university of groningen

faculty of science and engineering

Calculating single customer power consumption in a multi-tenant data center

Research Internship

Pieter Dekker

Supervisor:

Dr. Vasilios Andrikopoulos

Computer Science Department
Faculty of Science and Engineering
University of Groningen, The Netherlands

Second supervisor:

Christos Bantis

Product Manager
British Telecom Global Services Amsterdam, The Netherlands

Contents

1 Introduction **2**

2 Related work **3**
2.1 Previous work in this project . . . 3
2.2 Data center power consumption modeling . . . 4

3 Domain analysis and requirements **4**
3.1 Data center design . . . 4
3.2 Datacenter power model . . . 5
3.3 British Telecom Global Services' data center situation . . . 6
3.4 Project Requirements from British Telecom Global Services . . . 7

4 The model **8**
4.1 Preliminaries . . . 8
4.2 The model . . . 9
4.2.1 Calculating utilization level . . . 10
4.2.2 Calculating direct power consumption . . . 11
4.2.3 Calculating indirect power consumption . . . 11
4.2.4 Calculating Amenity Power . . . 12
4.2.5 Calculating total power consumption . . . 13
4.2.6 Note on data availability . . . 13

5 Discussion **13**
5.1 Completeness and correctness . . . 13
5.2 Accuracy . . . 14
5.3 Missing data . . . 14
5.4 Model evaluation . . . 14

6 Conclusion **14**

A Parameters to collect **15**

B Recommendations for implementation **16**

1 Introduction

In recent years, cloud based solutions for many business activities have become more and more favoured over their on-premise counterparts. These services are offered by running the required software on servers in large data centers and connecting these data centers to the Internet. This has led to an increasing number of such data centers being built, their designs and related technologies being improved, and the total power consumed by these data centers to increase. It is expected that the power consumption of data centers worldwide will grow to somewhere between 8 and 21 percent of global power consumption by 2030 [2]. It is therefore interesting to gain insights into the power consumption of the activities of a user of such a data center.

British Telecom Global Services is multi-national telecommunication services company that offers a number of cloud based or cloud related services, one of which is a contact center solution: Cloud Contact Cisco. This service provides customers with an omni-channel contact center, provided entirely with Cisco hardware. The customers therefore no longer have to own and maintain their own servers, nor do they have to have the required infrastructure to house these servers. One of the many benefits of cloud based services is that they allow for the sharing of infrastructure, which intuitively leads to lower power costs over time. British Telecom Global Services now wants to build a dashboard that provides insight for a customer into exactly how much power is used for the provision aforementioned contact center service. The dashboard should show the total power consumed in provision of Cloud Contact Cisco, for each customer. Next to that, the output of the dashboard should be auditable.

The problem that arises when attempting to compute such figures is that not all of the data center infrastructure is owned by British Telecom Global Services. The services are hosted in multi-tenant data centers. In the specific case of British Telecom Global Services and Cloud Contact Cisco this means that the networking infrastructure is owned entirely by British Telecom Global Services, while the building, cooling, lighting, et cetera are owned by a third party providing a service to British Telecom Global Services. Collecting data on the power consumption of the infrastructure owned by British Telecom Global Services is straight forward, although not trivial. The data collection on power consumption of the infrastructure that is not owned by British Telecom Global Services, however, is a complicated task. First of all, the data needs to be obtained from third parties. Secondly, with servers placed in a room together with servers of other companies, the power consumed by the cooling installation for that room may be known, but not all of that power is consumed on behalf of British Telecom Global Services. In this report we will attempt to develop a model for calculating the power consumption of a service deployed in multi-tenant data centers, and can account on a per client basis.

Below are some terms to ease the reading of the rest of this report.

- **Client**
  A client of British Telecom Global Services that uses the Cloud Contact

Cisco service. Each client may have many users that use the Cloud Contact Cisco service.

- **Infrastructure**
  The whole of servers, switches, routers, firewalls, etc. owned, maintained and operated by British Telecom Global Services to provide the Cloud Contact Cisco service.

- **Data center**
  The building that hosts the network infrastructure British Telecom Global Services uses to host Cloud Contact Cisco

- **Data center infrastructure**
  The whole of power consuming components of the data center that are needed for a data center in operation. Examples: heating, ventilation, air-conditioning and cooling (HVAC), lighting, closed circuit television (CCTV), et cetera.

- **Data center operator**
  The party that owns and manages the data center.

In this report we will develop a model of the power consumption of a service deployed in a multi-tenant data center, where some parts of the infrastructure are not owned by the provider of the service. The main goal is to develop a model to guide the calculation of the power consumed while serving a specific customer over a certain period of time in a particular data center.

The rest of this report is organized as follows: Section 2 mentions the related research in this area. In Section 3 we will explore the domain of application, identifying and describing the necessary concepts and components. After that, we will present the developed model that has its basis in the results of Sections 2 and 3. We discuss the model in Section 5. Finally, we draw our conclusions and discuss future work in Section 6.

2 Related work

2.1 Previous work in this project

This report is part of the second phase in a larger project with British Telecom Global Services. The first phase explored different metrics to express power usage, efficiency and environmental impact of a data center. The output of that phase was a selection of suitable metrics and a proof-of-concept dashboard to display these metrics. In [4] the author mentions power usage effectiveness (PUE) as an interesting metric. PUE compares the power consumption of IT devices in a data center such as servers, routers, and switches to the total power consumption of a data center. One important outcome of that phase of the project for this report is that it is important to make a distinction between power consumed by IT devices and power consumed by anything else in a data center.

2.2 Data center power consumption modeling

In this report we will be developing a model of the power consumption of the service provided by British Telecom Global Services. For this, we need to consider a complete model of the power consumption of a data center, as well as certain specific components.

A review of existing data center power models can be found in [1]. The authors explore power consumption models of devices on different levels of coarseness, going from individual digital circuits to devices as a whole. Knowing, as will be elaborated below, that the power consumption of the devices in question in this work is directly available, we will not concern ourselves with these models.

The authors also discuss models of all other aspects of power consumption of a data center, with of special interest [3] and [5]. In [3], we see that the power consumption of a computer room air conditioning (CRAC) installation is not fixed, but rather dependent on both the temperature of the inlet air as well as the temperature of the outlet air. In that work, a coefficient is derived, the coefficient of performance (COP) that can be used to calculate the power consumption of a CRAC installation based on air inlet temperature and air outlet temperature. The power consumption of a CRAC is researched in [5] which leads to an empirical model of CRAC power consumption based on only inlet air temperature and device power consumption, which may be useful in cases where the outlet temperature cannot be measured.

What the authors of [1] did not find in their review is a model for computing power consumption in a multi-tenant data center. This confirms the relevance for this work as filling a knowledge gap in the state of the science and the industry.

3 Domain analysis and requirements

3.1 Data center design

A data center can be viewed as having two parts. In the first place there is the computer room, in which racks are placed in rows. On these racks servers, switches, firewalls and routers are mounted, which provide the main functionality of the data center. Secondarily, there is the cooling, lighting, CCTV and power supply. These parts are essential, but do not directly create value for clients.

Cooling in a data center is done centrally, with an installation called computer room air conditioning (CRAC). Each rack is individually cooled by cool air that is blown upward along one side of the rack by the ventilation system. This air is taken in by the devices, heat is exchanged with the air inside the device and the hotter air is blown out of the device by fans inside the device. This air is then taken in from the other side of the rack by the ventilation system and transported to the CRAC, where the air is cooled and recirculated. Up-time is a key performance indicator for a data center, therefore the devices in the server racks are each redundantly powered by an uninterruptible power supply (UPS).

Next to that, the data center needs CCTV for security, lighting for the personnel that walks around in it and other such amenities. The power consumption is constant and independent of the activity of the devices that the data center houses.

3.2 Datacenter power model

With respect to the components mentioned above, we consider three categories of components: (1) components owned by British Telecom Global Services; (2) components owned by the data center operator whose power consumption is influenced by activity of components of category 1; (3) components owned by the data center operator whose power consumption is not influenced by the activities of category 1. It is important to note here that we do not consider components owned by other tenants of the data center. We will develop a model that only requires knowledge of the three categories mentioned above.

For components in the first category we can measure the immediate power consumption. In this category we find the following components:

- Server
  Performs the work needed to actually provide the service. It contains all the hardware needed to perform that work, namely CPU, storage, memory, and connectivity peripherals.

- Switch
  Connects servers to each other and to firewalls and connects firewalls to routers

- Firewall
  Restricts traffic based on filtering rules

- Router
  Routes traffic from and to servers

- Edge router
  Routes traffic between data centers and to the Internet

- Uninterrupted Power Supply
  Connects devices to the power net

For this category, to calculate the proper division of the power consumed over the clients served by the components in this category we need to consider the level of activity of each component. Any of these component performs some activity, like network switching, with some level of intensity, like total packet throughput. The total throughput is the sum of the throughput performed for each client individually. The power consumed by the component on behalf of a specific client is proportional to the ratio of the activity level for that client to the total activity level. Multiplying this ratio with the total power consumption yields the power consumed on behalf of a specific client.

In the second category, we find cooling. Cooling is dependent on the activity level of the devices operated by British Telecom Global Services. Every device generates a certain amount of heat that is transferred to the air around it, and that heat needs to be removed from the air so that the heat transfer may continue. The amount of heat generated is dependent on the amount of power consumed by the device which is in turn dependent on the amount of work the device has to perform, or the activity level. All the power consumed by a computing device is eventually turned into heat, be it by running through the semi-conducting circuits or friction of moving components.

In the last category we find amenities such as lighting and CCTV, we will refer to these as amenities from here on. These exist regardless of the level of work performed by the components owned by British Telecom Global Services, or those of any other tenant of the data center. The power costs in this category are constant and is to look at the ratio of floor area used by a single tenant of the data center to the floor area used by all data center tenants combined.

3.3 British Telecom Global Services' data center situation

British Telecom Global Services rents data center real estate in data centers across the world to provide the Cloud Contact Cisco service to customers around the world. In each of these data centers they own one or more server racks, in which the necessary devices are mounted. Each data center has other tenants that have their own devices installed and running there. The total cooling required for the entire data center is therefore compounded by the heat generated by devices owned by several parties. The cooling required for British Telecom Global Services' activity and the power consumed by that process is therefore a part of the total cooling required for the entire data center. It is typically not known what party caused the need for what portion of the total cooling required.

The devices British Telecom Global Services mounts in their server racks are all Cisco devices. The types of devices are servers, switches, routers, and firewalls. The servers run VMWare software running one or more virtual machines (VMs) that provide the Cloud Contact Cisco service to one client each. Each server may run VMs for different clients. Each server has some capacity reserved for a certain client of British Telecom Global Services. Each server is connected to two switches. Each switch is connected to a firewall, each firewall to a router. Each router then is connected to another switch which is connected to two MPLS routers, one for inbound traffic, one for outbound traffic. These MPLS routers connect the devices in each data center occupied by British Telecom Global Services to create a WAN.

All the devices in the racks owned by British Telecom Global Services are potentially shared by multiple clients of British Telecom Global Services. A server may host several virtual machines for different clients. Consequently, a

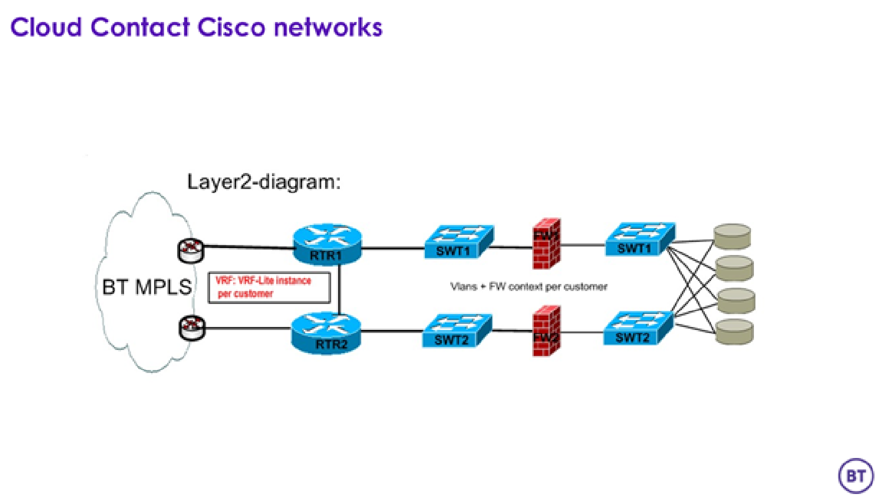

Figure 1: The topology of devices in one data center dedicated to providing Cloud Contact Cisco to one client.

switch may switch traffic for different clients, a router may route for different clients, and a firewall may filter traffic for different clients. A virtual machine, however, will always be dedicated to a single client.

The Cisco devices all have management interfaces that can be accessed either locally or remotely. Through this interface it is possible to extract power consumed, CPU utilization, memory utilization, uploaded bytes, downloaded bytes over 20 second intervals on a per client basis. Because the power consumption of a device is directly available it is not necessary to model de individual components of a device, such as CPU, memory, and storage.

To provide Cloud Contact Cisco to its clients, British Telecom Global Services runs VMs with reserved resources for each client. All resources are set up in a fault tolerant manner. Each client has two dedicated VMs on two separate servers within a single data center. Each server is connected to two switches. Each of these switches is then connected to a firewall. The firewall is connected to another switch which is connected to a router, which is connected to an MPLS router. The two VMs in one data center are continuously synchronized.

3.4 Project Requirements from British Telecom Global Services

The dashboard British Telecom Global Services wants to provide to its customers has three main requirements: (1) it should be possible to show the power consumption per client; (2) the dashboard in its entirety should be auditable, meaning that it should be possible what calculation output was shown to which client at what time and why that output was shown; (3) the confidence in the output shown should be shown alongside the output itself, in the cases where a figure is known to be less than 100% accurate.

To support the first requirement, we need calculations for each type of device to account the total power consumption of a device to the specific clients served by that device. For this purpose, the data extracted from the management interfaces can be used. To account power consumption of servers to specific clients, the power consumption per virtual machine may be used. To account power consumption of the other devices to specific clients, the utilization of that devices resources should be calculated as a ratio to the sum of utilization of its resources. An example would be to take the ratio of uploaded bytes to a devices by a specific client to the total bytes uploaded to that device.

For the second requirement, all inputs to calculations should be stored for the required duration of auditability. The third requirement requires all calculations that involve some uncertainty to include that uncertainty into the calculation in the form of an error term.

4 The model

4.1 Preliminaries

The total power consumption by and for British Telecom Global Services in a single data center is a sum of the power consumed per device owned by British Telecom Global Services, the power consumed by the cooling installation to remove the heat generated by British Telecom Global Services' devices, and British Telecom Global Services' share of the power consumed by amenities. Power consumed is always in Watts over a certain time period.

The power consumption per owned device can be extracted through management interfaces available in the Cisco devices. Through the same management interface it is possible to calculate the activity level per client as a ratio of the total activity over 10 second intervals. The power consumption for a specific client is obtained by multiplying the power consumption of a device by that ratio.

To derive the power consumed by the CRAC installation from the heat generated by the device, we need to know the coefficient of performance (COP) of that CRAC installation, that is the ratio of power consumed to heat removed. This coefficient is dependent on the outlet temperature of the CRAC and the temperature of the air flowing back to the CRAC. Together with the power consumed by the fans of the CRAC we can then derive he power consumed by the CRAC in providing cooling. Both heat and power can be expressed in the same unit, Watts, COP is then a ratio of Watts to Watts, or as a unit: Watt/Watt.

The power consumption for a UPS can be obtained by knowing the constant rate of that UPS and the power consumed by the device it is supplying power for. Again, to obtain the power consumption for a specific client this quantity should be multiplied by the activity level for that client on the device for which the UPS is supplying power. The provision of power by a UPS costs some power, which relates to the amount of power provided.

The coefficients of performance of CRAC and UPS are not constant, but dependent on environmental factors, such as supply air temperature, inlet air temperature and humidity in case of CRAC or amount of supplied power in case of UPS. These coefficients, which will be defined below, must be determined periodically. If possible, these must be updated as often as possible. However, this will require access to facilities that are not owned by British Telecom Global Services. Therefore, it is necessary to account for error in the model.

The power consumed for amenities is constant and independent of the activity in the devices. To obtain the total power consumption in this category the data center operator needs to be consulted. The power consumption of amenities is independent of the activities of the tenants of the data center. Therefore, to derive this power consumption for British Telecom Global Services we will take the floor area reserved for British Telecom Global Services and divide it by the total floor area of the data center and multiply the total power consumption of the amenities by this ratio to get British Telecom Global Services' power consumption of amenities. To then derive the power consumption per client, we take capacity reserved for that client and divide it by the the total capacity of British Telecom Global Services' devices. Multiplying British Telecom Global Services' power consumption of amenities by this ratio yields the power consumption of amenities for this client.

4.2 The model

Below we will present the model for calculating power consumption of the provision of Cloud Contact Cisco to one of British Telecom Global Services' clients. The calculations are performed on measurements obtained from the management interfaces of the Cisco devices used. The data obtained from these interfaces are averages over 20 second intervals. The direct power consumption of the devices can be obtained from these interfaces, and the only calculation necessary on it is to account the power consumption to specific clients.

In Table 4.2 we see a list of symbols used from here on and their meaning. We will use subscripts to further define the quantities a symbol refers to. For example: where P_D refers to the power consumed by device D, $P_{D,c}$ refers to the power consumed by device D for client c.

U_c	The utilization of a device for client c. Refers to how much of the devices activity was to serve that client as a fraction of the total activity.
P	Power consumed
PC	Power consumed to provide cooling
PU	Power consumed to provide power to a device
H	Heat produced
η_{CRAC}	COP of the CRAC
$\eta_{UPS,D}$	COP of the UPS of device D
$\epsilon_{CRAC,D}$	The error term in calculating the power consumed by the CRAC to cool the heat generated by device D
$\epsilon_{UPS,D}$	The error term in calculating the power consumed by the UPS of device D
Dn	Bytes downloaded during a time interval
Up	Bytes uploaded during a time interval
T	Throughput during a time interval: the sum of bytes uploaded and bytes downloaded during that interval
S	A server
X	A switch
F	A firewall
R	A router

When it comes to the indirect power consumption, the power consumption of UPS and CRAC, we need to make use of coefficients relating the performance of these components to the power consumed by the devices. These, as mentioned above, may be updated at a sub-optimal frequency. This will introduce a structural error in the calculations, for which we will account by introducing error terms.

4.2.1 Calculating utilization level

The utilization level U of certain client c for a device D is defined as follows:

- Switch:

$$U_{X,c} = \frac{T_{X,c}}{T_X}$$

$$T_{X,c} = Up_{X,c} + Dn_{X,c}$$

$$T_X = Up_X + Dn_X$$

- Firewall:

$$U_{F,c} = \frac{T_{F,c}}{T_F}$$

$$T_{F,c} = Up_{F,c} + Dn_F$$

$$T_F = Up_F + Dn_F$$

- Router:

$$U_{R,c} = \frac{T_{R,c}}{T_R}$$

$$T_{R,c} = Up_{R,c} + Dn_{R,c}$$

$$T_R = Up_R + Dn_R$$

The resulting quantities represent the ratios between 0 and 1 that represent the amount of power consumed for a specific client. If the device only performed work for one client, the utilization ratio for that client is 1 and it is 0 for all others and all power consumed by that device is accounted to that particular client. The sum of all utilization ratios over some interval on a particular device is always going to be 1.

4.2.2 Calculating direct power consumption

For servers, the power consumption $P_{S,c}$ for a client c can be obtained directly from the management interface, by taking the power consumption for each VM dedicated to that client. For the other devices, switches, firewalls, and routers, we need to take the total power consumption P_D of that device D and multiply it by the utilization level of c on D:

$$P_{D,c} = U_{D,c} \cdot P_D$$

4.2.3 Calculating indirect power consumption

The following calculations will make use of coefficients to derive power consumption that can not be directly measured. These coefficients are expected to be less than completely accurate, therefore the calculations involving these coefficients will need to include an error term to be correct.

The power consumed by the CRAC is related to the amount of heat generated in the data center computer room. We will use the CRAC power consumption model from [3]:

$$PC = \frac{Q}{\eta(T = T_{sup} - T_{adj})} + P_{fan}$$

This models the power consumption of a CRAC based on the total heat generated Q, the temperature of the cooled air supplied by the CRAC T_{sup}, the adjusted temperature arriving at the CRAC T_{adj} and the power consumed by the CRACs fans servicing the racks owned by British Telecom Global Services P_{fan}. $T_{adj} = T^{in}_{safe} - T^{in}_{max}$, the difference between the chosen safe temperature, that is the maximum temperature of the air permissible, and the maximum temperature observed. This difference is chosen over simply the temperature of the air transported to the CRAC because if T^{in}_{max} exceeds T^{in}_{safe}, T_{sup} will have to be lowered to ensure the temperature in the computer room and the devices therein does not exceed safe levels. η in the equation above is the coefficient of performance for the combination of T_{sup} and T_{adj}. To calculate the cooling power required for a specific device, we adapt the equation above to define the following function: $C(Q) = \frac{Q}{\eta(T=T_{sup}-T_{adj})} + P_{fan}$, which calculates the cooling power required for a device with a power consumption Q.

To calculate the cooling power required for a device D for a client c, we need $C()$ and the power consumed for that client. By multiplying the power consumption by the utilization level of c on D, we obtain the power consumed by the CRAC for that client on that device:

$$PC_{D,c} = U_{D,c} \cdot C(PD_D) + \epsilon_{D,c}$$

The error term $\epsilon_{D,c}$ is there to account for the difference between $U_{D,c} \cdot C(PD_D)$ and $PC_{D,c}$.

To calculate the power consumption of the UPS of a device D, we need to know the coefficient $\eta_{UPS,D}$, : $\eta_{UPS,D} = \frac{P_{UPS,D}}{P_D}$ The power consumed by the UPS of D is then

$$PU_D = \eta_{UPS,D} \cdot P_D - P_D$$

. In this calculation, $P_{UPS,D}$ is the total power consumed by the UPS for D and P_D is the power consumed by D. PU is then the power consumed by the UPS in the process of providing power to D. To account the power consumption to a client c, we need to multiply by the utilization level for that client on that device: $PU_{D,c} = U_{D,c} \cdot PU$

To calculate the power consumed by cooling for the heat generated by a UPS for D for a client c:

$$PC_{D,UPS,c} = C(PU_{D,c}) + \epsilon_{D,UPS,c}$$

Again, the error term $\epsilon_{D,UPS,c}$ accounts for the difference between $C(PU)$ and $PC_{D,UPS,c}$

The power consumption calculated for cooling is based on the direct power consumption, and is therefore calculated as frequent as direct power consumption is available.

4.2.4 Calculating Amenity Power

To calculate the power cost of amenities in a particular data center that can be attributed to a client c we must first determine the total power cost of amenities PA_{total}. Then, we must determine the ratio of floor area reserved for British Telecom Global Services, A_{BTGS}, to total floor area, A_{total} to get PA_{BTGS}, the power of amenitites attributable to British Telecom Global Services.

$$PA_{BTGS} = PA_{total} \cdot \frac{A_{BTGS}}{A_{total}}$$

Now, to calculate the amenity power attributable to a client c, we must know the capacity reserved for c, C_c and the total capacity that British Telecom Global Services has available in the data center, C_{total} and perform the following calculation:

$$PA_c = PA_{BTGS} \cdot \frac{C_c}{C_{total}}$$

PA_{total} is a metric that should be obtained from the data center operator periodically. It is useful to do this as often as possible, like the COP metrics mentioned above. The total will be an average power consumption, in watts, over a period of time. Whenever a time interval is calculated based on the data available for that time interval, the amenity power average available for that period of time should be used or, if no such data is available at that moment for that time interval, the most recent one should be used.

4.2.5 Calculating total power consumption

Let D_c be the set of devices involved in providing Cloud Contact Cisco to a client c. The total power for c can be calculated as such:

$$P_c = (\sum_{d \in D_c} PD_{d,c} + PC_{d,c} + PU_{d,c} + PC_{UPS,d,c}) + PA_c$$

This calculation must be run for each time interval for which data has been collected.

4.2.6 Note on data availability

Since the data required for the calculations is available at different frequencies, e.g. the direct consumption and the indirect consumption are available at 20 second intervals and the amenity power at completely uncertain intervals, it is possible that a complete calculation over a certain interval is not possible until some time after that interval passed. It is therefore important to store all the raw data required for the calculations persistently, so the total calculations can be made when all the data is available. This also makes the calculations auditable.

5 Discussion

5.1 Completeness and correctness

The model can only account for the power consumption of devices owned by British Telecom Global Services. The power consumed between the edge of the WAN and the end user cannot be known with the currently available technologies. It is therefore not included in the model. Heat from human bodies in the data center computer room is ignored, but is not expected to be so small that it can be completely ignored.

The omission of the effect of heat from human bodies on the total calculation is unknown at this point. Therefore, at this point, the model cannot be deemed to be actually complete. To get insight on this effect it is necessary to get empirical data on the total power consumption by British Telecom Global Services in a data center, the average time of human bodies spent in the computer room and the average heat dissipation from those bodies to get an indication if the presence of humans is of significant effect on the air temperature and therefore on cooling.

5.2 Accuracy

The calculations include error terms where appropriate. It is possible to get empirical data on these error terms by controlled experiments. With knowledge of the statistical average of the error terms it will be possible to include an accuracy term in the final calculation, indicating how likely it is for the true figure to be within a certain range of the calculated figure.

The accounting of amenity power is based on calculations involving reserved capacity and total capacity. These concepts are not well defined in the scope of power modeling. Although we do not have another method to present in this work, and although the account will be solid, it is possible that there is a different method for attributing the power consumption of amenities to the clients of British Telecom Global Services that is equally valid and leads to equally solid accounting.

5.3 Missing data

Missing data is not accounted for in the model at the moment. This will have to be added in later endeavors. Currently, if data is missing for a certain time interval, no calculation can be made for that time interval.

5.4 Model evaluation

To evaluate the model it is necessary to know the actual total power consumption caused by the activities of British Telecom Global Services in any particular data center. Some DC operators claim to be able to provide the average power consumption of a tenant in their DC. To be able to use this data to evaluate the model this figure needs to incorporate the same components, that is direct power consumption and cooling power, and power consumed by amenities. If the power consumption data provided by the data center operators is found to be reliable and accurate, such data may be used to get an empirical indication of the accuracy of the model.

The partial calculations in the model that contain error terms may be tested in isolated experiments to get empirical data on the error terms from which an expected accuracy with a certain confidence may be derived.

6 Conclusion

We have constructed a model of power consumption of a private cloud setup in a WAN where the points of presence are parts of multi tenant data centers. The model accounts for the power consumption within the WAN, power consumption of devices between the edges of the WAN and the end user is not accounted for.

Within the boundaries of the WAN, the model constructed is complete, with the exception of the inclusion of body heat in the cooling calculations. It is valid only for the specific situation of a private cloud set up in multi-tenant data centers where the WAN links are owned by and dedicated to the provider of the private cloud. The accuracy of the model cannot be determined at the time of writing, and should be determined in future work.

Future work

To determine the accuracy of the model it is necessary to obtain real world data on the total power consumption of British Telecom Global Services in each data center. It should also be determined how the utilization calculations for firewalls should be done. Another point of future work would be to determine the average heat dissipation of human bodies in the computer room and how they are proportional to the total power consumption of CRAC based on the heat generated by owned devices. Next to that, it would be interesting to find a way to determine the fairness of the attribution of amenity power and other power consumption that is not activity related, but presence related. Lastly, the correctness of the utilization calculations needs to be determined.

References

[1] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys and Tutorials*, vol. 18, pp. 732–794, 1 2016.

[2] N. Jones. How to stop data centres from gobbling up the world's electricity. [Online]. Available: https://www.nature.com/articles/d41586-018-06610-y

[3] J. Moore, J. Chase, P. Ranganathan, and R. Sharma, "Making scheduling "cool": Temperature-aware workload placement in data centers." 04 2005, pp. 61–75.

[4] R. Westerhof, "Exploring the factors of sustainability and creating a sustainability dashboard for bt global services," Research Internship, University of Groningen, 2021.

[5] X. Zhan and S. Reda, "Techniques for energy-efficient power budgeting in data centers," in *Proceedings of the 50th Annual Design Automation Conference*, ser. DAC '13. New York, NY, USA: Association for Computing Machinery, 2013. [Online]. Available: https://doi.org/10.1145/2463209.2488951

A Parameters to collect

Next to the measured data that is stored for each device, there are some parameters for the calculation of the total power consumption to be collected for each data center:

- CRAC COP
  The coefficient of performance of the CRAC installation of the data center.
- COP per UPS
  For each UPS, we need to know the COP of that UPS.
- Total capacity
  Some quantity expressing the total capacity available in the data center for British Telecom Global Services' clients. An example would be the number of CPUs available.
- Capacity reserved for a client
  Some quantity expressing the computing capacity reserved for a client. An example would be the number of CPUs available to this client
- Power consumed by amenities

B Recommendations for implementation

To implement the model in an auditable manner, we make the following recommendations. For the calculations, it is recommended to store all measurements for the entire period the calculations are supposed to be auditable. The constants that are used in the calculations to obtain the total power consumption should be stored as well, along with a date indicating when the constant with that value was taken into use in calculations and if the constant is replace, a date indicating when the constant with that particular value was retired. That way, a calculation that was done on some date d can be repeated by retrieving the constants that were used on date d at any time, as long as the data on these constants and the measured data for each device is stored.