Developing the Eco–Friendly/Renewable Energy Sustainable Data Centers of Cloud

LAKSHAY - 19BCE101 GAURAV - 18BCE176
Institute of Technology Nirma University Ahmedabad

December 2022

1 Abstract:

The abstract shall highlight the important details of the Internship work. It should consist of the objective of work, the scope of work, preliminary work carried out and important findings. Data center play vital part in information and communication technologies. Big companies like Google, Microsoft, and Amazon hold large amount of data center to provide cloud computing service. Due to the rapid growth in data center size and complexity, it is essential to highlight important design aspects and challenges of data centers. This project present data center network design, classification of the data center servers, recent developments, and future trends of the data center industry. Furthermore, the emerging paradigm of mobile cloud computing is debated for research issues.

2 Introduction:

2.1 General:

Data centers is facility that store and organize critical the data of big organization. The data centers design is based on a network of computing and storage resources that enables the application and data. To build data centers there are components which uses like router, switches, firewalls, storage system, servers and application delivery controllers. Data centers have experienced rapid growth in the number of hosted servers. Large Internet corporations and IT equipment manufacturers such as Google, Amazon, Microsoft, Dell, IBM, and HP are continually expanding their existing data centers and building new ones. Data centers is important because in the world, all the business application and activities like Email, Big data, machine learning, artificial Intelligence. It provide

2.1.1 Network infrastructure

To connect end-user to connect server, data center service.

2.1.2 Storage Infrastructure

To store data in data centers

2.2 Problem Statements

- Difficult to expand inter data center service. Data center networking are not Scalable and cost effective. It requires expensive high end switches. That's why it can't be afforded by small companies.
- When you have not enough space or enough resource then for that Data infrastructure management is key to solve the problem. It has to design system in place to find what is useful and what is useless
- One consistent source of datacenter networking challenges is security. A data breach could cost millions of dollars in lost intellectual property, private data leakage, and stolen personal information
- Data center required lots of equipment and while they all are running, the amount of heat generate can destroy system and it can be overheat.

Literature Review			
Author Name	Scheme	Pros	Cons
Ayman A Gouda	An Eco-Friendly Cloud Point Extraction for Preconcentration of Iron(III) in Water Samples prior to Spectrophotometric Determination	Decreases Energy Consumption	Variation in Outputs
Prerna Rawat,Prabhdeep Singh,Vikas Tripathi	Load Balancing in Cloud Computing Leading Us Towards Green Cloud Computing	Reduce Data Center Power Consumption,Performance	Expensive
Maha Al-Tameem	Eco-Friendly Method for Determination of Allopurinol Drug in pure form and Pharmaceuticals after Cloud Point Extraction	Using alkaline medium to form a micelles, Renewable energy efficiency increases	Difficult to implement
Rama Ran and Kenndy	Energy efficient task scheduling using adaptive PSO for cloud computing	Searches Optimal Solution	Can Fall into local optimum and low convergence iterative
Mueen Udeen,Muhammad Talha Azizha	Green Information Technology based Framework	Secure, Seamless,Eco Friendly Efficient	High Maintenance
Chee Shin Yeo, Rajkumar	Carbon Aware Green Cloud Architecture Energy Efficiency Environment Based	Environment Friendly On Cost	Not Private
Yuxiang Shi, Xiaohong Jiang	Cloud Broker Linear Predictive Technique,Reservation Reduced Technique	Better Response Time,Retrive useful Information Dynamic Allocation	Non Static Allocation
Li Hu and Jia Zhao	Power Consumption and Energy Efficiency	Decreases Energy Consumption	Decreases Less Carbon emission with hazardous human health
Chowdhury,Chatterjee	Virtualisation System and Controls Coefficent Constants	Fast Decision Making	Expensive,third-party resources that creates a security issue

Literature Review			
Author Name	Scheme	Pros	Cons
Muhammad Singh	DAG used to model process and Optimizes ECS using DVS	Fast Shreduling Time, Makespan Optimization	Shreduling performance limited due to local optimum
Ghribi et al.	Enables transferring memory images from overload host to destination host	Reduce Data Center Power Consumption,Performance	Algorithm relying on migration
Fujiwara et al, Lee et al	Multiple operating systems and applications can run on same machine	Allocates the resources Properly, Restricting CPU use below threshold	Short Market Life
Verma et al	Power and Migration Cost Aware, Application Placement in Virtualized Systems	pMapper minimizes power and migration cost, performance guarantees	Cant Handle Condition Constraints
Chase et al, Koseoglu and Karasan	Simulating an environment for energy aware cloud computing data centers	Reducing energy consumption by changing unused servers to power saving modes	Face the redundancy Requirements
Dumitru et al, Liu et al, Chiriac et al	Utilizing the prediction management plan for IT Workload	Makes the Data Center Environment more cooler,	Dependency Injection

2.3 Motivation

The Green Data center needs big space for the management, storage and distribution of data. For that, they need electrical, mechanical and computer systems designed to minimize the environmental impact and maximize energy efficiency. The main goal is to save energy and carbon-footprint reduction Moving basic software programs to the cloud can save electricity immensely. According to case study was done by Lawrence Berkeley National Laboratory that moving business software to the cloud can save enough electricity each year to power Los Angeles for 12 months This means the cloud would lower the total energy consumption of these software applications by 87 percent

2.4 Contribution

Data centers are energy-intensive critical infrastructures that provide a wide range of Internet-based services. Energy usage models are essential for creating and improving energy efficient processes in data centers to reduce excessive energy consumption. Cloud storage was invented by Rick Linder in 1960 by Dr. Joseph Karl Robnett. "Cloud computing" was first used in the modern context on August 9, 2006, when his then-CEO of Google, Eric Schmidt, introduced the term at an industry conference. The data center has its roots in his ENIAC giant computer room in the 1940s, one of his earliest examples of data centers.

2.5 Objective

Objective is to covers traditional cloud issues and green cloud uses, highlighting recent efforts in the area of green cloud computing to enable a healthy and green environment. Therefore, we conducted a comparative study in the field of green cloud computing. There are many directions for future work. This release addresses the issue of how to efficiently retrieve results from the cloud so that you can achieve all the features described in the release. Additionally, you can implement an approach that automates the green cloud manager that makes all service-related decisions.

3 Background

Today's clouds are designed to serve a large number of tenants over massive volume of data. The availability of large-scale, distributed storage systems lies the foundation of today's data centers.

3.1 Pv4 and IPv6 support :

Because many modern networks implement a dual stack environment that includes IPv4 and IPv6, this blueprint architecture provides support for both IP protocols

3.2 Data center Interconnect (DCI) :

The data center interconnect (DCI) provides the technology needed to send traffic between data centers. Routes are exchanged between devices in different data centers to allow for the passing of traffic between different data centers in this reference architecture.

3.3 DHCP relay:

The Dynamic Host Configuration Protocol (DHCP) relay allows the network to pass DHCP messages between a DHCP client and a DHCP server.

3.4 ARP Synchronization and Suppression:

When one of the IP gateways for a subnet learns about an ARP binding, this information is shared so the other gateways do not need to discover the same ARP binding independently.

With ARP suppression, when a leaf device receives an ARP request, it can check its own ARP table that has been synchronized with the other VTEP devices and respond to the request locally rather than flooding the ARP request. Such behavior reduces the amount of traffic in the network and optimizes processing for both network devices and end systems.

3.5 IP Fabric underlay Network :

The modern IP fabric underlay network building block provides IP connectivity across topology.

Here is a list of common benefits that come from data centers

- Reduce Data Center IT costs and risks. The need for IT staff should be reduced to match the difficulty of maintaining staff with IT skills
- Data Center offers best-in-class data capacity and state-of-the-art networking in the form of network reliability and scalability.
- As your business grows and requires more network resources. Data Center can easily meet new requirements.
- Data Center Managed Services provides readily available and competent IT staff for the operation, monitoring and maintenance of the system.
- Data Center managed services enable continuous operations. 24/7 continuous support, security and flexibility.

4 Carbon Evaluation

4.1 Computing and Analyzing Data Sets

- Data Directory Contains CSV files of Energy Production and Carbon Intensity datasets for regions like Germany, Great Britain and France of Entire Year.
  Cal ci.csv: Contains change in Carbon Intensity relative to every 5 minutes having more than 100000 Records of California. Similar for France, Great Britain, Germany.

Cal Production.csv: Initiates the production of Solar, Wind, Geothermal, Biomass, Biogas, Coal, Nuclear ,Natural Gas, Hydoenergy, Battery, Imports relative to every 5 minutes and have more than 100000 Records of California. Similar for France, Great Britain, Germany.

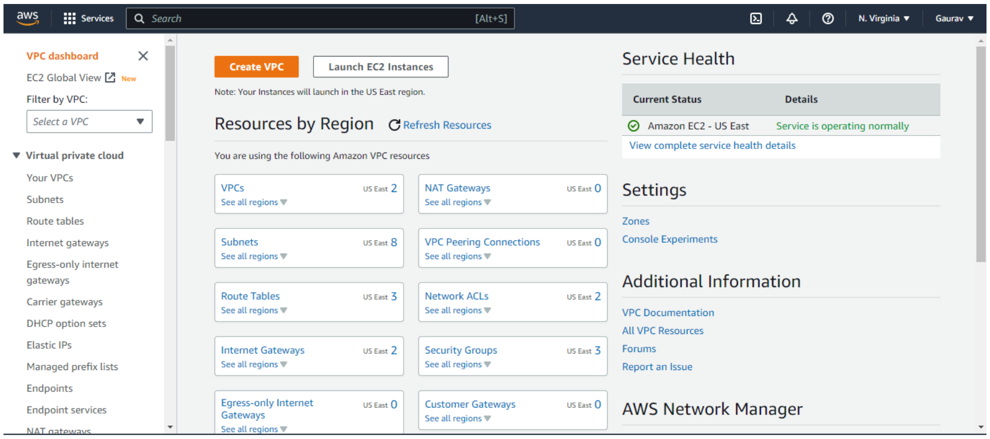
- Compute Carbon Intensity.py: It's a Script used to convert Energy production to Carbon intensity data using energy source provided by IPCC study. It's a literature review of life cycle of total energy sources CO2 emissions per unit electricity generation. Also stores the dictionary of values on Productivity Rate and CO2 emissions of Renewable energy like Bio power, Solar Power, Geothermal, Hydro power, ocean , wind and non Renewable Energy like Nuclear, Gas, Oil, Coal.
- Simulate.py: Simulates and Evaluates workload shifting approaches in Data Centers with the goal to consume low carbon consumption
- Analysis.ipnb: Used to analyse carbon intensity data
- Evaluation.ipynb: Used to analyse simulation result.
- Environment.yml: Includes prerequistes for installation and running
- Strategy.py: For workload to shift in both directions it schedules workload in the future.Its basic use is to maintain computation

4.2 Data Center

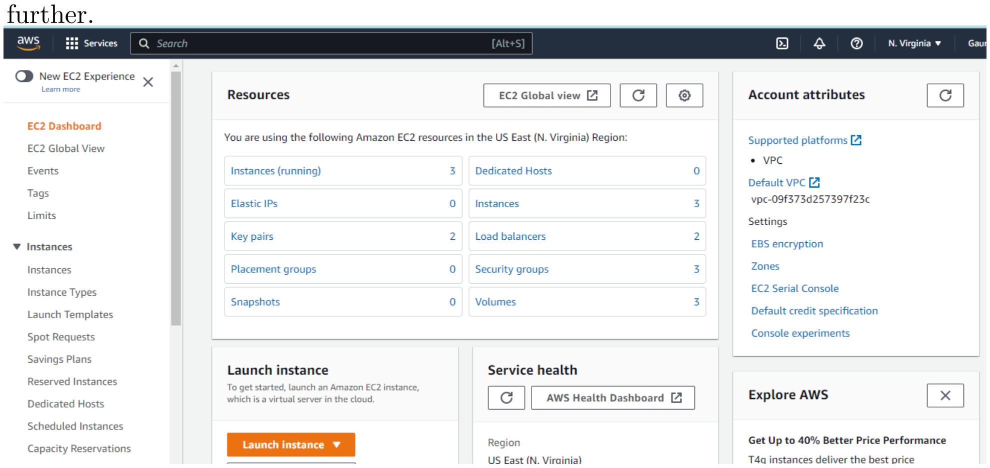
We have many platform to create private data center like Federate with Cloud Lab, Mass Open Cloud projects, The Open Cloud Testbed, Open Cirrus, TerraGrid, IBM/Google, and AWS. Some are paid to use and some gives us free trial for month. We are performing our experiment in AWS We will performing VPC, EC2, Load Balancing, Create instance, Subneting and etc

4.2.1 Component

VPC: Amazon Virtual Private Cloud (VPC) allows you to launch your AWS resources into a virtual network that you define. This virtual network is much like a traditional network running in your own data center with the benefits of leveraging the scalable infrastructure of AWS.

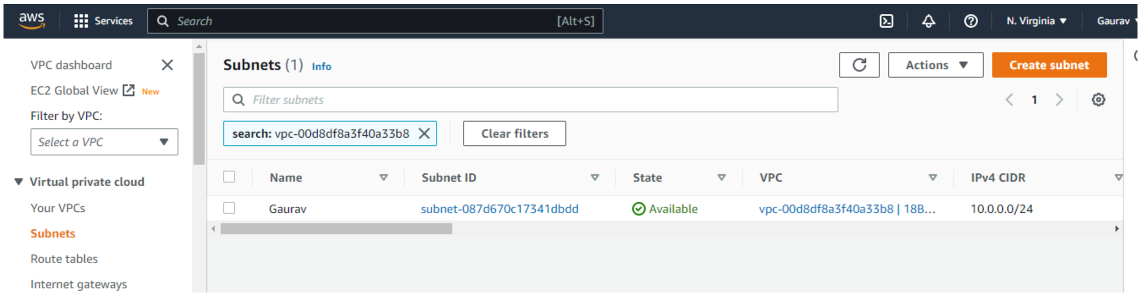


EC2: It is Amazon Elastic Compute Cloud. It's a web service which allow you to create compute capacity in cloud. It contain component which you will see further.

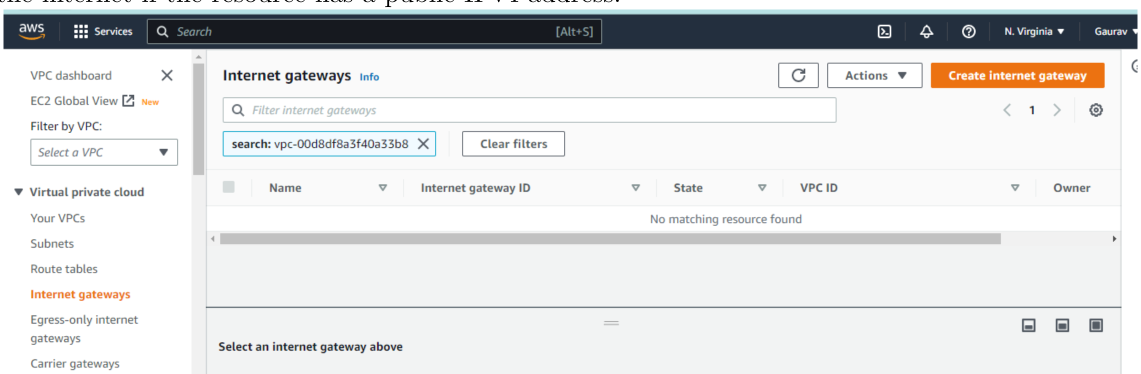


Subnet: When network consists more than one IP address range that called as Subnet.
AWS has 4 types of subnet:

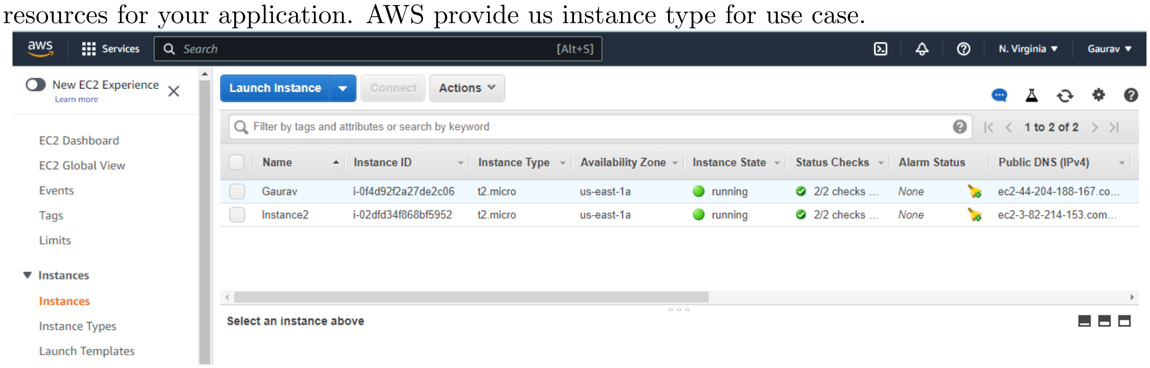
- A single public subnet only.
- A public and private subnets.
- A public and private subnets and AWS Site-to-Site VPN access.
- A private subnet only and AWS Site-to-Site VPN access



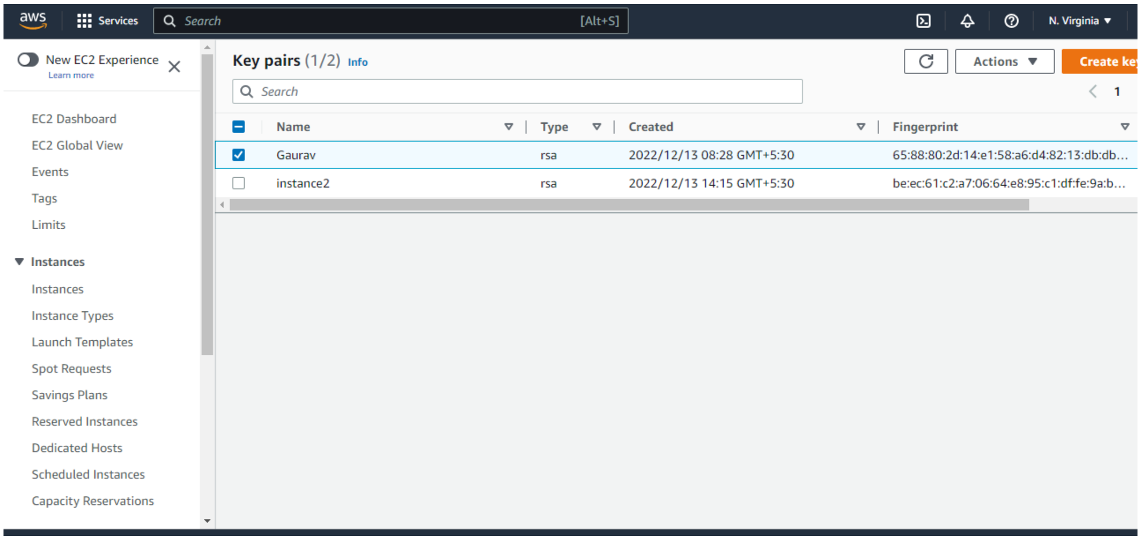
Internet Gateway: Internet Gateway is component that allow communication between VPC and Internet. It enables resources in subnets to connect to the internet if the resource has a public IPv4 address.



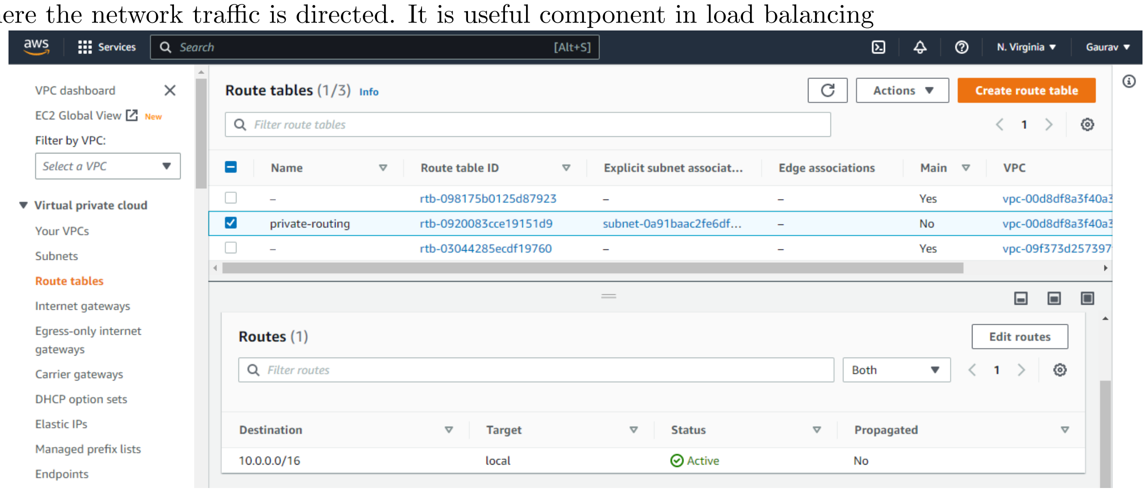
Instance: Instance types include different combinations of CPU, memory, storage, and network capacity, giving you the flexibility to choose the right mix of resources for your application. AWS provide us instance type for use case.



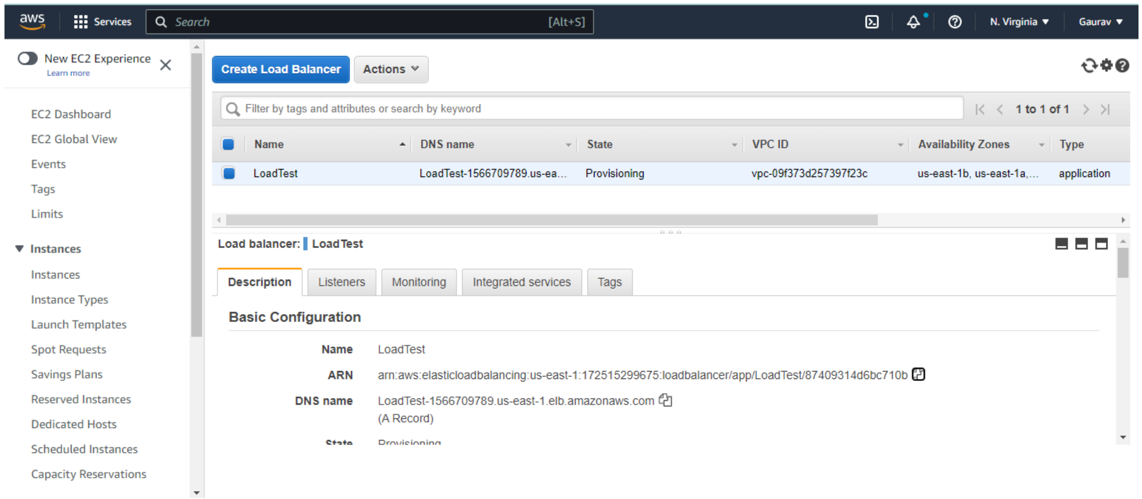
Key Pairs: Key Pair is component which store pubic and private key to conform user connect to AWS instance. When you download key pair file it have extension ppk or pem.



Route Table: It provide set of rules, called routes. It help to find out where the network traffic is directed. It is useful component in load balancing



Load Balancing: Load balancing is a core network solution used to distribute traffic across multiple servers within a server farm. The load balancer improves application availability and responsiveness and prevents server overload. Each load balancer sits between a client device and a backend server, receiving incoming requests before distributing them to any available server that can fulfill them. You can learn load balancing using AWS instances.



So basically load balancer work is to manage application server that if request is occure than it goes direct to server which will have less traffic. Using this method it reduce the traffic and fully utilize the server

Target Group is set of instance we created to perform load balancing. We will add one instance in target group which subnate route traffic is more and add inctance which has no traffic.So whenever client tried to opened, because of intance have more traffic instead of instance 2 will be activate.

5 Result Analysis:

5.1 Carbon Computation

Total Carbon Intensity from interval 01-01-2020 to 01-01-2021 in France Germany California Great Britain

	Germany	California	Great Britain	France
count	17569.000000	17569.000000	17569.000000	17569.000000
mean	313.396106	279.706220	212.942491	56.336777
std	110.454244	72.856504	65.291795	19.102152
min	100.657858	53.934563	63.689930	19.494886
25%	222.764276	230.823460	160.105196	41.593088
50%	312.549830	291.916133	210.428638	57.423619
75%	389.590156	338.124076	259.942319	69.287476
max	593.142244	427.009412	394.860422	116.394154

	Germany	California	Great Britain	France
Time				
2020-01-01 00:00:00	353.319875	342.092256	192.229224	56.033471
2020-01-01 00:30:00	351.252335	347.220539	194.251856	52.740640
2020-01-01 01:00:00	349.628409	344.967632	195.211861	48.375163
2020-01-01 01:30:00	346.373119	339.952418	190.238991	43.531716
2020-01-01 02:00:00	347.674818	336.563101	186.122673	40.306003
...	...	...	...	...
2020-12-31 22:00:00	426.575116	376.835143	251.141608	69.073480
2020-12-31 22:30:00	426.077061	374.563315	243.237501	68.195328
2020-12-31 23:00:00	427.580284	371.903162	233.778838	65.697676
2020-12-31 23:30:00	426.458266	368.503891	229.335018	63.288589
2021-01-01 00:00:00	424.369909	365.636659	230.679267	62.239152

17569 rows × 4 columns

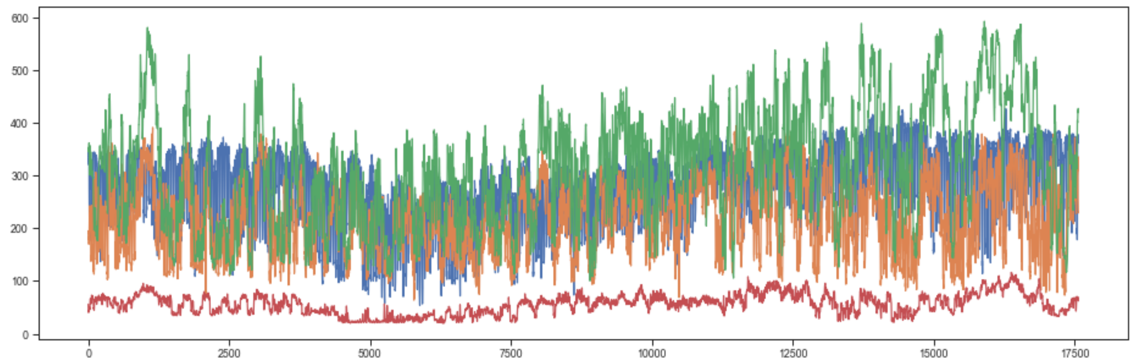


Figure 1: Time Series

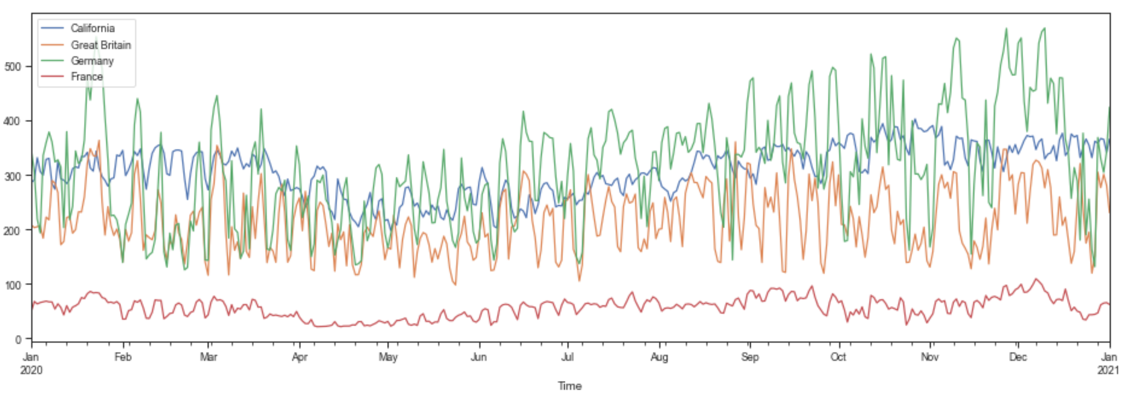


Figure 2: Time Series Smooted

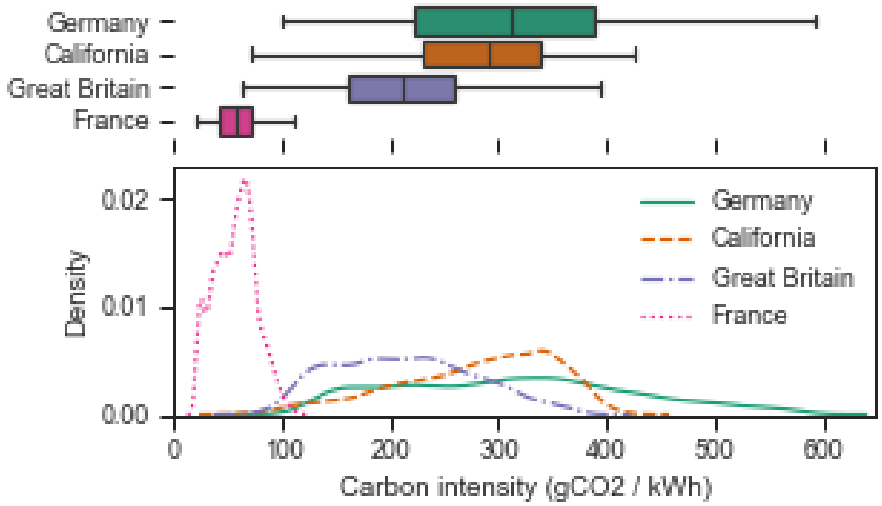


Figure 3: SNS plot of Population Density with carbon intensity:

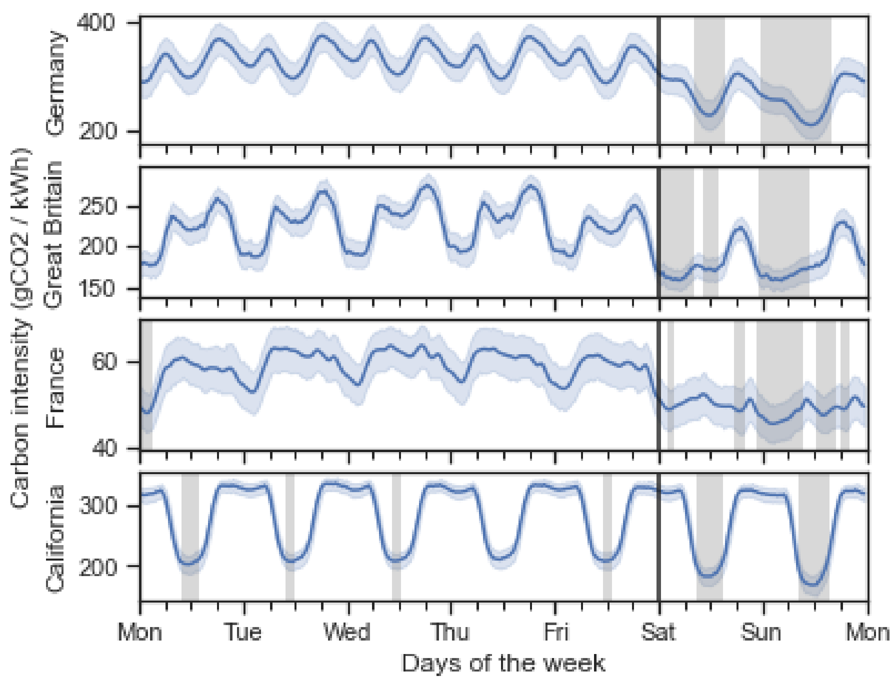


Figure 4: Weekly Analyse and Computing Carbon Intensity of different work-days and weekends:

```
Germany
[330.78029036]
[246.69745982]
25.42%

Great Britain
[222.64568388]
[172.15334183]
22.68%

France
[60.91000694]
[47.42044097]
22.15%

California
[289.73837365]
[271.76911269]
6.20%
```

Figure 5: Average Carbon Intensity in every country:

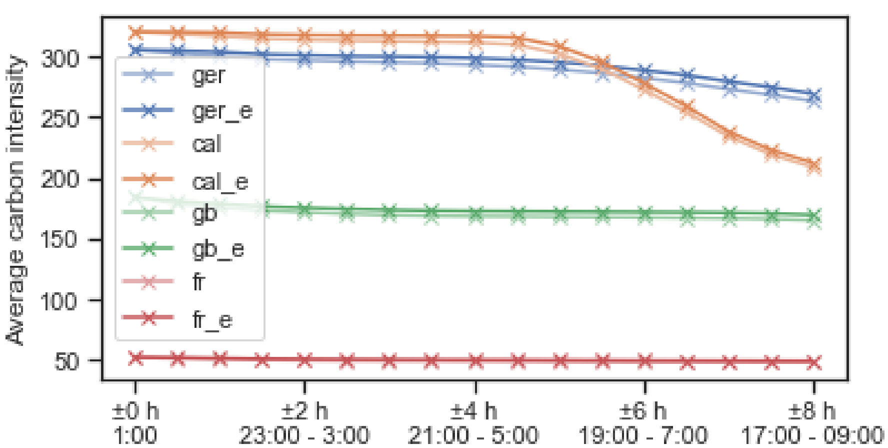


Figure 6: Average Carbon Intensity in every country:

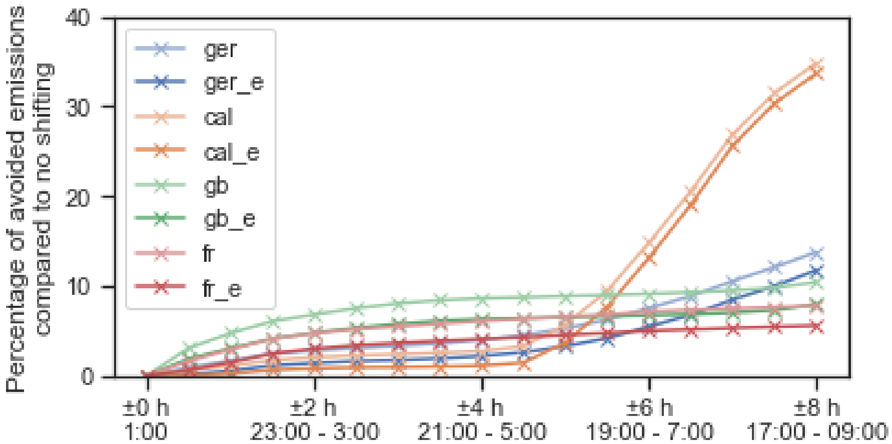


Figure 7: Percentage of Avoided Emissions in no shift.:

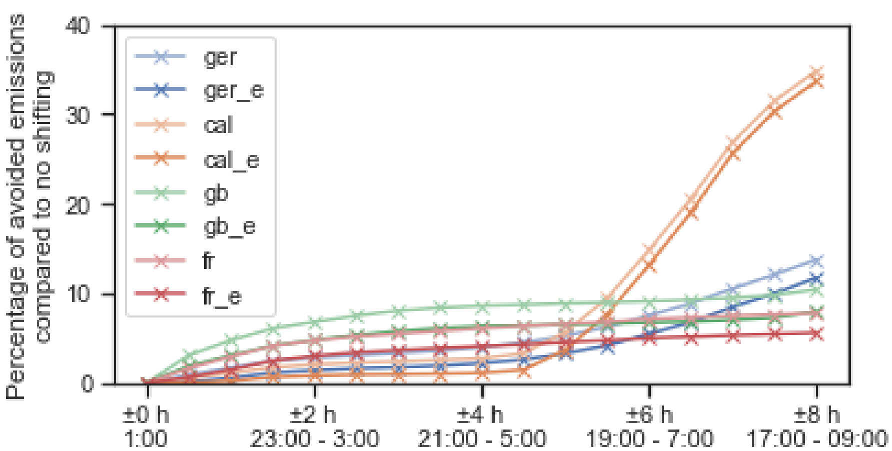


Figure 8: Percentage of Avoided Emissions in no shift.:

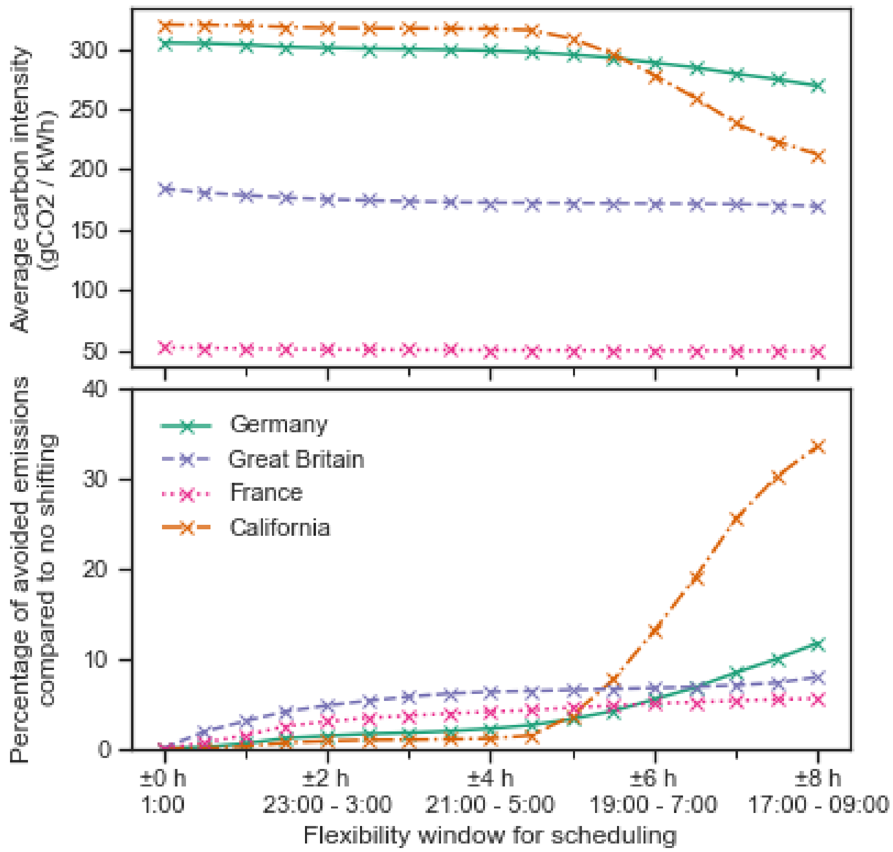


Figure 9: Percentage of Avoided emissions with Carbon Intensity in every country per flexibility window for scheduling.Having Optimal and 5 Percent Error :

```
cal_hist = [0.1, 0.1, 0.0, 0.0, 0.1, 0.1, 0.0, 0.1, 0.1, 0.1, 0.1, 0.1, 0.0, 0.5, 2.4, 11.4, 33.3, 89.8, 218.5, 0.0,
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 4.9, 0.5, 0.0, 0.2, 0.1, 0.1, 0.2, 0.2, 0.1, 0.3,
0.3, 0.1, 0.2, 0.0]
gb_hist = [20.1, 18.4, 13.8, 19.7, 20.1, 17.6, 20.2, 24.9, 24.9, 19.0, 13.0, 9.1, 8.4, 6.3, 6.3, 5.3, 8.4, 11.9,
21.9, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 7.1, 4.0, 1.1, 0.9, 0.9, 1.1, 2.0, 1.8,
3.1, 4.0, 5.0, 9.8, 14.4, 19.5]
ger_hist = [9.0, 13.3, 14.1, 14.9, 16.1, 14.5, 12.2, 10.2, 6.4, 3.5, 2.3, 2.3, 1.6, 2.2, 4.5, 8.4, 19.0, 43.4, 95.7,
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 14.9, 6.1, 4.4, 2.2, 2.4, 2.1, 2.8, 2.8, 2.0,
2.8, 5.3, 6.2, 7.7, 8.7]
fr_hist = [10.7, 14.3, 19.5, 24.1, 24.8, 25.0, 24.7, 15.5, 12.2, 8.0, 6.3, 4.9, 5.8, 5.5, 5.1, 5.2, 8.3, 7.6, 10.0,
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 10.5, 8.8, 10.9, 10.5, 12.7, 9.0, 8.5, 7.4,
6.2, 5.4, 7.2, 10.4, 9.0, 10.0]
```

Figure 10: Dataset for Demand of Jobs relative to supply

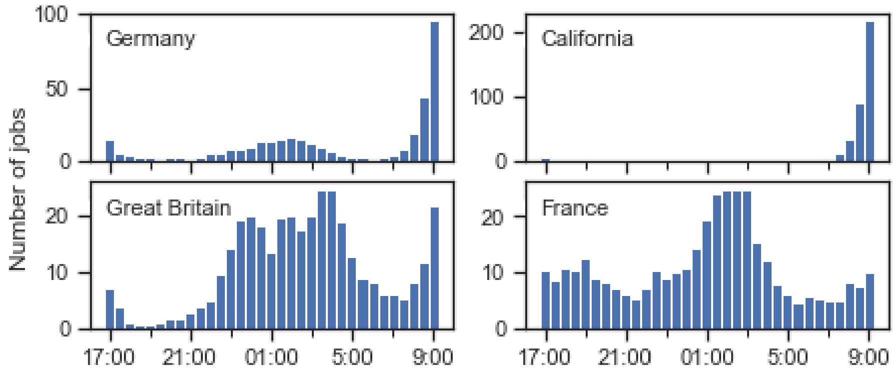


Figure 11: Number of Jobs availability

	active_jobs	ci	emissions
Time			
2020-01-01 00:00:00	0.0	342.092256	0.0
2020-01-01 00:30:00	0.0	347.220539	0.0
2020-01-01 01:00:00	0.0	344.967632	0.0
2020-01-01 01:30:00	0.0	339.952418	0.0
2020-01-01 02:00:00	0.0	336.563101	0.0

Figure 12: Active Jobs Head 5

```
e_b = baseline["emissions"].sum() / 2 / 1000000
e_x = x["emissions"].sum() / 2 / 1000000
e_xi = xi["emissions"].sum() / 2 / 1000000
Non-interruptible: {1 - e_x/e_b:.2%}
                                   Germany
                                   Non-interruptible: 5.91%
                                   Interruptible:     7.93%

                                   Great Britain
                                   Non-interruptible: 6.56%
                                   Interruptible:     9.03%

                                   France
                                   Non-interruptible: 4.77%
                                   Interruptible:     5.92%

                                   California
                                   Non-interruptible: 2.45%
                                   Interruptible:     5.71%
Interruptible:    {1 - e_xi/e_b:.2%}
```

Figure 13: Non Interruptible , Interruptible Emissions Calculation

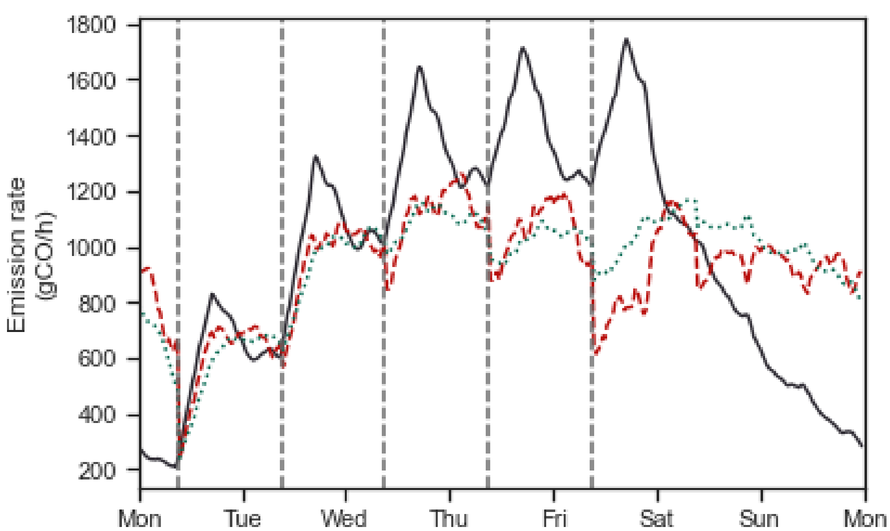


Figure 14: Continuous Integration on Emission Rate via Work Days. (– Interrupting , : Non Interrupting)

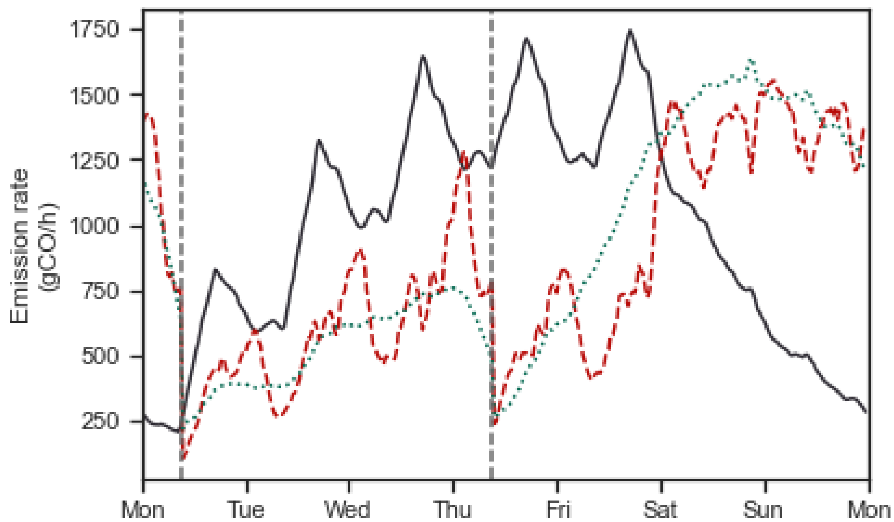


Figure 15: Continuous Integration on Emission Rate via Work Days. (– Interrupting , : Non Interrupting)

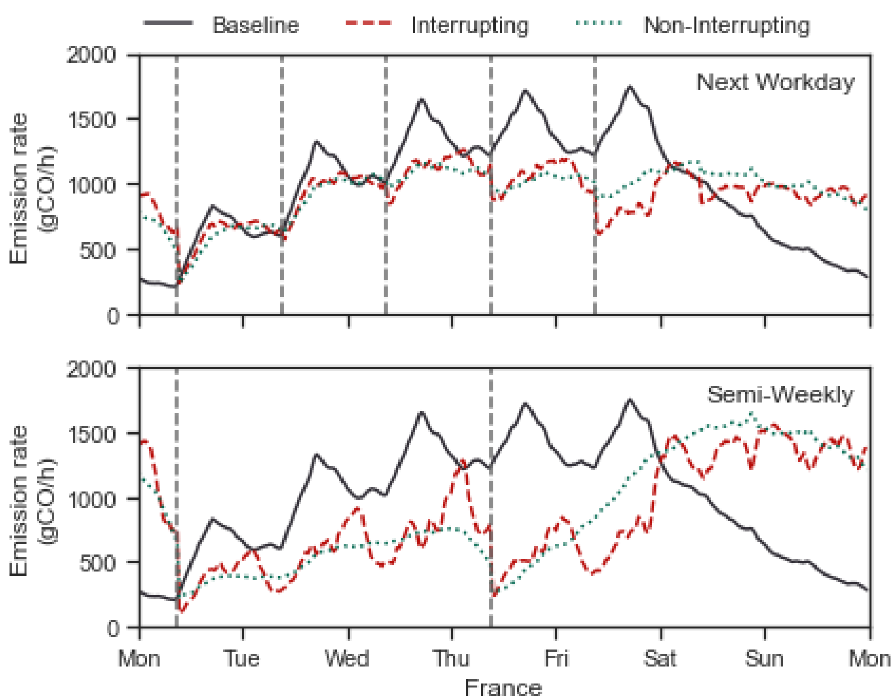


Figure 16: Emission Rate on Next Work Day , Semi Weekly in France

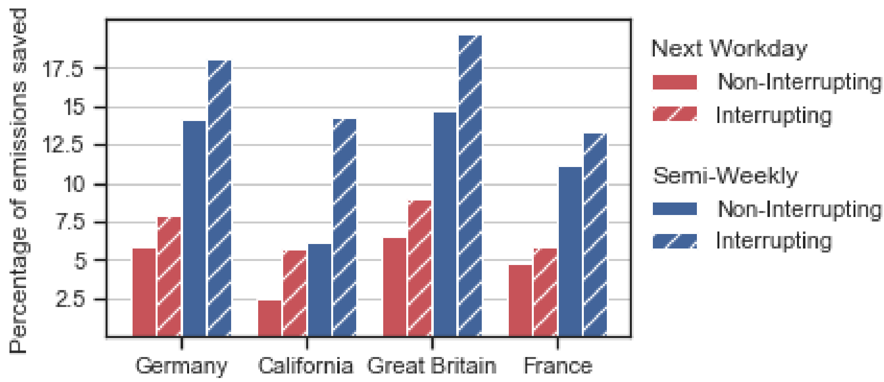


Figure 17: Carbon Emission Saved on Next Workday , Semi-Weekly in Germany,California,Great Britain,France (No Error)

Figure 18: Carbon Emission Saved on Next WorkDay having 5 percent error 10 percent error.

5.2 Methodology Used:

- Applies DVFS to provide a balance between power consumption and tasks deadline using GridSim toolkit.
- Virtual Machine Scheduling and Migration Algorithm is capable of performing the scheduling of Virtual Machines in non-federated homogeneous and heterogeneous data centers, and also improves power consumption in loads using CloudSim
- Load balancing Algorithm is good in reducing energy, pricing and time using AWS.
- Energy Consumption modelling, and analysis approaches, It helped to identify the relationship between energy consumption and running tasks in cloud environments, as well as system configuration and performance

6 Conclusion:

As explained above, all existing architectures have both constructive and destructive aspects. Buya et al. describe the Green Cloud architecture, stating that the main advantage of this architecture is the carbon footprint. Since this register measures the best service with the lowest CO2 emissions, you can quickly see that energy consumption is directly proportional to CO2 emissions that also reduce energy. Similarly, the downside is that CO2 emissions and energy are not the only factors to consider such as quality assurance and safety. This paper explores data center management considering utilization and related factors with the goal of increasing energy efficiency and improving performance. We have provided some background on AWS cloud data center and an overview of the usage rates supported by data in medium to large data centers in a university environment.We perform load balancing in AWS to see how traffic is controlled

7 References:

[1] Timoshenko, S. and Woinowsky-Krieger, S., Theory of Plates and Shells. Second Edition, McGraw-Hill, 1959. [2] Belytschko, T., Lu, Y.Y. and Gu, L., "Element-free Galerkin Methods". International Journal for Numerical Methods in Engineering; Vol. 37, 1994, pp. 229-256. [3] Ma, H.T. and Kanok-Nukulchai, W., "On the application of assumed strain methods", Proceedings of the second East Asia-Pacific Conference on Structural Engineering Construction (EASEC-2), Chiang Mai, Thailand, January 11-13, 1989

[8] Eppli. M., and Shiling. J. D., "How Critical is a Good Location to a Regional Shopping Center?" , Journal of Real Estate Research, Vol. 12, no. 3, pp.459-468,1996. [9] Kant, K., "Data center evolution A tutorial on state of the art, issues, and challenges", Computer Networks 53, pp.2939-2965, 2009.

[4] Cough, R. W. and Penzien, J., “Response to a general dynamic loading”, Chapter 7, Dynamics of Structures, McGraw-Hill, Singapore, 1975. [5] Alharthi, A., “Frameworkc For Selecting Data Center”, 2014. [6] Addis, B., Ardagna, D., Zhang, L., and Panicucci, B., “Automatic Management of Cloud Service Centers Availability Guarantees”, IEEE 3rd International Conference on Cloud Computing, 2010.

[8] Eppli. M., and Shiling. J. D., “How Critical is a Good Location to a Regional Shopping Center?” , Journal of Real Estate Research, Vol. 12, no. 3, pp.459-468,1996. [9] Kant, K., “Data center evolution A tutorial on state of the art, issues, and challenges”, Computer Networks 53, pp.2939-2965, 2009. [10] Al-Fares. M., Loukissas. A., and Vahdat. A., “A Scalable, Commodity Data Center Network Architecture”, SIGCOMM, pp.63-74, 2008. [11] Jackson, J.M., Koomey. J. G., Nordman. B., and Blazek. M., “Data center power requirements: measurements from Silicon Valley”, Elsevier, pp.837-850, 2003. [12] Koomey. J. G., “Growth In Data Center Electricity Use 2005 To 2010”, Analytics Press, August, 2011 [13] M. Kaur and P. Singh, (Eds.), “Energy Efficient Green Cloud: Underlying Structure”, Proceeding of the IEEE international conference of theEnergy Efficient Technologies for Sustainability (ICEETS), (2013) April 10-12, Nagercoil. [14] L. Hosman and B. Baikie, (Eds.), “Solar-Powered Cloud Computing datacenters”, vol. 2, no. 15, (2013).
[15] http://en.wikipedia.org/wiki/Cloud$_c$omputing.[16]*Cloudweaks*, *Cloudweaks, thefreeresearchportal*", (20
//*research.cloudtweaks.com/technology/networking/cloud$_c$omputing*.[17]*F.OwusuandC.Pattinson*, (*Eds.*),
27; *Liverpool*.[18]*R.Yamini*, (*Ed.*), *PowerManagementinCloudComputingUsingGreenAlgorithm*", *Procee*
InternationalConferenceonAdvancesinEngineering, ScienceandManagement(*ICAESM*)(2012), *March*30-
31; *Nagapattinam, TamilNadu*.[19]*S.K.GargandR.Buyya*, *GreenCloudcomputingandEnvironmentalSusta*
IEEEPressEbook(2012), *Edition*1, *no*.3, *pp*.76−87.[20]*D.H.Heo, X.LiuandT.Abdelzaher*, (*Eds.*), *Integratin*
Anemergingchallengein.