

FacebookAnalytics

Exploratory Data Analysis on the dataset of Facebook's Altoona Data Center

Thamizhiniyan Pugazhenthir - 200941620

22/12/2021

Introduction

Since IT operations are so important for business continuity, they usually incorporate redundant or backup components and infrastructure for power, data communication connections, environmental controls (such as air conditioning and fire suppression), and other security systems. A huge data center is a large-scale activity that consumes as much energy as a small town¹. Data center modernization and transformation improve performance and energy efficiency. Information security is also a concern, which is why a data center must provide a secure environment that reduces the risk of a security breach. As a result, a data center must maintain high standards for ensuring the integrity and functionality of the computer environment it hosts. The average age of a data center, according to industry research firm International Data Corporation (IDC) is nine years².

Business Understanding

The network traffic observed in the Altoona Facebook datacenter is the subject of this project. While Facebook uses typical datacenter services such as Hadoop, its core Web service and associated cache technology display a number of characteristics that differ from those mentioned in the literature. We discuss how network traffic in Facebook's datacenters differs in terms of locality, stability, and predictability, as well as the ramifications for network architecture, traffic engineering, and switch design.

Machines are grouped into racks and connected through 10-Gbps Ethernet cables to a top-of-rack switch (RSW). From cluster to cluster, the number of machines per rack varies. Each RSW is connected to four aggregation switches called cluster switches over 10-Gbps lines (CSWs). A cluster is defined as all of the racks served by a specific set of CSWs. Clusters can be homogeneous in terms of machines (for example, cache clusters) or heterogeneous (for example, frontend clusters with a mix of Web servers, load balancers, and cache servers). Fat Cats are another layer of aggregation switches that connect CSWs to each other (FC). Finally, for intra-site (but inter-datacenter) traffic, CSWs connect to aggregation switches, and datacenter routers for inter-site traffic. Web servers handle Web traffic, Database - MySQL servers hold user data and Hadoop servers do offline analysis. A pod is just like a layer3 micro-cluster in that it has nothing special about it. The pod is simply a standard "unit of network" on our new fabric, with no hard physical properties. Each pod is served by a set of four fabric switches, which maintain the benefits of our present 3+1 four-post architecture for server rack TOR uplinks while also allowing for future expansion. Each TOR presently has four 40G uplinks, allowing a rack of 10G-connected computers to access 160G of total bandwidth. What's distinctive about their new unit is its significantly smaller size – each pod has only 48 server racks, and this form factor is consistent across all pods. It's a proficient structure block that fits pleasantly into different server farm floor plans, and it requires just essential medium size changes to total the TORs. The more modest port thickness of the texture switches makes their inside design extremely straightforward, measured, and powerful, and there are a few simple to-find choices accessible from numerous sources.

Another prominent distinction is the means by which the cases are associated together to frame a server farm organization. For each downlink port to a TOR, they are saving an equivalent measure of uplink limit

on the case's texture switches, which permits them to increase the organization execution to measurably non-obstructing.

Data Understanding

There are three clusters in the data center i.e., Cluster A is for Database, Cluster B is for Web servers and Cluster C is for Hadoop servers. Since we are working only with Cluster A and Cluster C, both of them have same set of columns. Each has 273 bz2 files in compressed format. Upon decompressing, each file is of tsv type with columns

1. timestamp
2. packet length
3. anonymized source IP
4. anonymized destination IP
5. anonymized source L4 Port
6. anonymized destination L4 Port
7. IP protocol
8. anonymized source hostprefix
9. anonymized destination hostprefix
10. anonymized source Rack
11. anonymized destination Rack
12. anonymized source Pod
13. anonymized destination Pod
14. intercluster
15. interdatacenter

All the data fields are anonymized for confidentiality. The timestamp column has values of type 'Unix timestamp'. The IP protocol column has three values - 6, 17 and 58. The value '6' points to Transmission control protocol(TCP), '17' points to User Datagram protocol (UDP) and '58' points to Internet Control message protocol(ICMP) for IPV6(Internet Protocol version 6).The last two columns plays major role to instigate analysis. The value '0' in intercluster denotes that the transmission of packets is within the cluster and '1' denotes that the transmission was between cluster. Same goes with the column 'datacenter'. Since facebook uses TCP segmentation offload, packet length can be larger than 64 kb because outbound packets are sampled in the kernel. It is not guaranteed to be a 1:1 mapping since the content has been hashed and get a subset of the hash value. The prefix of a hostname is called the host prefix. A computer named "web102.prn1.facebook.com" has the hostprefix "web." It's a very rudimentary classification of machine types. However, keep in mind that numerous programmes can operate on the same machine.

Data Preparation and Analysis

For this analysis, we took two samples from each cluster as subset and we will analyse the dataset with respect to each column and derive results from many perspectives especially with intercluster and interdatacenter. First, we need to cleanse the data since some of the fields have corrupted values. After cleansing, we need to fix the values of timestamp column. The column has values of type 'unix timestamp' which need to converted to normal time value - BST value. Over the course of analysis, we will compare different characteristics of the data with respect to each column and analyse the data traffic in depth. Before step into the analysis, we will categorize the data into three types - Intra cluster and Intra datacenter, Inter cluster and Intra datacenter, Inter cluster and Inter datacenter. We made this classification with the values of intercluster and interdatacenter columns. The value '0' refers to the intra with respect to the column description and value '1' refers to the inter - between clusters or between datacenters.

Timestamp

In database cluster, We can see the level of traffic in the data center over a period of time from Figure 1. As we can see that the level of traffic is higher in machines within cluster and datacenter followed by the machines between clusters within datacenter. We cannot see much traffic in machines between clusters and between datacenters. Moreover, we can see the intensity of internet traffic is higher between 18.00 hrs and 24.00 hrs in a day in all three classifications. From this plot, we can understand that the usage of facebook is higher in this timeframe than the rest of the day.

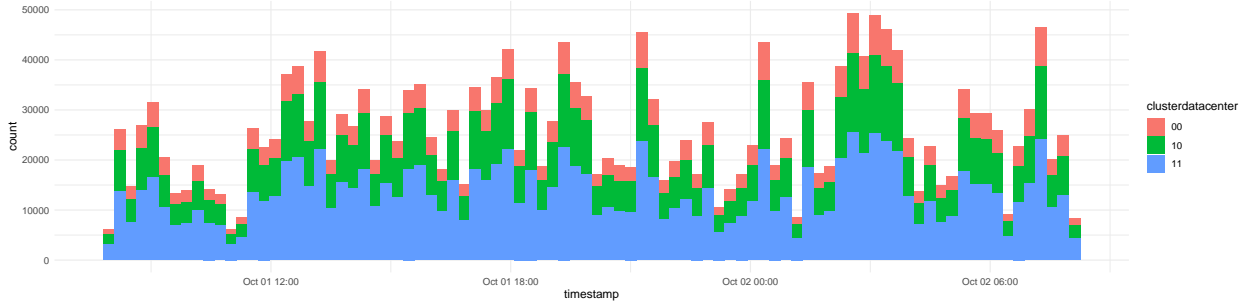


Figure 1: Traffic characteristics - Database cluster - Overall

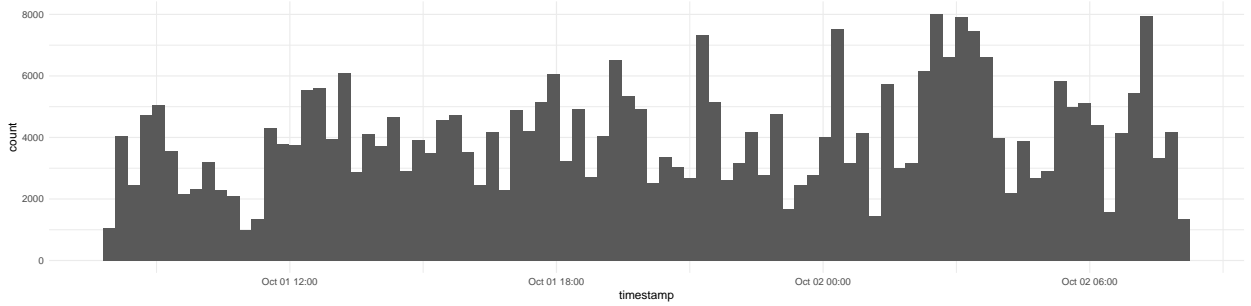


Figure 2: Traffic characteristics - Database cluster - Intra cluster - Intra datacenter

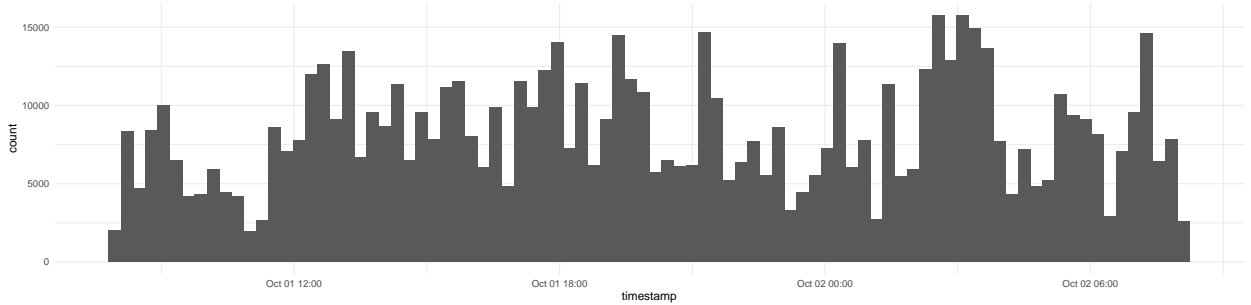


Figure 3: Traffic characteristics - Database cluster - Inter cluster - Intra datacenter

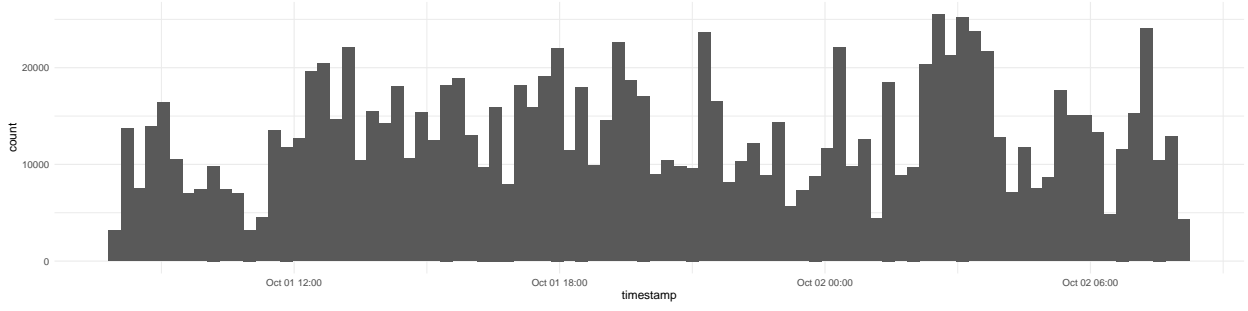


Figure 4: Traffic characteristics - Database cluster - Inter cluster - Inter datacenter

From Figure 2, Figure 3 and Figure 4, we can see the level of traffic in all three categorizations with separate plot for a clear picture. The plot legend represents the concatenated value of the columns - intercluster and interdatacenter. We were considering 24 hour timeframe from October 1 9.00 hours to October 2 9.00 hours for this analysis.

In Hadoop cluster, We can see the level of traffic in the data center over a period of time from Figure 5. As we can see that the level of traffic is very much higher in machines within cluster and datacenter followed by the machines between between clusters and between datacenters. We cannot see much traffic in machines between clusters within datacenter. From Figure 6, Figure 7 and Figure 8, we can see the level of traffic in all three categorizations with separate plot for a clear picture.

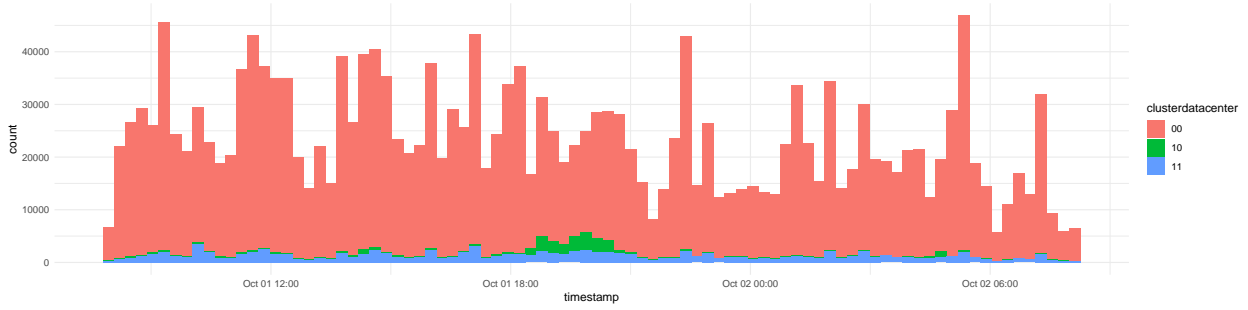


Figure 5: Traffic characteristics - Hadoop cluster - Overall

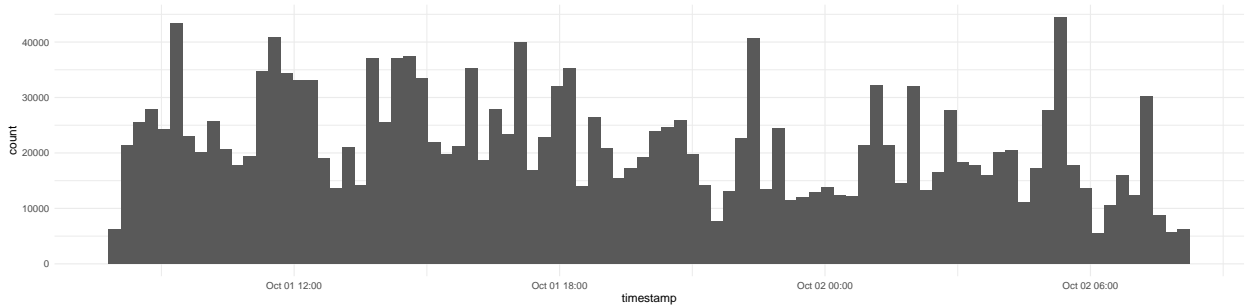


Figure 6: Traffic characteristics - Hadoop cluster - Intra cluster - Intra datacenter

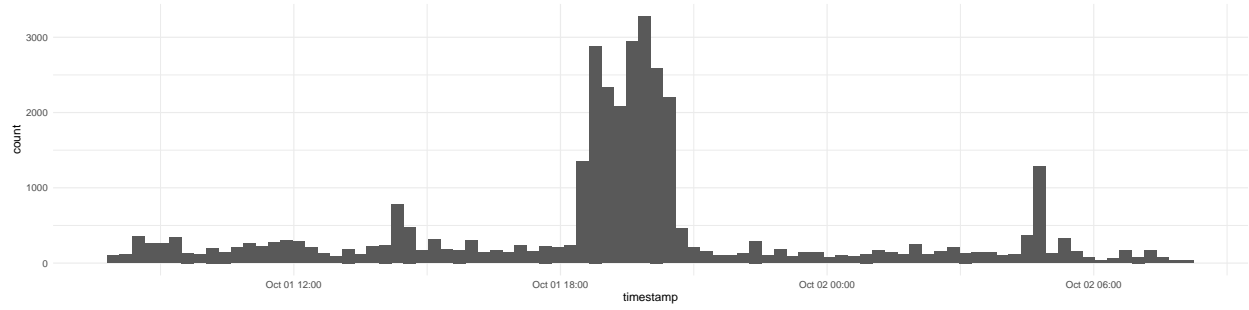


Figure 7: Traffic characteristics - Hadoop cluster - Inter cluster - Intra datacenter

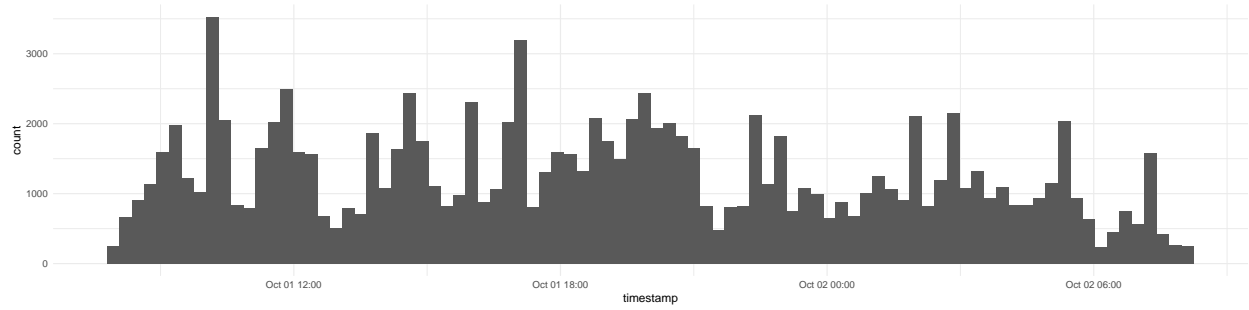


Figure 8: Traffic characteristics - Hadoop cluster - Inter cluster - Inter datacenter

In the second level of analysis, we will compare the level of traffic between database and hadoop cluster. It is evident that the traffic is higher in the database cluster than in hadoop cluster. We can identify this with the Figure 9.

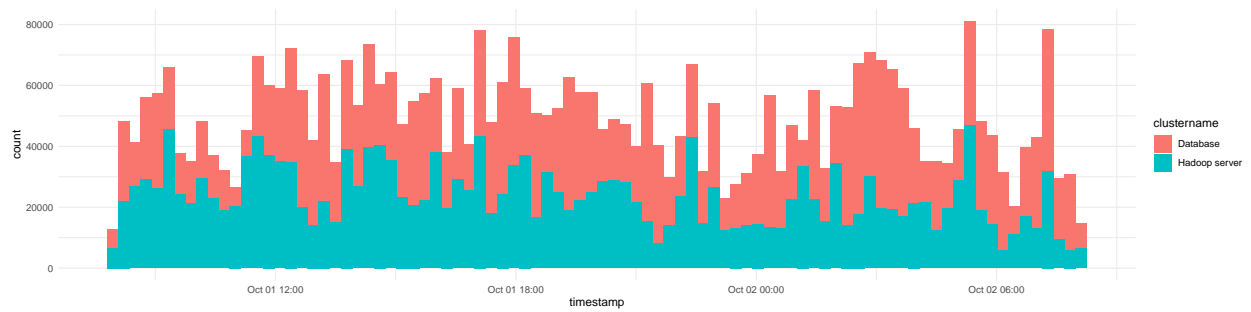


Figure 9: Traffic characteristics - Database and Hadoop server - Overall

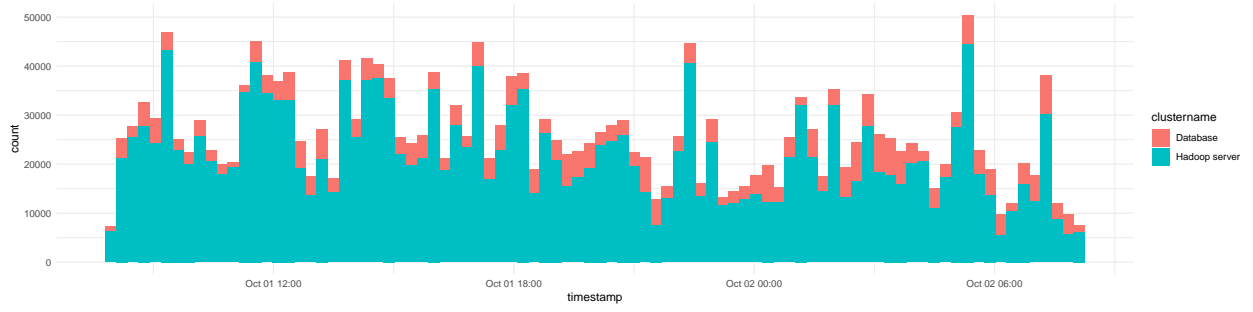


Figure 10: Traffic characteristics - Database and Hadoop server - Intra cluster and Intra datacenter

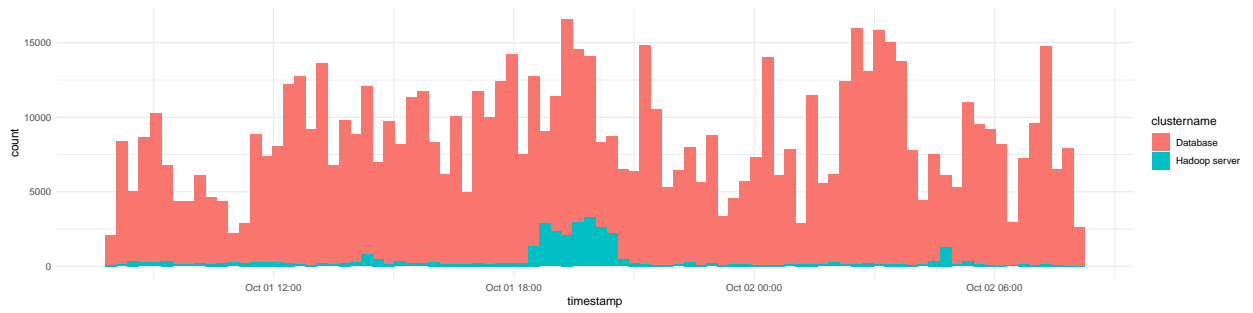


Figure 11: Traffic characteristics - Database and Hadoop server - Inter cluster and Intra datacenter

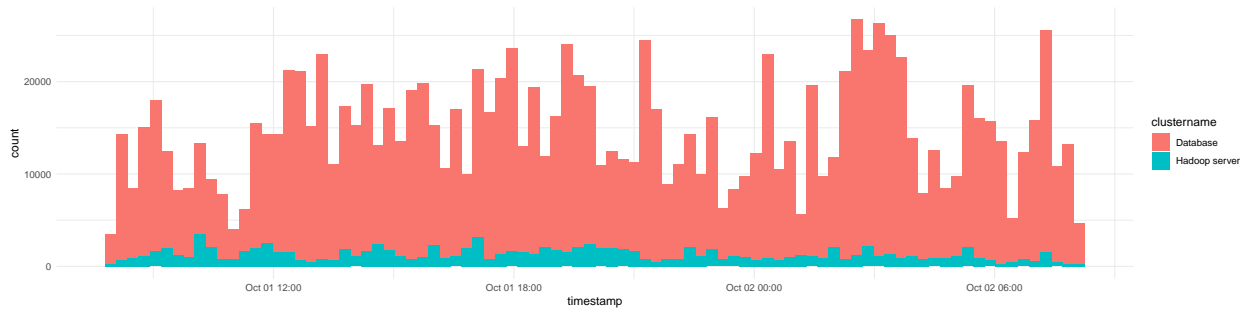


Figure 12: Traffic characteristics - Database and Hadoop server - Inter cluster and Inter datacenter

From Figure 10, Figure 11 and Figure 12, we can see the comparison between clusters in all three classifications with separate plot for each.

IP Address

We will analyse the data with respect to the source and destination IP address. From Figure 13, we can see the top 10 combination of source and destination IP address which contributes more to the traffic between the machines overall without any classifications in database cluster. With the count of the combination of anonymized source and destination IP values, the dataframe was created and the plot was designed.

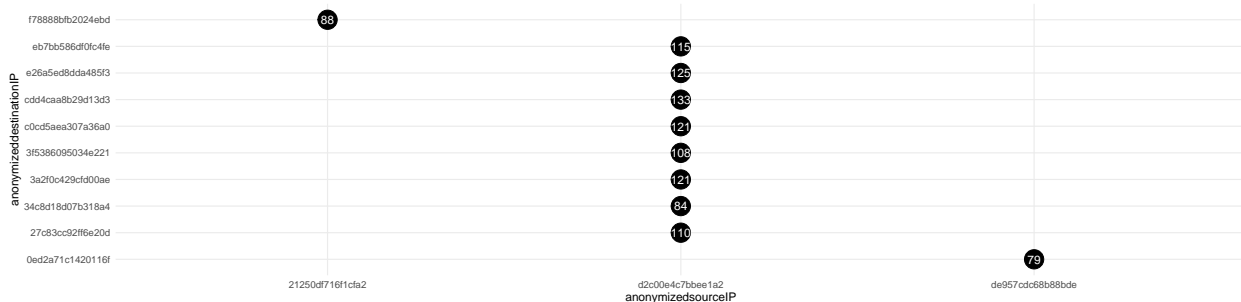


Figure 13: Traffic between source and destination IP - Database Cluster

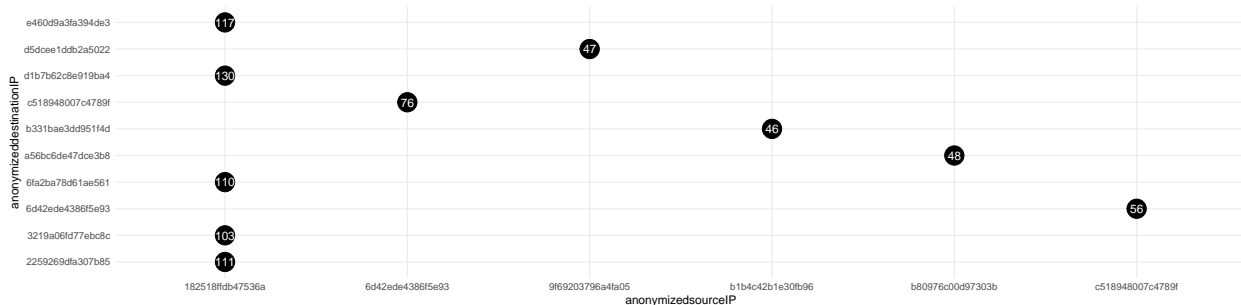


Figure 14: Traffic between source and destination IP - Hadoop Cluster

From Figure 14, we can see top 10 combinations of source and destination IP address which contributes more traffic in Hadoop cluster. To explain this plot, X axis denotes source IP address and Y axis denotes destination IP address with the number in the circle in align with the axis denotes the level of traffic between these addresses.

Rack and Pod

In Altoona data center, they divided the network up into small identical pieces – server pods – instead of huge devices and clusters, and created consistent high-performance connectivity between all pods in the data center. Each pod is served by a set of four fabric switches, which maintain the benefits of our present 3+1 four-post architecture for server rack TOR linkups while also allowing for future expansion. Each TOR presently has four 40G up links, allowing a rack of 10G-connected computers to access 160G of total bandwidth. Each pod has only 48 server racks, and this form factor is consistent across all pods. Another major difference is the manner in which the pods are linked to form a data center network. They constructed four different “planes” of spine switches to achieve building-wide connection, each scalable up to 48 independent devices within a plane. Within its local plane, each pod’s fabric switch links to each spine switch. Pods and planes combine to produce a modular network topology capable of supporting hundreds of thousands of 10G-connected servers,

growing to multi-petabit bisection bandwidth, and providing non-oversubscribed rack-to-rack performance across their data center facilities ³.

From Figure 15 and Figure 16, we can see the combination of source and destination Rack address which contribute more to the traffic between machines in Database and Hadoop clusters respectively. If we see the plot, we can identify that four source and three destination racks contributes more traffic to the entire system in Figure 16. This constitutes top 10 of the entire combinations. In Figure 17, we can see that it is 1-1 mapping in terms of traffic, each source and destination combination does not collide with any other combination within the cluster. Even this figure displays the top 10 combinations of the entire list. With the count of the combination of anonymized source and destination rack and pod values, the dataframes were created and the plots were designed.

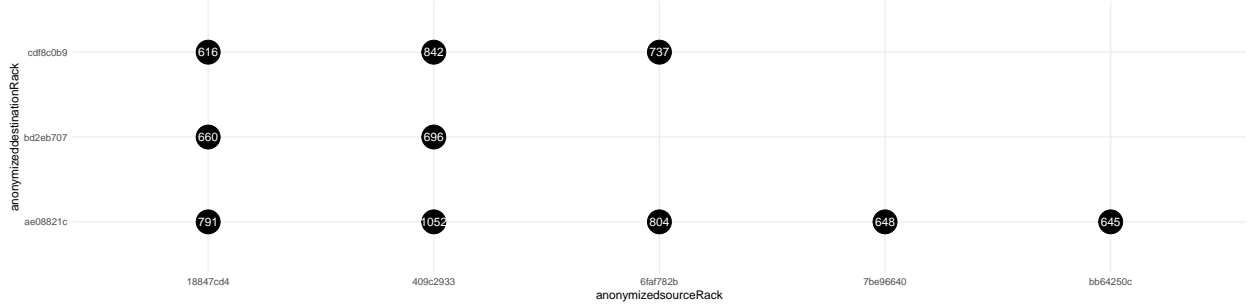


Figure 15: Traffic between source and destination Rack - Database Cluster

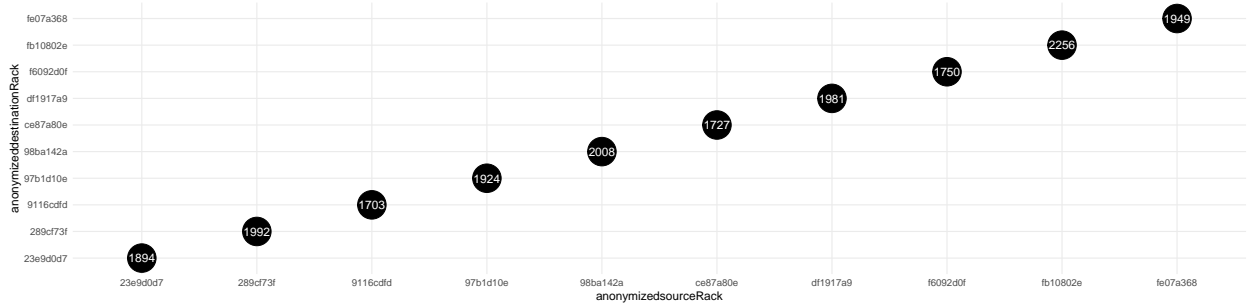


Figure 16: Traffic between source and destination Rack - Hadoop Cluster

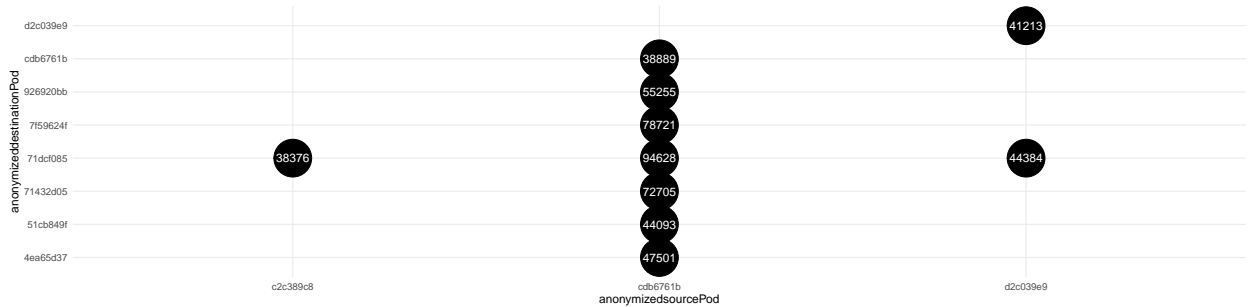


Figure 17: Traffic between source and destination Pod - Database Cluster

In Figure 17 and Figure 18, we can identify the combination of source and destination pairs of pod which contributes more to the traffic intensity. From Figure 17, we can see that only three source pod contributes more to the traffic but in Figure 18, there is an equal distribution of source and destination combinations in Hadoop cluster.

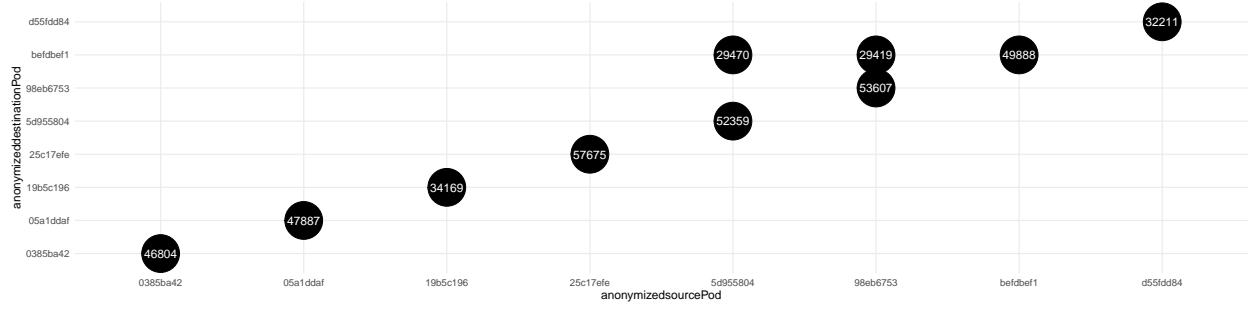


Figure 18: Traffic between source and destination Pod - Hadoop Cluster

We can see how pod and rack are interacting with each other in Figure 19 and Figure 20 for Database and Hadoop cluster respectively.

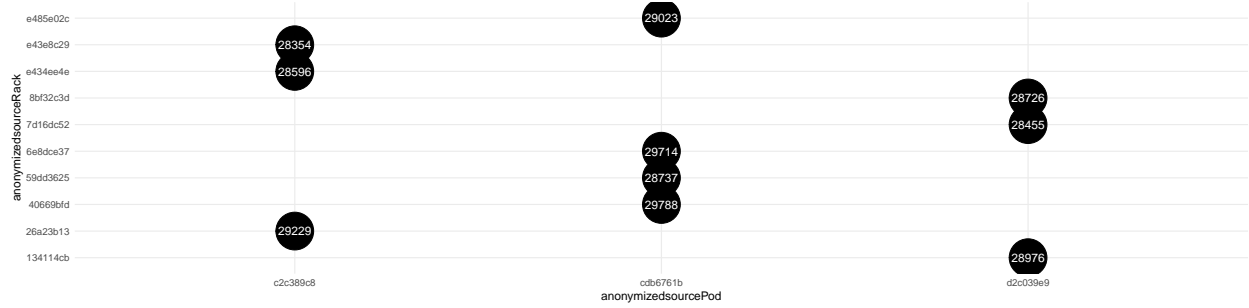


Figure 19: Traffic between source Pod and Rack - Database Cluster

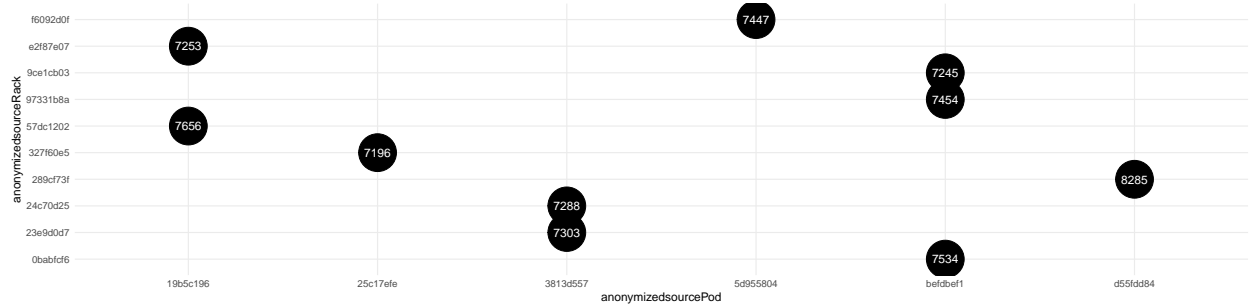


Figure 20: Traffic between source Pod and Rack - Hadoop Cluster

Host Prefix

From Figure 21 and Figure 22, we can see the top 10 combination of source and destination host prefix in database and Hadoop cluster respectively. As we can see three source host prefix and two source host prefix contributes more to the traffic in database and Hadoop cluster. The intensity of the traffic keeps increasing gradually from the plots of IP address to host prefix. With the count of the combination of anonymized source and destination hostprefix values, the dataframe was created and the plot was designed.

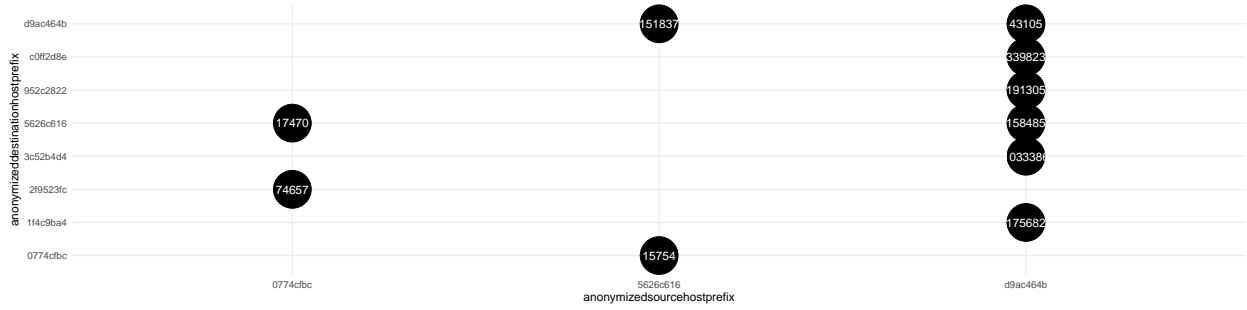


Figure 21: Traffic between source and destination hostprefix - Database Cluster

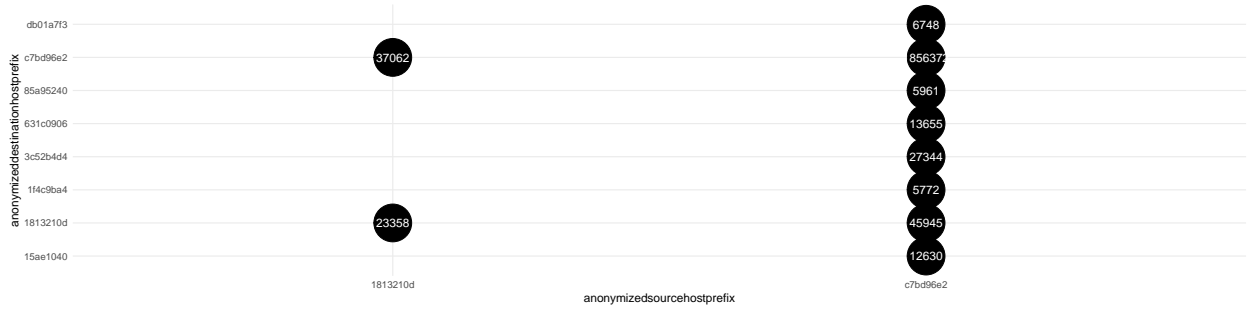


Figure 22: Traffic between source and destination hostprefix - Hadoop Cluster

L4 Port

From Figure 23 and Figure 24, we can see the top 10 combination of source and destination host prefix in database and Hadoop cluster respectively. As we can see three source host prefix and two source host prefix contributes more to the traffic in database and Hadoop cluster. With the count of the combination of anonymized source and destination L4 port values, the dataframe was created and the plot was designed.

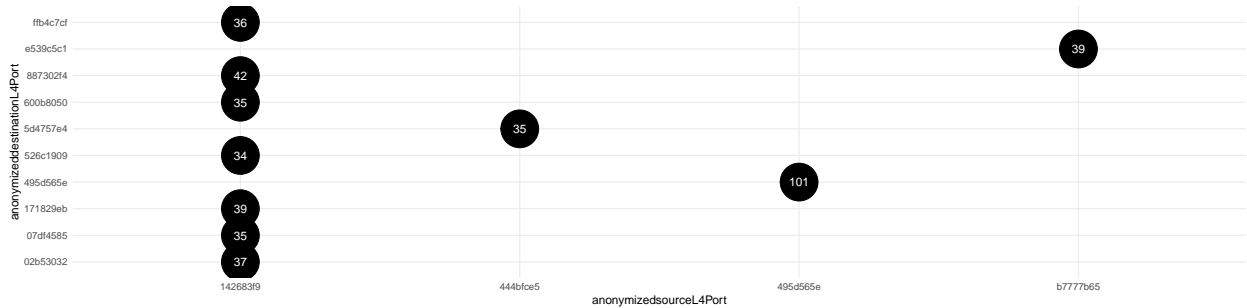


Figure 23: Traffic between source and destination L4 port - Database Cluster

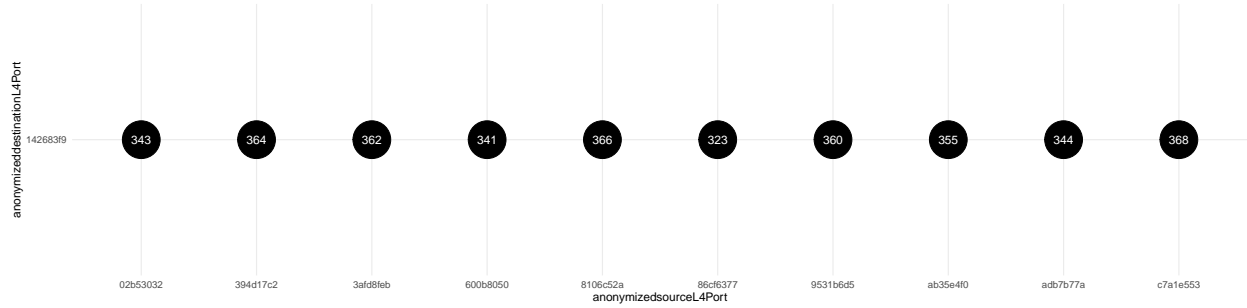


Figure 24: Traffic between source and destination L4 port - Hadoop Cluster

IP Protocol

All these request and reply happen within three IP protocol number - 6, 17 and 58. 6 is for Transmission Control protocol (TCP), 17 is for User Datagram protocol (UDP) and 58 is for Internet Control message protocol(ICMP) for IPV6(Internet Protocol version 6) (ICMP - IPV6). We can see that the lot of traffic intensity is from TCP which is way higher than the rest of the protocols. With the count of the combination of anonymized source and destination IP protocol, the dataframe was created.

```
##   IPprotocol      n
## 1          6 2257207
## 2         17   1468
## 3         58    101
```

```
##   IPprotocol      n
## 1          6 2058333
## 2         17   10445
## 3         58     81
```

Level of co-location of applications on servers

Database

With the Database, the level of co-location of applications is very low. It was uniformly distributed to a large extent. The x-axis of the below plot represents the count of the same spread of values from anonymized-sourceIP to anonymizeddestinationPod throughout the cluster with '1' being no matching values and the count increases according to the number of time the values have matched with the other values. Figure 25 represents the plot with co-location in the database cluster with the count it was repeated.

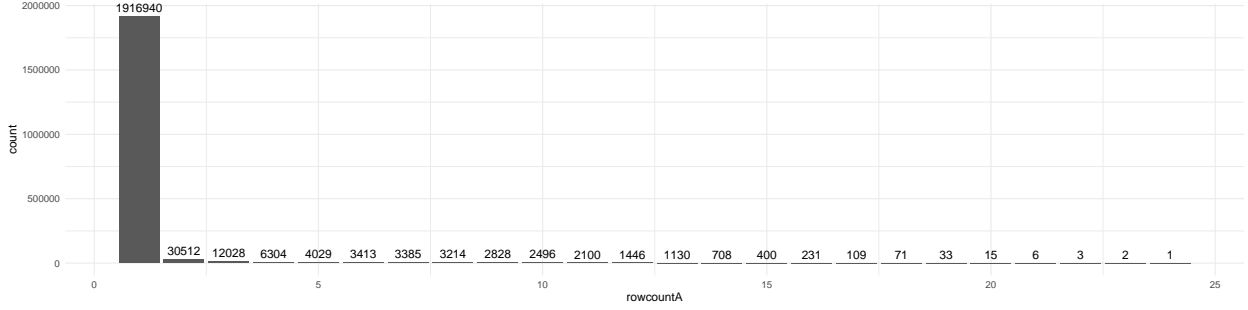


Figure 25: Co-location of applications on Database with the row count

Hadoop server

With the Hadoop server, the level of co-location of applications is extremely low. It was uniformly distributed to its maximum. The x-axis of the below plot represents the count of the same spread of values from anonymizedsourceIP to anonymizeddestinationPod throughout the cluster with '1' being no matching values and the count increases according to the number of time the values have matched with the other values. Figure 26 represents the plot with co-location in the database cluster with the number it was repeated.

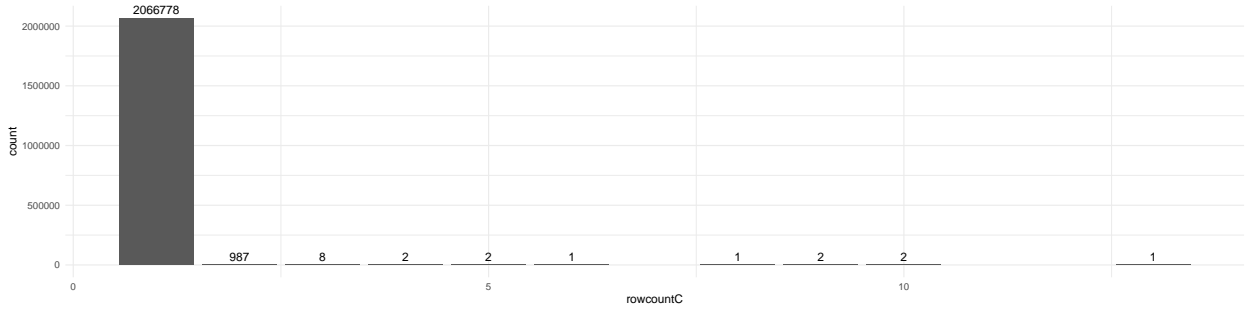


Figure 26: Co-location of applications on Hadoop server with the row count

Evaluation

- **How do traffic characteristics vary over time?**
 - From Figure 1 to Figure 12, we can see the traffic characteristics of both database and hadoop cluster varies over a period of time with various classification in terms of cluster and datacenter. The plots are self-explanatory to understand the level of machine to machine traffic in a time frame.
- **What patterns are present in the transmission of data? Are there particular source-destination pairs which represent a large quantity of traffic?**
 - From Figure 13 to Figure 22, we can identify the top source-destination pairs of IP address, Rack, Pod, L4 Port and Hostprefix which contribute the most to the large quantity of traffic. There is no specific pattern present in the transmission of data. Everything is uniformly distributed across.
- **How do different workload types (i.e. Database, Web, Hadoop) exhibit different traffic characteristics?**

- We were not able to work on web servers. But in terms of Database and Hadoop clusters, we found the traffic characteristics through plots. Figure 1 to Figure 4 represent the Database cluster and Figure 5 to Figure 8 represent the Hadoop cluster and we can see the comparison of both clusters in Figure 9.
- **What patterns are present with respect to rack- and pod-locality?**
 - We discussed about the patterns and level of traffic contribution within Rack and Pod in 4 different plots from Figure 15 to Figure 18. There are 48 top of rack switches in every pod which connected with 4 fabric switches. There will be X number of pods in a datacenter and interactions between rack and pod were depicted in Figure 19 and Figure 20.
- **What patterns are present with respect to inter-cluster and inter-datacenter communication?**
 - We classified each cluster into three different sub categories - intra cluster and intra datacenter, inter cluster and intra datacenter and inter cluster and inter datacenter. All plots were designed according to this three sub categories. From IP address to Pod, all characteristics were discussed on the basis of these three categories. We can find the patterns present with respect to inter cluster and inter datacenter communication in terms of each entity in the data set.
- **Is there any evidence within the data to suggest the level of co-location of applications on servers?**
 - From Figure 25 and Figure 26, we discussed about the level of co-location of applications of the Database and Hadoop servers respectively. The traffic intensity in both servers was uniformly distributed. There was not much of co-location found in both servers whereas the level is comparatively high in the Database more than the Hadoop server. These level of co-location is negligible since we were considering 24 hour time frame for the analysis.

Conclusion

From the analysis, we did analyse the dataset from various perspectives and measurements. From business understanding to evaluation, we worked across the various aspects in and around the dataset and provide some findings we found from the analysis. From IP address to hostprefix, we can see the level of intensity in traffic increases gradually. The people at Altoona data center should look at this carefully and do the necessary steps to maintain load balancing to reduce the machine to machine traffic. Since the machine to machine will be higher in Altoona data center compared to user to machine traffic⁴, they need to find out new strategy to tackle the traffic. Even though the facebook concentrates on renewable energy, they need to spend their time on the effect of the facebook data center to the environment since it needs tons and tons of electric power to run which will sure have negative impact on the environment. On the whole, we can say that they are heading in the right direction in terms of sustainability⁵.

References

1. https://en.wikipedia.org/wiki/Data_center
2. <http://www.mspmentor.net/2011/08/17/hp-updates-data-transformation-solutions/>
3. <https://engineering.fb.com/2014/11/14/production-engineering/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/>
4. <https://www.cs.unc.edu/xcms/wpfiles/50th-symp/Moorthy.pdf>
5. <https://datacenters.fb.com/wp-content/uploads/2021/10/Altoona-Data-Center.pdf>