob

2.12 Hadoop and Spark Clusters With Raspberry Pi

2.12.1 Abstract

Deployment of Hadoop and Spark on Raspberry Pi Clusters which involves: * Using CM-BURN command to burn multiple SD cards at once * Deploying Hadoop and Spark on Raspberry PI Clusters * Creating a cluster with as many nodes as we have SD cards for

2.12.2 Introduction

Majority of the data in today's world has been stored in HDFS. HDFS stands for Hadoop Distributed Storage System. The Raspberry Pi provides to the community a cheap platform with the ability to expose Linux and other operating systems to the masses. Due to its cost point, it is easy to buy a PI and experiment with it. As such this platform has been ideal to lower the entry barrier to advanced computing from the university level to highschool, middle school and even elementary school. However, the PI has also been used by universities and even national labs. Due to its availability and its convenient accessibility, it has become a staple of our educational pipeline. Due to its price point the PI can also be used to build cheap clusters putting forward a hardware platform ideal for experimenting with issues such as networking and cluster management as an educational tool. Many such efforts exist to use a PI as a cluster environment.

So it would be a good idea if we could somehow turn such a platform more powerful by deploying latest technologies such as Hadoop and Spark on it. Multi cluster Raspberry Pi, where one node can act as the master node and other nodes act as slaves and the master might be able to control the slaves.

2.12.3 Related Work

2.12.4 Architecture

2.12.5 Technologies used

- cm-burn
- cloudmesh-inventory
- Python
- HDFS
- Hadoop
- Spark

2.12.6 Work Breakdown

2.12.7 Progress

- Burning SD Cards using Etcher on Windows
- Read about battery operated clock for Raspberry Pis, NTP and RTC

2.12.8 Benchmark and Evaluation

- Develop a test program to review Hadoop and Spark on Clusters
- Series and parallel flashing of SD cards

2.12.9 Conclusion

2.12.10 References

- https://raspberrytips.com/install-raspbian-raspberry-pi/
- https://raspberrytips.com/raspberry-pi-cluster/
- https://www.raspberrypi.org/documentation/remote-access/ip-address.md
- https://dev.to/awwsmm/building-a-raspberry-pi-hadoop-spark-cluster-8b2
- https://dqydj.com/raspberry-pi-hadoop-cluster-apache-spark-yarn/
- https://www.mocomakers.com/building-a-raspberry-pi-cluster-with-apache-spark/
- https://tekmarathon.com/2017/02/16/hadoop-and-spark-installation-on-raspberry-pi-3-cluster-part-4/amp/
- https://medium.com/@glmdev/building-a-raspberry-pi-cluster-784f0df9afbd
- https://dev.to/awwsmm/building-a-raspberry-pi-hadoop-spark-cluster-8b2

2.13 AWS to Azure Data Pipeline: Using Cloudmesh in Airflow

Austin Zebrowski, fa19-516-159

azebrows@iu.edu

2.13.1 Abstract

Cloudmesh enables easy creation of virtual machines (VMs) in multiple clouds. Apache Airflow is an open-source data pipeline orchestration tool. By leveraging the strengths of each, files can be moved between clouds, and the files' status and history can be viewed through a user interface (UI). This can be achieved by creating a library of functions that uses Cloudmesh commands in an Airflow environment.

2.13.2 Introduction

Cloudmesh is a great tool for using the resources of multiple clouds. One of Cloudmesh's features is cloudmesh-storage, which enables cloud-to-cloud data transfer. This meets a key need, but often want to do more than simply move the files. This is where we can use Airflow - a python-based open source data pipelining project.

With Airflow, we can configure directed acyclic graphs (DAGs) that define a sequence of tasks in a workflow. We can then view these DAGs in a UI that displays their schedules, code definitions, histories, and execution logs. By creating a library of functions that uses Cloudmesh within Airflow, we can easily schedule and visualize the startup of VMs on multiple clouds and a data transfer between these clouds.

To make these capabilities broadly accessible, we create a RESTful API that exposes functionality such as starting VMs, listing files, and moving files.

2.13.3 Architecture

Architecture
Architecture

2.13.4 Implementation

- Technologies Used

    1. Apache Airflow

2. Azure Blob
3. AWS S3
4. Cloudmesh
5. Docker
6. Open API 3.0

2.13.5 Progress

10/18

Installed a fully operable instance of Cloudmesh. Previous Cloudmesh installation was not fully functional due to a problem in my environment.

10/20

Created Amazon AWS account. added "cloudmesh" user, and added access key information into cloudmesh.yaml. Attempted, unsuccessfully, to launch an AWS VM with Cloudmesh (can not refresh image list).

10/21

Created dockerfile and docker compose file to spin up a Docker container running Airflow.

10/24

Created architecture diagram.

10/28

Improvements to Docker Airflow setup:

```
1) Adding entrypoint script
2) Mounting local files to docker image
3) Creating test dag
4) Editing Airflow config file
5) File organization
```

11/4

Cloudmesh installation on Docker - first attempt (needs more work)

2.13.6 Work Breakdown

All of this work is to be completed by Austin Zebrowski. This is not a group project.

2.13.7 References

2.13.8 Results

TBD

- Deployment Benchmarks
- Application Benchmarks

2.14 Azure blob to / from Google cloud Storage Service (Cloudmesh Data Transfer Service) ☁

Shreyans Jain, fa19-516-160

2.14.1 Abstract

TBD * Azure blob to Google cloud storage and vise versa Cloudmesh Storage Provider for Virtual Directories: see cloudmesh-storage to start. develop OpenAPI REST services for it * Py Test

2.14.2 Objective

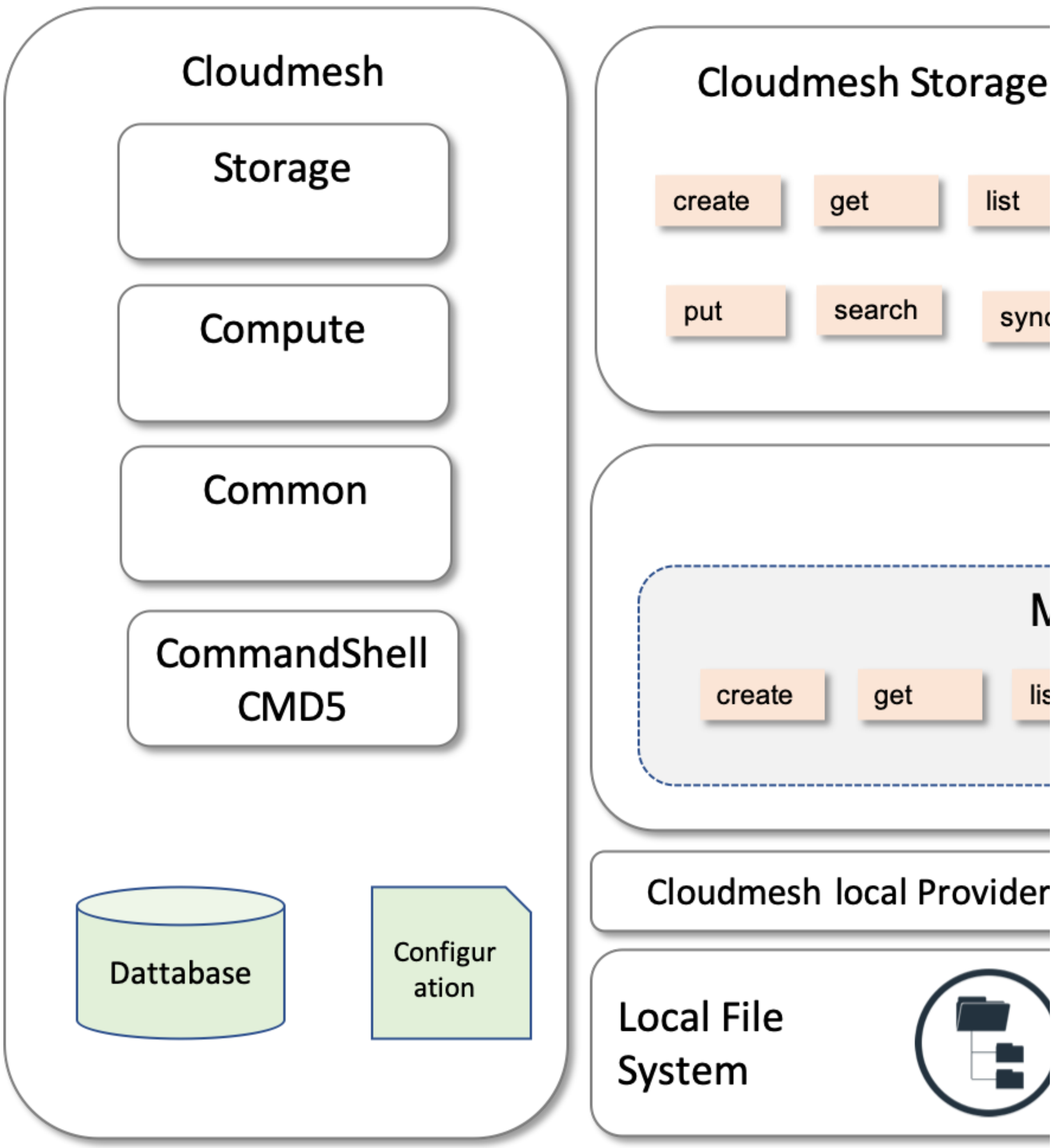
Provide cloudmesh users an API and REST service to transfer files, directories from data storage of one cloud service provider to other cloud service provider.
This packge will consider Azure Blob storage to Google cloud for current implementation.

2.14.3 Motivation

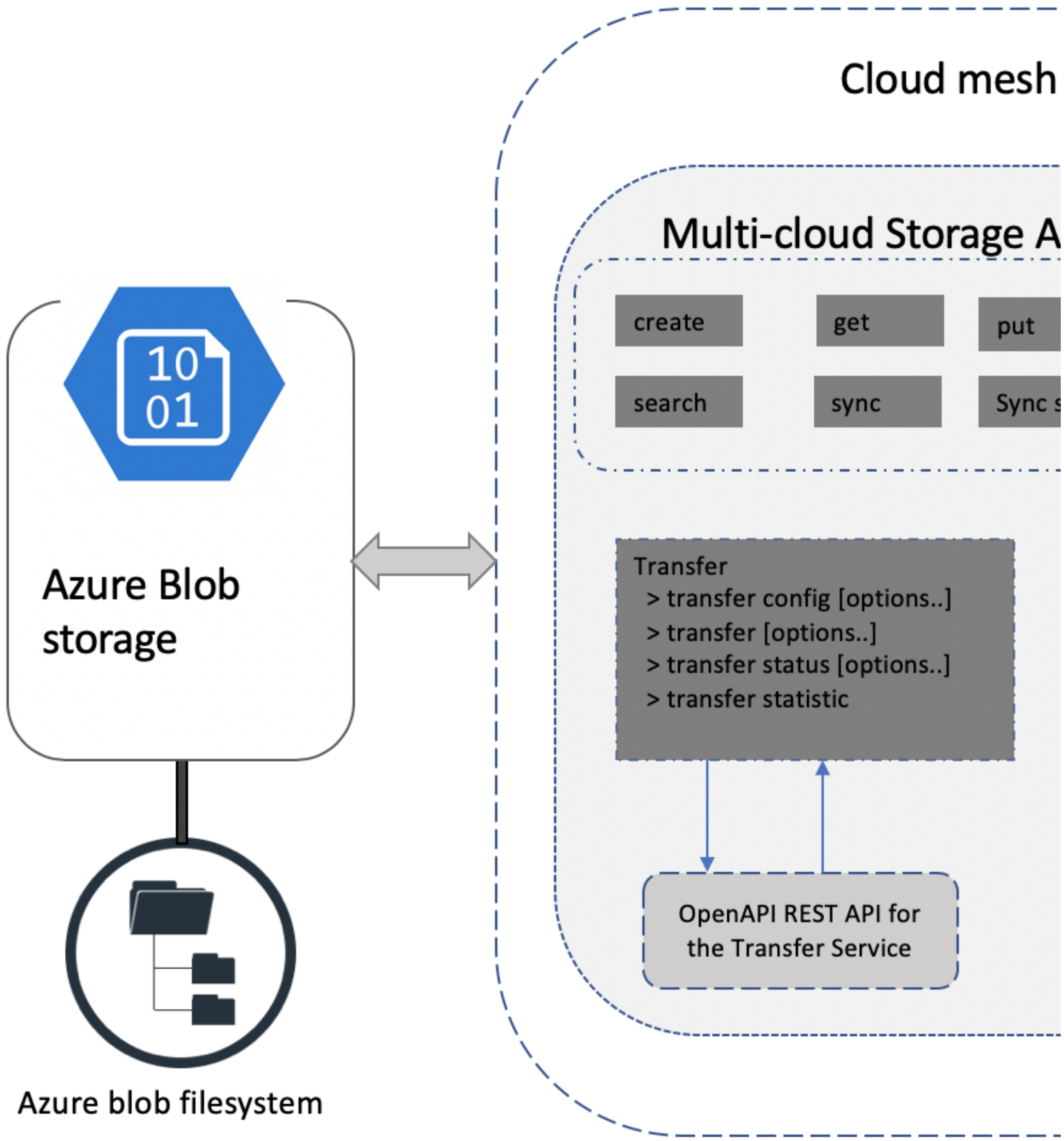
Cloud technology evolves at a very fast rate. Due to which, policies and facilities provided by cloud service providers change as well. There could be various practical scenarios in which users want to transfer the data currently stored in AWS S3 to Azure Blob. Such scenarios could be change in pricing policy or storage capacity rules of AWS S3 or Azure Blob.

Cloudmesh is a multicloud platform. With inclusion of data transfer service, a highly optimized and simple to use methos will be made available to cloudmesh users.

2.14.4 Architecture



Architecture



Architecture

2.14.5 Technology

- Python
- cloudmesh storage
- OpenAPI 3.0.2
- REST
- Azure blob storage
- Google Cloud Storage

2.14.6 Usage

- API:

```
Usage:
      transfer config [--file=ip_file]
      transfer --id=<transfer_id> --data=<file_name> [--copy=True|False]
      transfer status --id=<transfer_id>
      transfer statistic

This command is part of CloudMesh's multicloud storage service. Command allows users to transfer
files/directories from storage of one Cloud Service Provider (CSP) to storage of other CSP.
Current implementation is to transfer data between Azure blob storage and AWS S3 bucket.

Arguments:
    transfer_id   A unique id/name assigned by user to each transfer instance
    file_name     Name of the file/directory to be transferred
    Boolean       True/False argument for --copy option. When False, data will be removed from source location
    ip_file       Input file used to configure 'transfer' command

Options:
    --id=transfer_id        Specify a unique i/name to the transfer instance
    --data=file_name        Specify the file/directory name to be transferred
    --copy=True|False       Specify is the data should be kept in source location after the transfer
    --file=ip_file          Specify the file to be used for configuration of the transfer instance
    -h                      Help function

Description:
    transfer config [options..]
          Configures source/destination and authentication details to be used by transfer command

    transfer [options..]
          Transfers file/directory from storage of one CSP to storage of another CSP

    transfer status [options..]
          Returns status of given transfer instance

    transfer statistic
          Returns statistics of all transfer processes

Examples:
    transfer --id="Dummy transfer" --data=dummy_file.txt --copy=True
```

- REST service: TBD

2.14.7 Benchmarks

TBD - Benchmark report to be created * Benchmarks of what has been developed leveraging cloudmesh convenient stopwatch.

2.14.8 Testing

TBD - PyTest report to be created

2.14.9 References

- https://cloud.google.com/docs/compare/azure/
- https://cloud.google.com/docs/compare/azure/storage

2.15 Batch Processing using Kubernetes and Cloudmesh

Jim Nelson, fa19-516-161

please learn markdown, see section headers

Abstract

Introduction

Containerized applications have become the most efficient method for application deployment at dynamic scale. Containers allow for the simplified and separate functions of application development and functionality from deployment and cluster management. Containers also have the advantage over VMs of being lightweight and portable for use with different IaaS, PaaS, private/hybrid clouds and even local operating systems. Kubernetes is an open source project initially developed at Google. It allows for overall orchestration of cluster management including application/dependencies co-localization, automated app rollouts/batch processing, automated scalability, cluster health monitoring and multi-cloud mobility.

Objective

The goal of this project is to develop a work flow for efficient, large-scale, batch processing of applications using Kubernetes for use with Azure and Chameleon Cloud clusters. This project (along with other current projects) will also represents one the first uses of Kubernetes in the Cloudmesh architecture. Kubernetes will be utilized for cluster management, Cloudmesh will allow multi-cloud deployments. The SaaS application which will be batched will be time series forecasting using R packages.

Implementation

1. **Architecture**
   - Chameleon Cloud
   - Azure
   - Kubernetes
2. **Technologies**
   - Cloudmesh
   - Python
   - R
   - OpenAPI 3.0
   - Connexion
   - pytests

Results

1. **Benchmarks**

   - Cloudmesh command line installation
   - REST services
   - Kubernetes deployment API
   - Batch processing API
   - Cloud deployment
   - pytests

2. **Findings**

Discussion

2.16 Oracle Compute and Storage Service

Shivani Katukota, fa19-516-162

Insights: * https://github.com/cloudmesh-community/fa19-516-162/graphs/contributors * https://github.com/cloudmesh/cloudmesh-oracle/graphs/contributors

Example code at github location: https://github.com/cloudmesh/cloudmesh-oracle/blob/master/examples/examples.py

Code Directory: https://github.com/cloudmesh/cloudmesh-oracle

2.16.1 Abstract

- Oracle has released a free tier for Oracle cloud.

2.16.1.1 Cloudmesh Compute

- cloudmesh-compute project will identify Oracle's python API and develop its provider.
- cloudmesh-compute project will identify how to manage credentials in Oracle.
- cloudmesh-compute project will write and run pytests on Oracle cloud.

2.16.1.2 Cloudmesh Storage

- cloudmesh-storage project will add the feature to access storage from Oracle services.
- cloudmesh-storage project will provide a REST service based on OpenAPI that uses the cloudmesh API.
- cloudmesh-storage project will implement virtual directory from local.

2.16.2 References

2.16.3 Progress

2.16.3.1 Week 1 (30th September)

1. Setup Oracle cloud config file and installed python api for Oracle
2. Made following functions to work as examples on Oracle cloud using python
    - List images
    - Find image with given name
    - List flavors on cloud
    - Stop server with given name
    - Resume stopped server
    - List all servers
    - Delete/Terminate server
    - Reboot server
    - Rename server

2.16.3.2 Week 2 (7th October)

1. Figured out how to use cloudmesh.yaml for credentials
2. Installed mongo db on windows successfully

2.16.3.3 Week 3 (14th October)

1. Created python examples for following functions:
    - create vcn and subnet
    - launch an instance
2. Started integrating examples into Provider.py - In Progress, Have not checked-in yet

2.16.3.4 Week 4 (21st October)

1. Integrated the examples into Provider.py
2. Fixed issues with integrated examples to make them run via cmd
3. Fixed issues with saving data in mongo db via commands

2.16.3.5 Week 5 (28th October)

1. Integrating examples related to instance start and stop

2.16.3.6 Week 6 (4th November)

1. boot command worked but with minor defects

2.16.3.7 Week 7 (11th November)

1. ssh command

2.16.3.8 Week 8 (18th November)

1. boot and ssh work with default key
2. Issues running compute test cases
3. Started with oracle storage project

2.16.4 Work breakdown

2.17 TITLE MISSING

- Anish Mirjankar
- Siddhesh Mirjankar

2.17.1 Problem

In various enterprise data pipelines, there is a lack of multi-cloud architecture, often due to services like Spark being natively integrated into clusters such as AWS Elastic MapReduce. These data pipelines can benefit from a provider-agnostic solution that will encompass all their available options, rather than forcing them to choose a cloud platform over another. This can be especially beneficial to data teams that require dynamic storage solutions and want the flexibility to move between cloud platforms with ease.

2.17.2 Proposal

We will be exploring options for an implementation of Apache Spark that can be managed remotely from a multi-cloud orchestration service. We will abstract the storage and compute initalization within Spark to run parameterized jobs from this service. This will allow the performance bottlenecks of high-performance data transfer to be contained within the cluster itself, rather than a data source.

2.17.3 Action

In order to solve this problem, We will be implementing a Nomad cluster, and generating a standalone Spark image that will run parameterized jobs, utilizing all of the available multi-cloud options available to the orchestator as well as all of the compute instances. We will also be implementing a testing service that will provide the cluster with the access to compute resources and storage that the jobs will need to run.

2.17.4 Comments

LICENSE

- Apache

develop

- abstract API
- generate abstarct interface do deply a cluster (my be nomad on the backend, could be other technology)

- develop REST API with conexion introspection (we will teach you). e.g. Work on an abstarction API that can be used in the commandline implementation and the REST service

propose cloudmesh command(s) to

- Develop a proposed command in docopts from the commandline
- deploy nomad cluster
- generate spark image
- manage parameterized jobs

Testing

- develop pytests to test them and benchmark deployment (teardown) and execution of the pipeline sepeartely.
- Develop this report further with an Architecture diagram

Document

- use proper markdown ;-)
- add bibtex refernces … in report.bib and use in your document

2.18 Cloudmesh Compute Project with Google Cloud Platform (GCP) ☁

Zhi Wang, fa19-516-165, School of Public Health, Indiana University Bloomington

2.18.1 Abstract

2.18.2 Introduction

The purpose of this project is to implement related features to simplify compute interface. The project implementation will include the following clouds:

- Google Cloud Platform
- OpenStack

The benchmarks for this project will include the following components:

- virtual machine
- image
- flavor
- security groups

2.18.3 Progress

2.18.3.1 Week October 14

2.18.3.1.1 Solved Cloudmesh setting issues on macOS Catalina

2.18.3.1.2 Learned command line, bash file, and zsh in macOS Catalina

Since Catalina uses zsh as its default shell, installation process is a little different compared with previous bash file. Thanks to Dr. Laszewski I was able to install Cloudmesh smoothly. Besides this, I also studied a little further on shell on .bash_profile and .zprofile.

2.18.3.1.2.1 Why switch from .bash_profile to .zprofile

I found that it is even though there are bash v4 and v5 available, Apple does not update to newer versions because of the licenses (bash v3.2 with GPL v2. and v4/v5 with GPL v3.).

2.18.3.1.2.2 Solution to stick to bash without warning message

Bash file is not gone but will not last indefinitely either. For Catalina users, if you want to stick to bash, you can add follow environment variable to your .bash_profile or .bashrc.

```
export BASH_SILENCE_DEPRECATION_WARNING=1
```

For more information about zsh configuration, please visit oh-my-zsh for more information.

2.18.3.1.2.3 Complete Steps to reinstall/create Python virtual environment in macOS Catalina

- Why we need Python virtual environments?

For different python projects, developers need different python development environment to deploy codes. Virtual environments are needed to avoid destroy

- Steps to to create new environment for new projects by command lines.

  1. Create a new python 3 virtual environment and name it

     ```
     python3 -m venv ~/environment_name
     ```

  2. Activate your new virtual environment. This step is important since you need to activate your new python environment.

     ```
     source ~/environment_name/bin/activate
     ```

  3. Even though you can use pycharm to install a package, you still need to have a package manager to help yo manage your packages. Starting this point, you are able to install the packages you want.

     ```
     pip install pip -U
     ```

  4. To manage your environment easier, you can modify your zprofile and switch your environment easily.

     ```
     alias d="source ~/environment_1/bin/activate "
     alias env3="source ~/environment_2/bin/activate"
     ```

     Next time, you can just type aliases to activate your environments.

2.18.4 Results

2.18.4.1 Limitations

2.18.5 Conclusion

2.18.6 References

- Cloudmesh Development Projects- Cloudmesh Compute
- Using the Cloud Client Libraries for Python
- Github Cloudmesh Compute
- oh-my-zhs
- Scripting OS X

2.19 Cloudmesh Frugal With AWS, Azure, and Google Cloud ☁

Brian Funk

brfunk@iu.edu

fa19-516-166

Contributors

2.19.1 Abstract

Cloudmesh command for comparing the cost of compute for AWS, Azure, and GCP in various regions. It will compare price relative to the hardware specifications of the machines, an provide the VM with the best value.

2.19.2 Introduction

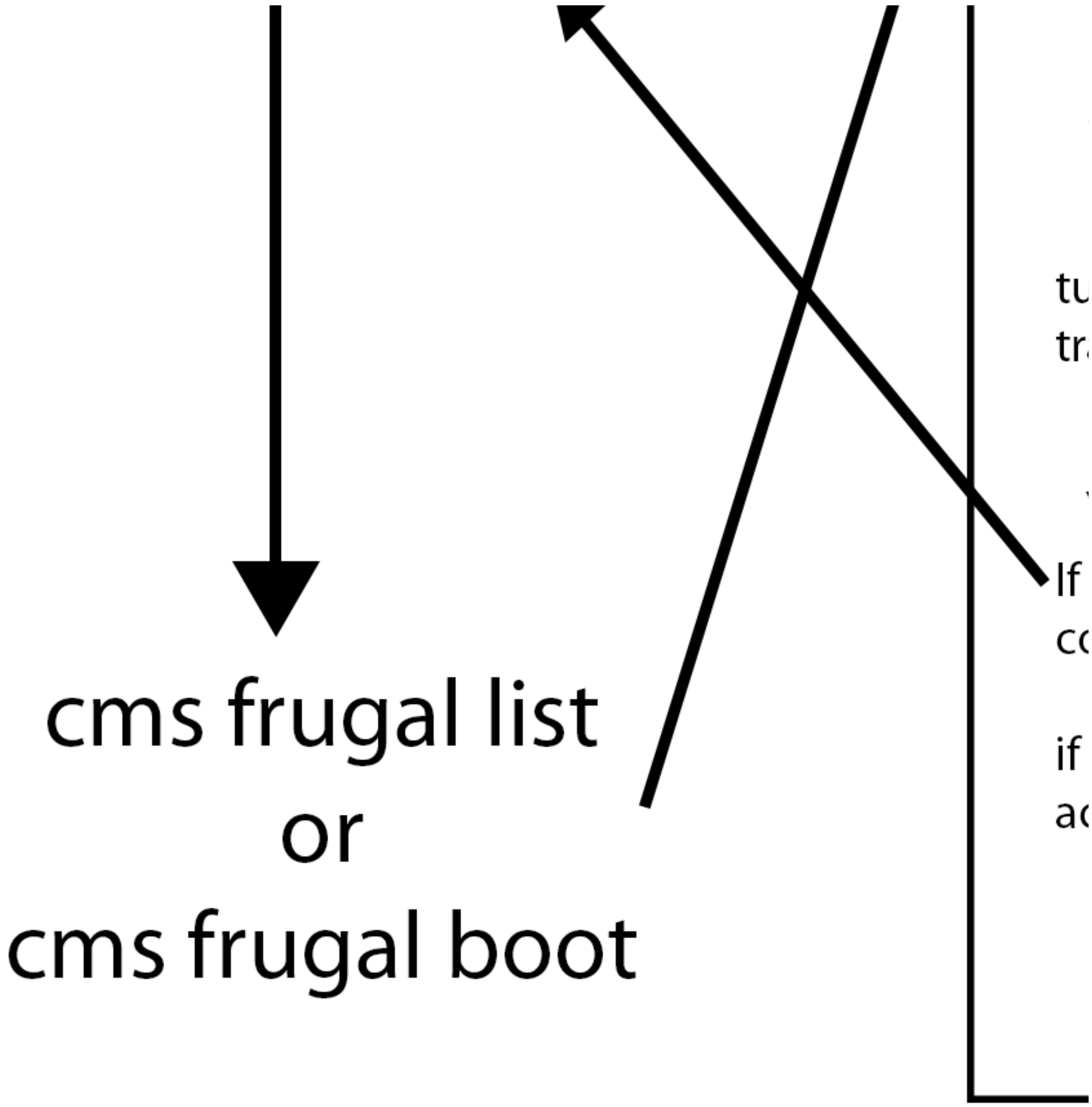
Cloudmesh Frugal will collect pricing information on all of the availble VMs for AWS, GCP, and Azure. Those prices will be compared to the performance of the machine in benchmarking, which will then all be compared against each other. There is already pricing information for AWS in Cloudmesh, which will be extended to GCP and Azure. The frugal benchmarks will compare them based on hardware specs/price

2.19.3 Design

Calling the frugal command (not yet designed/finalized) will first check to see if pricing information exists in the local mongodb for AWS, GCP, and Azure flavors. As of of now, only the information only exists for AWS. If the information does not exist, then it is pulled (and stored back into db? -join for if vm exists but not pricing?).Data will then be joined into a single numpy area/pandas frame (depending on calculations tbd). Many of the dimensions of the VMs will be transformed by price for benchmarking. The transformed table should also then be saved to the mongodb for memory? Or calc script will rerun since pricing will stay in mongo? Once again tbd. Anyway, the best value vm will be returned.

2.19.3.1 Architecture

Below is an early sketch of the logic flow of cloudmesh frugal. It will likely be changed in final implemenetation, as this is just an overview.



2.19.4 Implementation

TBD

2.19.4.1 Technologies Used

TBD

add Docker information (see Piazza post 354)

2.19.5 Results

TBD

2.19.5.1 Deployment Benchmarks

TBD

2.19.5.2 Application Benchmarks

TBD

2.19.6 Limitations

TBD

2.19.7 Conclusion

2.19.8 Work Breakdown

2.19.8.1 TODO

- add Azure proper to frugal
- frugal boot - need to actually boot instead of just limit
- pytest for how long it takes to do various frugal cmd commands
- some kind of benchmark for a booted machine that measures the actually speed
- lastly I need to do some kind of Docker test or something?

2.19.8.2 Questions moving forward

- When to write back to mongodb? IE should GCP and Azure flavors be written back to mongodb? Compute for those two clouds are in development so might have to wait until those are further along before worrying about them
- Keep everything in frugal.py or move GCP and Azure get pricing functions to other directories much like AWS currently is? Once again this will depend on other developments to Cloudmesh
- GCP has an api to get pricing but it needs a key…how to use with cloudmesh/is this okay?

2.19.8.3 Weekly Work Updates

2.19.8.3.1 Week of 9/29/19

Cloudmesh not entirely working at this point, began to explore AWS frugal example. Tested some urls to obtain the pricing info from GCP and Azure. Needs to be cleaned, and will also likely need to be pushed into mongodb (not running yet). See AWS example. Code so far at [frugaltesting.py]

2.19.8.3.2 Week of 10/6/19

Most of the week was spent getting Cloudmesh to run properly after bouncing back and forth from using Windows to Windows Subsystem. Success on Window so will be using that for development entirely. Was able to get Chameleon VM up and running after some trials, so which helped in understand of Cloudmesh command flow. No code updates

2.19.8.3.3 Week of 10/13/19

Currently reading documentation on OPENAPI and thinking about how that service will be integrated in frugal project. Updated report with all new progress, and Google pricing is completed in frugaltesting.py, and it is done without using the API. It uses a simplified JSON from their calculator. The file is parsed into the general structure that will be used for benchmarking, a 2d numpy array with features [provider, machine name, region, cores, memory, and price]. Next up will be doing the same with Azure.

2.19.8.3.4 Week of 10/20/19

Azure is complete as well, read into same template as GCP. Started Mongo as well but ran into issue with YAML file for the MODE of the Mongo - ended the week…

2.19.8.3.5 Week of 10/27/19

Amazon info connected from Mongo, is now parsed into matrix. going to start working on Mongo refresh checks/downloads

2.19.8.3.6 Week of 11/03/19

Prep for demo at end of week with Gregor. Started to actually pull everything together. Flipped the frugal testing file into an base class with a list method. Now have separate frugal files for aws, gcp, and azure that return a price matrix for each one. Built in logic for it to save back to the db, as well as an argument for a refresh. Wrapping up integration for it to be called from the cloudmesh console. Console implementation completed. All that is left is frugal boot, which will require some checks

2.19.8.3.7 Week of 11/10/19

Was not able to make too much progress this week because of other assignments. Frugal boot now has a command line interface. For now it doesn't boot anything, but I worked on logic so that it filters the table to providers that can actually be booted. For now this is only AWS but it should work with Azure once I incorporate a few more things. Otherwise I am now looking at the cloudmesh code to figure out what information I need to boot a vm. Hope to have all of that done by the end of next week so I can begin to add tests.

2.19.8.3.8 Week of 11/17/19

Resolved an issue with not being able to boot up aws vms, which was hindering the frugal boot code. Now frugal boot works with aws, and the structure is there for it to work with azure with a few tweaks. Currently having a few issues with getting an azure vm up and running, but I'll try to figure that out during class time this week. After that I'll get Azure running and then start working on PyTests and documentation. Hope to have that all done by the end of the weekend to give time to check over things

2.20 Cloudmesh Virtual Directory Life Cycle Service

Bill Screen

2.20.1 Abstract

TBD

2.20.2 Objective

Integrate Cloudmesh with cloud service providers' object storage lifecycle management services to effectively manage storage costs throughout their lifecycle. This service is intended to be added to the Cloudmesh storage project https://github.com/cloudmesh/cloudmesh-storage Implementation of lifecycle policies will be provided for AWS and Azure cloud providers, and locally if time permits.

2.20.3 Motivation

As the volume of data being generated and stored by cloud service providers (CSP) continues to grow, the cost to store the data needs to be managed in a thoughtful manner. Without a data lifecycle management strategy, objects sent to a CSP could be stored indefinitely; even if the objects (data) are infrequently or never accessed, the customer will likely be charged for storage (beyond a certain 'free storage limit'). Thus applying a lifecycle storage policy will help to mitigate cost overruns by applying 'lifecycle rules' to efficiently manage objects stored in the cloud. For example, a lifecycle rule may schedule an object for deletion after 30 days or automatically move the object to a lower-cost storage tier after 90 days.

2.20.4 Architecture

TBD

2.20.5 Technologies

- cloudmesh storage
- OpenAPI 3.0.2
- REST
- MongoDB (if time permits)

```
Usage:

    storage_lifecycle_configuration [--storage_provider=<SERVICE>] put lifecycle_rules STORAGE_BUCKET_ID, LIFECYCLE_RULES
    storage_lifecycle_configuration [--storage_provider=<SERVICE>] delete lifecycle_rules STORAGE_BUCKET_ID, LIFECYCLE_RULES
    storage_lifecycle_configuration [--storage_provider=<SERVICE>] get_available_subresources STORAGE_BUCKET_ID
    storage_lifecycle_configuration [--storage_provider=<SERVICE>] load lifecycle_rules STORAGE_BUCKET_ID

Arguments:
    STORAGE_BUCKET_ID: Unique identifier of the storage bucket
    LIFECYCLE_RULES: An XML document used to specify the lifecycle of the STORAGE_ID. For example:

    <LifecycleConfiguration>
        <Rule>
            <ID>id2</ID>
            <Filter>
                <Prefix>my_logs/</Prefix>
            </Filter>
            <Status>Enabled</Status>
            <Expiration>
                <Days>365</Days>
            </Expiration>
        </Rule>
    </LifecycleConfiguration>

Options:
  -h --help
  --storage_provider=<SERVICE> Cloud storage provider service name (i.e. aws, azure, google, box)

Description:
    Manage Amazon S3 objects so that they are stored cost-effectively
    throughout their lifecycle.

    storage_lifecycle_configuration put lifecycle_rules
        Creates a new lifecycle configuration for the bucket (STORAGE_ID)
        or replaces an existing lifecycle configuration.

    storage_lifecycle_configuration delete lifecycle_rules
        Deletes the lifecycle configuration from the specified bucket
        (STORAGE_ID). Removes all the lifecycle configuration rules in the
        lifecycle subresource associated with the (STORAGE_ID).

    storage_lifecycle_configuration get_available_subresources
        Returns a list of all the available sub-resources for this Resource.

    storage_lifecycle_configuration load
        Returns the lifecycle configuration information set on the bucket.

Example:
    set storage_provider=aws|azure
    storage_lifecycle_configuration put STORAGE_BUCKET_ID, LIFECYCLE_RULES
```

2.20.6 Progress

Week of 10/06: After many struggles, finally got Cloudmesh and MongoDB to install on Windows 10 (call 'cms init' twice!). Automated the install process via a batch script. Also added logic to validate connectivity to chameleon Cloud by executing 'image' and 'flavor' list commands.

Week of 10/13: Reviewing the cloudmesh storage source code to understand the Python coding design pattern to follow for this project. Also trying to understand how to use VS Code debugger with CM source code, and how to setup pytests.

2.21 Mutiple source cloud based datawarehouse

deepak deopura fa19-516-168

O please leran markdown

O this project has significant issues as tehchnology such as snowflake is use that does not provide a free tier. The 30 day free licensis is insufficient. Please chose a different technology for testing. However you can in addition to some other artifact certaily use snowflake, but it can ot be your main objective. When reading your comment you actually do understand the issue …. Maybe you can use MariaDB or something like that. Also remember you need to technically compare streaming vs backu and upload. What we found n AWS is that streaming for data takes 45 min while backing up and reload 2 minutes …. however we have not spend any time optimiszing this. This is storing images from aws into our cloudmesh mongodb

2.21.1 Abstract

- AWS to/from Azure data transfer using APIs.
- Extended version can be push data in SQL base warehouse (for example snowflake warehouse. It may be out of scope for now for this project purpose.

2.21.2 Objective

Develop APIs to transfer data between AWS and Azure.

2.21.3 Motivation

Motivation of this project is from “Snowflake cloud warehouse” (https://www.snowflake.com) . I am using snowflake cloud base warehouse form some of my clients. Snowflake is currently providing most of their services in AWS. They are also increasing their footprint in Azure and will soon start on Google cloud. I would like to use this project opportunities when there are multiple clouds in an organization and they want to build warehouse based on various data sources across clouds.

2.21.4 Architecture

TBD

2.21.5 Technology

- AWS
- Azure
- Python
- Cloudmesh storage
- REST API

2.21.6 Testing

TBD

2.22 Cloudmesh Compute Project for Google Cloud Platform (GCP)

Harshawardhan Pandit, fa19-516-165

2.22.1 Abstract

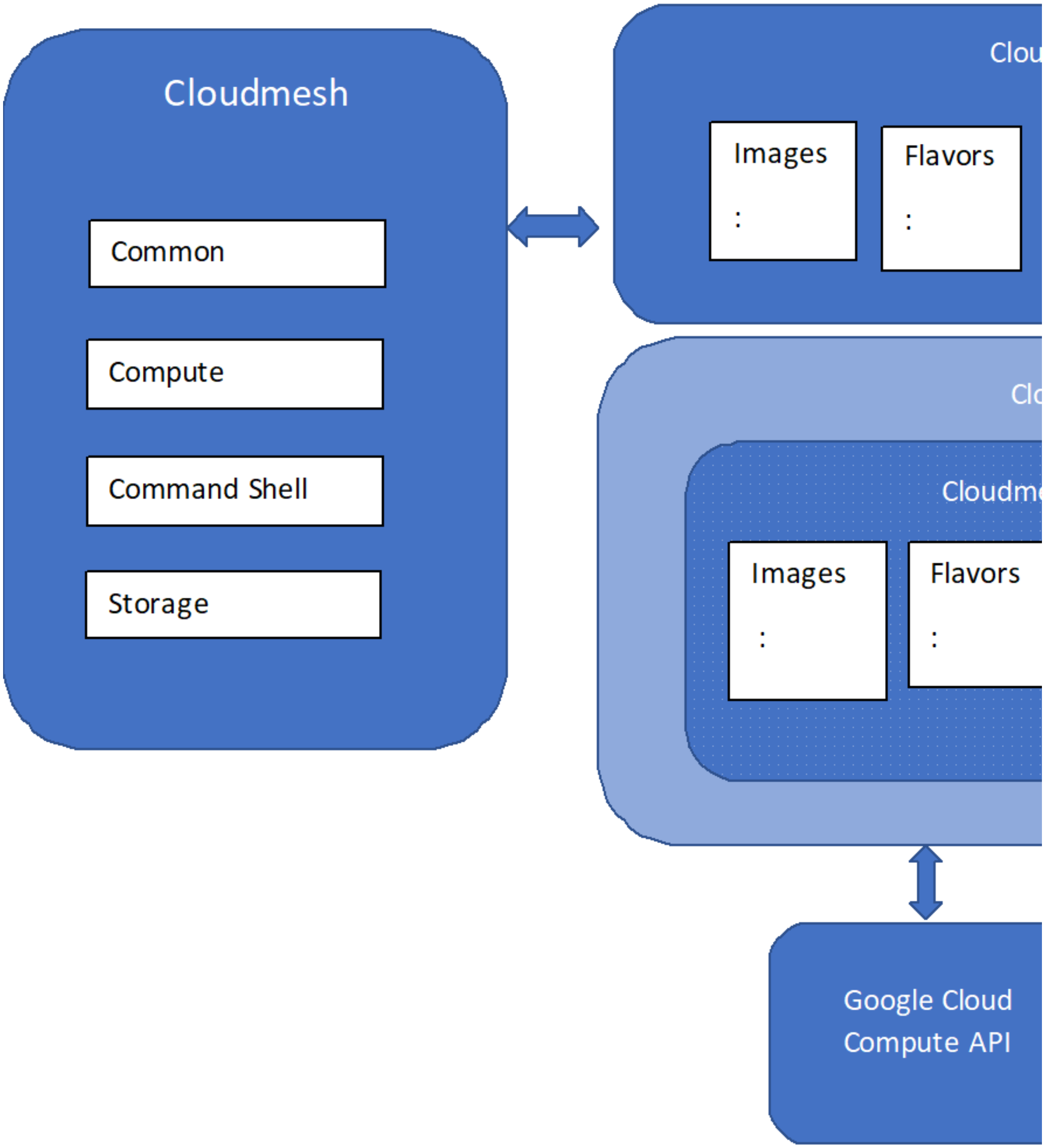
Cloudmesh enables you to access multi-cloud environments such as AWS, Azure, Google, and OpenStack Cloudsvery easily. The purpose of this project is to implement identified features for the **Google Cloud Platform**. The two cloud interface will be implemented using:

1. Google Cloud Platform
2. AWS

Selected APIs will be added for the following features:

1. Images
2. Flavors
3. Virtual machines
4. Keys
5. Security groups

2.22.2 Architecture



Architecture

2.22.3 APIs

Typical List of APIS to be developed may include in the follow categories:

2.22.3.1 VM

1. start

2. reboot
3. stop
4. resume
5. suspend
6. info
7. status
8. list
9. create
10. create_vm_parameters

2.22.3.2 Flavors

1. flavors
2. flavor

2.22.3.3 image

1. image

2.22.3.4 keys

1. keys
2. key_uplo
3. key_delete

2.22.3.5 Security groups

1. ssh
2. get_resource_group
3. set_server_metadata
4. delete_server_metadata
5. list_secgroups
6. list_secgroup_rules
7. add_secgroup
8. add_secgroup_rule
9. remove_secgroup
10. upoad_secgroup
11. add_rules_to_secgroup
12. remove_rules_from_secgroup

2.22.4 Technologies

- Python 3.7.4
- REST API
- Cloudmesh

2.22.5 Progress

2.22.5.1 Week of 22nd Sep

- Project Definition and approval

2.22.5.2 Week of 7th Oct

- Rest API ramp up

2.22.5.3 Week of 14th Oct

- Cloudmesh Installation on Windows. Solving the laptop issues and finally succeeded in Cloudmesh installation on Windows 10 environment.

2.22.5.4 Week of 21st Oct

- Study of existing Cloudmesh Compute Project/source code structure.
- Initial Architecture Diagram (In Progress).
- Listing down potential APIs for Cloudmesh Compute GCP Interface.
- Initial study of Compute Engine API from Google.
- Creation of WBS. Plan.

2.22.6 Work Breakdown Structure

This is intended WBS and schedule. This can change as we proceed as the proposed plan below and will evolve.

1. Week of 28th Oct: Architecture, Design finalization.
2. Week of 4th Nov:

- API signature definition and scope freeze (This should define how many APIs will be)
- Environment creation and Code Framework Creation (Skeleton, stub and dummy test cases)

1. Week of 11th Nov: Selected APIs for VM
2. Week of 18th Nov: Selected APIs for Image
3. Week of 25th Nov: Selected APIs for Security

4. Week of 2nd Dec:

    - Buffer
    - Documentation completion
    - Pending PyTest (if any) completion

2.22.7 Results

2.22.8 Benchmark

TBD Benchmarking will be added as the interfaces are defined

2.22.9 Testing

TBD Pytests will be added to the interfaces, as the APIs are developed.

2.22.10 References

1. https://cloudmesh.github.io/cloudmesh-manual/projects/project-compute.html
2. https://github.com/cloudmesh/cloudmesh-cloud/tree/master/cloudmesh/compute

2.23 Abstract Streaming interfaces for the NIST Big data Architecture

Jagadeesh Kandimalla, fa19-516-171 O

O Original title proposal was “Ingest Live Streaming Data/Replicating Database using Logs” whcih is too limited and does not showcase applicability on multiple clouds. However, this is easy to fix by developing a “register” like specification in OpenAPI just as we did for databases in the NIST document. YOu will have to look at 2 cloud providers to derive some common features. You could even describe how to much such dat between clouds. There are APIs and services that doe this. Provide a survey.

O I like the idea of the queue, but you need to move forward, with an architecture diagram and code as well as potential tests to do this

The streaming service initial architecure is when a Object gets added to S3 ,An AWS Lambda function will get triggered and will call the Google cloud endpoint which inturn will call the APP engine where we write the data to Google cloud Storage

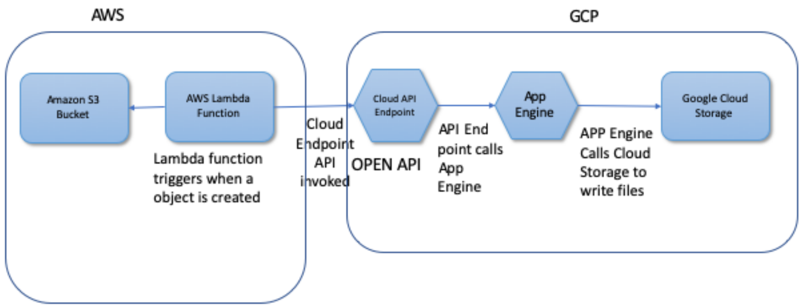
2.24 Objective

The streaming service initial architecure is when a Object gets added to S3 and S3 event will trigger AWS Lambda function and the lambda will call the Google cloud endpoint which inturn will call the APP engine where we write the data to Google cloud Storage.This way the streaming is automated ,anytime a object gets detected in S3 it is loaded to Google Storage and all the components are loosely coupled so that anytime we can replace the destination like Google Big query by just switching the app name in Google App Engine.

2.25 Technologies

AWS S3,
AWS Lambda,
OpenAPI 3.0.2,
GCP Cloudendpoint,
GCP App Engine,
GCP Coud Storage

2.26 Architecture



architecture

2.27 Progress

1.) Created a AWS account/S3 bucket - Done.
2.) Updated the .cloudmesh.yaml file with awss3 key pair - Done .
3.) Debugged the Cloudmesh-Storage awss3.provdier.py and StorageABC.py and got the put and list commands working for AWS S3- Done .
4.) Uploded the files to S3 files using Cloudmesh commands(PUT and LIST) – Done .
5.) Created the GCP account and set up gsutil on the mac - Done .
6.) Created the project and bucket and created the google application credentials - Done
7.) Download the Google app credentials and set the authentication in the environment variables - Done .
8.) Developed a flask application to read the data and load the data to google cloud storage -Done
9.) Debugged the flask application in the local using POSTMAN - Done .
10.) Extend the flask application to read from the URL and get the data downloaded from URL and load in to Google cloud storage - In Progress.
11.) Development of AWS Lambda application - Not Started .

2.28 Group Key Management for Cloudmesh

- Nayeem Baig
  - email: nayeemullahbaig.93@gmail.com
  - repo: fa19-516-172

please move this document into project dir for 172

- Contributors
- Forked Branch

2.28.1 Introduction

In the cloud we need to give access to VMs to multiple users.
The management of keys need to be automated and integrated with mongoDB.
Cloudmesh is missing functionality to easily add keys and control the access policies related to key management. Functionality to utilize mongo DB have

already been developed for the Security Rules and Security Groups functions.
We can add Key Groups that are defined by both the related cloud provider and
collection of related keys to fine tune access control for all connected machines.

After addressing the completion of this elected task the students ahs the opportunity to work on additional other cloud security aspects of his chosing if desired.

2.28.2 Implementation

2.28.2.0.1 Automating Key Management

2.28.3 References

O Missing

2.28.4 Tasks

2.28.4.1 CMS Key Command

Task Lead: Nayeem
Status: In Progress

Last Update: Implemented `cms key add --source=FILE_PATH`

2.28.4.2 CMS KeyGroup Command

Task Lead: Nayeem
Status: In Progress

Last Update: Added KeyGroup.py file based on SecGroup.py

2.28.5 Progress

- not reported

2.29 AI Services With Raspberry Pi

Sahithi Ancha, sancha@iu.edu, fa19-516-174

2.29.1 Abstract

Creating REST services using open API.

2.29.2 Introduction

We set up a Raspberry pi cluster and integrate Kubernetes. We then implement REST services using open API and deploy AI services on them.

2.29.3 Related Work

2.29.4 Architecture

2.29.5 Technologies used

- cm-burn
- cloudmesh-inventory
- Python
- REST

2.29.6 Progress

- Burning SD Cards using Etcher on Windows
- Setting up the cluster

2.29.7 Benchmark and Evaluation

- Develop a test program to review
- Use PyTest

2.29.8 Conclusion

2.29.9 References

- https://raspberrytips.com/install-raspbian-raspberry-pi/
- https://raspberrytips.com/raspberry-pi-cluster/
- https://dev.to/awwsmm/building-a-raspberry-pi-hadoop-spark-cluster-8b2
- https://dqydj.com/raspberry-pi-hadoop-cluster-apache-spark-yarn/

3 DATACENTER

3.1 Datacenter fa19-516-140

3.2 DataCenter fa19-516-141

- fix markdown

- use proper lists

3.2.1 E.DataCenter.2

- DataCenter: Digital Reality, SFO, CA.
- Year Build: 2004
- IT Load: 4800kw
- Cost per kwh in CA: 14.47cents ($0.1447)
- Cost per Year: 4800*0.1447*24*365: $6.08M
- Corbon Foot Print: 4800*24*365*16.44 : 691.269,120 pounds : 313,554 Tons
- Equivalent Cars: 66,572

3.2.2 E.DataCenter.3

My own carbon footprint is 15,952lbs of CO2

3.2.3 E.DataCenter.4

Wind Energy: Wind energy reduces the CO2 emissions.

a. Wind energy produced in US by end of Q2-2019 is nearly 97.96 gigawatts.
b. 30GW of energy is used by worldwide 8 Million dataCenters as of 2017.
c. Data Centers powered by Wind energy are listed below: c1) Other World Computing, Illinois - OWC may be the first one to use renewable energy in US. c2) Baryonyx Corp.(Texas) c3) Green House Data(Wyoming) c4) Microsoft Virtual Earth(Colorado)

3.2.4 E.DataCenter.5

a. InUS: Out of total energy produced in US, 17.1% from renewable energy sources. Among renewable energy, the main sources were 7%hydropower, 6.6% wind energy and 1.6% solar power.

b. Alabama: As per 2019 data from the below referenced website, Electricity from renewables: 9.2%(12.8 Million MWh) Total electricity generagtion: 140.0 Million(6th highest in US) Largest renewable energy source: Hydroelectric(9.2 Million MWh) Renewable energy growth in 10 years: 61.8%(18th lowest in US)

3.2.5 E.DataCenter.8

Data Centers effected by outages are appoximately: 33.33% in 2018 & 25% in 2017. Top three causes of down time are give below: 1)Power outages 33% 2)Network Failures are 30% 3)Software errors are 28%. As per the 80% of Data Center Managers, down times can be minimised with appropriate planning, good processes in place and robust design. According to IT Intelligence Consulting Report of 2018, average outage cost of $260,000/Hour and a five-minute outage costs $2,600.

3.2.6 References:

4.a) https://www.awea.org/wind-101/basics-of-wind-energy/wind-facts-at-a-glance 4.b) https://www.telehouse.com/2017/01/any-way-the-wind-blows/ 4.c) https://www.datacenterknowledge.com/wind-powered-data-centers 5.a) https://www.prescouter.com/2019/04/2018-was-a-record-year-for-renewable-energy-2019-could-be-the-same/ 5.b) https://www.usatoday.com/story/money/2019/07/26/renewable-energy-hydro-wind-solar-power-produced-by-each-state/39801879/ 8) https://www.datacenterknowledge.com/uptime/lessons-years-data-center-outages-focus-fundamentals

3.3 Datacenter fa19-516-143

3.3.1 E.Datacenter.2 Data center energy usage

- data center - University of Arkansas, ADSB Data Center

As per the Data Center Facilities Manager, Charles M. Dwyer:

- HPC systems usage = 239kva
- All other systems usage = 128kva

This does not include the required cooling.

"Unfortunately, the Data Center is not metered separately and is just included with the entire ADSB building usage. We don't have access to the building stats but this may be available from FAMA. Sorry, all this was setup back in the early 1980's and there is not a good way to measure or meter what our IT power consumption is." - Charles Dwyer

This DC is powered by coal through SWEPCO (as per David Chaffin of AHPCC) and produces ~1900lb CO2/MWh.

I don't have a power factor rating, so I chose 1 (most efficient use of power), which converts the HPC system usage to 239KWh (0.239MWH) and 454.1lb CO2 to run just the HPC, without cooling (which requires a LOT of power).

As a note, I started doing this for IU's data center - but saw in lecture that someone else had covered it. However, they're assumption that the two generators would be a good estimate is not accurate - the data center must have 100% redundant systems, so one generator could power the backup. However, backup is only for a subset of the systems and does not account for all of the power draw on a regular basis, making the calculated number inaccurate.

Also, the IU DC is getting second power in this year, doubling capacity. Meaning a generator twice the size of the current ones is being installed to maintain this redundancy. If I get more accurate numbers from the manager, I will report them, for the interest of including information in the course book.

3.3.2 E.Datacenter.3

Reported on webdoc - 36,817 lbs of Carbone (I fly a lot for work)

3.3.3 E.Datacenter.4 Wind (7 mod 6)

Data centers powered by wind come in two varieties - those that are powered by utility power that is sourced from wind generators and those that are powered directly by on-site wind generators (1).

Data centers currently generated by on-site wind generators include:

- Other World Computing in Woodstock, Illinois (2009) - first data center to be entiredly powered by wind in U.S. (1)

- Baryonyx Corporation in Stratford, Texas (2009) - powered by wind turbine leases on and off shore (1)

Data centers planned with on-site wind generators:

- Google in Land of Lakes, MN (announced 2019) - 300MW data center with two wind farms (2)

Data centers currently powered by utility power from wind generators:

- Green House Data in Cheyenne, Wyoming (2007) - 10,000 sqft data center run entirely by wind (1), with an average PUE of 1.25 or lower (3). Data center includes other optimizations such as hot/cold aisles, power and temperature tracking, renewable energy credits (3)
- Microsoft Virtual Earth Boulder, Colorado 2008 (1,4) - using three shipping containers placed near energy sources as a model for future efficiency (1, 4), the data center is "100 percent wind powered" through offsets (1, 4)

A continuation of Microsoft's concept is seen in Germany's WasfalenWind-Group, a company planning to put small data centers inside wind turbines (5). This project was announced in 2018, and is project aimed primarily at cloud services, using 62U server racks in each turbine, with 92% of the power coming from home turbine (5).

1. Wind-Powered Data Centers. (2010) Retrieved from https://www.datacenterknowledge.com/wind-powered-data-centers
2. Google said to be planning $600m wind powered data center in the land of lakes. (2019) Retrieved from: https://www.datacenterknowledge.com/google-alphabet/google-said-be-planning-600m-wind-powered-data-center-land-lakes
3. Green Data Centers. Retrieved from: https://www.greenhousedata.com/green-data-centers
4. Microsoft unveils wind-powered containers. (2008) Retrieved from: https://www.datacenterknowledge.com/archives/2008/04/18/microsoft-unveils-wind-powered-containers
5. WindCores project deploys small data centers inside wind turbines. (2018) Retrieved from: https://www.datacenterdynamics.com/news/windcores-project-deploys-small-data-centers-inside-wind-turbines/

3.3.4 E.Datacenter.5 Germany

Germany was the largest energy consumer in Europe and the 7th in the world in 2015 [1], but required imports [1] to drive it's coal and nuclear power structure [2]. The government set a goal to obtain 65% of its electricity from renewable power by 2030 [2] and 80% by 2050 [1], partly characterised by it's Energiewend project to phase out coal and nuclear [1], including the closing of all nuclear plants by 2022 [1].

These projects have been largely successful, with 30% of electricity coming from renewable sources in 2015 [1], 40.4% in 2018 [2], and 65% in 2019 [2]. Of this, wind was the largest contributing source (48.4%), followed by biomass (7.6%), solar (5.1%), and hydropower (3.5%) [2]. In contrast, power from lignite coal was down to 12% in 2019 from 24% in 2018 [2].

Germany is working on providing cloud data centers inside of individual wind turbines [3], which have 92% of their power coming from the host turbine (with the rest coming from redundant power sources and other turbines) [3].

3.3.5 E.Datacenter.8 Data Center Outage: Reddit outage 8/31/19

Amazon Web Services hosts the popular online forum Reddit (1). Unlike Twitter's outage on the same day (2), this outage was not related to high traffic on a U.S. holiday. Reddit was down in Northern Europe, parts of Asia, North America, and Southern Australia (map at http://www.bleepingcomputer.com/news/technology/reddit-experiencing-outage-due-to-amazon-aws-issue/). The outage lasted approximately six hours (1) and impacted seven Reddit companies - including desktop and mobile sites as well as comment processing (1). Downdectector.com reported 15,531 reports of the site being down within one hour and over 91,000 reports reported within the six hour outage (graph at https://downdetector.com/status/reddit/news/264740-problems-at-reddit-14). As reddit does not have a sales section like facebook, so financial impact to users was limited. However, Reddit is largely monitized by ads (4), but no value per hour was able to be found.

1. Reddit experienceing outage due to Amazon AWS Issue. (2019) Retrieved from: https://www.bleepingcomputer.com/news/technology/reddit-experiencing-outage-due-to-amazon-aws-issue/
2. Twiter Down, Thousands Reporting Outages. (2019) Retrieved from: https://popculture.com/trending/2019/08/31/twitter-down-thousands-reporting-outages/
3. Problems at Reddit. (2019) Retrieved from https://downdetector.com/status/reddit/news/264740-problems-at-reddit-14
4. How Reddit makes money. (2019) Retrieved from https://www.investopedia.com/articles/investing/093015/how-reddit-makes-money.asp

3.4 Datacenter fa19-516-144

3.4.1 E.Datacenter.2.a

```
Data Center: Condorcet Data center
Location: Paris, France
Year: 2009
Electricity Cost: 0.14
IT Load: 6400
Yearly Cost: 15,681,000
Yearly CO2 Footprint: 9307
Equivalence in cars: 2051
```

3.4.1.1 Note: PUE of Condorcet Data center

Electricity Cost PUE is assumed to be 2.0. No resource with a specific PUE for the Condorcet data center was discovered. Equinix (TelecityGroup) stated in reference three that the average annual PUE falls under category three it is assumed their efficiency is average. Thus 2.0 was the selected value.

This assumption doubles the total input power thus drastically increasing power cost in comparison to a simple multiplication of `electricity_cost*load*hours_in_year` Yearly cost, CO2 footprint, and equivalence of cars are calculated by the "Data Center Carbon Footprint Calculator".

3.4.1.2 References

1. http://94.247.169.249/wp-content/uploads/2014/11/cs_datacenter_2011_eng.pdf
2. https://www.equinix.com/data-centers/design/green-data-centers/pue-metrics/

3.4.2 E.Datacenter.2.b

Data has been placed within the file

3.4.3 E.Datacenter.3

Submitted information to URL.

3.4.4 E.Datacenter.4

Recyclers use the excess heat produced from the data center and reuses it in another productive fashion. The Condorcet Data Center utilizes the excessive heat from the data center to warm the arboretum throughout the year. This has the double positive environmental impact of utilizing the excess energy and preserving trees. The Energy Systems Integration Facility utilizes the wasted energy to heat the water and ventilation systems throughout the data center.

3.4.4.1 References

1. https://www.datacenterknowledge.com/data-centers-that-recycle-waste-heat
2. http://94.247.169.249/wp-content/uploads/2014/11/cs_datacenter_2011_eng.pdf
3. https://www.equinix.com/data-centers/design/green-data-centers/pue-metrics/
4. https://www.smithgroup.com/perspectives/2018/waste-heat-recovery-in-data-centers

3.4.5 E.Datacenter.5

Germany is a major contributor to global solar energy production. In 2017 Germany produced 36.1% of its energy from solar power. In 2016 Germany at one point met the energy demands of its nation with renewables. Two years later it achieved another acomplishment of meeting 100% of the energy demand with renewable technology. The extensive solar market within Germany was supported by subsidies for installation of solar units. As reported by Fortune the price of solar installation fell faster than anticipated which created a large incentive to quickly build and install solar power within Germany.

3.4.5.1 References

1. https://www.technologyreview.com/s/601514/germany-runs-up-against-the-limits-of-renewables/
2. https://www.cleanenergywire.org/news/renewables-cover-about-100-german-power-use-first-time-ever
3. https://fortune.com/2017/03/14/germany-renewable-clean-energy-solar/
4. https://www.cleanenergywire.org/factsheets/solar-power-germany-output-business-perspectives

3.4.6 E.Datacenter.8

During three days August of 2016 Delta Airlines Atlanta data center experienced a power outage in their Atlanta data center. This outage required Delta to ground all flights for three days. 451 total flights were cancelled due to the outage. This outage cost Delta over $150,000,000 due to the loss of potential tickets which excludes the number of travellers that were given cash and hotel compensation due to the cancelled flights.

3.4.6.1 References

1. https://www.datacenterknowledge.com/archives/2016/08/08/delta-data-center-outage-grounds-hundreds-flights
2. https://www.datacenterknowledge.com/archives/2016/09/08/delta-data-center-outage-cost-us-150m

3.5 Data center fa19-516-146

3.5.1 E.datacenter.2.b

3.5.1.1 The Dalles data center - Google

Located in the Dalles City, Oregon right off the Columbia river, Google's Dalles data center opened in 2006 and was the first Google designed and built itself.

As of the 2019 Quarter 2 performance report, the main facility has a PUE of 1.11.

According to baxtel.com the total power for the site was 70.0 megawatts. Other sources online stated that Google does not disclose details on energy consumption on individual data centers, but for the sake of this analysis, we will assume this data is correct.

The industrial electricity rate for the Dalles is 4.2¢/kWh. This is considerably lower than average rates in the United States and Oregon according to electricitylocal.com. Assuming Google pays this rate, that would mean this site costs about $70,560 per day to operate and $25,754,400 a year.

```
70,000 kW * 4.2¢/kWh * 24 hours = 70,560
```

```
70,560 * 365days = 25,754,400
```

Google claims that its data centers are carbon neutral. This is not done by running the data centers off renewable energy directly, rather it accomplishes this by purchasing carbon offsets.

In a 2018 report Oregon's net emission rate was `0.675 lb. CO2 per kWh`. If we assume since Google is probably pulling energy from this same system we can estimate carbon footprint without offsets to be `229,720.05 tons`

```
70,000kW * 1.11PUE * 0.6751bs. * 24 hours * 365 days = 459,440,1001bs.
```

```
459,440,1001bs ÷ 2000 = 229,720.05 tons
```

3.5.1.2 References

- https://www.google.com/about/datacenters/inside/locations/the-dalles/
- https://www.google.com/about/datacenters/efficiency/internal
- https://baxtel.com/data-center/google-the-dalles-oregon
- https://www.electricitylocal.com/states/oregon/the-dalles/
- https://static.googleusercontent.com/media/www.google.com/en//green/pdfs/google-carbon-offsets.pdf
- https://www.oregon.gov/energy/Get-Involved/rulemakingdocs/2018-03-21-CO2-RAC-Background.pdf

3.5.2 E.datacenter.3

My footprint is 12240lbs. of CO2 according the http://carbonfootprint.c2es.org. Data record to spreadsheet.

3.5.3 E.dataceneter.4

3.5.3.1 Hydro

Hydroelectricity is the process of using water power to generate electricity. This is accomplished by using dams to direct water to turbines that deliver power generators. Those generators convert that power to electricity.

The Dalles data center presumably pulls power from the Dalles Hydroelectric dam This dam is one of the largest in the United States and spans the Columbia River a few miles away for the Google data center.

3.5.3.2 References (E.dataceneter.4)

- https://www.nwp.usace.army.mil/The-Dalles/

3.5.4 E.dataceneter.5

3.5.4.1 Indiana renewable energy

Indiana produces a large amount of biofuels from corn and soy accounting for 8% of total production in the United States. Burning ethnol fuel still releases carbon, despite being a renewable source. It can be argued that since source of the carbon is from a natural part of the carbon could be considered neutral. However is not as clean a other renewable sources, such solar or wind.

This state is also 8th in the county when it comes to the number of windmills. This makes up about 5% of Indiana's electricty.

3.5.4.2 References(E.dataceneter.5)

- https://www.eia.gov/state/?sid=IN
- https://www.eia.gov/state/analysis.php?sid=IN
- https://www.wlfi.com/content/news/Indiana-sees-surge-in-wind-power-despite-lack-of-standards-559829491.html

3.5.5 E.datacenter.8

Google data center outage on June 2, 2019

On Sunday, June 2, 2019, a server configuration change caused a 4 hour outage for Google services in the eastern United States. According to the Inside Google Cloud article about the incident this resulted in some service being or not available for that region. Google reports YouTube by itself saw a 2.5% drop in views during this time. Other sources, such as an article from 9to5Google reports that YouTube's global views went down by 10%, which is much higher than Google. This would be a significant loss in ads revenue. And that is only one of a many services that were affected. Others reported not being able to control their own thermostat, since Nest relies google services affect in the outage. I imagine the loss in customer good will is as much, if not more of a concern.

3.5.5.1 References (E.datacenter.8)

- https://www.datacenterdynamics.com/news/google-suffers-sunday-outage-impacts-cloud-youtube-gmail-and-more/
- https://9to5google.com/2019/06/03/google-outage-cause/
- https://cloud.google.com/blog/topics/inside-google-cloud/an-update-on-sundays-service-disruption

3.6 Datacenter fa19-516-147

3.6.1 E.Datacenter.2a

Carbon footprint of a data center

- Data Center : Volkswagen
- Location : Rjukan , Norway
- Year Build : 2019
- IT load ($/kW): 2750
- Yearly CO2 Foot Print : 0
- Equivalent in car : 0

As per volkswagenag.com news(Rjukan/Wolfsburg, 2019-06-18), this data center is "100 percent powered by hydropower – resulting in a savings of more than 5,800 tons of CO2 per year"

3.6.1.1 References:

1. https://www.volkswagenag.com/en/news/2019/06/volkswagen_norway_data_center.html
2. https://www.cnet.com/roadshow/news/volkswagen-carbon-neutral-data-center-norway/
3. https://www.volkswagen-newsroom.com/en/press-releases/green-computing-performance-volkswagen-opens-carbon-neutral-data-center-in-norway-5090

3.6.2 E.Datacenter.2b

Data available online has been entered in this file .

3.6.3 E.Datacenter.3

My Carbon footprint is 11266 lbs.

3.6.4 E.Datacenter.4

Solar Energy

Renewable energy is the alternate form of energy which is from natural sources and filled up naturally e.g Solar Energy. Solar energy is light and heat from the sun which is naturally available . Solar energy is converted to electricity using solar panels which are installed in the open places where they can receive direct sunlight . Photovoltaic cells in the solar panel absorbs sunlight and create direct current (DC) which again gets converted to alternate current (AC). Two most common way of installing solar panels to generate solar energy are Solar Rooftop and Solar Farms. In Solar rooftop solution ,solar panels are installed on top of the house/building where direct sun light can be received while Solar farms are the open space of land where Solar panels are installed.

An example of data center with Solar renewable energy is Apple data center in Maiden NC.

3.6.4.1 References

1. https://en.wikipedia.org/wiki/Solar_energy
2. https://news.energysage.com/solar-panels-work/
3. https://fortune.com/2012/09/15/aerial-photo-apples-massive-solar-farm-in-maiden-n-c/

3.6.5 E.Datacenter.5

North Carolina State

North Carolina’s renewable energy consumption is growing fast , as per Arcadia Power blog on North Carolina renewable energy “In 2015, approximately 7% of North Carolina’s energy mix came from renewable sources such as hydroelectric, biomass, and solar. Though mainly producing hydroelectric energy, North Carolina ranks 3rd in the country in installed solar capacity.” This blog further stated that “North Carolina has already installed over 2,436 MW of solar energy, and is expected to more than double its capacity over the next 5 years by adding 3,656 MW of solar electric capacity. Wind power, on the other hand, just recently took off. In 2015, North Carolina started construction on the state’s first commercial-scale wind farm, which will be the largest wind farm in the Southeast to date.”

North Carolina’s energy conservation efforts have helped state rank 2nd position in installed solar power generating capacity which is over 4,400 megawatts.

3.6.5.1 References

1. https://www.eia.gov/state/?sid=NC
2. https://blog.arcadiapower.com/north-carolina-renewable-energy/

3.6.6 E.Datacenter.8

Prosa, Santa Fe, Mexico data center outage

August 2019 , there was a data center outage in Mexico which caused entire banking system to halt ,impacting several customers. The data center outage happened at Prosa, Santa Fe, Mexico. Prosa is an electronic transaction services firm which process Credit/Debit card payments in Mexico. Major banking systems of the country use this electronic payment service . The services were interrupted when an electrical fault in data center caused outage and shut off making any payment , purchase or ATM use. This outage remained for three hours and had impact across the Mexico where customers of major banks and supermarket were not able to make payment during the rush hours of weekend.

Details of revenue impact and customer count of this outage are not available online .

3.6.6.1 References

1. https://www.datacenterdynamics.com/news/banking-services-across-mexico-go-down-due-prosa-data-center-outage/
2. https://techerati.com/news-hub/data-centre-outage-mexico-banks-banking-prosa/

3.7 Datacenter fa19-516-150

3.7.1 E.Datacenter.2a - FrankFurt 3

The required values for the various attributes has been added to the file

3.7.2 E.Datacenter.2.b

FrankFurt 3 Data center is handled and managed by a company called e-shelter whose parent organization is NTT communications. It is the 3rd Data Center set up by e-shelter in FrankFurt after the FrankFurt 1 and FrankFurt 2. It encloses a total area about 12000 square meters(4000 in the first and approximately 8000 in the second phase of construction). FrankFurt 3 Data center was completed by the end of 2015. It utlises a electric Power capacity of 90000 Kilo Watts.[1]

The main advantage of FrankFurt 3 data center is that it uses a method of hybrid cooling instead of the traditional free coolin methods. This has contributed to a reduction in the cooling costs by one-third. Furthermore, since it makes use of the renewable energy sources it has a C02 and radioactive emission of 0mg/KWh. The cost of per KW/H is around 0.2$ [2]. Since the Data Center operates for 24 hours a day, it comes to around less Half a million per day and costs about 165 million per year. All these facts are as per the official website of e-shelter.

3.7.3 E.Datacenter.3

Currently reside in Bloomington(woodbridge apartments) in a town house Travelling around 40 miles a week (through campus bus shuttle) Do not have my own vehicle Will usually take one long flight, once a year My carbon foot-print turns out to be around 9165 lbs.

A carbon foot-print has been created and has been addded to the file file. Carbon foot-print has been calculated using [3].

3.7.4 E.Datacenter.4

others

Intrestingly I could find out a data center which runs or may be ran on Bio mass. Biomass can be described as a fuel which comes out when the so called waste plants and animals products are burned. These plants and animal products store the energy in the form sunlight and when it is burned chemical energy in it is released in the form of heat [4].It is one of the oldest forms of energy whose usage began several decades ago. The vineyard data center park, located downtown of the colorado spring managed to run a data center on bio mass.

It aimed at enivironmental friendly practices, with 100% usage of renewable energy, protection of surrounding wildlife, recycled building materials, etc. It’s area is around 80000 square feet with a power cost of around 4.4 cents/KWH and a PUE OF 1.21. It was estimated that the data center would be run on 50 MW Bio mass plant. [5] [6]

Vineyard Data center park was in great news in 2011 and it was estimated to be completed by 2014. The facts and figures given above are also mentioned in the official website(as per 2011) but the website appears to be completely outdated. There is a very little information of whether the vineyard data center park is active or not. It appears that it ran for very days and later it tanked.

3.7.5 E.Datacenter.5:India

I would like to discuss about India's effort towards renewable sources and the changes it has bought along with it. Renewable soucres refers to the process of utilising the natural sources such as sun, wind, water etc to generate energy from. Conversion of energy from one form to another form forms a significant part. For example heat energy can be obtained form electric energy and vice versa. The replacement of non renewable energy sources with renewable energy sources has both environmental as well as economic benefits. When it comes to a nation as a whole, this benefit is highly commendable.

India's renewable sector is the 4th attractive renewable energy market in the world. India's contibution towards the renewable sources is heading towards a siginificant increase as the time passes by. As per the stastics till April 2019, total renewable power installed in the country stands at 78.53 GW. There are various notable investements and developments in the Indian renewable energy sector. It has invested funds in many projects as well received investments from other countries and investors.

To name a few we have, 1MW of solar capacity every hour(2018), world's largest solar park named 'shakthi shithala' was launched with an investment of 16,500 crores.The consequence of using the renewable energy sources had a tremeduous impact on people. Besides, contributing to a decerease in global warming, there was a tremendous improvement in the public health and also created many employment oppurtunities. There are various initiatives taken by the indian government such as GO Green, hydropower policy,custom benefits to houses which has solar rooftop sector, etc.[7]

3.7.6 E.Datacenter.8: The Endurance Outage(2013)

On 2013 there was a failure in 2 network switches in a data center near Provo, Utah. Though on the outer apperance it seems that its effect would be very minimal, there was very huge consequence. The basis for all these was the concept of roll-up and convergence. This means convergence of infrastructure into few data centers. On elaborating, suppose that there are 3 different brands under a company X. All these brands are hosted under the same hosting platform, which means that failure at one point can cause the breakage of all the 3 brands at once, affecting the services of millions of customers. This is exactly the same which happened in Provo.

The company called endurance manages the hosting of many of the top companies. When the network switch failed it affected many of the brands including the 4 national brands namely BLUE HOST, HOST GATOR, HOST MONSTER and JUST HOST. The role of endurance here can given as analogy to the presently trending social media comapny Facebook. Facebook manages many of the other pages and even promotes the advertisments of many of the companies and organizations. This is also one of the sources for its high economy.

The problem with these companies is that they dont give an insight to the customers of what's actually going on i.e they aren't very transparent. When the incident happened, there was an attempt by endurance to resolve the customer queries through Twitter. Every single attempt to resolve the situation was reported.Endurance Outage had a significant impact and showed to its counterparts and to the world the need to address single points of failure.All these past and previous experiences is what the present experts and engineers in the data center consider to avoid a data center outage.[8]

3.7.7 References

[1] https://www.e-shelter.de/en/location/frankfurt-3-data-center

[2] https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-carbon-footprint-comparison-calculator/

[3] http://carbonfootprint.c2es.org/

[4] https://www.eia.gov/energyexplained/biomass/

[5] http://www.vineyarddatacenterpark.com/

[6] http://biomassmagazine.com/articles/5733/proposed-colorado-data-center-to-be-powered-by-biomass

[7] https://www.ibef.org/industry/renewable-energy.aspx

[8] https://www.datacenterknowledge.com/archives/2013/08/05/how-did-the-failure-of-network-switches-at-a-little-known-data-center-in-provo-utah-knock-four-major-services-and-millions-of-web-pages-offline

3.8 Datacenter fa19-516-151

3.8.1 E.Datacenter.2a

"1 megawatt (MW) of power consumed by a supercomputer today typically requires another 0.7 MW of cooling to offset the heat generated—and **each megawatt of electric power costs approximately US1 million per year.**The Japanese Earth Simulator, for example, ranked as the top supercomputer on the TOP500 list from 2002 to 2004, **consumed 12 MW of power, resulting in US$10 million per year** in operating costs just for powering and cooling." — Green Supercomputing Comes of Age

Given the information from the above source, using the carbon footprint calculator. The results are as follow:

IT Load (kW):	12000 KW
Yearly Cost ($):	10 Million
Yearly CO2 Footprint (tons):	10,958,760
CO2 equivalent in cars:	4,226,880

Reference:

http://seec.cs.vt.edu/pubs/papers/feng-itpro-green.pdf

https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-carbon-footprint-comparison-calculator/

3.8.2 E.Datacenter.2b

datacenter.md was submitted to https://github.com/cloudmesh-community/fa19-516-151

3.8.3 E.Datacenter.3

The carbon emission for a person who lives in a 1b1b, and only travels 14 miles a week by the city bus, in the mean time has no car is about 21550 lbs.

3.8.4 E.Datacenter.4

One of the renewable energy used by data centers is the wind energy. As the definition on wiki says, "Wind power or wind energy is the use of air flow through wind turbines to provide the mechanical power to turn electric generators and traditionally to do other work, like milling or pumping."

Now the containers of Microsoft located at Colorado, Microsoft Virtual Earth, are wind-powered.

Reference:

https://www.datacenterknowledge.com/archives/2008/04/18/microsoft-unveils-wind-powered-containers

https://en.wikipedia.org/wiki/Wind_power

3.8.5 E.Datacenter.5

China has already put a lot of effort to developed the renewable energy. So far China has been the world's leading country using renewable energy. For example, according to Dong, "In 2017, renewable energy comprised 36.6% of China's total installed electric power capacity, and 26.4% of total power generation, the vast majority from hydroelectric sources."

Besides, China is planning to reform the tax system and "transforming the pollution charge into tax" in order to control the carbon emission. Furthermore, this is not the first time China tries to solve the environmental issues. Currently China is subsidizing the new energy vehicles, so people who decide to buy electric cars will receive a decent discount. In the meantime, people who are going to buy a high emission cars will be heavily taxed due to the higher carbon emission, which will cause more air pollutions.

In conclusion, China is becoming more active on the climate or environmental issues, and has started to address those problems. The whole society therefore cannot have the equal amount of carbon emission permits as USA, e.g. people are encouraged to buy low-emission cars.

Reference:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2695612

https://www.brookings.edu/2018/05/18/utility-of-renewable-energy-in-chinas-low-carbon-transition

3.8.6 E.Datacenter.6

The air-flow cooling system will eventually reach its limit in volume. Instead, utilizing the liquid-cooling system would be a better option as the power-use of data center is dramatically increased. The liquid-cooling system is passive and it serves the extraction of the heat generated by the air flow.

The rear door heat exchanger, which using the technology, has an obvious advantages on reducing the operation cost due to its designed fixed parts that reduces the maintenance cost, in contrast to the computer room air conditioning units. Besides, This technology can even eliminate the chiller operation because its great performance at chilled water set-points.

References:

https://datacenters.lbl.gov/sites/all/files/rdhx-doe-femp.pdf

3.8.7 E.Datacenter.7

As the article from nature mentioned, "Smoother traffic flow and easier carpooling would reduce the total energy demand of the transport industry by 60%." Now the machine learning is used to help optimize traffic and reduce pollution. The research is led by the Berkeley Lab researchers, Alexandre Bayen. He believes that flow smoothing will be standard practice.

As the article from nature says, the power consumption of data center is becoming larger and larger. The total consumption in a year even higher than some small countries. It is worthwhile to address the power issue, because the it is becoming harder by boosting the computer performance. More computers are required to generate the computational power. There needs a structural change, that is, how the hardware is built.

To recap, to preventing the power consumption of data centers grow too fast than the system can control, measure should be taken to reduce the power usage.

Reference:

https://newscenter.lbl.gov/2018/10/28/machine-learning-to-help-optimize-traffic-and-reduce-pollution/

https://www.nature.com/articles/d41586-018-06610-y

3.8.8 E.Datacenter.8

Last week an outage occurred at Amazon AWS located in North Virginia. Even though the backup generator came to work, they failed after about two hours.

The outage "impacted 7.5% of EC2 instances and EBS volumes in the Availability Zone." The outage ended up with the data stored in the instances remained unrecoverable. A company of BleepingComputer whose services relies on the Amazon AWS was heavily affected so that their data became unrecoverable.

This outage reminds all of the users to invest in a secondary backup provider regardless of where data is stored.

References:

https://www.bleepingcomputer.com/news/technology/amazon-aws-outage-shows-data-in-the-cloud-is-not-always-safe/

3.9 Datacenter fa19-516-152

- [ ] use lists in refernces

3.9.1 E.Datacenter.2.b

Microsoft's Columbia data center at Quincy Washington is a 470,000 sqft. facility built in 2007. Located close to the Columbia River, it benefits from the hydro electric energy generated from the river. The location also has other advantages such as low land costs, abundant data fiber and low cost electricity.

Reference : Columbia Data Center Wiki

IT Load - In Microsoft Ignite presentation, Microsoft Azure CTO Mark Russinovich sampled out Columbia data center and mentioned it has a capacity of 64MW. He specified that there are two datacenters consisting four independent 8MW colocations totaling to a 64MW (64000kW) capacity, hence this value is chosen as the IT Load for the data center.

Reference: Microsoft Ignite Presentation

Cost per kWH - The Grant PUD county rate for large industrial consumers consuming greater than 15MW is $0.03/kwH but according to a portofquincy.org web article, Microsoft is receiving a rate of $0.019 hence this cost is being used.

Reference:

- Port of Quincy Blog post
- Rates for GrantPUD

Electricity Cost ($/kW) - The yearly cost is calculated with power consumption rate of 64MW at the cost of $0.019 assuming that the data center is 100% available in a year.

64000 * 24 * 0.019 * 365 = $10,652,160

Yearly CO2 Footprint (tons) and CO2 equivalent in cars - Using the Carbon emission calculator the CO2 emission value and C02 equivalent in cars is derived. The average PUE Microsoft's data center is 1.125 is used here and cost is 0.019/kWH is used for this calculation.

CO2 Emission - 62250.72 Tons

CO2 equivalent in cars - 13724

Tool : Schneider electric Carbon Emission Calculation link

Reference - Microsoft PUE

3.9.2 E.Datacenter.4

Hydro Energy

- Britannica definition states that "Hydroelectric power, electricity produced from generators driven by turbines that convert the potential energy of falling or fast-flowing water into mechanical energy."
- Hydro energy is one of the oldest sources for generating mechanical and electrical energy. This was used years ago to turn paddle wheels to grind grains. Hydro electricity is largely produced at major rivers or large dams.
- In Hydro power plants, Water conserved in a reservoir gains potential energy just before it flows downhill. As water gets released, the potential energy is converted to kinetic energy. This is used to turn the blades of a turbine in order to generate electricity which is in turn distributed to power plants.
- Hydro power is environment friendly compared to fossil fuels, coal or oil.
- A Hydro power plant can be used a peaking power plant when demand increases. When needed, the water stored in reservoir can be released and energy can be produced quickly.
- Top providers of hydroelectricity include China, United States, Canada, Brazil and India. Of all renewable energy, 71% of renewable electricity is from Hydro energy.

Reference -

- National Geographic Reference
- Hydropower Reference from eia.gov
- Britannica Reference
- Wikipedia Reference

3.9.3 E.Datacenter.5

Germany

- Germany's major renewable sources are from Wind energy, Solar energy and Biomass.
- In April 2019, Germany's renewable energy accounted for 77% of Germany's net public power supply with the advantage from strong winds and abundant sunshine. On the whole, Wind power provided 40 percent, solar 20 percent, and biomass 10 percent of renewable energy.
- According to wikipedia, The German energy policy is framed within the European Union which has a mandatory energy plan that requires a 20% reduction of CO2 emissions before year 2020 and the consumption of renewable energies to be 20% of total EU consumption.
- Germany has one of largest wind turbine in the world, and its wind farm operates in averages of 10,500MWh a year. Through such efforts, 113.35TWh of electricity in generated in Germany in 2018.

- Germany's renewable energy innovations have opened up an entirely new job sector based on technology development, production, installation and maintenance.
- A post from iass-potsdam states that "An additional aspect of local value creation is linked to the emergence of citizens as renewable energy producers and energy providers. An estimated 47% of the overall installed renewable energy capacity in Germany as of 2013 — adding up to an installed capacity of 33.5 GW — is in the hand of citizens, mainly through privately owned solar rooftop systems and citizens' wind farm cooperatives. Those projects provide approximately 1.6 million Germans with additional income or reduced spending for external electricity."

Reference:

- Wikipedia Reference
- Clean Energy Wire Reference
- Iaas-potsdam Reference

3.9.4 E.Datacenter.8

Google Cloud Data Center Outage

- Google Cloud Data Center encountered an outage on June 2nd 2019 12.53 PST. This outage caused slow performance and increased error rates on several Google services, including Google Cloud Platform, YouTube, Gmail, Google Drive and others.
- Google Cloud status mentions that "Multiple US regions noticed a elevated packet loss due to a network congestion for duration between 3 hours 19 minutes - 4 hours 25minutes."
- The cause was found to be due to an incorrect configuration change that was applied to a larger number of servers in several regions but originally intended for small group of servers in a single region. Due to this, the impacted regions were running 50% of their available network capacity.
- The impacted regions were "us-central1, us-east1, us-east4, us-west2, northamerica-northeast1, and southamerica-east1."
- From a user impact perspective, Google Cloud's blog states that "YouTube measured a 2.5% drop of views for one hour, while Google Cloud Storage measured a 30% reduction in traffic." Around 1% of active Gmail users, which would represent a million users couldn't send or receive emails.
- Services that were impacted include Cloud Endpoint, Cloud Interconnect, Cloud VPN, Cloud Pub /Sub, Cloud Spanner and Cloud Storage.

Reference:

- Google Status Reference
- Google Blog Reference

3.10 Datacenter fa19-516-153

3.10.1 E.Datacenter.2.a

3.10.1.1 NSA Bumblehive - Bluffdale, UT

Electricity Cost	IT Load	Yearly Cost ($)	Yearly CO2 Footprint (tons)	CO2 equivalent in cars
$0.071/kWh	65,000 kWh	$40,000,000	477,360 metric tons	103,773 cars

Given values of 561,600,000 kWh per year, avg 1.7 lbs/kWh

3.10.2 E.Datacenter.2.b

The NSA Bumblehive datacenter, otherwise known as the Utah Data Center, is a large government data center used by the NSA to intercept and analyze up to 100 years of communications data. The data center performs scientific computing instead of on-demand computing compared to public data centers, which allows it to operate at peak performance all of the time, making it much more efficient. The total operating cost is upwards of $40 million a year, in addition to the cost of 1.5 million gallons of water a day being pumped in to cool the servers. This data center has the 8th highest power consumption of data centers worldwide. The proposal and development of the data center faced immediate backlash in 2013/2014, due to growing economic and environmental concerns about the value of the data center being used for “spyware”. Many of the actual details and blueprints regarding the development are classified, but various estimators and experts have guessed that the facilty can hold/process anywhere between 3-12 exabytes of data. The data center campus as a whole takes up 1m sq. ft. of space, however the server structure takes up only 10% of that space.

References

NSA Bumblehive Energy Consumption Statistics
State of Utah Energy Consumption/Emissions Statistics (table 7)
NSA Utah development information

3.10.3 E.Datacenter.3

25,685 lbs, (“5,183 lbs. of your carbon footprint comes from home energy use, 20,502 lbs. comes from transportation.”)

3.10.4 E.Datacenter.4

3.10.4.1 Thermal Energy

Geothermal energy is used pretty rarely in data centers, but many data centers that use geothermal energy end up with a LEED (Leadership in Energy and Environmental Design) platinum certification.

An example of a unique geothermal use is the American College Testing in Iowa, which doesn’t use geothermal energy, but instead thermal cooling in order to remove waste heat from the servers. Many sources involved with the project identified it as one designed for sustainability.

Another example of this type of geothermal cooling is present in Prairie Bunkers, a 760-acre data center in Nebraska with geo-thermal cooling. It, however is especially unique as it is one of the first projects designed in a nuke-proof data bunkers.

References

Geothermal Data Centers

3.10.5 E.Datacenter.5

3.10.5.1 Google’s Renewable Energy Efforts

Google is making a large effort to introduce and upgrade their renewable energy sources in order to cut their environmental impacts. They have invested upwards of $3 billion in their renewable energy projects and plan to grow the industry in a renewable manner. One particular way that this has been implemented is using a PPA (Power Purchase Agreement) with clean energy producers in the region of their data centers. Google’s successes with this project include their milestone in 2017, reaching 100% renewable energy to match their consumption by data centers.

References
[Google Data Centers Renewable Energy Report]

3.10.6 E.Datacenter.8

3.10.6.1 Azure Data Center Outage

In 2018, a lightning strike caused over 40 data center services to go down in the South-Central region of the US. Luckily, this outage happened about a year after Azure implemented their multi-zone availability strategy that Google and AWS had implemented much earlier. This allowed the services to replicate themselves for protection for an instance like this. This still affected an undisclosed number of customers, as well as several Office 365 cloud-based services, which caused it to impact a mixture of corporations, organizations, and home users. One of the largest impacts of this outage was Azure DevOps, lasting over a day. The outage demonstrated some large failures with their system, as they didn't implement seamless failovers or backups.

References

Azure Data Center Outage

3.11 Datacenter fa19-516-155

3.11.1 E.Datacenter.2.a: Carbon footprint of data centers

- Table updated at : 'https://docs.google.com/spreadsheets/d/1gh869zfjA4sVxL8-ga0af2_HLTTuOoD1IReuRSrbq4I/edit#gid=0'

3.11.2 E.Datacenter.2.b: Carbon footprint of data centers

- Data center : Prineville, OR
- Organization : Facebook
- Location : Prineville, OR
- Year Build : 2011

Facebook targets 100% usage of renewable energy by 2020. In 2018, this percentage was 75%. Following are the details of 2018 data:

Total energy consumption of the data center was 488,000 MWh in the year 2018, which translates to 55707.76 kW [4].

Using following formula the 'IT load' comes out to be:

- PUE = Total facility energy / IT equipment energy
- 1.11 = 55707.76 / IT equipment energy
- IT equipment energy = 50,187.17 kW

Electricity cost in Prineville, OR is $ 0.09/kWh. Maximum available value of the IT load in the Schneider carbon foot print calculator is 1/5th of the above load. Hence using 10,000 kW as the IT load and then multiplying the results by 5 we get following results -

Attribute	Value
Electricity Cost ($/kW)	0.09

Attribute	Value
IT Load (kW)	50,187.17 kW
Yearly Cost ($)	43.35 M
Yearly CO2 Footprint(tons)	58,394
CO2 equivalent in cars	12,873

Facebook has provided an interesting visualization which shows variation in PUE of this data center in an interactive manner [5]:

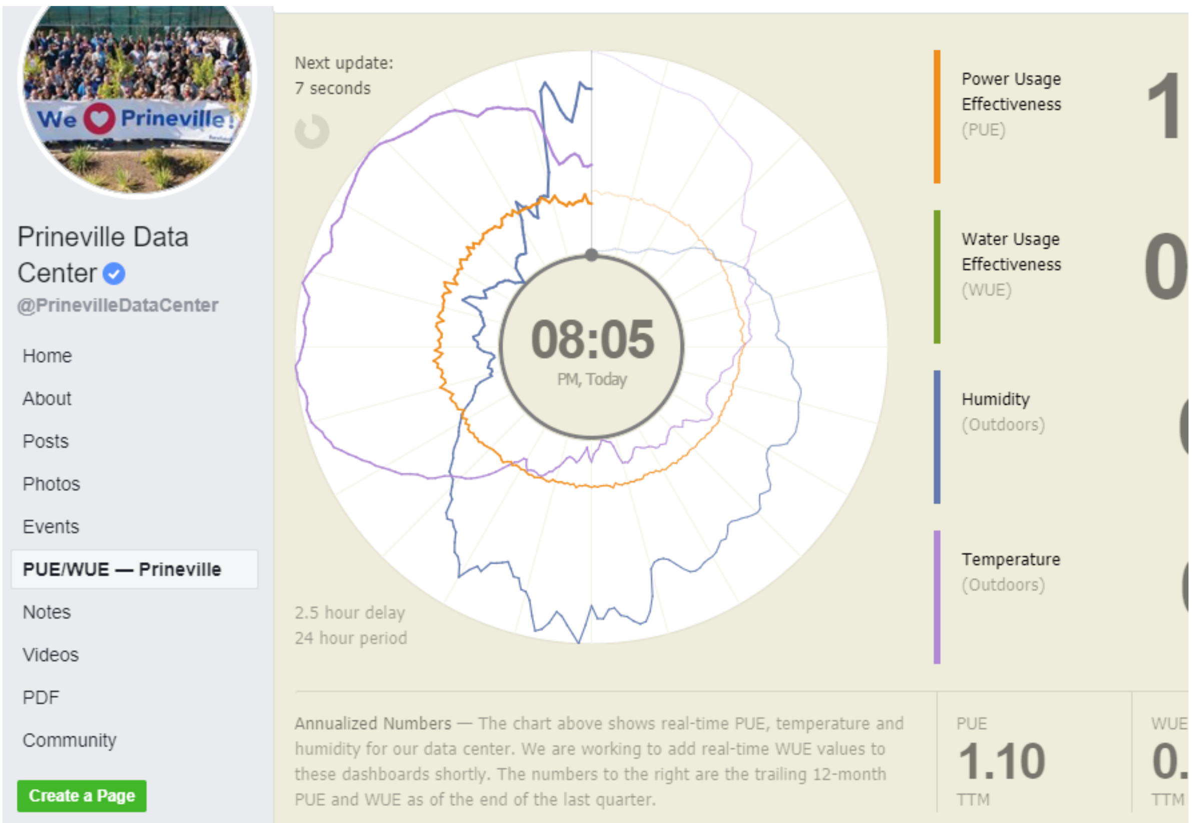


Figure 1: Facebook Prineville Datacenter [5]

3.11.2.1 References

1. [6] https://sustainability.fb.com/sustainability-in-numbers/
2. [7] https://sustainability.fb.com/wp-content/uploads/2019/08/2018-Sustainability-Data-Disclosure.pdf
3. [8] https://www.facebook.com/PrinevilleDataCenter/app/399244020173259/
4. [9] https://sustainability.fb.com/sustainability-in-numbers/#section-GreenhouseGasEmissions

3.11.3 E.Datacenter.3: Your own Carbon footprint

My Footprint `11927 lbs of CO2`

3.11.4 E.Datacenter.4: Thermal energy usage at Data Centers

- Thermal energy the energy available in the form of heat. It is utilized by convection, conduction or radiation. Examples of thermal energies are solar thermal energy, geothermal energy.

- Thermal energy can be converted in electricity and used for data center usage. Another way of using thermal energy is by recycling the head generated by data center and using it for various purposes such as heating residential or industrial areas.
- Thermal energy made available by data centers in this way, indirectly reduce usage of the conventional fuels for heating purposes.
- An interesting example of such reuse of thermal energy is observed at the **Facebook datacenter of Odense, Denmark** [10]. This data center has deployed waste heat recovery system that captures the excess heat generated by data centers and then distributes it to the local community using a heat pump.
- This is thermal energy is used for heating of local houses reducing dependency on conventional energy sources. This also helps to reduce greenhouse gas emission.
- Heat captured at the data center is conducted to local houses through district heating network.
- Best part is that this data center itself is also powered by renwable wind energy.

3.11.4.1 References

1. [11] https://sustainability.fb.com/innovation-for-our-world/sustainable-data-centers/

3.11.5 E.Datacenter.5: Efforts towards renewable energy

- Continuing the example of **Facebook datacenter of Odense, Denmark** [10], this center uses wind energy instead of conventional energy sources demonstrating Facebook's efforts toward renewable energy.
- In addition to this, this data center also employs a excess heat recovery system which captures and recycles the heat generated by data center cooling process.
- This captured thermal energy is transferred to district heating network where it is put to good use by local houses.
- Denamrk has a national goal to phase out coal usage by 2030. By using the thermal energy generated by data center Odense city hopes to achieve this goal by 2025 itself.
- Once fully functional, this data center aims to save 100,000 MWh of energy per year.

3.11.6 E.Datacenter.8: Data center outage

This is a summary of a data center outage effecting AWS3 [12] [13].

- Data center outages occur due to reasons ranging for power supply failures to natural calamities. Such outages disable one or many services provided by the data center causing unavailability of services relied on such data center.
- One such curious case occurred on Feb 28th, 2017 at AWS US-East-1 data center in Ashburn, Virginia.
- This outage mainly **affected S3 service on US-East-1** region of Amazon Web Services.
- Affected AWS S3 service was in use by oranganizations such as Quora, Coursera, Docker, Medium. As these are all customer facing services, the impact of this outage was directly felt by the end users.
- This is a peculiar case of data center outage because it was in fact **cuase**ed by **human error**. As later investigated by AWS, it was found out that while analizing an issue in the billing system, the technical team brough down unexpected number of servers. Which led to unavailability of varios other services. A complete restart was required in order to restore all these services. This increased the down time of dependant services.
- Details of the impact of this outage in dollars was not found.
- Since then Amazon has implemented multiple methods **to avoid** this scenario in future. Few of such methos include greater control on human errors, improvement of the avaialbility monitoring system and setting up a threshold value below which the technical team won't be able to bring down any server capacity.

3.12 Datacenter fa19-516-156

O use lists

3.12.1 E.Datacenter.2.b - Cost of Data Center

The Facebook/Fort worth Data center is mainly using Wind energy (details can be find in .5 section below) and I could find ceratain data from usage of 2018 yearly report from Facebook page and not able to determine all the facts at this point from this center. Also I have reached Facebook/Datacenter through their facebook page and hope they may give some specifics within in couple days. Here is my finding so far from Facebook sustainibility report, Ref sustainibility web page / DOWNLOAD OUR ENERGY, GHG, AND WATER DATA

Facebook Datacenter, Fort worth, TX . Established 2015-16. Electricity Use (2018 yearly) : 461,000 MWh . Operational Greenhouse Gas Emissions : 1,000 (market-based Scope 1 & 2 metric tons CO2e)

CO2 equivalent in cars: 212 cars 212 cars = (1000 tons/yr) / (4.71 metric tons CO2E/vehicle /year) PUE : 1.11 (avaerage)

Surprisingly this Data center Power usage Effectiveness is very efficient, Ref 42u webpage

3.12.1.1 Reference :

https://www.42u.com/measurement/pue-dcie.htm https://sustainability.fb.com/sustainability-in-numbers/ https://sustainability.fb.com/wp-content/uploads/2019/08/2018-Sustainability-Data-Disclosure.pdf

3.12.2 E.Datacenter.4 - Solar

Solar energy refers to capturing the energy from the Sun and subsequently converting it into electricity. “Solar energy is radiant light and heat from the Sun that is harnessed using a range of ever-evolving technologies such as solar heating, photovoltaics, solar thermal energy, solar architecture, molten salt power plants and artificial photosynthesis” Wikipedia

Skybox Data center in Houston, TX implemented Solar renewable energy through NRG Energy provide with 1.34 PUE which is consider as efficient according to PUE calculator mention by Cloudcomputing book by Gregor Von Lazewaski 4.1.5.4, as per blog from skyxbox.xom “The Skybox Houston One facility is a 96,129 square-foot purpose-built datacenter located in the fast-growing Energy Corridor of Houston, Texas. Skybox Datacenters worked closely with NRG to develop and accommodate each unique requirement – including a 1.34 PUE rating at full load, highlighting design efficiency in the Houston climate – and delivered a bespoke solution for their customized data hall. In addition to collaborating on the project, NRG also signed a long-term lease for a unique private data hall in the datacenter and has secured a multiyear wind electricity contract to support Skybox’s efforts to operate a 100% green energy datacenter”

3.12.2.1 References:

https://www.skyboxdatacenters.com/skyblog/skybox-datacenters-push-energy-efficiency-with-houston-first https://cloudmesh-community.github.io/book/vonLaszewski-cloud.pdf https://en.wikipedia.org/wiki/Solar_energy

3.12.3 E.Datacenter.5: - Facebook, Fort Worth, TX

In August 15th Facebook announced it operating 100% wind energy in Fort worth Data center location.Ref : Facebook/Fortowrth Data center, August 15 date page

In September 3rd 2019 Facebook released news to media about 200MW Wind energy deal with Apex for operating renewable energy in Fortworth Coke county location and to achive 100% in 2020. “Facebook Signs 200MW Wind Deal With Apex Clean Energy in Texas” Greentechmedia,P1.

The study from RTI.org webpage finds Facebook’s four operational domestic data centers, including Fort Worth Data Center RTI.org,P1 “Facebook data centers have contributed a cumulative $5.8 billion in gross domestic product (GDP) to the U.S.economy from 2010-2016, or $835 million per year" also "RTI estimates that, from 2011-2016, Facebook avoided over 2.5 million MWh of carbon-intensive electricity consumption through energy efficiency and renewable energy investments. This resulted in CO2 emissions reductions of over 1.2 million metric tons."

Also cosidering community contribution as from Facebook data center hom page "Facebook is committed to investing in the long-term vitality of Fort Worth and its residents by supporting groups that meet the community needs. In that spirit, Facebook's Fort Worth Data Center will be awarding competitive grants for projects and/or organizations based in the City of Fort Worth or within 10 miles of the data center." Ref, Facebook/notes

3.12.3.1 Reference:

<https://www.greentechmedia.com/articles/read/facebook-signs-200mw-wind-deal-with-apex-clean-energy-in-texas?fbclid=IwAR20r9OW4zbB0npewQ3Hx687CWq_mj8DRRc810BU2MghEZBCzQpltlnQ39Y#gs.0nrp50 > https://www.facebook.com/FortWorthDataCenter/ <https://www.facebook.com/notes/fort-worth-data-center/2019-facebook-fort-worth-community-action-grant-application-is-now-available/1980545985372461/ > https://www.rti.org/sites/default/files/facebook_data_centers_2018.pdf?fbclid=IwAR2SyuzMv-IRPq7kYjFPf7LNdRySMZS6oJNpRPdUHLow0YR0YI5bW0vI4HI

3.12.4 E.Datacenter.8: Major Data Center Outage

Delta Data Center Outage

I have did my research on internet and compared couple Data center outages, found Delta airline Data center outage is significant because it made direct impact on daily users life as people lost their flight connections and other plans about work and trips are getting delayed. In this Data Center outage more than 1000 flights got cancelled and as per study by the Ponemon Institute the "average cost of a single data center outage today is about $730,000" in 2016 .Ref, Datacenterknowledge web page

"The Delta Airlines data center outage that grounded about 2,000 flights over the span of three days in August cost the company $150 million, the airline's representatives told the audience of a transportation industry conference in Boston Wednesday. The number is extraordinarily high, illustrating that major airlines have a lot more at stake when designing and managing critical infrastructure than most other data center operators." Ref, Datacenterknowledge webpage

Another details reported provide by five9sdigital as "Delta Air Lines says the total bill for its devastating computer outage will come to $150 million. The problem occurred when the company lost power at its data center in Atlanta early on the morning of Aug. 8, causing computers needed to book in passengers and fly jets to be down for nearly five hours. The airline eventually canceled about 1,000 flights on the day of the outage and ground an additional 1,000 flights over the following two days. It also agreed to give affected customers refunds and vouchers for future travel. The cost of the outage was disclosed in a presentation Delta made to investors Wednesday. The losses came out of pre-tax profits, but the airline did not provide a break down of the various costs." Ref five9sdigital web page

3.12.4.1 References:

https://www.datacenterknowledge.com https://five9sdigital.com/knowledge/delta-airlines-our-5-hour-data-center-outage-cost-us-usd150m/

PS: Fixed list using Pycharm

3.13 DATACENTER FA19-516-157

3.13.1 E.DataCenter.2

The Apple data center in Maiden, North Carolina announced its construction in 2009. Although there is no information online stating the exact year that this data center went into operation, there are plenty information about the year it started construction.

There are two sources online WorldTopDataCenters and Baxtel that have different information about this data center's energy consumption. WorldTopDataCenters states that this facility has IT load of 20,000kW and Baxtel states that this facility has IT load of 100,000kW. The report from WorldTopDataCenters is from 2014 and Apple has announced its plan of expanding the Maiden facility since then, though the publish date of Baxtel's report is unknown, there is good chance that Baxtel's report is after the facility's expansion.

According to Power usage effectiveness in data centers: Overloaded and underachieving from The Electricity Journal. Apple's Maiden datacenter has PUE of 1.1 which is quite efficient. The local electricity cost in Maiden, North Carolina is $0.08/kWh. Calculation shows that Apple's Maiden data center has yearly cost of 81.8M dollars and 542,510 tons of yearly CO2 footprint which is equal to 119,600 cars.

3.13.1.1 Reference

- http://worldstopdatacenters.com/apple-maiden-north-carolina/
- https://baxtel.com/data-center/apple-maiden-north-carolina
- https://cedmcenter.org/wp-content/uploads/2017/10/Power-usage-effectiveness-in-data-centers-overloaded-and-underachieving.pdf
- https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-carbon-footprint-comparison-calculator/

3.13.2 E.Datacenter.3

My carbon footprint is 28190 lbs, added to link

3.13.3 E.Datacenter.4

Wind

Wind energy is harvested from wind turbines, the gigantic turbines are oftern planted in areas such as plains that have winds throughout the year. The wind will push the propellars which then turns the electric generator to generate electricity. It is a form of converting mechanical power to electricity.

According to Savvas' article Wind Turbines Set To Deliver The Edge Data Centre Needs Of The Future? Green IT- Das Systemhaus has a data center called WindCORES in Paderborn, Germany and 92% of the data center's energy comes from wind turbines yet the facility is still able to offer service with reliability level of 99.98%.

3.13.3.1 Reference

- https://www.awea.org/wind-101/basics-of-wind-energy
- https://data-economy.com/wind-turbines-set-to-deliver-the-edge-data-centre-needs-of-the-future/

3.13.4 E.Datacenter.5

According to this Article from South China Morning Post. China has aimed for renewable energy to account for at least 35 percent of electricity consumption by 2030. The government also required non-compliant firms to pay fines to energy providers, the fines are used to pay for the government subsidy to those energy providers. The government also required some provinces to increase their target of non-hydro power consumption.

3.13.4.1 Reference

- https://www.scmp.com/news/china/politics/article/2165831/china-steps-green-energy-push-revised-renewable-target-35-2030

3.13.5 E.Datacenter.6

The rear door heat exchangers are great alternatives to using chillers in data centers. According to Data Center RackCooling with Rear-doorHeat Exchanger from U.S. Department of Energy, the circulating water within the device will absorb the heat from server fans before it is discharged into the data center. According to the same article, the device will increase reliability and reduce cost since it is a passive device meaning it has no moving parts.

3.13.5.1 Reference

- https://datacenters.lbl.gov/sites/all/files/rdhx-doe-femp.pdf

3.13.6 E.Datacenter.8

In May 27 2017, British Airways suffered an outage at a data center near Heathrow airpot in London. According to Patrizio from NetworkWorld, more than 400 flights canceld and 75,000 passengers stranded in one day because of a simple human error. An engineer in the data center disconnected a power supply in the data center, when the power was reconnected, it caused a major power surge in the data center.

3.13.6.1 Reference

- https://www.networkworld.com/article/3200105/british-airways-outage-like-most-data-center-outages-was-caused-by-humans.html

3.14 Datacenter fa19-516-158

- [ ] fix urls

3.14.1 E.Datacenter.2.a

The Carbon Emission for "Inner Mongolia Information Park" is 0 because China has moved into reduction of CO2. This is to ensure eco friendly environment. Since this data centre is being built, it plans on reducing all emissions.

3.14.2 E.Datacenter.2.b

Abundant power supply, favorable climatic conditions, availability of skilled staff and preferential policies are major reasons why the data centre is popular. Inner Mongolia Information Park Data Centre is located in Inner Mongolia, China. It is owned by China Telecom. This is the largest data centre in Hohhot, the capital of Inner Mongolia.

It was established in 2016. "In 2016, the region designated the big data and cloud computing industries as new engines for local development, vowing that the regional big data industry's output value will exceed 100 billion yuan (14.5 billion U. S. dollars) in 2020." [1]

The IT Load is 150MW = 150 * 1000 = 150,000kW. The data centre is operated for 24 hours. So, the electricity cost would be 150,000 * 24 = $241,200,000. The yearly cost projected is around $14.5billion+ by 2020. [5]

Since the yearly CO2 footprint is being minimized every year, it is hard to estimate the exact figure but it is nearly 6 tons. [2] [3]

3.14.3 E.Datacenter.3

- My residential home in Bloomington is in an apartment building with more than 6+ units.
- It is an air-conditioned home with electricity usage used to power the fuel.
- I do not own a vehicle.
- I travel about 10 miles in the campus bus and 10 miles in the inter city bus per week.
- I generally take one long flight once a year.
- According to my usage, my carbon footprint is 18,724LBS of CO2.

3.14.4 E.Datacenter.4

All major data centres at Facebook, Google and the ones in China use Hydro as their renewal energy. Electricity is the main source for data centres to power their systems. It is essential to keep this in check at all times. Infact, all major data centres including Microsoft and other IT related companies are moving towards sustainable environment and using energy-efficient sources. Most of these data centres are looking at doing this to ensure they are not harming the environment. They ensure that by doing so, they are well-known and qualified vendors all over the world. In the wake of the changing environment and pollution due to various emissions in daily life, renewal energy is looked at serious by all data centres.

“Mongolia Information Park can make use of free air cooling for up to eight months a year. Power suppliers are also taking a keen interest in the Hohhot Cloud Computing Base. Beijing Energy Investment Holding said it will build a 2×350 MW and 2×600 MW cogeneration plant with 8.5bn CNY. While Inner Mongolia Electric Power Co said it will invest 1.11bn CNY for building one 500kV substation and two 220kV substations.” [7]

“Hohhot gets 7.7m kW from a mixture of thermal and hydroelectric sources, but is projected to gain another 1.8m kW after work is finished on a new thermal power facility.” [8]

Hohhot generates so much hydroelectric power beacuse yearly average of over a foot of rain adds to its 1.2 billion cubic meter reserve of water.

3.14.5 E.Datacenter.5

“China’s hyperscale data center industry is the second largest in the world, comprising 8% of the global market. Researchers estimate that in 2018 China was home to 1.2 million server racks from large and ultralarge scale data centers and 1.5 million server racks from small and mid-scale data centers. China’s data center industry consumed 161 TWh of electricity in 2018.35% of the country’s total electricity consumption. Researchers found that electricity consumption from the industry is on track to increase by 66% in the next five years, reaching 267 TWh by 2023, more than Australia’s total 2018 electricity consumption. In a typical data center in China, 80% of electricity is consumed by servers and cooling equipment.” [4]

Renewable Energy and impact of country’s dynamics on the environment have been major factors for China over the past decade. Several initiatives have been taken to ensure that China is technology driven in eco friendly ways. But, electricity is one thing that was not paid much attention to. Now, significant efforts have been taken at data centres to ensure that there is cost cutting even in the electricity sector.

3.14.6 E.Datacenter.8

In May 2017, British Airways filed a lawsuit against its data centre supplier CBRE. Around 700 flights were cancelled and 75,000 people were left stranded. It costed them 80 million Euros. This could be a factor that the ticket prices of British Airlines came down in 2018.

“Simon Allen, Membership Secretary at the UK Data Centre Interest Group (DCIG), and one of the key responsible for the DCIRN, said:”The airline industry has an enviable record of continuously improving flight safety by industry-wide sharing of accident and potential accident information," he said.

"However, the same is not the case in the data centre industry where it is common practice to cover up failures or potential disasters in a misguided attempt to protect reputations.

“Root cause investigation findings are normally secret and bound by NDA which has resulted in The Data Centre industry being at a disadvantage in learning from failures.” [6]

Similarly, in August 206, Delta Airlines cancelled 2,000 flights in 3 days. It costed them $150million. The reason was that an electrical component which provided power to the airline’s data centre failed and caused the shutdown of the data centre.

Therefore, this shows that data centre outages play a significant part in Airlines. Passengers waiting for their flights, check-in passengers would face major difficulties and hit the media everytime. People would tend to book flights in other airlines.

A piece of electrical component at its Atlanta headquarters failed which led to a shutdown of the transformer providing power to the airline’s data centre.

3.14.7 References

[1] http://www.xinhuanet.com//english/2017-08/05/c_136502192.htm

[2] https://data.worldbank.org/indicator/en.atm.co2e.pc

[3] https://www.insight.com/content/dam/insight/en_US/pdfs/apc/apc-estimating-data-centers-carbon-footprint.pdf

[4] https://secured-static.greenpeace.org/eastasia/PageFiles/299371/Powering%20the%20Cloud%20_%20English%20Briefing.pdf

[5] https://community.spiceworks.com/topic/1290566-why-are-so-many-data-centers-in-inner-mongolia

[6] https://data-economy.com/british-airways-dreadful-data-centre-outage-costs-airline-80m/

[7] https://www.datacenterdynamics.com/news/inner-mongolia-an-emerging-region-for-cloud/

[8] https://community.spiceworks.com/topic/1290566-why-are-so-many-data-centers-in-inner-mongolia

3.15 FA19-516-159

3.15.1 Week 1

3.15.1.1 E.Datacenter.2.b

The Microsoft Redmond Ridge datacenter is located in Redmond, Washington. It was created in 2009. The datacenter can generate approximately 17.3 MW.[1] One of the goals of the datacenter is to improve Microsoft’s energy efficiency. The results are apparent - the datacenter has lowered the company’s carbon footprint by 12,000 tons yearly, which is about 30% of the datacenter’s total footprint.[2] This total amount comes to about 40,000 tons, which is equivalent to the carbon output of 8421 vehicles.[3] Using the electricity cost estimate of 10 cents/kW[4], the datacenter costs about $15,154,800 per year.

3.15.1.1.1 Sources:

1. https://news.softpedia.com/news/Disparate-Microsoft-Server-Labs-Shift-to-Redmond-Ridge-1-121721.shtml
2. https://www.greenbiz.com/blog/2009/09/09/inside-microsofts-new-purpose-built-data-lab
3. https://www.reference.com/science/much-pollution-average-car-produce-7f2a3d41ed9d2689

4. https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-capacity-and-growth-planning-calculator/

3.15.1.2 E.Datacenter.3

My carbon footprint is 32,235 lbs. I was surprised by how much “single family housing” affected this metric.

3.15.1.3 E.Datacenter.4

Thermal energy is the energy contained within an object due to the movements of atoms within that object. Thermal energy, therefore, can depend on a variety of factors, including the object’s substance, temperature, and size. For objects where the molecules are loosely bound, such as in gas or liquids, there is likely more thermal energy than a solid at the same temperature. There is likely more thermal energy in objects that are hotter than in objects of a similar size and nature. Objects that are larger will contain more thermal energy than their smaller counterparts.[1]

Google plans to use thermal energy storage to cool the servers in one of its datacenters in Taiwan.[2]

3.15.1.3.1 Sources:

1. https://www.chegg.com/homework-help/definitions/thermal-energy-2
2. https://serverlift.com/blog/thermal-energy-storage-technology-in-data-centers/

3.15.1.4 E.Datacenter.5

The purpose of a rear door heat exchanger (RDHx) is to remove the hot air that comes off of server racks. One implementation of an RDHx is called a “passive” RDHx. The term “passive” comes from the fact that these heat exchangers do not have any moving parts (although they do require circulating water). As a result, passive RDHx’s demand less energy to operate and less maintenance. So, how does it work? Basically, server fans pump their hot air through the passive RDHx, which contains cold circulating water. Heat is exchanged between these two components, and the air temperature is thereby reduced. The air temperature can be reduced by anywhere from 10-35 degrees Fahrenheit. The exact amount of air temperature change depends primarily on the size of the RDHx and the server’s workload, but this cooling technique has been shown to decrease air temperature enough such that having dedicated “hot” and “cold” aisles is no longer necessary. Due to the simplicity of a passive RDHx, the installation is generally less complicated than the installation of other cooling units.[1]

3.15.1.4.1 Source:

1. https://datacenters.lbl.gov/sites/all/files/rdhx-doe-femp.pdf

3.15.1.5 E.Datacenter.8

In August of 2016, a fire broke out in a datacenter in Atlanta, Georgia. The primary user of this datacenter, Delta Airlines, was heavily affected. The fire affected both the primary and the secondary power generators, which forced Delta to reboot approximately 500 of their servers.[1] Delta representatives reported that the outage cost the company $150 million. Moreover, the outage affected over 2,000 flights over the course of three days.[2]

It appears that Delta did not have a redundant power architecture in their datacenter, which is part of what made this outage possible. In other words, the datacenter did not have a reliable backup for each of its components. The datacenter lost power when the Automatic Transfer Switch (ATS) - the unit that switches between primary and secondary power generators - caught fire. However, this should not have crippled the datacenter, as it should have had a secondary ATS. It is suspected that if the datacenter *did* have a secondary ATS, that it did not properly activate.[3]

3.15.1.5.1 Sources:

1. https://arstechnica.com/information-technology/2016/08/data-center-disaster-disrupts-delta-airlines/
2. https://www.datacenterknowledge.com/archives/2016/09/08/delta-data-center-outage-cost-us-150m
3. http://up2v.nl/2016/08/16/a-reconstruction-of-the-delta-airlines-datacenter-outage/

3.16 Datacenter fa19-516-160

- please use proper markdown

all images must be in an images directory, see other students examples

-

3.16.1 E.Datacenter.2.a: Table for Carbon footprint of data centers

3.16.2 E.Datacenter.2.b: Carbon footprint of data centers

- Data center : Forest City
- Organization : Facebook
- Location : Forest City, NC
- Year Build : 2012.

Forest City is home to about 7,500 residents, as well as a major campus for Facebook. As a result, it has joined the ranks of rural towns that are populated by more servers than people. This is company's second data center on its North Carolina campus.

Facebook Forest City Datacenter

3.16.2.1 References

- https://sustainability.fb.com/wp-content/uploads/2019/08/2018-Sustainability-Data-Disclosure.pdf
- https://www.facebook.com/notes/forest-city-data-center/study-finds-facebook-data-centers-have-contributed-nearly-6-billion-in-gdp-suppo/2228046683887349/
- https://www.rti.org/sites/default/files/facebook_data_centers_2018.pdf
- https://www.datacenterknowledge.com/inside-facebooks-north-carolina-data-center
- https://www.facebook.com/ForestCityDataCenter/

3.16.3 E.Datacenter.3

3.16.3.1 Reference

- http://carbonfootprint.c2es.org/

3.16.4 E.Datacenter.4

3.16.4.1 EQUINIX GREEN INITIATIVES

Equinix was founded in 1998 with the goal to become a leader in data centers, they pioneer in green data centers

They constantly work on their green technology. They always work on design power usage effectiveness (PUE).

Following are few examples of site-specific solutions incorporated by them:

- Rooftop solar panels
- Deep Lake Water Cooling (DLWC)
- Green rooftops (plants and vegetation)
- Water-side economizers And more.

SG3 (Singapore), one of newest sites, runs on solar energy power

O filenames must not have spaces

Equinix Data Center SG3

3.16.5 Here is the reference for more details :

- https://www.equinix.com/www/resources/infopapers/Green_Data_Center_Brochure_US_EN.pdf
- https://www.equinix.com/locations/asia-colocation/singapore-colocation/singapore-data-center/sg3/

3.16.6 E.Datacenter.5

- Largest Solar Farms ever Built for Google to Power Its Southeast Data Centers Two solar projects in Tennessee and Alabama will supply two data centers under construction with renewable energy (solar) most hours of the day.

Google South East USA DataCenter

- world's largest data center is planned for Ballangen, a small Norwegian fjord town inside the Arctic Circle, adjacent to the Lofoten Islands. Countries like Island are working towards using natural ways of colling data centers.

- China's Greenpeace: Data Centers on Track to Use More Energy than All of Australia. Biggest data center under construction in Guizhou Province of China. There is lot of push to use renewable source of energy part of program.

3.16.6.1 References

- https://www.datacenterknowledge.com/energy/largest-solar-farms-ever-built-google-power-its-southeast-data-centers
- https://e360.yale.edu/features/energy-hogs-can-huge-data-centers-be-made-more-efficient
- https://www.cnbc.com/2017/08/15/worlds-largest-data-center-to-be-built-in-arctic-circle.html
- https://www.datacenterknowledge.com/archives/2016/08/11/how-renewable-energy-is-changing-the-data-center-market
- https://www.datacenterknowledge.com/asia-pacific/greenpeace-china-s-data-centers-track-use-more-energy-all-australia
- https://www.researchgate.net/publication/270758358_Best_practice_case_studies_for_energy_efficient_IT_anc
- https://lifelinedatacenters.com/data-center/green-data-centers-looking-beyond-renewable-energy-cooling-efficiencies/

3.16.7 E.Datacenter.6

3.16.7.1 The ChilledDoor

(A Whole New Cooling Experience)

The ChilledDoor® Rack Cooling System not only remove 100% of the heat from server rack, it is a way to change the dynamic of how data center is cooled. It uses advanced technology "Active" rear door heat exchanger technology, by this cooling system becomes a dynamic entity, reacting to minute by minute changes in compute loads of up to 75kW. It can cool advanced High Performance Computers, high end storage, or simple switchgear, the ChilledDoor works to keep entire computing environment "heat neutral". Active rack cooling can redefine PUE and are ultra efficient, scalable and rack agnostic.

ChilledDoor

3.16.7.2 Reference

- https://www.chilleddoor.com/

3.16.8 E.Datacenter.7

Pretty much what we do in our daily life has some impact on our planet. There are so many ways human adds to carbon footprint. Like wise large corporations adds to carbon footprint indirectly all that is driven by human (consumer) demand. When we look at the modern computing or data driven companies they seems to be very clean apparently these companies use large amount of energy. Typically energy used by their data centers. Every time we click a link to search, upload a picture or watch a video we use energy starting form network, electronic device and data centers. Article points to variety of ways we add to use of energy.

3.16.8.1 Reference

- https://www.nature.com/articles/d41586-018-06610-y

3.16.9 E.Datacenter.8

Facebook, Instagram, WhatsApp suffer global outage started around 9:00 PDT on 13 March 2019 and continued to affect some services until 14 March 2019. This outage impacted enterprise accounts as well. This outage remained for around 14 hours. Facebook reported this is because of server configuration changes and impacted many users. These outages had also affected Facebook's ad-buying system, many brand marketers had tweeted to report issues.

3.16.9.1 References

- https://www.datacenterdynamics.com/news/facebook-instagram-whatsapp-suffer-global-outage/
- https://downdetector.com/status/facebook/map/

3.16.10 E.Datacenter.2.b: Tulip Data City

The Tulip Data City is located in Bangalore, India. It is operated by Tulip Telecom Ltd and was created in partnership with IBM, and several other tech firms, at a cost of ~$700 million (1). When it opened in 2012, it was the 3rd largest data center in the world. Today it ranks as the 10th largest in size at nearly one million sq.ft as well as 10th in power consumption at 90 MW/yr (2). Power to the Tulip Data City is supplied by 2 kV substations with a load capacity of 40 MVA (3). There is full back up capacity. The PUE is currently ~1.5 which climbs to 1.94 at full capacity (3). The annual IT load is 60MW (2). The current cost of one unit of electricity in Bangalore is Rs 7.55, which equates to $0.1 (4). Using these figures, the annual cost to operate the Tulip Data City is roughly $78.5 million (5, 6). The Tulip Data City does not utilize any renewable energy and thus has a rather large carbon footprint of 477, 770 tons which equates to 105, 330 CO2 car equivalents (6).

1. https://www.firstpost.com/business/biztech/tulip-telecom-unveils-asias-largest-worlds-third-largest-dc-in-bengaluru-1886925.html
2. http://worldstopdatacenters.com/tulip-data-city/
3. https://tulipdatacity.wordpress.com/about/
4. https://www.thenewsminute.com/article/electricity-bill-bengaluru-households-increase-about-5-june-102691
5. https://www.quora.com/Data-Centers-What-is-the-cost-per-MW-to-power-a-datacenter
6. https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-carbon-footprint-comparison-calculator/

3.16.11 E.Datacenter.4: Wind Energy

Wind turbines which supply power to the grid are very large, typically 80 meters or more tall, and are often clustered into large groups at a single location (1). These turbine groups have become known as “wind farms”. Wind farms can be located on land or in the ocean. Turbines located in the ocean tend to be bigger than those on land. Besides being a clean, renewable and widely available resource, wind energy is also very cost efficient at around 2-6 cents per kilowatt-hour (1). In the US wind power accounts for 6.5% of the total energy production (2). In contrast, Denmark, which has a high proportion of Data Centers, creates 39% of their energy needs via wind power (3). As shown in the table, currently there are relatively few wind-powered data centers worldwide, however some of the major Cloud providers such as Google and Microsoft have made commitments to offset their carbon footprint by purchasing renewable energy credits from wind energy providers. In addition, there a many wind-powered data centers planned or under construction throughout the world.

Name	Company	Location	Status
Other World Computing HQ	MacSales.com	Woodstock, IL	Active
Wyoming Data Center	Green House Data	Cheyenne, WY	Active*
Oklahoma Data Center	Google	Mayes County, OK	Active*
Iowa Data Center	Google	Council Bluffs, Iowa	Active*
Papillion Data Center	Facebook	Papillion, NE	Active
Fort Worth Data Center	Facebook	Ft. Worth, TX	Active
Virtual Earth Data Center	Microsoft	Boulder, CO	Active
San Antonio Data Center	Microsoft	San Antonio, TX	Active*
Wyoming Data Center	Microsoft	Cheyenne, WY	Active
Dublin Data Centers (4)	Microsoft	County Kerry, Ireland	Planned
WindCORES	Green IT	Paderborn, Germany	Planned
Dallas Data Center	Akamai	Dallas, TX	Planned
Denmark Data Center	Facebook	Odense, Denmark	Planned
Sweden Data Center	Facebook	Lulea, Sweden	Planned
Midwest Data Center	Facebook	New Albany, OH	Planned
Viborg Data Center	Apple	Viborg, Denmark	Planned
Aabenraa Data Center	Apple	Aabenraa, Denmark	Planned
Aabenraa DC	Google	Aabenraa, Denmark	Planned
Fredericia DC	Google	Fredericia, Denmark	Planned
Minnesota DC	Google	Becker, MN	Planned

*Operated via Renewable Energy Credits (REC) purchased from wind energy providers through long term Power Purchase Agreements (PPA)

1. https://www.energy.gov/eere/wind/advantages-and-challenges-wind-energy
2. https://www.awea.org/2018-market-report_us-wind-power-grew-8-percent-in-2018
3. https://en.wikipedia.org/wiki/Wind_power_by_country
4. https://www.datacenterdynamics.com/news/microsoft-adopts-googles-approach-to-buying-wind-power/

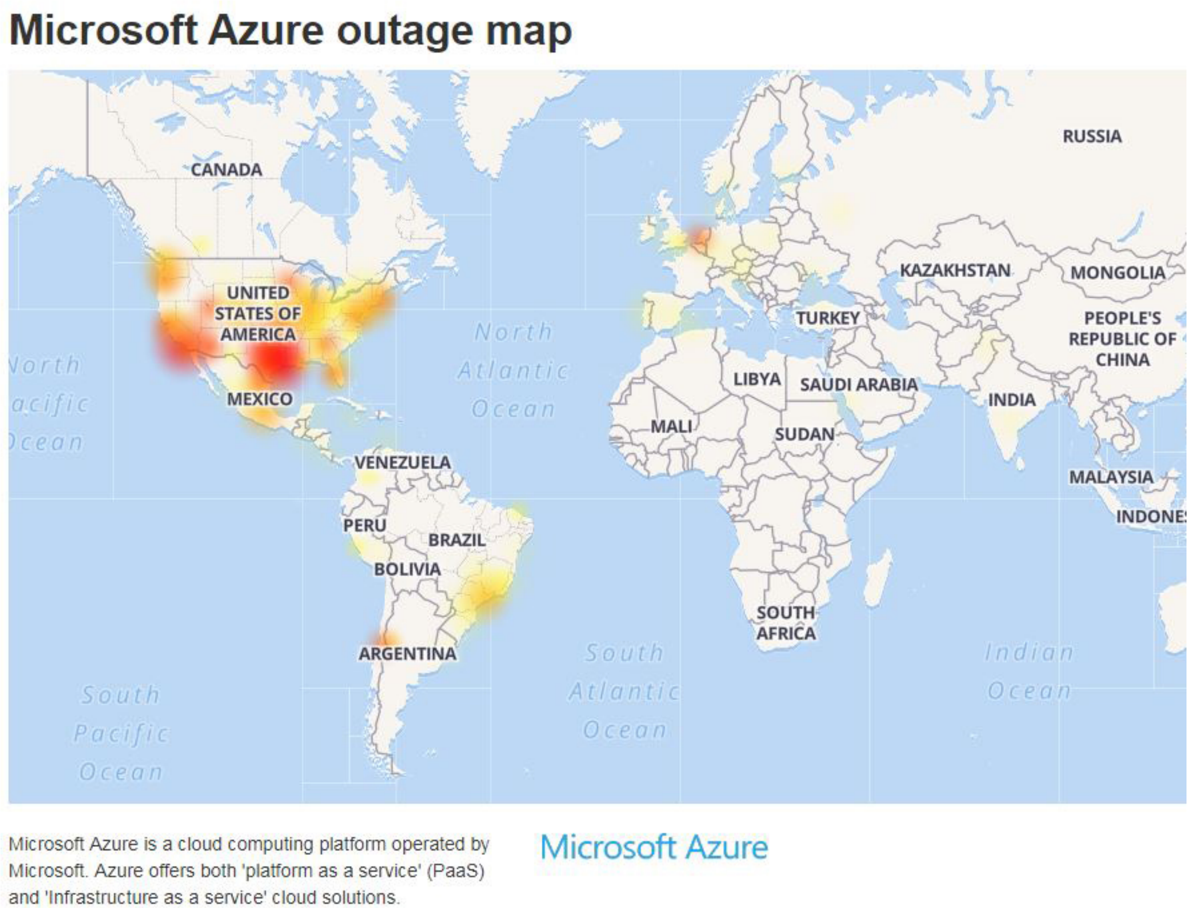
3.16.12 E.Datacenter.5: Google renewable energy effort

In 2012 Google made a commitment to eventually utilize 100% renewable energy for the entire company worldwide, including offices and data centers (1). It took Google five years to achieve this goal and in doing so it has become the largest corporate buyer of renewable energy in the world, spending $3.5 billion dollars to purchase 2.6 gigawatts over 7 years (1). While Google does directly utilize much of the renewable energy it purchases, it is not always possible in all its data center and office locations to connect to renewable energy sources. To offset this limitation the renewable energy industry has developed a system where they provide a Renewable Energy Credits or certificate (REC) for the amount of energy the sell. Google purchases an equivalent value of RECs to that of the fossil fuel energy consumed throughout the entire corporation (2). In this way Google can claim they are able to utilize 100% renewable energy. By far the largest amount of renewable energy Google purchases is wind energy and has many long-term purchase agreements with wind energy providers in the US Midwest, Northern Europe and Chile (1, 2). Recently, Google has announced major investments in renewable energy data centers in Minnesota, Finland and Denmark worth close to $2 billion (3-5). Lastly, it is not surprising that Google, a leading developer of AI algorithms that power its search engine, has committed major resources to use machine learning to optimize its data centers (6). The result of these efforts were a 40% reduction in cooling system energy use and 15% overall reduced energy consumption (6, 7).

1. https://static.googleusercontent.com/media/www.google.com/en//green/pdf/achieving-100-renewable-energy-purchasing-goal.pdf
2. https://sustainability.google/projects/ppa/
3. https://www.datacenterknowledge.com/google-alphabet/google-said-be-planning-600m-wind-powered-data-center-land-lakes
4. https://www.reuters.com/article/us-alphabet-denmark/google-to-invest-data-center-and-green-energy-in-denmark-idUSKCN1NP0GC
5. https://www.blog.google/around-the-globe/google-europe/accelerating-europes-clean-energy-transition/
6. https://sustainability.google/projects/machine-learning/
7. https://blog.google/outreach-initiatives/environment/deepmind-ai-reduces-energy-used-for/

3.16.13 E.Datacenter.8: Microsoft San Antonio Data Center failure

Date: 9/4/18 – 9/7/18 **Data Center:** Microsoft Azure South Central Data Center in San Antonio Tx **Category:** Weather-related **Description:** Lightning strike caused a voltage spike which damaged a cooling system leading to automatic shutoff of hardware (1,2).
Duration: Some Azure services were restored the following day, however it took 80.5 hours for full recovery (2). **Impact:** More than 40 Azure services, including Azure Active Directory, Azure DevOps and Resource Manager were unavailable (2,3). In addition, Office 365 SaaS including Skype, Exchange, SharePoint, Power BI and Teams were also unavailable (2,3). In total it was estimated that >30% of North American Azure users as well as users as far away as western Europe were affected (see map 2, 4, 6). **Cost:** Estimates of the economic impact ranged from $2.1 – $4.5 billion (5).



enter image description here

1. https://www.datacenterknowledge.com/uptime/microsoft-lames-severe-weather-azure-cloud-outage
2. https://www.datacenterdynamics.com/news/microsoft-azure-suffers-outage-after-cooling-issue/
3. https://www.datacenterknowledge.com/microsoft/azure-outage-proves-hard-way-availability-zones-are-good-idea
4. https://www.onmsft.com/news/azure-outage-update-services-still-down-as-outage-hits-san-antonio-data-center
5. https://www.lloyds.com/news-and-risk-insight/press-releases/2018/01/failure-of-a-top-cloud-service-provider-could-cost-us-economy-$15-billion
6. https://rcpmag.com/articles/2018/09/04/~/media/ECG/redmondmag/Images/2018/09/0904red_outage_b.ashx

3.17 Datacenter fa19-516-162

3.17.1 E.Datacenter.2: Carbon footprint of data centers

3.17.1.1 E.Datacenter.2.a:

The yearly cost and CO2 footprint of Google’s data center at Singapore have been added to the file.

3.17.1.2 E.Datacenter.2.b:

The 230,455 sq. ft. [1] Google data center in Singapore is the first Google data center in South East Asia [2]. It came online in December 2013. Google is developing two more data centers at the same location.

This data center uses an estimated 25 MW of power with a PUE of 1.15 [3].

Then the IT Load of the center can be calculated as

PUE = Total Load / IT Load [4]

IT Load = Total Load / PUE

IT Load = 25 * 1000/1.15 = 21740 kW

On putting these values in the Schneider Electric Data Center Carbon Footprint Calculator, the yearly cost and carbon footprint of the data center were calculated as follows:

Data center	**Electricity Cost ( )**	**Yearly CO_2 Footprint (tons)**	**CO_2 equivalent in cars**		
Google Singapore Jurong West	0.13	21,740	29.6M	160,089	35,293

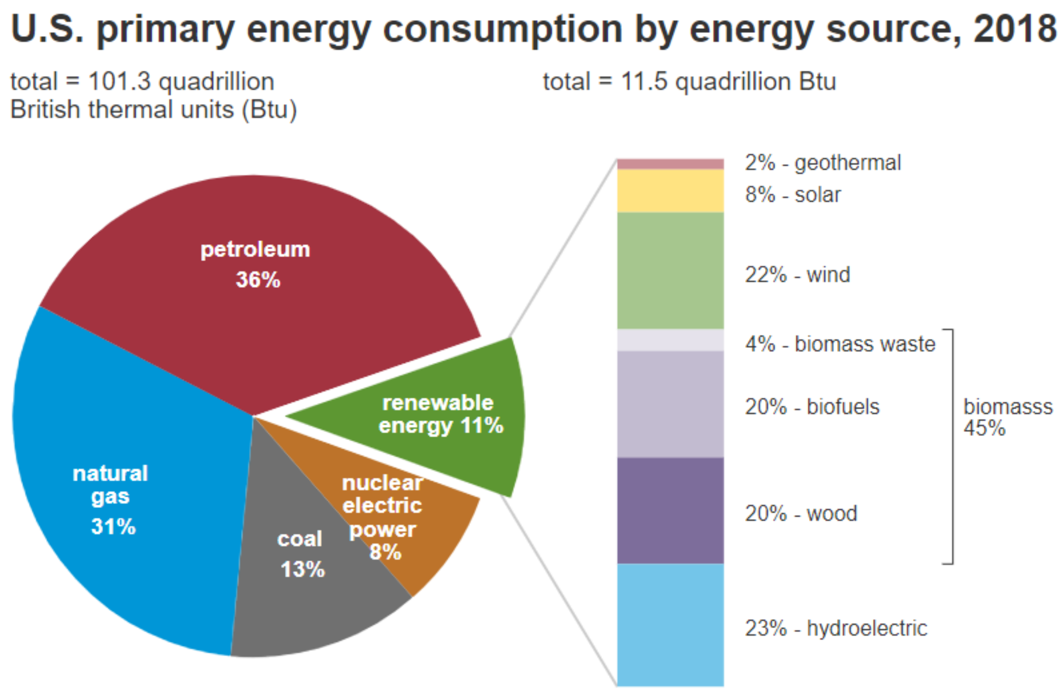
3.17.2 E.Datacenter.3: My Carbon footprint

My carbon footprint has been added to the file.

3.17.3 E.Datacenter.4: Hydro Power

Hydro power is energy in moving water[5]. It is one of the oldest sources for producing electricity. It is generated by rotating big turbines in heavy flowing streams or rivers or using a dam that releases water when necessary. The turbines then spin a generator that produces electricity.

Hydro electricity accounts for nearly 7% of the total electricity produced and 23% of the electricity produced using renewable resources in the US.



Hydro Power Consumption Source

3.17.3.1 Data centers that run on hydro power

- Near the Linth river in Switzerland, Data Center Light has established one of a kind of data center that boasts of an in-house hydro-power plant. It uses electricity generated by 100% renewable resources. It produces over 8000 MW/h of energy per year[6].
- Facebook opened a data center in Lulea, Sweden that is mostly powered by locally generated hydroelectric power. The center has an excellent PUE rating due to factors such as using natural air to cool the systems and using efficient hardware[7].
- Hydro66, a British corporation, has opened a 1000 sq. m. data center in North Sweden close to the Facebook's data center. This facility also uses the hydro electric power generated by the local plant.

3.17.4 E.Datacenter.5: California

- In September 2018, California passed a bill to move the entire power grid to renewable energy [11].
- California also has an issued order to remove as much carbon dioxide from the atmosphere as it emits [12].
- "The California Energy Commission estimates that 32 percent of retail energy sales were powered by renewable sources last year" [Camila Domonoske, 2018][12].
- "California has abundant energy resources, being among the top producers of oil, hydroelectricity, solar, biomass, and geothermal energy in the United States" [13].
- Due to strict emission laws, no coal-fired power plants operate in California.
- California leads the nation in electricity generation from non-hydroelectric renewable energy sources, including geothermal power, wind power, and solar power.
- The site https://www.dsireusa.org/ has several state incentives for residential to big corporations to promote the use of renewable energy.

3.17.5 E.Datacenter.8:

In August 2019, a power outage at Nissan data center in Denver, Colorado led to four days of shut down of most of the Nissan and Infiniti dealerships across North America, Canada and Mexico [10].

Reason for data center outage

Extreme heat in Colorado led to multiple power outages in the Denver area. The Nissan data center was also affected by this power outage. The power outage disabled the backup systems and made the recovery even more difficult [8].

Effects of data center outage

- The power breakdown resulted in their communication system called NNANet to shut down. This resulted in the dealers not being able to order new cars , parts, file warranty claims, check the inventory of other dealers or customer payoff information [9].
- The dealers couldn't make any sales [9].
- The data crash also brought down production at Nissan's factories in Tennessee and Mississippi [10].

How it could have been avoided

It took four days for the data center to recover from this outage. This was because the backup systems were co-located and were affected by the same power outage. If there are multiple data centers for a big corporation, a backup of data, applications, and systems should be kept in a different region because it has a very less chance of affecting different locations [8].

3.17.6 References:

[1] https://baxtel.com/data-center/google-singapore-jurong-west

[2] https://www.google.com/about/datacenters/inside/locations/singapore/

[3] https://www.google.com/about/datacenters/efficiency/internal/#measuring-efficiency

[4] https://en.wikipedia.org/wiki/Power_usage_effectiveness

[5] https://www.eia.gov/energyexplained/hydropower/

[6] https://datacenterlight.ch/en-us/cms/hydropower/

[7] https://www.greenbiz.com/blog/2013/06/13/facebook-flicks-hydropowered-data-center-sweden

[8] https://www.autonews.com/retail/computer-glitch-crashed-nissan

[9] https://carbuzz.com/news/nissan-and-infiniti-dealers-are-going-crazy-right-now

[10] https://www.automotiveit.com/cyber-security/data-center-power-outage-shuts-down-nissan-and-infiniti-dealer-and-production-systems-in-north-america/38945.article

[11] https://thehill.com/policy/energy-environment/405903-california-passes-law-committing-state-to-100-renewable-energy-use

[12] https://www.npr.org/2018/09/10/646373423/california-sets-goal-of-100-percent-renewable-electric-power-by-2045

[13] https://en.m.wikipedia.org/wiki/Energy_in_California

3.18 DATACENTERS FA19-516-163

3.18.1 E.Datacenter.2.b: Carbon Footprint

Data center: IU Bloomington
Organization: Indiana University
Location: Bloomington, IN, USA
Year Built: 2009 [1]

IT Load: `3300 kW.`

The IU Datacenter has 2, 2200 horsepower diesel generators for backup power [1]. Lets assume these can provide sufficient power to keep the Datacenter up and running. This can give us a very rough idea of the Datacenter load.

```
2 generators * 2200 hp/generator * 0.75 kW/hp = 3300 kW.
```

Electricity Cost: `$.10/kWh.`

I estimated this value from the latest published Duke Energy rates for high load factor consumers in Indiana [2].

Yearly Cost: `$2.89MM/yr.`

Annualize the aggregate IT Load and the Eletricty Cost:

```
3300 kW * $0.10/kWh * 24 hr/day * 365 day/yr = $2.89MM/yr.
```

Yearly CO2 Footprint (tons): `26336 ton/yr.`

The U.S. EIA has determined that Indiana's aggregate CO2 emissions are 1822 lbs/MWh (1.822 lbs/kWh) [3].

```
3300 kW * 1.822 lbs/kWh * 0.0005 ton/lb * 24 hr/day * 365 day/yr = 26336 ton/yr.
```

Yearly CO2 Footprint (cars): `5726 car/yr.`

The U.S. EPA estimates that the average passenger car emits 4.6 tons of CO2 per year [4].

```
26336 ton/yr / 4.6 ton/car = 5726 car/yr
```

3.18.1.1 References

1. Nolan, B. (2017, Oct 23). Indiana University Data Center serves as university's technology hub. https://news.iu.edu/stories/2017/10/iu/inside/23-data-center.html
2. Duke Energy Indiana. (2018, August 22). Rate HLF – Schedule for High Load Factor Service. https://www.duke-energy.com/_/media/pdfs/for-your-home/rates/electric-in/ratehlf.pdf?la=en
3. U.S. Energy Information Administration. (2019, January 8). Indiana State Electricty Profile, 2017. https://www.eia.gov/electricity/state/indiana/
4. U.S. Environmental Protection Agency. (2018, March). Greenhouse Gas Emissions from a Typical Passenger Vehicle. https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle

3.18.2 E.Datacenter.4: Geothermal Energy

Geothermal energy production uses steam or hot water from beneath the earth's surface to drive electrical turbines and generate electricity [1]. The spent water is pumped back under the surface where it is reheated naturally. This form of energy production is renewable because most/all of the heat energy is dervied from the liquid upper mantle of the earth's core.

Like many other renewables, the economic viability of geothermal energy production is highly location-dependant. Producers must drill wells in order to access and return geothermally heated water. The shallowest wells can be drilled where the liquid mantel rises furthest towards the surface. This phenomenon predominantly occurs near seismic boundaries.

As of 2018 the United States leads the world in geothermal energy production, with an installed capacity of ~3600 MW [2]. As a share of national energy production Kenya leads with just over 50%, followed by Iceland with 30%.

3.18.2.1 References

1. Union of Concerned Scientists. 2014. How Geothermal Energy Works.
2. Wikipedia. 2019. Geothermal power.

3.18.3 E.Datacenter.5: Iceland

Iceland is an excellent case study in the use of geothermal energy. Iceland is located on the Mid-Atlantic Ridge, the divergence point of the North American and Eurasian plates [1]. The island itself is a product of seismic activity. As of 2013 geothermal sources account for nearly one-third of Iceland's electrical energy production [2]. Geothermal heat sources have a number of other uses in Iceland, such as heating homes and swimming pools, melting snow on roads and sidewalks, and drying fish.

Iceland's subarctic climate provides a natural supply of cold air that can be used in cooling. These resources make Iceland an ideal location for datacenter construction, with Power Usage Effectiveness (PUE) approaching 1.07 and a Total Cost of Ownership (TCO) 40% lower than comparable datacenters in the United States [3].

3.18.3.1 References

1. Wikipedia. 2019. Geology of Iceland.
2. Orkustofnon (National Energy Authority). Electricy Generation.
3. National Power Company of Iceland. Data Centers in Iceland.

3.18.4 E.Datacenter.8: Google Outage, 2019 June 2.

On Sunday, June 2, 2019, an errant configuration change on the Google Cloud Platform resulted in network congestion and outages of many Google services [1]. Youtube views were reduced by 10% for about four hours. Roughly 1% of Gmail users were unable to access their account. The Uptime Institute has determined that over a quarter (28%) of data center outages are the result of configuration or software errors such as the one experience by Google [2].

Google has not released figures on the financial impact of this outage. However, Youtube's business model is based on advertising and views, so we can estimate the cost of a traffic reduction. Youtube has an estimated annual revenue of $16-25 billion. If we assume the income is constant, is an hourly revenue of $1.8-2.5 million. 10% of that is $180-250 thousand. Over four hours, the cost to Google would have thus been nearly `$1 million.

3.18.4.1 References

1. Tung, Liam. (2019 June 3) Google: Here's what caused Sunday's big outage.
2. Uptime Institute. 2019. Data center outages are common, costly, and preventable.
3. New York Times. (24 July 2019). YouTube Is a Big Business. Just How Big Is Anyone's Guess.

3.19 DATACENTER FA19-516-164 ☁

3.19.1 E.Datacenter.2.b : Tahoe Reno 1

Tahoe Reno 1 is the first datacenter owned by the company Switch in Reno, Nevada. This datacenter is located in the Tahoe Reno Industrial Center in a 2,000 acre campus known as the Citadel Campus. This data center was built in the year 2017. Since January 1, 2016, Switch decided to run its data centers on 100% renewable energy. This practise was even implemented with the Tahoe Reno 1. The data center itself occupies a space of 7.2 million square feet and runs on a 650,000 kW power which is generated at a price of $0.040 per kWh. Moving the servers to the Switch campus at the Tahoe Reno Industrial Center cost a one time investment of $550,000. The annual cost for maintaining these servers and renting the server space is $125,000. Up till the year 2015, Switch would contribute to the greenhouse gas emissions with a CO2 emission of 110460 metric tons which is about 121761 tons in 2015. From 2016, since Switch decided to use renewable sources of energy, there were no CO2 emissions from Switch data centers. Therefore Tahoe Reno 1 that was built in 2017, does not contribute to any CO2 emissions.

3.19.2 References:

1)https://baxtel.com/data-center/switch-reno-campus-the-citadel/news/city-of-reno-s-data-servers-will-be-moved-to-switch-campus-next-year

2)https://www.switch.com/switch-tahoe-reno-data-center-now-open/

3)https://www.switch.com/sustainability/

3.19.3 E.Datacenter.3 : Carbon Footprint

My carbon footprint is 15625 lbs of CO2

3.19.4 E.Datacenter.4 : Recyclers

Generally, the heat generated from data centers range between 80 and 115 degrees Fahrenheit. This heat which is wasted should be recycled in order to conserve energy. There are a few ways in which this can be done. Heat generated from data centers can be diverted to nearby homes and businesses, thereby utilizing the waste as a source of energy and providing heat during colder temperatures. Recycling the heat generated has benefits. It contributes to energy conservation. The amount of money spent in order to reduce the amount of heat generated can be used in other areas of the data center in order to make it function better.

3.19.5 References:

1)https://lifelinedatacenters.com/data-center/heat-generated-by-data-centers-can-recycled/

3.19.6 E.Datacenter.5 : Iceland

Iceland has probably set a record in the modern era. This is due to its use of only renewable sources of energy mainly hydro and geothermal for generating electricity. Iceland uses a district heating system that uses geothermal energy in order to generate heat. This heat generated satisfies 87% of the demand for hot water. 81% of the energy generated in Iceland for electricity, heat and transportation is generated from hydro and geothermal energy. Apart from providing energy for itself, Iceland also plans to use renewable sources of energy in order to provide electricity to the British Isles in what people say is an ambitious “$2 billion project” that proposes to provide Britain’s energy market with 5 TWh of Iceland’s renewable electricity.

3.19.7 References:

1)https://reneweconomy.com.au/iceland-a-100-renewables-example-in-the-modern-era-56428/

3.19.8 E.Datacenter.6 : Data Center Rack Cooling

Data Center Rack Cooling is done using a Rear-door Heat Exchanger. A Rear-door Heat Exchanger is a passive device which has no moving parts. It optimizes energy efficiency in a data center. It does not use electricity from the infrastructure in order to operate and is able to eliminate the chiller energy depending on the climate. This device has a simple installation and operation. Moreover it requires low maintenance. Rear-door Heat Exchangers are placed in the airflow outlet of a data center server rack. The heat generated by the data center is forced through this device by the use of server fans. This heat is passed through a cooling tower in order to exchange the hot air with circulating water. This reduces the temperature of the server rack.

3.19.9 References:

1)https://datacenters.lbl.gov/sites/all/files/rdhx-doe-femp.pdf

3.19.10 E.Datacenter.8 :

Google suffered a catastrophic cloud outage on June 2, 2019. This outage was caused due to a change in the configuration that was intended for a small group of servers of a particular region but ended up being applied to a large number of servers in that region and the neighboring regions. This cloud outage disrupted google services for about four and a half hours. Applications like Youtube, Gmail, Snapchat and Vimeo were affected due to the networking issues caused by this outage. Even the "Google Cloud Platform Services" such as App Engine, Bigquery, etc. were affected by this incident. Benjamin Treynor Sloss, the VP of engineering at Google apologized to the customers for this outage. He said that rectifying the damages caused by this outage took a lot more time than the company had anticipated. He also stated that this outage, which affected 1% of a billion people that use Gmail, resulted in a 10% reduction in Youtube traffic and a 30% reduction in the Google Cloud Storage traffic.

3.19.11 References:

1)https://www.zdnet.com/article/google-details-catastrophic-cloud-outage-events-promises-to-do-better-next-time/

3.20 Data Center fa19-516-165

3.20.1 E.Datacenter 2

- Data center: the Google Dublin Data Center
- Organization: Google
- Location: Dublin, Ireland
- Year: 2012 (in operation)
- Electricity Cost ($/KW): 0.13

  Calculation:

  The average cost per kW was EUR 0.117, which was $0.13 kW (based on 9/3/2019 currency rate between EUR and US Dollar).[1]

- Yearly cost $50 M. Between 2011 and 2017, the company spent $350 M in Ireland data center. [1]
- IT Workload: 22,187.5 KW.

Calculation:

Google centers consumed 710 MW between 2010 to 2018 in Europe.[1] Google has four data centers in Europe. For each, on average the 22.1875 MW.

3.20.2 E.Datacenter 3

- Based on the estimation, my carbon footprint is 71,718.

3.20.3 E.Datacenter 4

Biomass is a method to provide energy to data centers other than solar, wind, hydro, or recyclers. It utilizes the wastes from cities to generate electricity as power to support the operation of data centers.[2]

According to Vineyard Data Center website and Biomass Magazine, HP Lab Vineyard Data Center Park uses biomass as their main fuel.[2,3]

3.20.4 E.Datacenter 5

3.20.4.0.1 Apple Solar Panel Farms in China

Apple Inc. sets up the goal of using 100% renewable energy for its global operation. In China, they operated various wind and solar projects across six regions of China [4]. These operating projects in China compensates its manufacturing CO2 emissions [4]. Apple created a micro biosystem in their solar panel farms in China. The company sets their solar panels high enough to grow grass in these panels farms. These grass on the other hand is can feed yaks.

3.20.5 E. Dataceter 8

On Jan 23, 2017, United Airlines suffered from a IT data center outage. This outage prevented United Airlines pilots from getting key data and passengers could not check their flights' information from internet. Thus, the company on that day cancelled over 200 flights due to the Federal Aviation Administration's (FAA) ground stop for the its scheduled flights.[5] It assumed that on average, 200 passengers on one domestic flights. In total, over 4,000 passengers had to reschedule their travel. On average, the net profit for one passengers for the top seven airlines was about $17.75 [6]. Thus, due to this outage, United Airlines lost at least $71,000 net profit from cancelled flights. At the same time, the company also delayed 4,000 flight. This number did not count for the costs of reschedule passengers' travel, compensated passengers' time, reputation damages, and potential fine from the FAA.

3.20.6 References

[1] European data centres How Google's digital infrastructure investment is supporting sustainable growth in Europe Country case: Ireland. Retrieved from https://www.copenhageneconomics.com/dyn/resources/Filelibrary/file/9/109/1525764693/copenhagen-economics-2018-european-data-centres-case-study-ireland.pdf

[2] Biomass-Powered Data Centers. Retrieved from http://biomassmagazine.com/articles/5808/biomass-powered-data-centers

[3] Vineyard Data Center Park. Retrieved from http://www.vineyarddatacenterpark.com/

[4] Apple now globally powered by 100 percent renewable energy. Retrieved from https://www.apple.com/newsroom/2018/04/apple-now-globally-powered-by-100-percent-renewable-energy/

[5] United Says IT Outage Resolved, Dozen Flights Canceled Monday. Retrieved from https://www.datacenterknowledge.com/archives/2017/01/23/united-says-it-outage-resolved-dozen-flights-canceled-monday

[6] Here's How Much Airlines Are Profiting Off Your Plane Ride. Retrieved from http://money.com/money/5158363/airline-profit-per-passenger/

3.21 DataCenter fa19-516-166

3.21.1 E.Datacenter.2.b

There are two Google data centers in Council Bluffs, Iowa, the first of which opened in 2008, and the second announced in 2012. Both data centers have roughly the same PUE (1.09 vs 1.11), though the following calculations are for the second one. In this area of the country, the dollar per kWh is $.09. The IT workload of the center was not available online, but it was estimated to be around 9000 kW based off of comparisons to other data centers. When all of this is put into the CO2 calculator hosted by Schneider Electric, the annual electricity cost is $7.86 million and the total CO2 output is 77,429 tons (17070 equivalent in cars).

One thing worth pointing out as well is Google's outreach to the community in Council Bluffs since they built the data center there. The Google page for the site mentions that Google has awarded $2 million to local schools and nonprofits, as well as creating a free WiFi network for everyone in the area. Of course these are probably part of some agreement with the city and Google to benefit Google on breaks, its worth applauding Google for their efforts.

3.21.1.1 Sources

- https://www.google.com/about/datacenters/inside/locations/council-bluffs/
- https://www.google.com/about/datacenters/efficiency/internal/
- https://www.schneider-electric.com/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-carbon-footprint-comparison-calculator/

3.21.2 E.Datacenter.4

Solar power is the conversions of energy produced by the sun into electricity through the use of solar panels.

The article provided in the book describes the massive solar farms that Google is building in Alabama and Tennessee to power it's data centers in those respective states. The two farms together will be composed of 1.6 million solar panels. Some other data centers that use solar power are:

- Facebook's Henrico County data center in Virginia (under construction)
- Intel's New Mexico data center
- Apple's North Carolina and Nevada data centers
- Amazon also has six solar farms in Virginia to power its data centers

3.21.3 E.Datacenter.5

AWS is currently in the process of transitioning to 100% renewable energy to power its data centers by 2040. The sustainability section on their website mentions that they achieved over 50% in 2018. Currently they have six solar farms in Virginia, as well as wind farms in Indiana, Ohio, and North Carolina. In April they announced plans to build three more wind farms in California, Ireland, and Sweden. Additionally in August they announced another wind farm for Ireland, and another solar farm in Virginia. All of these projects are expected to be completed and generating clean electricity by the end of 2020.

Amazon as a company has 53 active wind and solar projects across the world now, and they have announced company-wide initiatives for net zero carbon for all deliveries by the end of 2030. They are seen as a leader in renewable energy by the Solar Industries Energy Association, having the second most active solar panel installations (behind Apple).

3.21.3.1 Sources:

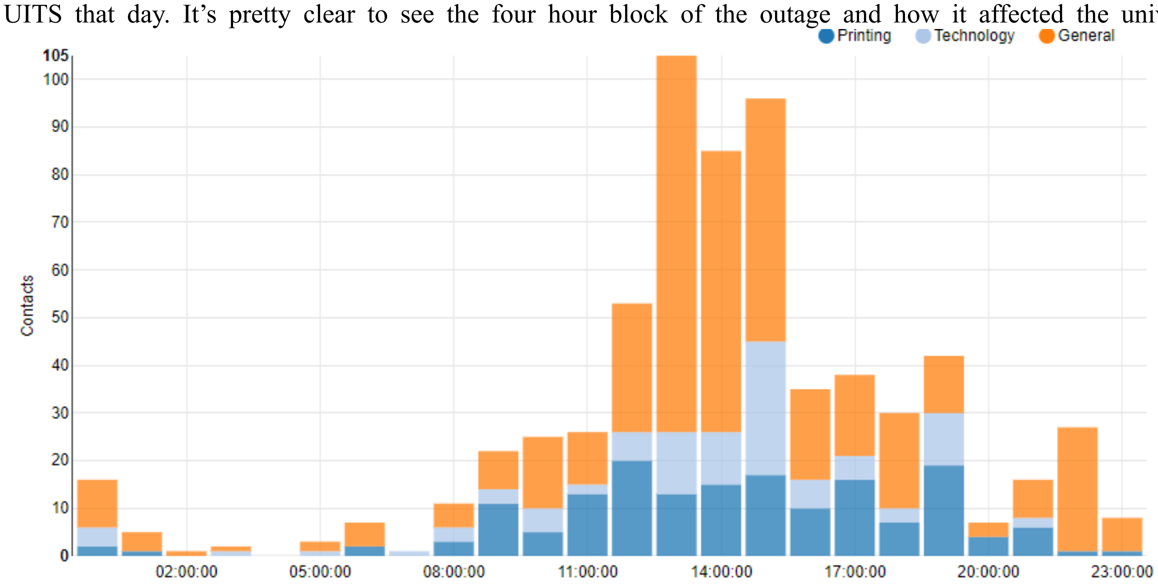
- https://aws.amazon.com/about-aws/sustainability/
- https://www.businesswire.com/news/home/20190801005260/en/
- https://www.businesswire.com/news/home/20190408005471/en/Amazon-Announces-New-Renewable-Energy-Projects-Support

- https://www.cnbc.com/2019/08/02/amazon-announces-new-renewable-energy-projects-in-the-us-and-ireland.html

3.21.4 E.Datacenter.8

On February 28th 2017, the AWS data center in northern Virginia experienced a service disruption for roughly four hours during the afternoon. It kind of broke the internet (both literally and figuratively). The outage affected S3 buckets for many companies/services, including Slack, Venmo, Gizmodo, Apple Cloud, and many more. Some of these services were down completely, while others just slowed down. It was later reported by Amazon that the outage was caused by an employee doing debugging accidentally widened his scope too big and it dominoed the servers.

As for the impact it had, obviously there were a large number of services that went down with the server. The Wall Street Journal reported that the total outage cost S&P 500 companies $150 million dollars in lost revenue, as many retailers saw their website performance drop. IU hosts all of its own services locally, but it just so happened that its Canvas instance was hosted on that AWS server. So Canvas was down for roughly four hours, and it caused quite the increase in calls to UITS. Students were freaking out they couldn't submit assignments, professors were complaining they couldn't administer quizzes/tests and so on. The graph below shows number of users assisted by UITS that day. It's pretty clear to see the four hour block of the outage and how it affected the university.



3.21.4.1 Source

- https://www.npr.org/sections/thetwo-way/2017/03/03/518322734/amazon-and-the-150-million-typo

3.22 Datacenter fa19-516-167

Carbon emission of a data center

3.22.1 E.Datacenter.2.a - Ashburn VA3

- Data center: Ashburn VA3
- Organization: Raging Wire Data Centers
- Location: Ashburn, Va
- Year Build: 2017
- Electricity Cost ($/kW): $.0808/kWh

* IT Load (kW): 16,000kW
* Yearly Cost ($): $11,324,928 MM/yr

.0808/kWh) * (365 days) * (24 hours/day)

"Yearly CO2 Footprint(tons)": ~75212 tons/yr

```
~75212 tons/yr = [(16,000kW) * (0.759 lbs/kWh) * (0.000707 metric tons CO2/kWh) * (365 days) * (24 hours/day)]
```

CO2 equivalent in cars: ~15969 cars

```
~15969 cars = (75212 tons/yr) / (4.71 metric tons CO2E/vehicle /year)
```

3.22.2 E.Datacenter.4 - Wind Energy

Wind energy operates on the principle of using moving air (wind) to spin a properller, which is connected to a rotor that turns the arm of a generator to produce electricity.

According to Sebastian Moss of DatacenterDynamics.com, Google and Facebook are partnering with Norwegin wind farms to help power its data centers. Additionally, Salesforce is partnering with a Canadian company to help power data centers.

3.22.3 E.Datacenter.5 - Rhode Island

In 2004, the Rhode Island's Renewable Energy Standard (RES) was established with the mission to produce 38.5% of the state's retail electricity sales from renewable resources by 2035. This includes electricity supplied by the state's retail electricity providers and non-regulated power producers and distribution companies.

Additionally, according to US EIA: "Rhode Island was the first state in the nation to implement an electricity restructuring plan, separating power generation from transmission and distribution, and the first operational offshore wind farm in the nation, the 30-megawatt, 5-turbine Block Island project" [7]

If Rhode Island power producers can stay on target with their annual renewable power supply targets and the Rhode Island Public Utilities Commission (PUC) confirms there is an adequate renewable energy supply, then by 2035 the state will have achieved its renweable energy goal.

3.22.4 E.Datacenter.8 - Wells Fargo data center outage

On February 7, 2019, Wells Fargo published a press release confirming a large-scale outage in one of its data centers which was initiated after smoke was detected following routine maintenance. According to the press release the impact of the outage affected the following services [8]:

* Wells Fargo ATM services.
* Mobile and Online Banking systems
* Using Wells Fargo credit and debit cards for purchases
* Bank branches
* Contact center systems

3.22.5 References

1. Ragingwire ashburn virginia data center specs. https://www.ragingwire.com/sites/default/files/ragingwire-ashburn-virginia-data-center.pdf
2. Commercial Electricity Rates & Consumption in Virginia. https://www.electricitylocal.com/states/virginia/
3. U.S. Energy Information Administration, Virginia Electricity Profile 2017. https://www.eia.gov/electricity/state/virginia/

4. Greenhouse Gases Equivalencies Calculator - Calculations and References. https://www.epa.gov/energy/greenhouse-gases-equivalencies-calculator-calculations-and-references
5. U.S. Dept. of Energy. https://www.energy.gov/maps/how-does-wind-turbine-work
6. DatacenterDynamics.com. https://www.datacenterdynamics.com/news/windcores-project-deploys-small-data-centers-inside-wind-turbines/
7. Rhode Island Profile State Profile and Energy Estimates https://www.eia.gov/state/analysis.php?sid=RI
8. Wells Fargo Provides Update on Restoring Services for Customers https://newsroom.wf.com/press-release/corporate-and-financial/wells-fargo-provides-update-restoring-services-customers-0

3.23 DATACENTER

O please use proper markdown with section headers for assignments

O please use proper urls

E.Datacenter.2: Carbon footprint of data centers

Microsoft San Antonio Texas data center

Metric Value Size (sq. ft.) 477,000 Investment ($) 550 million Power Consumed (MW) 300

Used carbon footprint calculator 4402379

Added inofrmation in link https://docs.google.com/spreadsheets/d/1gh869zfjA4sVxL8-ga0af2_HLTTuOoD1IReuRSrbq4I/edit#gid=0

Reference http://worldstopdatacenters.com/microsoft-san-antonio-tx/ https://www3.epa.gov/carbon-footprint-calculator/ https://www.fujifilmusa.com/products/tape_data_storage/case_studies/pdf/Hyperscale_Heat_Wave.pdf

E.Datacenter.2.b: Table Added information of in datacenter.md file

E.Datacenter.3: Your own Carbon footprint My Carbon footprint 14675

Added information in link https://docs.google.com/spreadsheets/d/1gh869zfjA4sVxL8-ga0af2_HLTTuOoD1IReuRSrbq4I/edit#gid=314181983

E.Datacenter.4: renewal energy - Others

Bio Energy

"Biomass is organic material that comes from plants and animals, and it is a renewable source of energy.

Biomass contains stored energy from the sun. Plants absorb the sun’s energy in a process called photosynthesis. When biomass is burned, the chemical energy in biomass is released as heat. Biomass can be burned directly or converted to liquid biofuels or biogas that can be burned as fuels.

Examples of biomass and their uses for energy

Wood and wood processing wastes—burned to heat buildings, to produce process heat in industry, and to generate electricity Agricultural crops and waste materials—burned as a fuel or converted to liquid biofuels Food, yard, and wood waste in garbage—burned to generate electricity in power plants or converted to biogas in landfills Animal manure and human sewage—converted to biogas, which can be burned as a fuel"

reference : https://www.eia.gov/energyexplained/biomass/

Waste heat in datacenters https://en.wikipedia.org/wiki/Green_data_center#Reusing_waste_heat

"Reusing waste heat Data centers use electric power, releasing more than 98 percent of this electricity as heat. Waste heat can be actively reused, and a data center becomes a closed-loop heating system with no waste. Examples include:

The IBM Reusing Data Center in Switzerland, where the heat warms a local swimming pool In Finland, the Yandex and Academica data centers replacing the heat used by 500-1,000 homes with data-center energy. Amazon reused heat from a nearby data center for a biosphere project in Seattle.[17]"

E.Datacenter.5: Pick a country, state, or company from Section 4.1.11 and summarize their efforts towards renewable energy and impacts within the society

“Microsoft showed its dedication to green energy when it invested in a 20 year agreement RES Americas wind farm for 100% of its energy. By the end of 2015 the Keechi Wind Farm Project will contain 55 turbines capable of handling the 300 megawatts of power that the San Antonio, Texas data center will need. This data center was modeled after the Microsoft facility in Quincy, Washington. At this location Microsoft hosts applications such as Microsoft.com, Hotmail, and Windows Live Messenger.”

https://www.fujifilmusa.com/products/tape_data_storage/case_studies/pdf/Hyperscale_Heat_Wave.pdf
http://worldstopdatacenters.com/microsoft-san-antonio-tx/

E.Datacenter.8:

major data center outages

Cost of datacenter outage and impact

https://www.nefiber.com/blog/the-9000-per-minute-cost-data-center-outages/

recent data center outage in Nissan https://www.datacenterdynamics.com/news/power-outage-denver-data-center-took-down-services-nissan-north-america/

Some major data center outage https://www.computerworld.com/article/3412199/11-data-centre-disasters.html

https://www.computerworld.com/article/3412199/11-data-centre-disasters.html#slide3

3.24 DATACENTER: FA19-516-169

3.24.1 DataCenter.2b

DuPont Fabros Technology (Digital Reality) has one big data center in Ashbur, Virginia. With 2.1 million sq.ft area and 10500 servers, this is one of the biggest data centers. This can run yearly capacity of the 200 MWatts. The data center features a calculated annualized Power Usage Effectiveness (PUE) of 1.15 by virtue of a highly efficient cooling plant and medium voltage oil-filled PDUs (power distribution units). These data points (reference below) combined with Schnider Electric Data Center Carbon Footprint Calculator, Yearly Cost, Yearly CO2 and Yearly Equivalent CO2 in cars could be computed.

Data Center	Location	Year	Electricity Cost*	IT Load	Yearly Cost	Yearly CO2 Footprint	CO2 Equivalent in Cars
Virginia Data Center,Dupont Fabros (Digital Reality)	Ashburn, Virginia	2014	0.08 $/kWh	200 M Watts	16.4MN	106380	23450

3.24.1.1 Reference

1. Dupont Data Center Information
2. Dupont PUE Information
3. Schnider Electric Data Center Carbon Footprint Calculator

3.24.2 DataCenter.3

I have computed my own carbon foot print using the CarbonFootPrint Calculator. Taking into account, my family's daily car travel of 10 miles and my air travel trips yearly (10 short and 2 long), the yearly carbon foot print assuming my residential area in Santa Clara- comes to 7939 lbs. However, since I actually reside in India, I decided to use another calculator which is specific for India. Using the calculator, my carbon foot print comes to 7187 lbs (3.26 tonnes). Currently, I have reported the first value in the spreadsheet on google drive.

3.24.3 DataCenter.4

According to Wikipedia:

Renewable thermal energy is the technology of gathering thermal energy from a renewable energy source for immediate use or for storage in a thermal battery for later use. The most popular form of renewable thermal energy is the sun and the solar energy is harvested by solar collectors to heat water, buildings, pools and various processes.

Another example of Renewable Thermal is a Geothermal or ground source Heat Pump (GHP) system, where thermal stored in the ground from the summer is extracted from the ground to heat a building in another season. This example system is "renewable" because the source of excess heat energy is a reliably recurring process that occurs each summer season.

One way data centers can make use of Renewable Thermal Energy is by using Geothermal cooling systems. Geothermal cooling systems include an array of v ertical holes drilled into the earth's surface that house a closed-loop piping system filled with water and/or coolant. The cool temperatures underground allows the piping system to serve as a heat exchanger.

Some examples of data centers using GeoThermal Energy include: 1. American College Testing Data Center 2. Prairie Bunkers DataCenters

3.24.3.1 Further References for GeoThermal Energy:

1. Future of Data Center Cooling
2. GeoThermal Data Centers

3.24.4 DataCenter.5

Renewable Energy in China

China has done tremendous investment in Renewable Energy forms and it's utilization. China has currently taken lead in renewable energy and is world's largest producer, exporter and installer of solar panels, wind turbines, batteries etc. China renewable energy sector is growing faster than fossil or nucelar power capacity. China sees renewables as a energy security rather than only for reduction carbon emission. However, inspite of being leader in renewables, only a quarter of the actual electricity is currently happening from renewables.

Hydro Power in China:Hydroelectricity is currently China's largest renewable energy source. At the same the big dams have caused environmental damages as well.

Wind Power in China: China has set goal for 15% all elecricity from renewables by 2020. With really big coastline, China has exceptional wind power resources. China is leading windpower generation. There are research estimate that China could meet all of their electricity demands through wind.

Solar Power in China: China is world's largest market for solar thermal energy and photovoltaics. Solar water heating also is extensively used. There is a concern though, because the byproduct of first generation solar cells - silicon tetrachloride, can have environmental impact, due to in-suffecient recycling.

With Chinese government providing subsidies as well as creating the policies, China is bound to continue it's leadership position in Renewable Energy.

3.24.4.1 Reference:

1. Forbes- China Renewable Energy Super Power
2. China's Dam Disaster
3. China and Renewable Energy- Wiki

3.24.5 DataCenter.8

On January 24th and 25th 2019, the online productivity suite experienced multiple outage issues throughout the last two days, with users reporting that they couldn't access their mailboxes through multiple protocol. Furthermore, some user reports also mentioned slow e-mail sending and receiving (with more than 3-hour delays), missing e-mails, as well as sending or receiving multiple repeated e-mails at one time.

Impact was USA, UK, France

The problem was due to a subset of Domain Controller Infrastructure becoming un-responsive, resulting in user connection timeouts. Microsoft later blamed DNS provider CenturyLink which acknowledged a software defect affecting connectivity to customer's cloud resources.

3.24.5.1 Reference:

1. Microsoft Outage in January 2019
2. Microsoft Office365 outage in January 2019
3. Microsoft outage one of the biggest in 2019

3.25 Datacenter fa19-516-170

3.25.1 E.Datacenter.2.a

3.25.1.1 Lakeside Technology Center - Chicago, IL

Electricity Cost	IT Load	Yearly Cost ($)	Yearly CO2 Footprint (tons)	CO2 equivalent in cars
$0.15/kWh	100,000 kW	$131.4M	670,140 metric tons	147,736 cars

Assuming the data center works 24/7 at 100 MW, then resulting in 876,000,000 kWh per year,
CO2 footprint: avg 1.7 lbs/kWh.
Yearly Cost is calculated based on electricity cost and IT Load.

3.25.2 E.Datacenter.2.b

The Lakeside Technology Center (350 East Cermak) is a 1.1 million square foot multi-tenant data center hub owned by Digital Realty Trust. Originally developed by the R.R. Donnelley Co. to house the printing presses for the Yellow Book and Sears Catalog, 350 East Cermak was converted to telecom use in 1999, and today it is one of the world's largest carrier hotels and the nerve center for Chicago's commodity markets, housing data centers for financial firms attracted by the wealth of peering and connectivity providers among the 70 tenants.

The industrial strength infrastructure includes four fiber vaults and three electric power feeds, which provide the building with more than 100 megawatts of power.

One of the most distinctive features of the facility is its cooling system, which is supported by an 8.5 million gallon tank of a refrigerated brine-like liquid. The huge tank serves as thermal energy storage for the Metropolitan Pier and Exposition Authority (MPEA), including the nearby McCormick Place Exposition Center and Hyatt Regency Hotel as well as 350 East Cermak. Thermal energy storage can reduce costs by running chillers during off-peak hours when power rates are cheaper.

References

Strange Marriage by the Lake
Digital Realty- Lakeside
Massive South Loop data center ready for expansion
World's Largest Data Center: 350 E. Cermak
Chicago's Data Fortress for the Digital Economy

3.25.3 E.Datacenter.3

12343 LBS of CO2, 5834 LBS comes from home.

3.25.4 E.Datacenter.4

Solar power is energy from the sun that is converted into thermal or electrical energy. Solar energy is the cleanest and most abundant renewable energy source available, and the U.S. has some of the richest solar resources in the world. Solar technologies can harness this energy for a variety of uses, including generating electricity, providing light or a comfortable interior environment, and heating water for domestic, commercial, or industrial use.

According to Solor-Powered Data Centers, solor power hasn't been widely used in data centers because a decent amount of energy required by data centers demands a very large installation of photovoltaic(PV). However, there are still some data centers that successfully implements solar power. For example, AISO is a nature intended web hosting company. It claimed that "Every hosting account with AISO is green and powered by on-site solar". Due to this philosophy, the company eliminate the production of 34488 lbs of co2 per year.

References

Solor-Powered Data Centers
AISO

3.25.5 E.Datacenter.5

According to China's data centers emit as much carbon as 21 million cars, China's data centers produced 99 million metric tons of carbon dioxide in 2018, which is equivalent to about 21 million cars on the road. And the article also anticipates that co2 emission will climb to 163 million metric tons, which is equivalent to 35 million cars. To reduce the amount of co2 emission produced by data centers, China's data centers take actions to decouple their electricity consumption from their carbon footprint by relying more on wind and solar energy. In 2015, Chinese government launched a green data center pilot program. In the same year, one of China's largest cloud computing providers–Alibaba started a data center located in Hangzhou, relying on solar and water energy. With the use of renewable energy, Chinese data centers combines the best of natural resources with modern city and advanced technology. Moreover, operational efficiency in the delievery of cloud computing and big data services is improved by the smart use of renewable energy.

References

China's data centers emit as much carbon as 21 million cars
AliCloud Launches New Energy-Efficient Qiandao Lake Data Center

3.25.6 E.Datacenter.8

3.25.6.1 Disruption of Google Cloud Service

3.25.6.1.1 Incident

Google is one of the most widely-used and reliable service providers on the Internet. Once it experiences an outage, the consequences can be far-reaching. In June 2nd, 2019, Google experienced an outage.

According to Google Cloud Networking Incident #19009, The outage was initiated by two misconfigurations and a specific software bug, then leading to software maintenance events and resulting in network congestion. The congestion lasted for between 3 hours 19 minutes and 4 hours 25 minutes varied from region to region. Users who lives in eastern US were severely affected by the outage. Fortunately, most users who lives in western US and all the users in Europe and Asia were not affected. During the congestion, Google services including computing engine, App engine, cloud platform were disrupted.

3.25.6.1.2 Impact

According to incident report YouTube measured a 2.5% drop of views for one hour, while Google Cloud Storage measured a 30% reduction in traffic. Approximately 1% of active Gmail users had problems with their account, which represents millions of users who couldn't receive or send email.

3.25.6.1.3 Discussion

After the Google Outage, many people pays more attention to cloud network reliability. At the time cloud services are disrupted, people realizes these services have been an integral part of life, and the integral part of life relies on the ability of companies who manage these services to operate. Thus, people claim for cloud service should include reater dynamic visibility, more accountability and a better cost recovery plan when problems occur.

References

Google Cloud Networking Incident #19009
An update on Sunday's service disruption
Google cloud is down, affecting numerous applications and services
Latest Google Outage Impacted Way More Than Its Own Apps
Google Outage Sharpens Focus on Cloud Network Reliability
Google's Cloud outage reveals the holes in cloud computing's atmosphere

3.26 Datacenter: fa19-516-171 ☁

3.27 E.DataCenter.2b:

The Switch SuperNAP held the title of largest data center in the world for quite some time, boasting 3.5 million square feet. It is located in Las Vegas, Nevada. This can run yearly capacity of the 587914 MWatts. The data center features a calculated annualized Power Usage Effectiveness (PUE) of 1.28 . These data points (reference below) combined with Schnider Electric Data Center Carbon Footprint Calculator, Yearly Cost, Yearly CO2 and Yearly

Equivalent CO2 in cars could be computed. Data Center Location Year Electricity Cost* IT Load Yearly Cost Yearly CO2 Footprint CO2 Equivalent in Cars Switch SuperNA Lasvegas,Nevada 2000 0.04 $/kWh 587194 M WattsH 23.4MN 275092 60861

https://www.switch.com/sustainability/

Schnider Electric Data Center Carbon Footprint Calculator

https://www.racksolutions.com/news/data-center-news/top-10-largest-data-centers-world/

3.28 E.DataCenter.3:

I have calculated my footprint using http://carbonfootprint.c2es.org/ calculator based on where I live and what heating mechanism I use and what I drive. My footprint is 8304 LBS OF CO2.

3.29 E.Datacenter.4:

Solar energy is a renewable free source of energy that is sustainable and totally inexhaustible, unlike fossil fuels that ar finite. It is also a non-polluting source of energy and it does not emit any greenhouse gases when producing electricity. Solar electricity can supplement your entire or partial energy consumption. Using solar power means reducing your energy bills and saving money. Low maintenance and unobtrusive, installing solar panels adds value to your home. https://www.energymatters.com.au/components/renewable-energy/ Solar and water are the renewal energies I have picked and these are used in the Switch data center as part of green initiative.

https://www.switch.com/sustainability/ In 2016, Switch began construction of two solar power stations in Las Vegas, Nevada with a combined 179MW of capacity. These facilities have the ability to - Power all Switch data centers with 100% renewable energy - Remove the equivalent of 50,000 cars off the road - Eliminate 265,000 carbon tons of emission from the environment

3.29.1 Water

Switch also applies sustainability initiatives to protecting the world's most precious resource: water. Through industry partnerships, Switch helped pioneer and develop proprietary technology that eliminates chemicals from our cooling systems, while increasing efficiency by over 400% and saving more than 155 million gallons of water in just 3 years.

3.30 E.Datacenter.5:

India is one of the countries with the largest production of energy from renewable sources.Four of the top seven largest solar parks worldwide are in India including the second-largest solar park in the world at Kurnool, Andhra Pradesh, with a capacity of 1000 MW. The world's largest solar power plant Bhadla Solar Park is being constructed in Rajasthan with a capacity of 2255 MW.

Installed grid interactive renewable power capacity (excluding large hydropower) as of 30 June 2019 (RES MNRE)[1][12][13] Source Total Installed Capacity (MW) 2022 target (MW) Wind power 36,368 60,000 Solar power 29,549 100,000 Biomass power (Biomass & Gasification and Bagasse Cogeneration) 9,806 *10,000 Waste-to-Power 138 Small hydropower 4,604 5,000 TOTAL 80,467 175,000

https://en.wikipedia.org/wiki/Renewable_energy_in_India

3.31 E.Datacenter.8:

3.32 DataCenter Exercises: ☁

3.32.1 E.Datacenter.2

Data center: Appling Data Center, Memphis, TN.

Year of inaugaration: 2004

IT load: 1500 KW

Cost per kWh in TN: 11.36 cents = $0.1136

Estimated Total Cost per Year: 1500 * 0.1136 * 24 * 365 = $1.49M

Carbon foot print: 1500 * 24 * 365 *4.61 = 7726 tons, this has been added to the file

Equivalent Cars: 1703

3.32.2 E.Datacenter.3

My carbon footprint has been calculated using the carbon foot print calculator and added to the file.

Apparently the more people live in a house changes carbon footprint a lot more than expected.

3.32.3 E.Datacenter.4

Hydro energy: Hydro energy is generated by converting energy generated from flowing water near dams.

a. Hydro energy produced in US by the end of 2018 is nearly 79,893 megawatts (MW) or 79.89 gigawatts (GW).

b. Data Centers powered by hydro energy are listed below:

   • Apple – Maiden, North Carolina, USA • Green Mountain - Schneider Electric owned data center in Norway • Facebook – Lulea in Sweden • Datadock – Strasbourg in France

3.32.4 E.Datacenter.5

Although China is outpacing US in renewable energy, most of the data centers in China rely about 75% of their power consumption on traditional sources like coal. Whereas US datacenters are quickly moving to renewable energy sources for power consumption. For example, all of Apple's datacenters in the US are currently running on renewable energy since 2013.

a. In US:

• Out of total energy produced in US, 17.1% from renewable energy sources.

• Among renewable energy, the main sources were 7% hydropower, 6.6% wind, and 1.6% solar power.

b. In China:

• Out of total energy consumbed by dataceters in China, about 25% from renewable energy sources.

• Among renewable energy, the main sources were 20% hydropower, 4% wind, and 1.84% solar power.

3.32.5 E.Datacenter.8:

• Data Centers effected by outages approximately: 33.33%(2018), 25%(2017) & 22%(2016)

• I'm a frequent user of Amazon for online shopping, which internally uses AWS as its hosting infrastructure. The AWS outage on Amazon Prime Day sale in 2018 for 63 minutes is estimated to have caused a loss of $100 million. A 4 hour outage in 2017 affected websites like Coursera, Medium, Quora, Docker, Expedia, and US Financial services companies lost around $160 million. An outage of 6 hours in 2015 caused downtime for Netflix, IMDB, Reddit, and a few more websites. The affected services were Amazon CloudWatch, AppStream, CloudSearch, Cognito, EC2 Cloud, EC2 Containers, Elastic Load Balancing, Elastic MapReduce, Elastic Transcoder, ElastiCache, Glacier, Kinesis, Machine Learning, to name a few

3.32.5.1 References:

2. https://www.zayo.com/services/data-center-colocation/facilities/memphis-7620-appling-ctr/

4.a) https://www.eia.gov/energyexplained/hydropower/where-hydropower-is-generated.php

4.b) https://www.colocationamerica.com/blog/top-five-greenest-data-centers

5. https://www.cnn.com/2019/09/10/asia/china-data-center-carbon-emissions-intl-hnk/index.html

6. https://en.wikipedia.org/wiki/Renewable_energy_in_China

7. https://www.cbronline.com/news/amazon-outage-lost-sales https://www.datacenterknowledge.com/archives/2017/03/02/aws-outage-that-broke-the-internet-caused-by-mistyped-command https://www.datacenterknowledge.com/archives/2015/09/21/amazon-data-center-outage-affects-netflix-heroku-others

3.32.5.2 E.Datacenter.3:

1. I stay in a house in Bloomington, Indiana along with three other people.
2. The house is centrally air-conditioned.
3. The electricity usage has amounted to 0.36 metric tons of carbon emission.
4. Two international trips every year sum up to 7.50 metric tons.
5. I also travel approximately 20 miles a week by the city bus. The carbon emission here amounts to 0.01 metric tons.
6. Secondary activities such as buying food, electronics, furniture, etc, all sum up to 2.99 metric tons.
7. My carbon footprint is 10.85 metric tons, i.e., 23920.155lbs of carbon emission.

3.32.5.3 E.Datacenter.4:

4. Recyclers

This can be desribed as a process of reusing materials that would otherwise simply go to waste, into useful products. In the context of datacenters, this implies resuing the energy that has been relased from it. This not only reduces costs, but also helps cut down on the carbon emissions. This energy recycling can be done in many forms, one of the most popular ones being reusing the heat from the hot aisles to warm nearby homes, offices, etc.[1]

One such example is Intel recycling the waste heat its datacenter produces, thus eliminating the need to use boilers to provide heating to the entire building and hot water throughout the entire year. This not only reduces the carbon emissions and energy usage, but also reduces costs drastically. These savings have been pegged at USD 235,000 annually. To recycle heat, the company employs the use of 'Heat Recovery Chillers (HR Chillers)'. The facilty has been designed and built to specifically accomodate this system from day-1. However, this posed a problem as the intial heat output from the datacenter would only bw 0.3 MW whereas the building requires 1.3 MW of heat energy. To accomodate this, they recover heat from offices, communication and meeting rooms and reuse it again as a part of the HR chilling system.[2]

3.32.5.4 E.Datacenter.5:

Efforts toward renewable energy and impacts within the society The California state government has been putting in plenty of effort duing the last few decades to combat climate change and has been gravitating towards cleaner energy since then. Offering consumer rebates for intsalling and using wind and solar energy, passing a legislation to ensure that those who sell electricity increase their renewable energy purchases by at least one percent every year are some efforts made by the state to name a few.[3]

California now fulfills a huge part of its electricity needs from these renwable energy sources and aims to power atleast 50% of their energy needs from renewable sources by 2025 and increase this percentage to 60 by 2030 and eventually 100% by 2045. Since the generation of this kind of energy can get inconsistent, large-scale energy storage systems can be used to overcome this problem.[4]

While these efforts are going to pay off in the long term, the immediate impact on the society can have a detrimental affect. The people in the state are already paying 50% over the national average for electricity. This aggressive shift towards green sources can end up being expensive as well. Irregular generation of energy means that this excess energy would have to be stored for future use. In another scenario, if no energy is generated at all, they would need to have oil or natural gas based power plants on standby for such emergencies. However, the public at large still does support these policies while probably not fully understanding the consequences.[5]

3.32.5.5 E.Datacenter.8:

A major datacenter outage that happened was in Sydney. It affected most of the services in Australia and life was almost brought to a standstill. 1. Date: June 2016 2. Datacenter: AWS 3. Category: Severe Weather 4. Description: Storms led to a power outage and eventually failure of service instances. 5. Duration: 10 hours 6. Impact: Severe A large number of businesses were affected by this datacenter outage. Some of the big names include Domino’s Pizza, Foxtel, Westpac, Bank of Queensland, Me Bank, IMB Bank, etc. While some recovered quick enough, banking services were disrupted and bank cards couldn’t be used temporarily. This led to a major part of the citizens lives coming to a standstill as they were unable to pay for any services provided.[6] The online services disrupted include food delivery and streaming services as well.

3.32.5.6 References:

- [1]https://www.datacenterknowledge.com/data-centers-that-recycle-waste-heat
- [2]https://media10.connectedsocialmedia.com/intel/09/5601/Intel_IT_Green_IT_Sustainability_Data_Center_H
- [3]https://ww2.energy.ca.gov/renewables/
- [4]https://www.npr.org/2018/09/10/646373423/california-sets-goal-of-100-percent-renewable-electric-power-by-2045
- [5]https://www.forbes.com/sites/arielcohen/2018/09/21/californias-new-100-green-energy-target-may-do-more-harm-than-good/#6d03ba04675e
- [6]https://www.smh.com.au/technology/web-chaos-mostly-over-after-amazon-web-services-hit-by-power-outage-during-sydney-storm-20160606-gpc707.html

4 REFERENCES

[1] US Energy Information Adminitstration, “Eia - beta: Germany.” Web Page, Aug-2016 [Online]. Available: https://www.eia.gov/beta/international/analysis.php?iso=DEU

[2] “Renewables generated a record 65 percent of germany’s electricity last week.” Web Page, Mar-2019 [Online]. Available: https://e360.yale.edu/digest/renewables-generated-a-record-65-percent-of-germanys-electricity-last-week

[3] S. Moss, “WindCores project deploys small data centers inside wind turbines.” Web Page [Online]. Available: https://www.datacenterdynamics.com/news/windcores-project-deploys-small-data-centers-inside-wind-turbines/

[4] Facebook, “Facebook sustainability data 2018.” Web Page, Aug-2019 [Online]. Available: https://sustainability.fb.com/wp-content/uploads/2019/08/2018-Sustainability-Data-Disclosure.pdf

[5] Facebook, “Prineville, or data center.” Web Page, Sep-2019 [Online]. Available: https://www.facebook.com/PrinevilleDataCenter/app/399244020173259/

[6] Facebook, “Sustainable data centers – facebook sustainability.” [Online]. Available: https://sustainability.fb.com/innovation-for-our-world/sustainable-data-centers/

[7] Facebook, “2018 sustainability data disclosure.” [Online]. Available: https://sustainability.fb.com/wp-content/uploads/2019/08/2018-Sustainability-Data-Disclosure.pdf

[8] Facebook, “(42) prineville data center.” [Online]. Available: https://www.facebook.com/PrinevilleDataCenter/app/399244020173259/

[9] "Sustainability in numbers – facebook sustainability." [Online]. Available: https://sustainability.fb.com/sustainability-in-numbers/#section-GreenhouseGasEmissions

[10] Facebook, "Denmark data center to warm local community." [Online]. Available: https://sustainability.fb.com/wp-content/uploads/2019/06/Waste_Heat_Recovery_Final_Jun2019.pdf

[11] Facebook, "Sustainable data centers – facebook sustainability." [Online]. Available: https://sustainability.fb.com/innovation-for-our-world/sustainable-data-centers/

[12] Y. Xu, "AWS s3 outage likely caused by internal network issue." Web Page, Feb-2017 [Online]. Available: https://blog.thousandeyes.com/aws-s3-outage-likely-caused-by-internal-network-issue/

[13] C. Newton, "How a typo took down s3, the backbone of the internet - the verge." Web Page, Mar-2017 [Online]. Available: https://www.theverge.com/2017/3/2/14792442/amazon-s3-outage-cause-typo-internet-server