HEALTH MONITORING OF ATLAS DATA CENTER CLUSTERS

AND

FAILURE ANALYSIS

by

MEENAKSHI BALASUBRAMANIAN

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2018

Acknowledgements

I am deeply indebted to Professor David Levine, my supervisor who gave me an opportunity to work under him and encouraged me throughout the work. It is his consistent support and motivation that helped me to complete this thesis and his contribution towards my professional growth is immense.

I am extremely grateful to my committee member, Patrick McGuigan, whose suggestion and guidance in every stage of this thesis, directed me towards the completion of this thesis.

I would like to express my deepest gratification to Dr. Farhad Kamangar, my committee member for his support and time spent on my work review.

I would like to extend my sincere thanks to my parents, who are always there supporting and motivating me. I am also deeply thankful to my husband, Karthikeyan Rajamani, who is beside me all times, inspiring to do a remarkable work and my special thanks to my daughter, Shreya, who stood by me throughout with patience and tolerance.

Finally, I would like to thank all my friends, who helped me throughout this academic journey.

November 16, 2018

Abstract

HEALTH MONITORING OF ATLAS DATA CENTER CLUSTERS AND FAILURE ANALYSIS

Meenakshi Balasubramanian, MS

The University of Texas at Arlington, 2018

Supervising Professor: David Levine

Monitoring the health of data center clusters is an integral part of any industrial facility. ATLAS is one of the High Energy Physics experiments at the Large Hadron Collider (LHC) at CERN. ATLAS DDM (Distributed Data Management) is a system that manages data transfer, staging, deletions and experimental data on the LHC grid. Currently, the DDM system relies on Rucio software, with Cloud based object storage and No-SQL solutions. It is a cumbersome process in the current system, to fetch and analyze the transfer, staging and deletion metrics of a specific site for any regional center.

In this thesis, a web-based cluster health monitoring framework is designed to monitor the health of the sites at the Tier 2 facility in the Southwest region of US, which eases these problems. A large volume of data flows in and out of each of these sites. If the transfer / deletion rate of files goes below the user-defined threshold at any source or destination site, the data center monitor is alerted automatically. This thesis also analyses the failures that have happened between any two performing sites. A machine learning algorithm finds the pattern of transfer / deletion with the existing data and detects the sites that may possibly fail due to diminishing transfer / deletion of files.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

“Data Never Sleeps” [1] is the online behavior today, leading to data generation of roughly 2.5 quintillion bytes (10^{18}) a day. The application programs which solves complex computational problems in scientific experiments, utilizing vast amount of data and performing modeling and simulation require High Performance Computing (HPC) [2]. Various monitoring software, to track and measure data, is available for managing the HPC clusters, to ensure they are running without any failures. ATLAS (which we will discuss in the next section) is one of the biggest scientific experiments, which creates an enormous flow of data. The ATLAS Distributed Data Management (DDM) system, based on Rucio (which we discuss in next chapter), is responsible for handling multi-petabyte volume of ATLAS data. The ATLAS DDM helps to store, manage and process the experimental data in a heterogeneous distributed environment [3].

The main goal of this Thesis is to provide an enhanced web-based monitoring solution for the ATLAS experimental sites, UTA_SWT2, SWT2_CPB, OU_OSCER_ATLAS at US-South Western Tier 2 Regional Center. In addition to the efficient monitoring, the administrators at UTA ATLAS operations will receive alerts when the transfer / deletion rates fall behind a threshold, which can be configured within our monitoring solution. The failures at CERN may not be discovered for a long time, because of the volume, velocity, variety and veracity (4 V’s of big data) [4] of data. This thesis provides a solution by analyzing the transfer of files that are happening between two performing sites. A machine learning algorithm finds the pattern of transfer / deletion with the existing data and detects the sites that may possibly fail due to diminishing transfer / deletion of files.

1.1 CERN and ATLAS Background

CERN, a European Organization for Nuclear Research, in which the world's most complex scientific experiments are done to study the fundamental structure of the universe [5] and its main area of research is particle physics. The Large Hadron Collider (LHC) [6] is the world's largest and powerful particle collider and is the largest machine in the world. ATLAS is the biggest experiments at LHC at CERN. The data generated from LHC are sent to CERN Data Center for reconstruction. CERN has grid-based infrastructure (which we discuss in next chapter), called Worldwide LHC Computing Grid (WLCG), to perform computation.

WLCG [7] is a distributed computing infrastructure, used to store, distribute analyze enormous amount of data. It is composed of four levels or Tiers (Tier 0, Tier 1, Tier 2, Tier 3). Each Tier has a collaboration of computer centers and perform specific services. The Southwest Tier 2 is an ATLAS tier 2 facility and is a consortium among the Physics department of University of Texas at Arlington (UTA), University of Oklahoma (OU) and Langston University (LU) [8].

1.2 My Motivation for Thesis

CERN Data Center had crossed the milestone of 200 petabytes of permanently archived data by the of June 2017. With such massive data and the volume still increasing, data management and monitoring the health of clusters is always challenging. CERN uses Ganglia (which we discuss in next chapter), a distributed monitoring system for high performance computing systems. As ATLAS experiment creates non-trivial amount of data, it uses a Distributed Data Management system based on Rucio (which we discuss in next chapter) for extreme scalability. Motivated by the complex scientific experiments, data center facility in UTA and tremendous data generated by the world's largest machine, a solution was developed to monitor the cluster health and detecting anomalies of the sites at US Southwest Regional Center.

1.3 Goal for Thesis

The main objective of this thesis is to implement an enhanced web-based cluster health monitoring solution, for ATLAS operations. The administrators at the ATLAS will receive alerts if the data deletion/ transfer between the sites fail and this would help them to resolve issues quickly. To know more details on failures, this thesis provides an analysis on failures between two performing sites. To further improve the quality, a machine learning algorithm is implemented which finds the pattern of transfer / deletion with the existing data and detects the sites that may possibly fail due to diminishing transfer / deletion of files, so they could take preventive steps and the keep the clusters always running. This thesis would help in monitoring the cluster health effectively, analyze failed sites and predict the site failures that would happen in future.

1.4 Organization of Thesis

Chapter 1- Starts with the introduction, explains about CERN and ATLAS background, and mentions the goal of the thesis.

Chapter 2- This chapter explains Grid computing, Worldwide LHC Computing Grid, Storage Resource Manager that handles space and data management in ATLAS experiments. It also explains the existing monitoring systems like Ganglia Monitoring System and ATLAS Distributed Data Management system.

Chapter 3- The drawback of ATLAS DDM and the problems faced by the administrators to monitor the health of clusters are discussed in this chapter.

Chapter 4- This chapter explains the technical environment used for the proposed solution , architecture, design and data collection process.

Chapter 5- Implementation of the proposed solution, about how the cluster health is monitored with the file deletion and transfer metrics is explained in this chapter.

Chapter 6- This chapter explains how the site failures are analyzed during file transfers and the implementation of identifying sites having failed transfers.

Chapter 7- This chapter contains the summary and conclusion of this thesis.

Chapter 8- This chapter explains the future work.

Chapter 2

The current monitoring system

2.1 Grid Computing

CERN passed the milestone of archiving 200 petabytes of data, by June 2017. Even after a huge data reduction performed by the experiments, CERN Data Center computes an average of 1 petabyte of data every day [9]. See Figure 2-1 for the data transfer throughput of different CERN experiments. On any hour of the day, the transfers are happening in an average of 20GB for ATLAS. To compute such vast data, grid computing is used to share the computing burden among different centers.

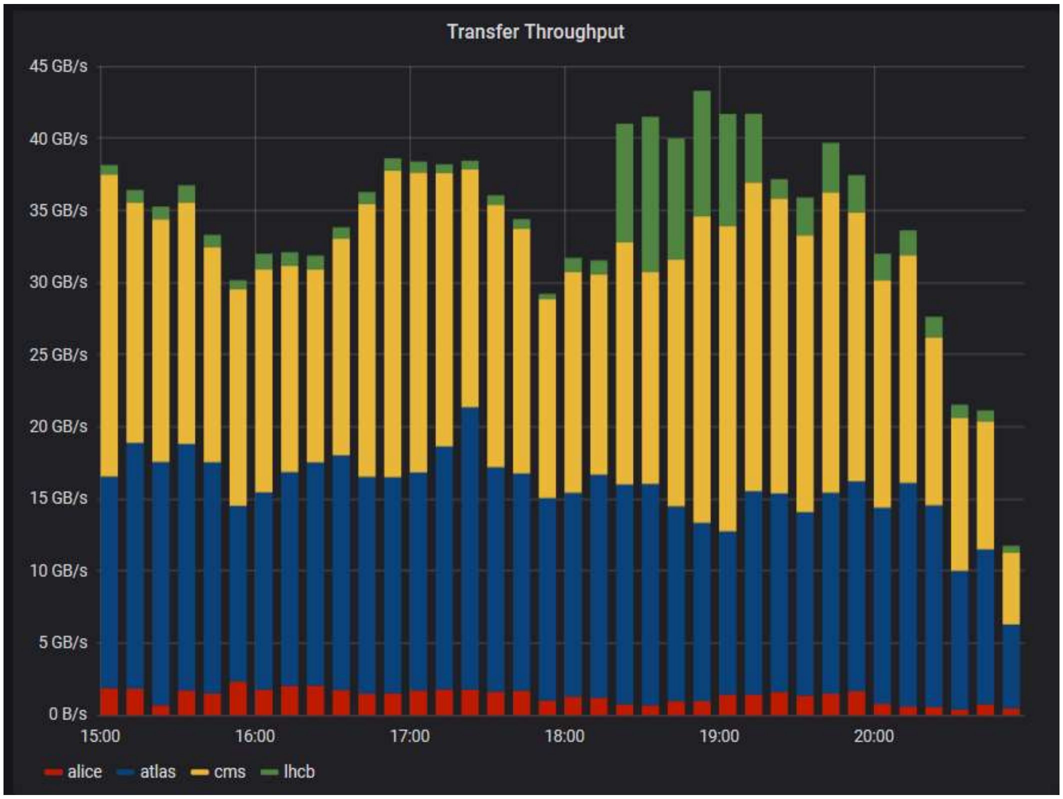


Figure 2-1 Transfer rate of CERN experiments

Image src: http://monit-grafana-open.cern.ch/d/000000306/wlcg-transfers-dashboard?orgId=16

The grid-based infrastructure offers many advantages over centralized system and is most effective for data analysis and management at LHC. Multiple copies of data are kept at different centers around the world, to ensure there is no single point of failure. It also helps scientists access data independent of geographical locations across multiple time zones, providing flawless access to computing resources.

Users can make job requests, without worrying from where they are using the computing resources. A job request can be storage, processing or analysis. The computing grid authenticates the identity of the user and redirects them to the available sites, that can provide their requested resources.

2.2 Worldwide LHC Computing Grid Architecture

The goal of the Worldwide LHC Computing Grid (WLCG Mission) [10] is to provide computing resources to store, distribute and analyze the petabytes of data generated from the LHC at CERN. WLCG has more than 170 computing centers in 42 countries.

WLCG is the world's largest computing grid and is based on two main grids. The computer centers in the grid are arranged in tiers [10].

a) Tier 0 is CERN DC (Data Center) and is the heart of all LHC experiments and stores the first copy of raw data. The data from LHC is passed to this DC, but the data center has only 20% of the grid capacity. Tier 0 transfers the raw data to Tier 1 and reprocesses the data, when LHC is down.

b) Tier 1 has 13 large computer centers and they provide enough capacity and around the clock grid support. They store a proportional share of both raw and reconstructed data. It distributes data to Tier 2 and keeps a share of simulated data produced at Tier 2.

c) Tier 2's are usually scientific institutions and universities which provides sufficient computing power for specific task analysis and store required data. There are currently around 160 Tier 2 centers. The data centers in UTA comes under Tier 2. See Table 2-1 for Southwest Tier 2 centers in US.

Site Name	**Location**
SWT2_CPB	University of Texas, Arlington
UTA_SWT2	Fort Worth
OU_OSCER_ATLAS	University of Oklahoma

Table 2-1 Southwest Tier 2 Centers

d) Tier 3 is used by individual access to computing resources, either in their systems or the local clusters in the University. See Figure 2-2 for WLCG architecture.

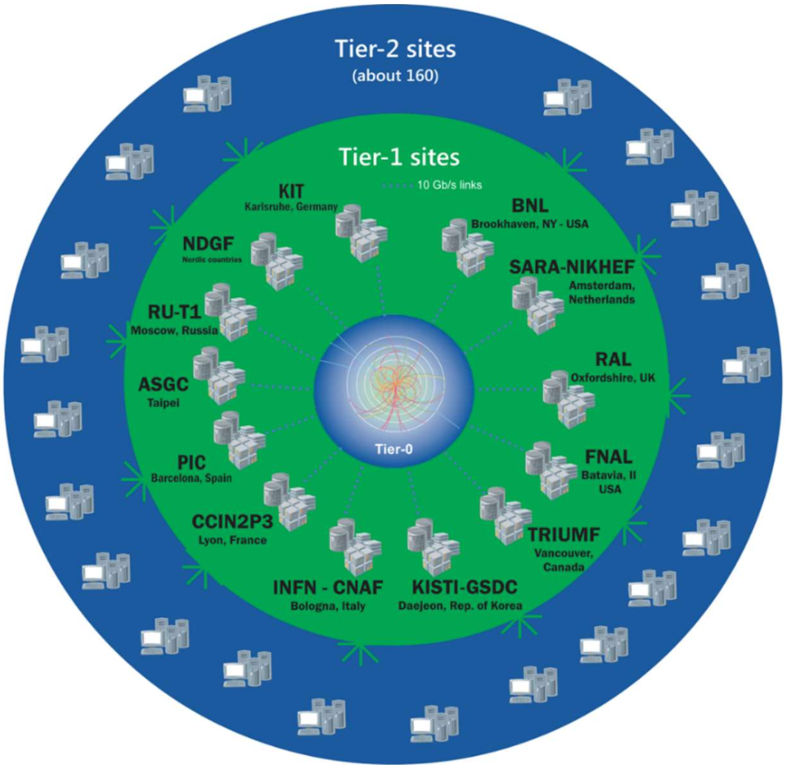


Figure 2-2 WLCG Architecture

Image src: http://wlcg-public.web.cern.ch/tier-centres

2.3 Storage Resource Manager

Storage services are important components of WLCG to serve the computing and storage resources of High Energy Physics experiments at CERN. To cater the needs of the growing large data sets, WLCG uses an interface called Storage Resource Manager (SRM) ,which ensures prevention of data loss, decrease the task analysis time and decrease error rates in data replication. As SRM's come with multiple disk arrays, parallel files systems, they are predominantly used in petascale computing.[11]

SRM is a middleware component, that provides dynamic space allocation and file management in Grid. The main functions provided by SRM interface are space management functions and data transfer functions [12]. Space management functions allows user to reserve, manage or release any space on their computation needs. Data management functions allows user to send/ receive files to/from their computer to the remote storage in the grid.

Grid FTP [13] is a secure, reliable data transfer protocol used to transfer files to / from the grid. Files can be downloaded simultaneously from multiple sources. FTP does not allow a portion of file to be transferred, but Grid FTP allows a subset of file to be transferred. Grid FTP also provides a fault tolerant implementation [14], so the transfers can start automatically if any problem occurs.

Monitoring of resources and data is very crucial in such complex experiments, which requires massive data transfers between multiple centers and enormous use of resources for task analysis. The monitoring systems used by ATLAS are explained below.

2.4 Ganglia Monitoring System

The LHC cloud resources are monitored using Ganglia [15]. Ganglia is a distributed monitoring system for high performance systems like grid and clusters. It is highly scalable and uses XML for data representation, portable data transport and RRD tool [16] for data storage and visualization. RRD (Round Robin Database) tool is a system to store and display time-series data [17] such as network bandwidth, temperatures and so on. The tool extracts and processes data and display the results in meaningful graphs. The implementation is very robust and is used on thousands of clusters around the world.

In Ganglia, a cluster is a group of hosts, that has a similar configuration. The Ganglia Monitoring System consists of three components [18].

a) Ganglia Monitoring Daemon (gmond): It runs on each cluster host and monitors changes in the host state. It collects the metric changes and distributes within each cluster.

b) Ganglia Metadata Daemon (gmetad): It polls every cluster, receives metric data and stores the data in to the database using RRDtool.

c) Ganglia Web Frontend (gweb): It is the web-based user front end for the Ganglia Monitoring System. This interface accesses the database and provides the metric result as graphs to the user.

See Figure 2-3 for Ganglia Monitoring System architecture.

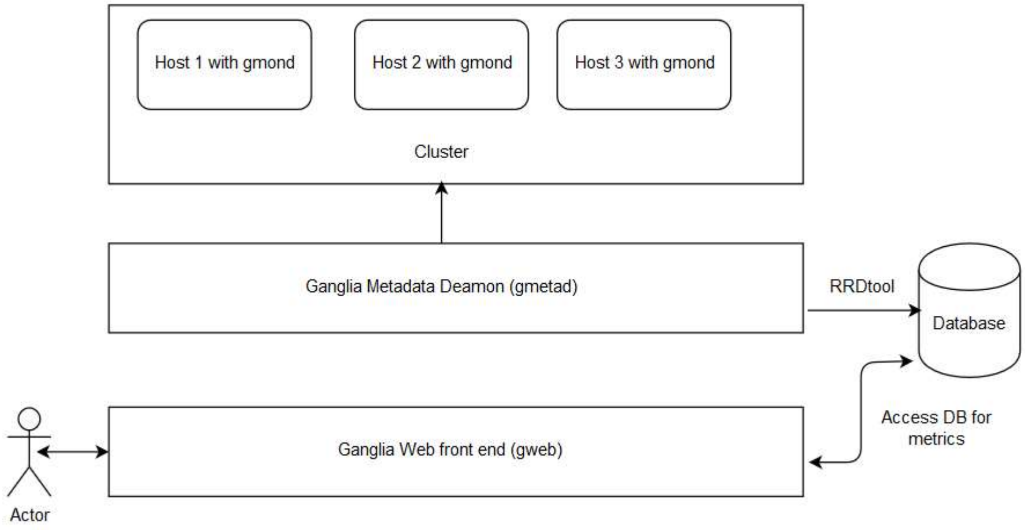
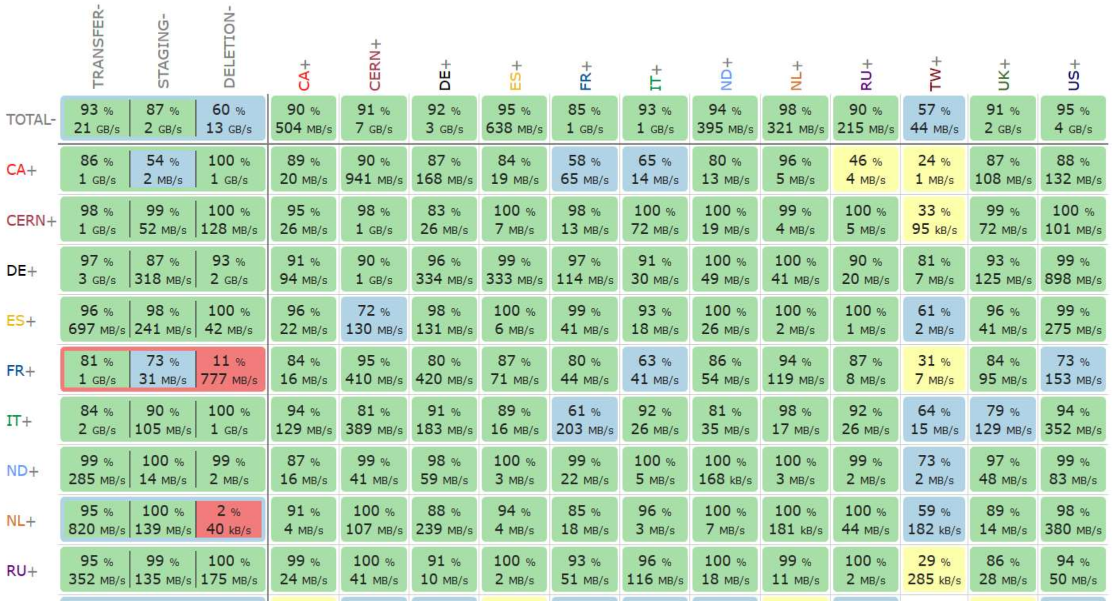


Figure 2-3 Ganglia Monitoring System Architecture

2.5 ATLAS Data Monitoring System

ATLAS Distributed Data Management (DDM) is a system used by ATLAS for managing large volume of data. It is based on the Rucio [19] project, which is an evolution from the ATLAS DDM system Don Quixote 2 (DQ2). Rucio ensures system scalability and addresses the needs of complex scientific experiments like ATLAS. Rucio manages the entire life cycle of the experimental data from the raw data to the derived data. It uses parallel and distributed mechanism to ensure safety and performance of data [20]. One of the main components of Rucio is the user interface. The users can view details about the experiment data, data transfers and deletions. See Figure 2-4 for ATLAS DDM.



	TRANSFER-	STAGING-	DELETION-	CA+	CERN+	DE+	ES+	FR+	IT+	ND+	NL+	RU+	TW+	UK+	US+
TOTAL-	93 % 21 GB/s	87 % 2 GB/s	60 % 13 GB/s	90 % 504 MB/s	91 % 7 GB/s	92 % 3 GB/s	95 % 638 MB/s	85 % 1 GB/s	93 % 1 GB/s	94 % 395 MB/s	98 % 321 MB/s	90 % 215 MB/s	57 % 44 MB/s	91 % 2 GB/s	95 % 4 GB/s
CA+	86 % 1 GB/s	54 % 2 MB/s	100 % 1 GB/s	89 % 20 MB/s	90 % 941 MB/s	87 % 168 MB/s	84 % 19 MB/s	58 % 65 MB/s	65 % 14 MB/s	80 % 13 MB/s	96 % 5 MB/s	46 % 4 MB/s	24 % 1 MB/s	87 % 108 MB/s	88 % 132 MB/s
CERN+	98 % 1 GB/s	99 % 52 MB/s	100 % 128 MB/s	95 % 26 MB/s	98 % 1 GB/s	83 % 26 MB/s	100 % 7 MB/s	98 % 13 MB/s	100 % 72 MB/s	100 % 19 MB/s	99 % 4 MB/s	100 % 5 MB/s	33 % 95 kB/s	99 % 72 MB/s	100 % 101 MB/s
DE+	97 % 3 GB/s	87 % 318 MB/s	93 % 2 GB/s	91 % 94 MB/s	90 % 1 GB/s	96 % 334 MB/s	99 % 333 MB/s	97 % 114 MB/s	91 % 30 MB/s	100 % 49 MB/s	100 % 41 MB/s	90 % 20 MB/s	81 % 7 MB/s	93 % 125 MB/s	99 % 898 MB/s
ES+	96 % 697 MB/s	98 % 241 MB/s	100 % 42 MB/s	96 % 22 MB/s	72 % 130 MB/s	98 % 131 MB/s	100 % 6 MB/s	99 % 41 MB/s	93 % 18 MB/s	100 % 26 MB/s	100 % 2 MB/s	100 % 1 MB/s	61 % 2 MB/s	96 % 41 MB/s	99 % 275 MB/s
FR+	81 % 1 GB/s	73 % 31 MB/s	11 % 777 MB/s	84 % 16 MB/s	95 % 410 MB/s	80 % 420 MB/s	87 % 71 MB/s	80 % 44 MB/s	63 % 41 MB/s	86 % 54 MB/s	94 % 119 MB/s	87 % 8 MB/s	31 % 7 MB/s	84 % 95 MB/s	73 % 153 MB/s
IT+	84 % 2 GB/s	90 % 105 MB/s	100 % 1 GB/s	94 % 129 MB/s	81 % 389 MB/s	91 % 183 MB/s	89 % 16 MB/s	61 % 203 MB/s	92 % 26 MB/s	81 % 35 MB/s	98 % 17 MB/s	92 % 26 MB/s	64 % 15 MB/s	79 % 129 MB/s	94 % 352 MB/s
ND+	99 % 285 MB/s	100 % 14 MB/s	99 % 2 MB/s	87 % 16 MB/s	99 % 41 MB/s	98 % 59 MB/s	100 % 3 MB/s	99 % 22 MB/s	100 % 5 MB/s	100 % 168 kB/s	100 % 3 MB/s	99 % 2 MB/s	73 % 2 MB/s	97 % 48 MB/s	99 % 83 MB/s
NL+	95 % 820 MB/s	100 % 139 MB/s	2 % 40 kB/s	91 % 4 MB/s	100 % 107 MB/s	88 % 239 MB/s	94 % 4 MB/s	85 % 18 MB/s	96 % 3 MB/s	100 % 7 MB/s	100 % 181 kB/s	100 % 44 MB/s	59 % 182 kB/s	89 % 14 MB/s	98 % 380 MB/s
RU+	95 % 352 MB/s	99 % 135 MB/s	100 % 175 MB/s	99 % 24 MB/s	100 % 41 MB/s	91 % 10 MB/s	100 % 2 MB/s	93 % 51 MB/s	96 % 116 MB/s	100 % 18 MB/s	99 % 11 MB/s	100 % 2 MB/s	29 % 285 kB/s	86 % 28 MB/s	94 % 50 MB/s

Figure 2-4 ATLAS DDM System for Data Monitoring

Image src: http://dashb-atlas-ddm.cern.ch/ddm2/#

Chapter 3

The Current problem with ATLAS DDM

In the current dashboard of the ATLAS DDM, it is laborious to view the metrics such as the successes, failures, efficiency, throughput of the data transfers / deletions of the files within any site in a Regional Center. By default, the dashboard shows the metric details for all the destinations. See Figure 3-1 for a sample chart to view the number of files deleted for all the clouds.

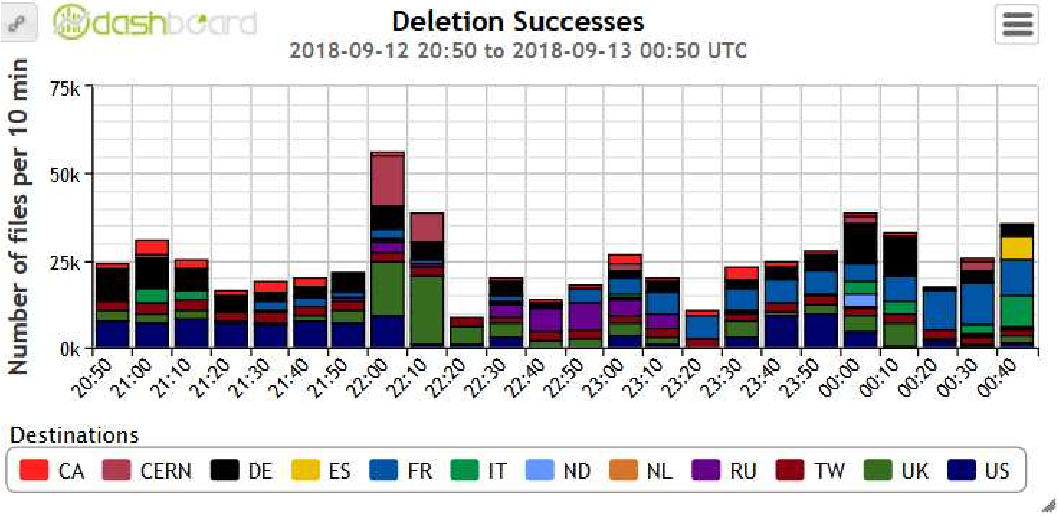


Figure 3-1 Successful deletion details for all clouds

Image src: http://dashb-atlas-ddm.cern.ch/ddm2/#tab=deletion_plots

The administrator needs to set various filters to fetch the visualizations of the above metrics for any individual site and it is really daunting to drill down to the exact chart. See Table 3-1 for various filters used in the dashboard.

Serial Number	**Filter Description**
1.	Time Interval, E.g. Last hour, Last 4 hours
2.	Activities, E.g. Staging, Recovery, Deletion
3.	Sources, E.g. Tiers, Countries, Cloud, Sites
4.	Destination, E.g. Tiers, Countries, Cloud, Sites
5.	Transfer, E.g. Efficiency, Throughput, Success, Errors
6.	Deletion, E.g. Efficiency, Throughput, Planned, Success, Errors

Table 3-1 Filters in ATLAS DDM

Each cloud has several regional centers and every regional center has many sites. See Figure 3-2 for a sample chart, which shows the number of files deleted in the US, after setting the required filters. (Filter: Clouds: US)

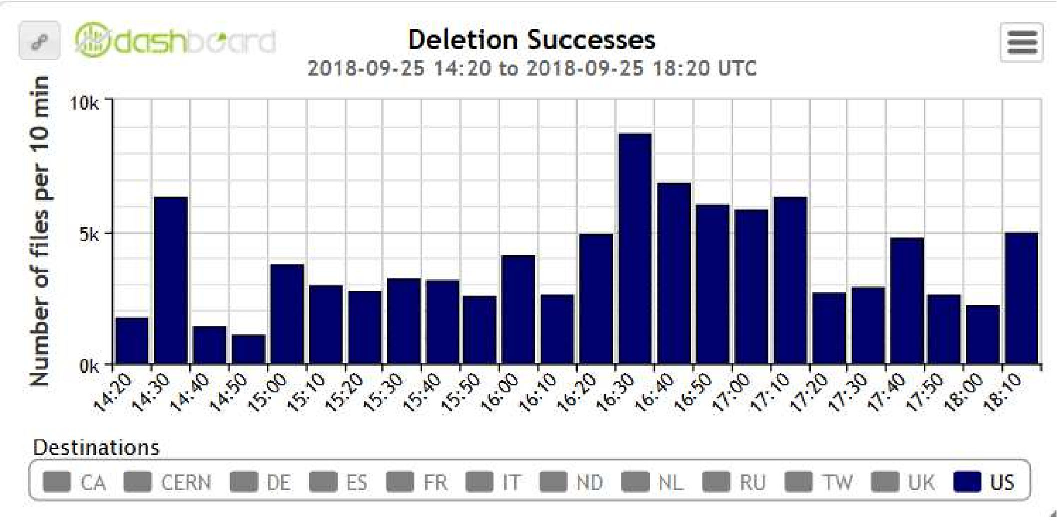


Figure 3-2 Successful deletion details at US cloud

Image src: http://dashb-atlas-ddm.cern.ch/ddm2/#tab=deletion_plots

There is no option in the current dashboard to visualize both deletion and transfer details in a wholistic view for all sites in any regional center. See Figure 3-3 for a sample chart, which shows the deletion failures in SWT2_CPB site, after setting the required filters. (Filter: Clouds: US; Sites: US SWT2_CPB DATADISK; Interval: Last 12 hours).

The chart below shows on the disk basis but does not provide a whole picture of the entire SWT2_CPB site. To overcome the hard process of visualization and to address the current problem, a web-based interface is developed to monitor the health for all the Tier 2 sites of US South West Regional Center.

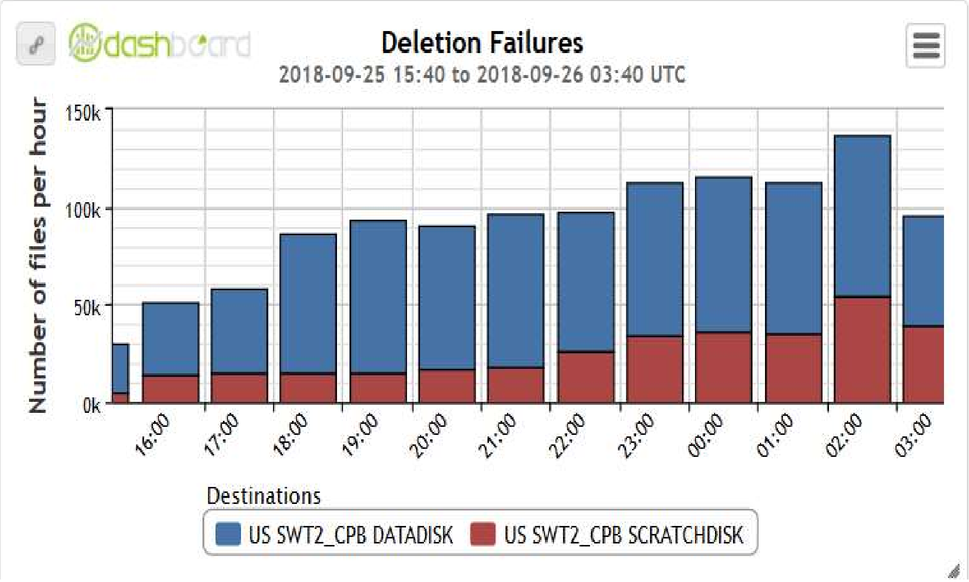


Figure 3-3 Deletion failures at the SWT2_CPB site

Image src: http://dashb-atlas-ddm.cern.ch/ ddm2/ dst.cloud=(%22US%22)& dst.site=(US,SWT2_CPB,DATADISK)&grouping.dst=(cloud,site,token)&m.content=(d_dot, d_eff,d_plf,s_err,s_suc,t_eff,t_thr)&s.state=DELETION_FAILED&tab=deletion_plots

Chapter 4

Cluster Health Monitoring Web Framework

4.1 Technical Environment

The web-based monitoring framework is developed on an environment, which uses the following:

a. Programing Language – Python

b. Web Framework for Python – Flask

c. Visualization libraries – matplotlib, plotly

d. Configuration settings – Easy settings library

e. JavaScript

f. Framework for HTML and CSS – Bootstrap

g. Cron – Task Scheduler

h. Shell script for cron job

i. Database - MySQL

4.2 Design

The web framework is built with web API's and web resources. Various processes are handled for the framework design and are explained below.

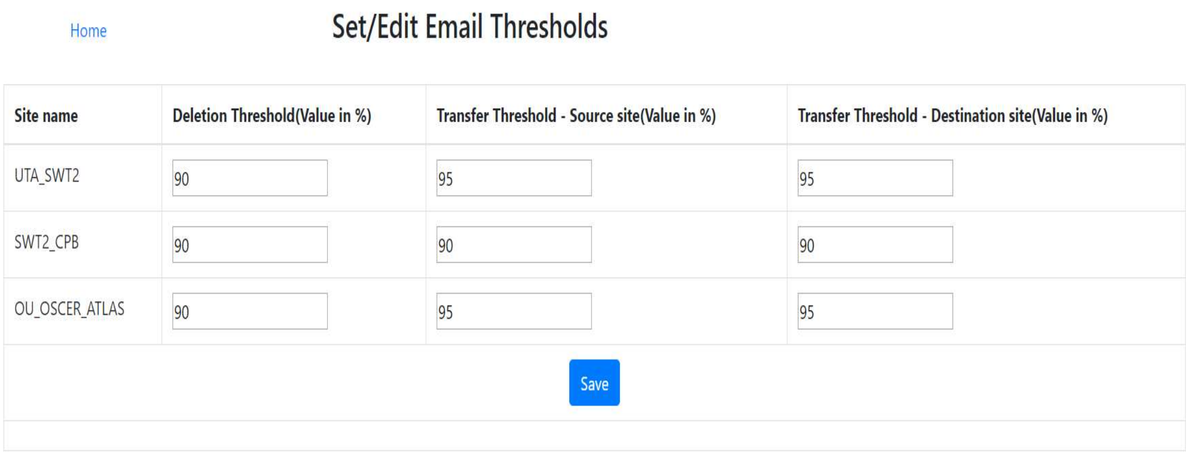
4.2.1 ETL (Extract, Transform and Load)

ETL [21] is a process, which has three phases. The first phase is 'Extract', in which the data is fetched from a single or multiple source. The second phase is 'Transform', in which the data is cleaned, validated and made into a proper storage format. The final phase is 'Load', in which the cleaned data is inserted into a working database. The framework uses ETL process to fetch data from ATLAS DDM API, which is received in JSON (JavaScript Object Notation) [22]. JSON is a language independent data format and the data objects are represented in key-value pair and array data type.

The JSON data contains the details of deletions, transfers and staging processes for all the sites related to ATLAS experiments. The thesis looks at the deletions and transfers, which are more important in health monitoring and focuses on the sites of the US Southwest Regional Center. The data is then serialized for deletion and transfer metrics and inserted into the respective tables in MYSQL.

4.2.2 Define threshold for email alerts

The inserted data explains the number of successful and failed deletions at each site. This is similar for transfer details, in which each site can be either a source or a destination. The administrators must be alerted at this point, when the deletions and transfers are failing below a defined threshold. Thresholds for deletion and transfer (both source and destination) for each site can be configured within the web framework. See Figure 4-1 for email threshold configuration for the sites at US SW Regional Center.



Set/Edit Email Thresholds

Site name	Deletion Threshold(Value in %)	Transfer Threshold - Source site(Value in %)	Transfer Threshold - Destination site(Value in %)
UTA_SWT2	90	95	95
SWT2_CPB	90	90	90
OU_OSCER_ATLAS	90	95	95

Figure 4-1 Email Threshold configuration for US SW Regional Center

4.2.3 CRON Scheduler

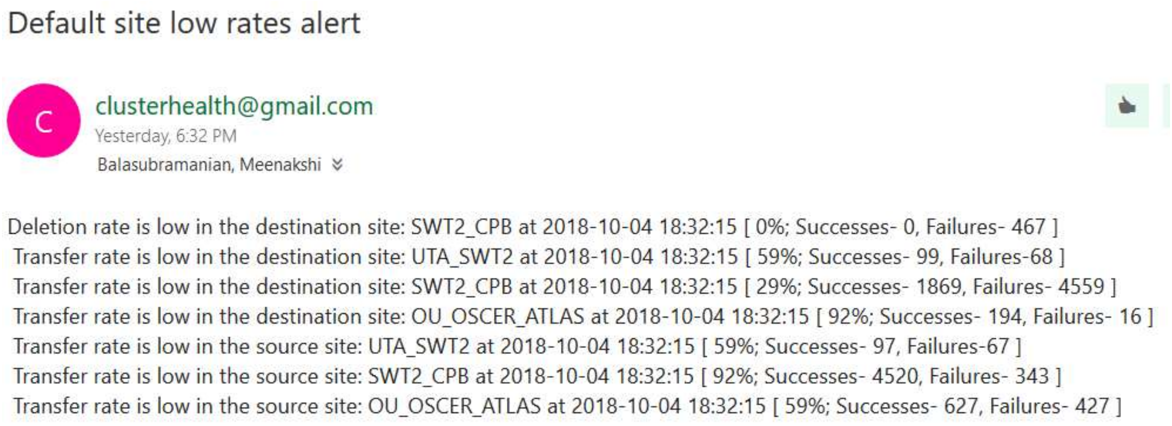
Cron [23] is a software utility in Unix based systems for running time-based job schedulers. Cron jobs are used to automate a task repeatedly at specific date or time or interval. In this thesis, a cron scheduler is set to run at a time interval of 1-hour; if cron is set for 30 mins, the alerts would be too frequent, and it is not significant to look the data so frequently. Also, a 2-hour interval is very large and there are chances that administrators fail to receive email alerts promptly.

The cron scheduler runs the script every 1 hour, which does the ETL process and sends an email alert to administrator for low thresholds. If the deletion / transfer success rates are below the defined threshold, an email alert is sent to the administrator with the details of the site and the failure point, with the count of the number of successes and failures. See Table 4-1 which shows how to find the success percentage in deletion and transfer. This is calculated on each site basis.

Serial Number	Description	Success Percentage (By site)
1.	Success percentage in deletion	Total number of done files / (Total number of done files +Total number of failed files)
2	Success percentage in transfer	Total number of transferred files / (Total number of transferred files +Total number of failed files)

Table 4-1 Success percentage in deletion and transfer

See Figure 4-2 for a sample email alert received when deletion and transfer of files failed.



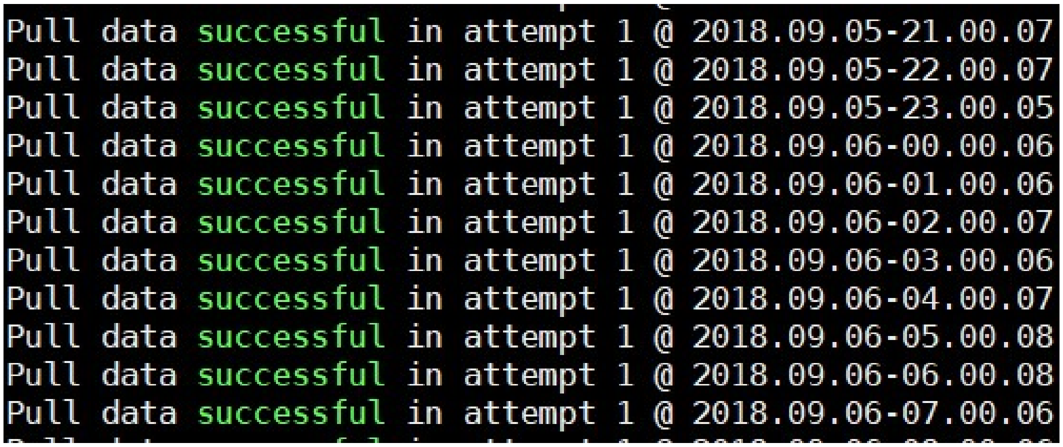
Default site low rates alert

clusterhealth@gmail.com
Yesterday, 6:32 PM
Balasubramanian, Meenakshi

Deletion rate is low in the destination site: SWT2_CPB at 2018-10-04 18:32:15 [0%; Successes- 0, Failures- 467]
Transfer rate is low in the destination site: UTA_SWT2 at 2018-10-04 18:32:15 [59%; Successes- 99, Failures-68]
Transfer rate is low in the destination site: SWT2_CPB at 2018-10-04 18:32:15 [29%; Successes- 1869, Failures- 4559]
Transfer rate is low in the destination site: OU_OSCER_ATLAS at 2018-10-04 18:32:15 [92%; Successes- 194, Failures- 16]
Transfer rate is low in the source site: UTA_SWT2 at 2018-10-04 18:32:15 [59%; Successes- 97, Failures-67]
Transfer rate is low in the source site: SWT2_CPB at 2018-10-04 18:32:15 [92%; Successes- 4520, Failures- 343]
Transfer rate is low in the source site: OU_OSCER_ATLAS at 2018-10-04 18:32:15 [59%; Successes- 627, Failures- 427]

Figure 4-2 Email alerts for deletion and transfer failures

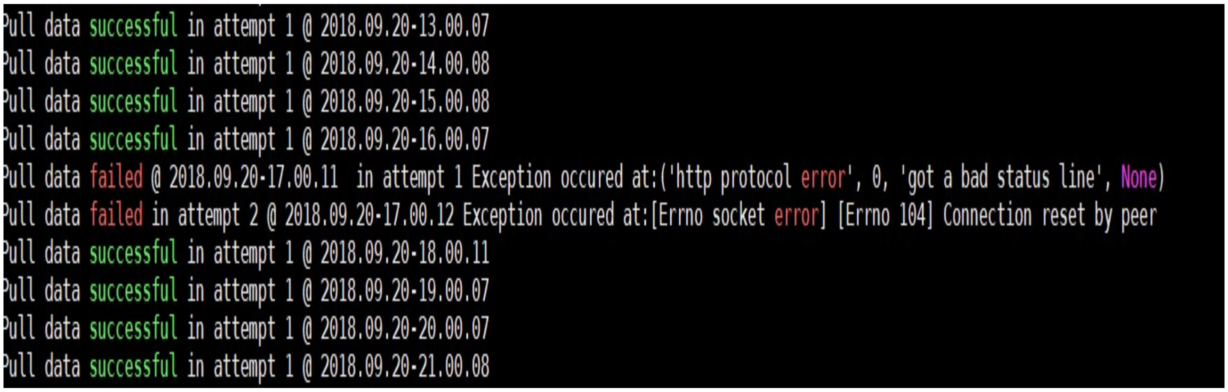
The alerts help the administrators to take corrective actions effectively. The success and failure of cron job is also logged automatically in the server. See Figure 4-3 for the sample logs generated when the cron job is successful.



```
Pull data successful in attempt 1 @ 2018.09.05-21.00.07
Pull data successful in attempt 1 @ 2018.09.05-22.00.07
Pull data successful in attempt 1 @ 2018.09.05-23.00.05
Pull data successful in attempt 1 @ 2018.09.06-00.00.06
Pull data successful in attempt 1 @ 2018.09.06-01.00.06
Pull data successful in attempt 1 @ 2018.09.06-02.00.07
Pull data successful in attempt 1 @ 2018.09.06-03.00.06
Pull data successful in attempt 1 @ 2018.09.06-04.00.07
Pull data successful in attempt 1 @ 2018.09.06-05.00.08
Pull data successful in attempt 1 @ 2018.09.06-06.00.08
Pull data successful in attempt 1 @ 2018.09.06-07.00.06
```

Figure 4-3 Log generation for successful cron jobs

When a cron job fails, the cause of failure is also logged, which helps in debugging and fixing the issues efficiently. See Figure 4-4 for the sample logs generated when the cron job fails.



```
Pull data successful in attempt 1 @ 2018.09.20-13.00.07
Pull data successful in attempt 1 @ 2018.09.20-14.00.08
Pull data successful in attempt 1 @ 2018.09.20-15.00.08
Pull data successful in attempt 1 @ 2018.09.20-16.00.07
Pull data failed @ 2018.09.20-17.00.11  in attempt 1 Exception occured at:('http protocol error', 0, 'got a bad status line', None)
Pull data failed in attempt 2 @ 2018.09.20-17.00.12 Exception occured at:[Errno socket error] [Errno 104] Connection reset by peer
Pull data successful in attempt 1 @ 2018.09.20-18.00.11
Pull data successful in attempt 1 @ 2018.09.20-19.00.07
Pull data successful in attempt 1 @ 2018.09.20-20.00.07
Pull data successful in attempt 1 @ 2018.09.20-21.00.08
```

Figure 4-4 Log generation for failed cron jobs

4.2.4 CRON Architecture

If the cron job fails for the first time, it is triggered to run again automatically. If the job fails subsequently for the second time, an email alert is sent to the administrator, with the reason of failure. See Figure 4-5 for Cron job process.

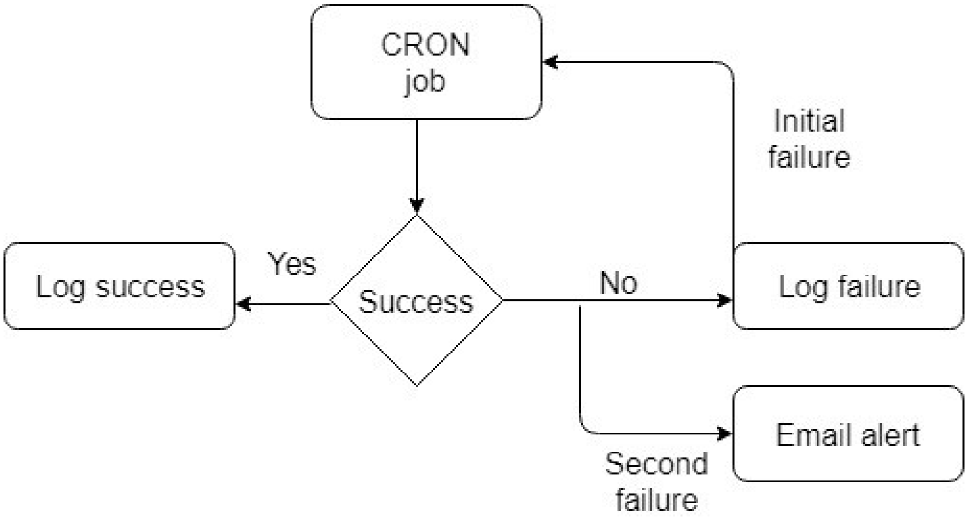


Figure 4-5 Cron job process

See Figure 4-6 for the architecture of the design process used in this thesis.

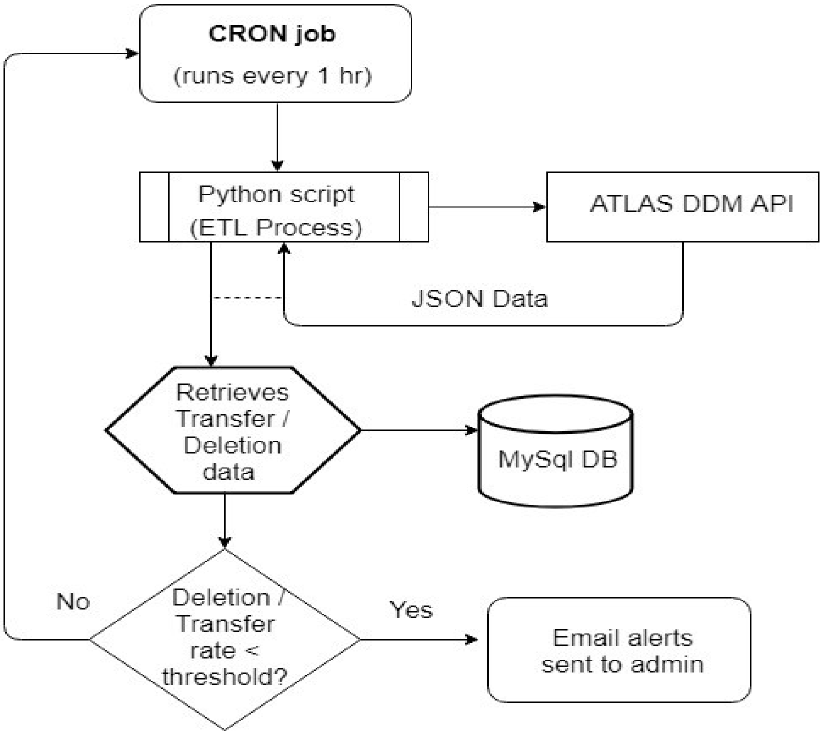


Figure 4-6 Design process architecture

The design process used in this thesis helps to collect data and reports failures effectively. With this data, we will discuss in the next chapter, how do we monitor the health of the sites.

Chapter 5

Cluster Health Monitoring Implementation

Cluster Health Monitoring [24] involves tracking and measuring data. The monitoring solution in this thesis gives the administrators an insight of how the clusters are working by providing the deletion and transfer details of files within each Tier 2 site in US South west Regional center. The deletion and transfer metrics, which are more important for health monitoring, are visualized in graphs in this thesis. This thesis uses Matplotlib and Plotly for monitoring the health of clusters in real time.

Matplotlib [25] is a visualization library which produces quality charts across any platforms. One of the important advantages [26] of matplotlib is, it can be used with any operating systems and support different output types. Plotly [27] is an open source interactive graphing library for python. It is a data visualization toolbox [28] and under every graph is a JSON object, with which any type of graph can be constructed.

5.1 Data visualization

Everyone is aware of the expression," A picture is worth a thousand words" [29], a statement which references accurately about Data visualization. We would stare a table with numbers and would never see any results, but the same are immediately obvious in visualization. Our brains interpret through verbal processing, which becomes visible and understandable when communicated visually. This is the power of "data visualization."[30]. Data visualization [31] is a general term, which means, the representation of data in a visual form that people can easily understand. Any patterns or trends cannot be detected in a text data but can easily be recognized through visualization.

5.1.1 Advantages of Data Visualization

The important advantage is, huge volume of data can be accessed easily. The human brain [32] processes visual information easily than the written information. The visual charts allow administrators to analyze critical data and move in the right direction, taking required action on time, which reduces the risk of large failures.

5.2 Implementation

A cluster [33] is a collection of datacenters. A data center is a collection of racks. A rack is a collection of servers. A node is the data storage layer within a server. The cluster manages the resources of all servers associated with it. A cluster health monitoring system [34] analyzes the cluster resources and provides real time metrics of cluster failures. The number of metrics varies from 30 to 40 [35] ,which includes basic checks like network, load, memory, disk, in addition to hardware monitoring. The metrics in this thesis, checks for deletion and transfer of files, for ATLAS Tier 2 operations in US South west region.

The cluster health monitoring solution in this thesis, monitors the following sites:

a) UTA_SWT2

b) SWT2_CPB

c) OU_OSCER_ATLAS

See Table 5-1 for the metric details monitored for the above sites.

Serial Number	Site name	Metrics Monitored
1.	UTA_SWT2	File deletions
2.	SWT2_CPB	File deletions
3.	OU_OSCER_ATLAS	File deletions
4.	UTA_SWT2	File transfer as source site

5.	SWT2_CPB	File transfer as source site
6.	OU_OSCER_ATLAS	File transfer as source site
7.	UTA_SWT2	File transfer as destination site
8.	SWT2_CPB	File transfer as destination site
9.	OU_OSCER_ATLAS	File transfer as destination site

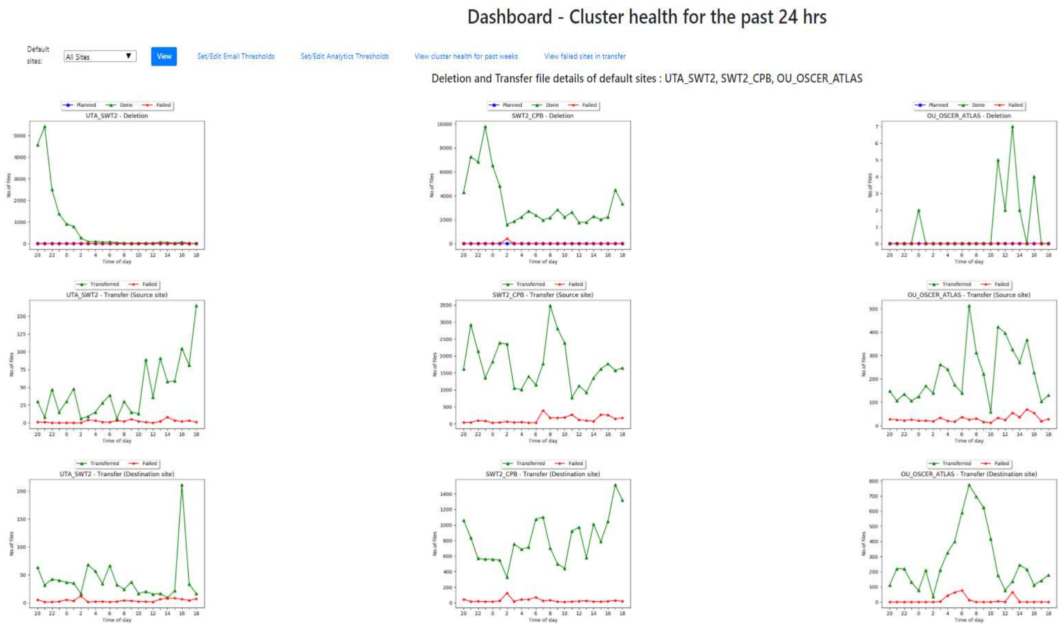
Table 5-1 Metrics monitored for the sites

See Table 5-2 for monitoring frequency in each of the above sites.

Serial Number	Monitoring Frequency
1.	Last 24-hrs
2.	Last 7 days, with 24-hrs on each day
3.	Last 14 days, with 24-hrs on each day

Table 5-2 Monitoring frequency for the sites

This thesis provides the solution to monitor the metrics of all sites in a single point of view. The overall dashboard provides file deletion metrics, file transfer metrics with both source and destination. See Figure 5-1 for the overall 24-hour monitoring dashboard for all sites.



Dashboard - Cluster health for the past 24 hrs

Default sites: All Sites ▼ View

Set/Edit Email Thresholds | Set/Edit Analytics Thresholds | View cluster health for past weeks | View failed sites in transfer

Deletion and Transfer file details of default sites : UTA_SWT2, SWT2_CPB, OU_OSCER_ATLAS

Planned
Done
Failed
UTA_SWT2 - Deletion
SWT2_CPB - Deletion
OU_OSCER_ATLAS - Deletion
Transferred
Failed
UTA_SWT2 - Transfer (Source site)
SWT2_CPB - Transfer (Source site)
OU_OSCER_ATLAS - Transfer (Source site)
UTA_SWT2 - Transfer (Destination site)
SWT2_CPB - Transfer (Destination site)
OU_OSCER_ATLAS - Transfer (Destination site)
No.of files
Time of day

Figure 5-1 Overall monitoring dashboard for all sites

The administrators can click on any chart and view the enlarged graph for a detailed view. There can be situations when the admin wanted to view a site for a clear picture. This thesis provides solution, to choose the required site from the dropdown and view the metrics on per site basis. See Figure 5-2 for deletion and transfer metrics related to OU_OSCER_ATLAS.

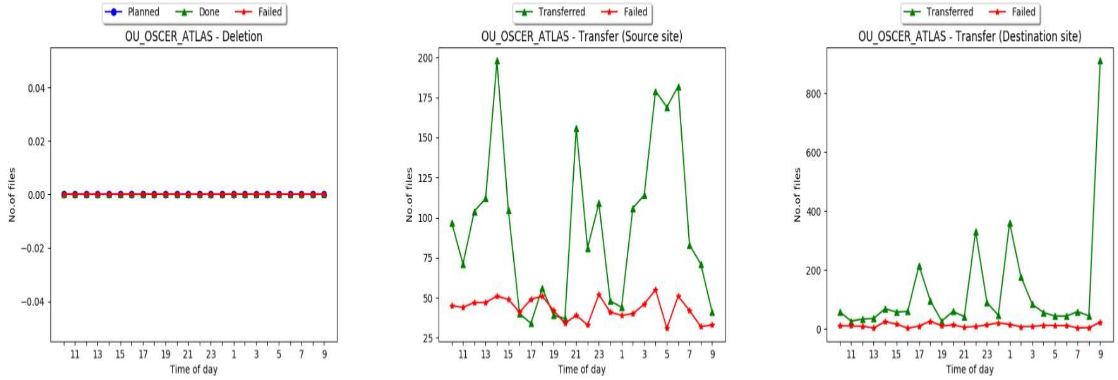


Figure 5-2 Deletion and transfer metrics in OU_OSCER_ATLAS

Similarly, the deletion and transfer metrics can be viewed separately for UTA_SWT2 and SWT2_CPB sites.

5.2.1 File deletion metrics

The file deletion metrics shows the following details for each site:

a) No.of files Planned

b) No.of files done

c) No.of files failed

The last 24-hour cluster health file deletion shows how many files are planned, how many files are deleted successfully and how many files failed deletion for any time of the day. See Figure 5-3 for the last 24-hour file deletion in SWT2_CPB.

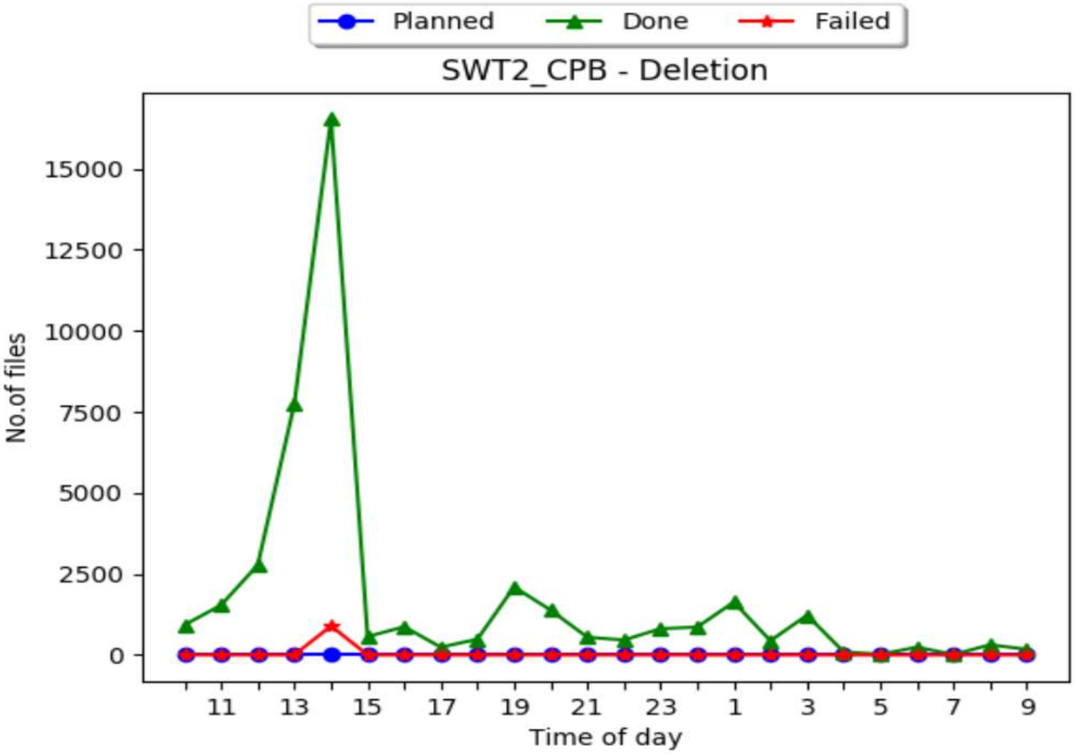


Figure 5-3 Last 24-hour file deletions in SWT2_CPB

In the above graph, it’s very explicit, that the graph is generated at 9.00am of the day, and it shows the metrics till 10.00am of the previous day. It is clear for the admin, when the graph was generated and the file deletion metrics for any hour of the day. The failure metrics makes the admin to take corrective actions on the respective clusters. There are cases, in which file deletions would not have happened for a site for the past 24-hours, although it is not a failure. Here the number of done files, failed files and planned files are zero. See Figure 5-4, when there are no file deletions for the last 24-hours.

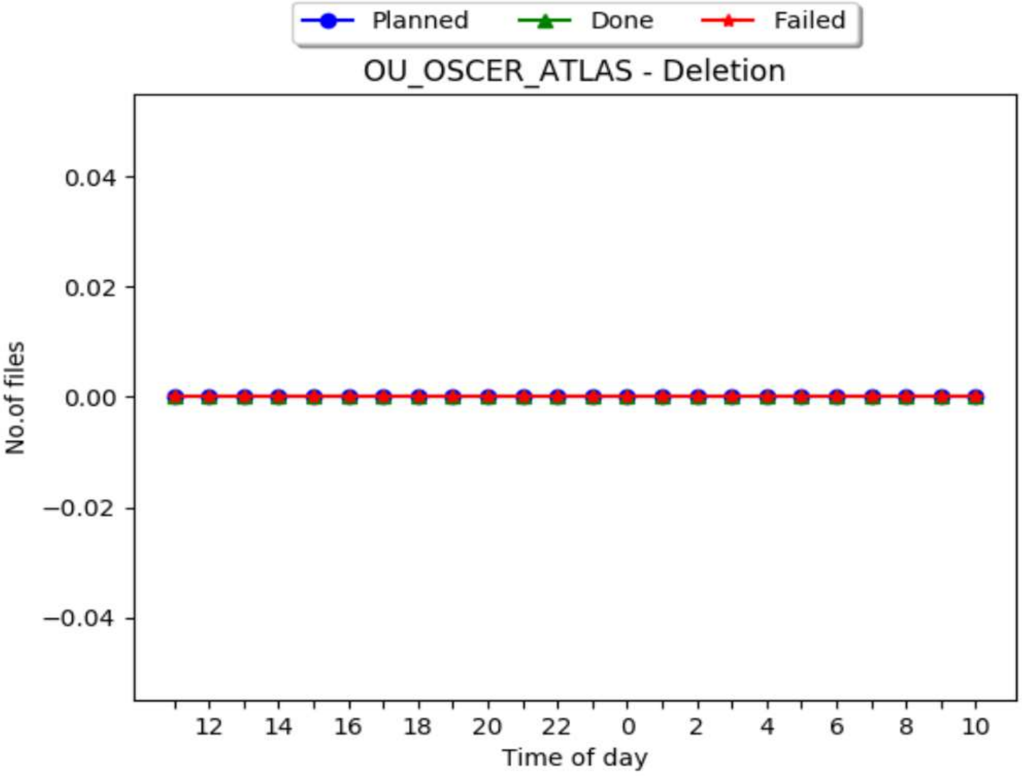


Figure 5-4 No file deletions for the last 24-hours in OU_OSCER_ATLAS

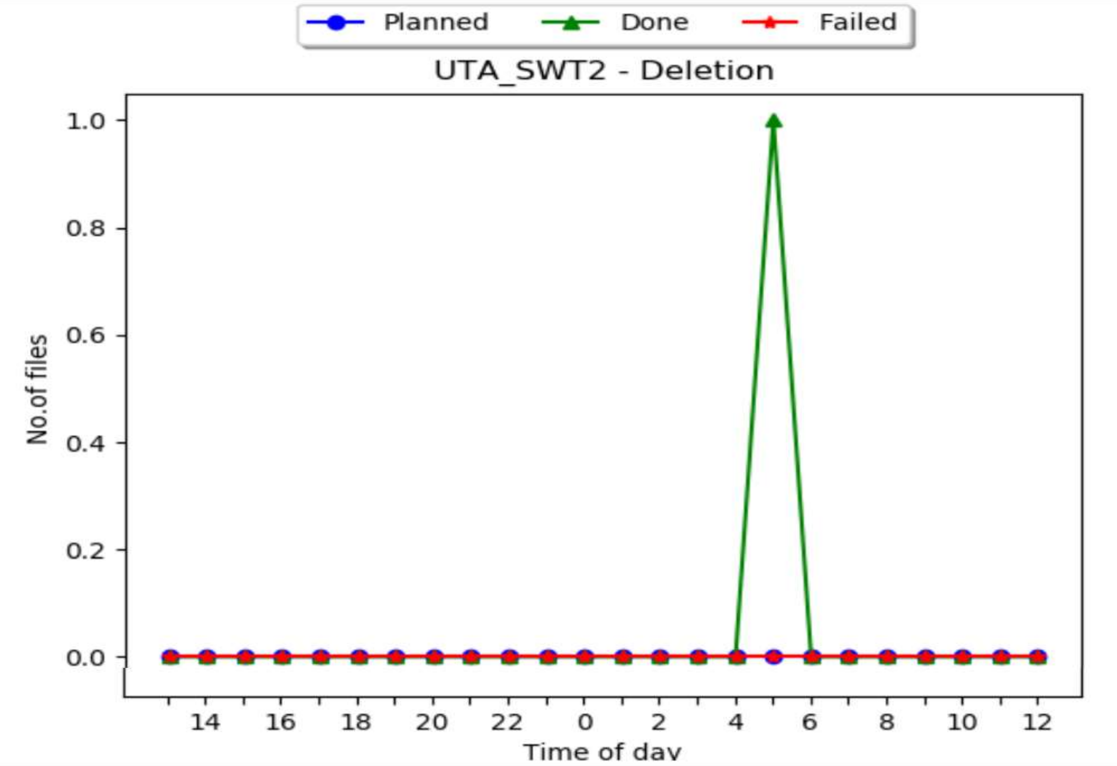


Figure 5-5 Last 24-hour file deletions in UTA_SWT2

This thesis provides admin to view the cluster health not only for the past 24-hours, but also for the past one week and past two weeks for any site. In addition to the past one- and two-weeks cluster health, the admin also monitors the cluster health for every day of the week. See Figure 5-6 for Cluster health dashboard for previous weeks.

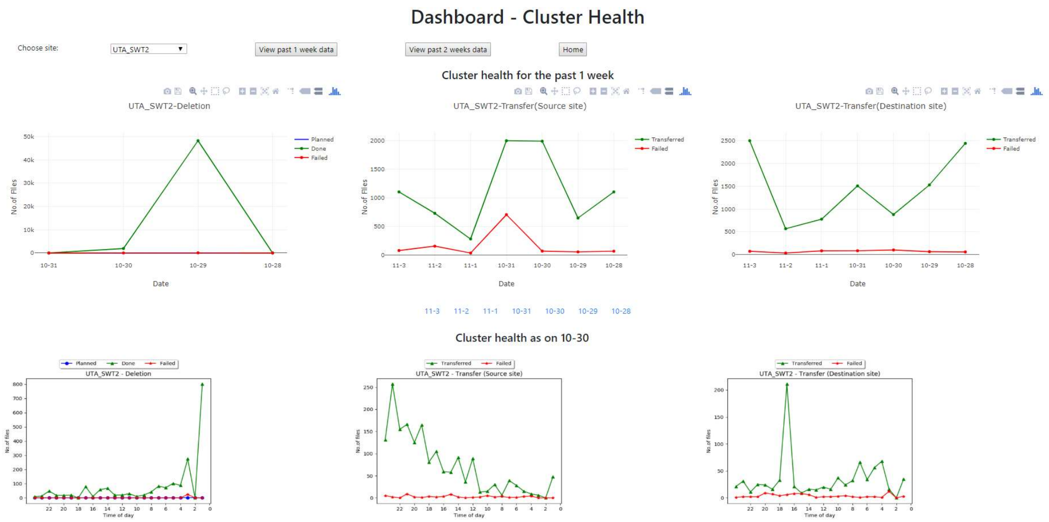

Dashboard - Cluster Health
Choose site:
UTA_SWT2
View past 1 week data
View past 2 weeks data
Home
Cluster health for the past 1 week
UTA_SWT2-Deletion
UTA_SWT2-Transfer(Source site)
UTA_SWT2-Transfer(Destination site)
Planned
Done
Failed
Transferred
Failed
No.of Files
Date
11-3 11-2 11-1 10-31 10-30 10-29 10-28
Cluster health as on 10-30
UTA_SWT2 - Deletion
UTA_SWT2 - Transfer (Source site)
UTA_SWT2 - Transfer (Destination site)
No of files
Time of day

Figure 5-6 Cluster health dashboard for the previous weeks

See Figure 5-7 , which shows the cluster health for the past one week in UTA_SWT2.

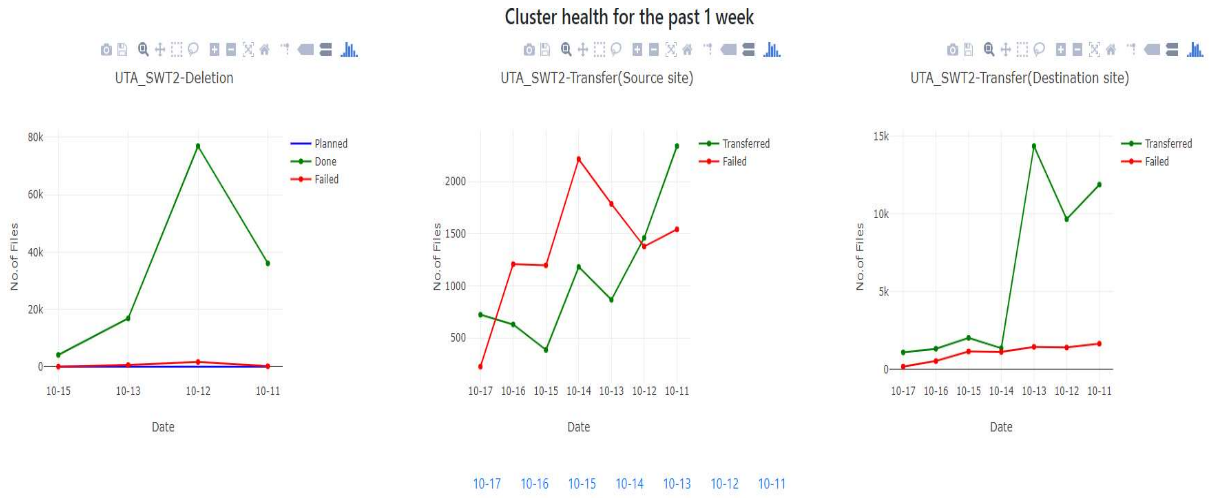


Figure 5-7 Cluster health for the past one week in UTA_SWT2

The weekly cluster health monitoring uses plotly library for visualization. See Figure 5-8 for the weekly deletion metrics for UTA_SWT2, as on 10/20/2018.

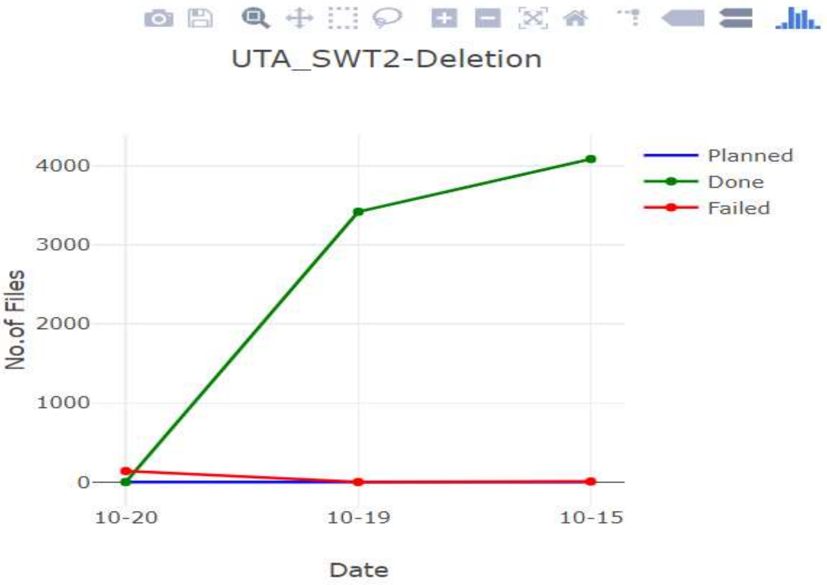


Figure 5-8 Weekly deletion metrics for UTA_SWT2

In the above chart, the other dates within the seven-day range are not displayed, which indicates no deletion operations happened on those dates. The chart provides an interactive visualization [36], which enables the manipulation of chart images with the data and color. In this thesis, the admin can view the number of deletions planned, number of deletions done successfully and number of failed deletions individually for a unique view. By clicking on the ‘Failed’ legend, the chart shows only the failure metrics. See Figure 5-9 for number of deletions failed in SWT2_CPB.

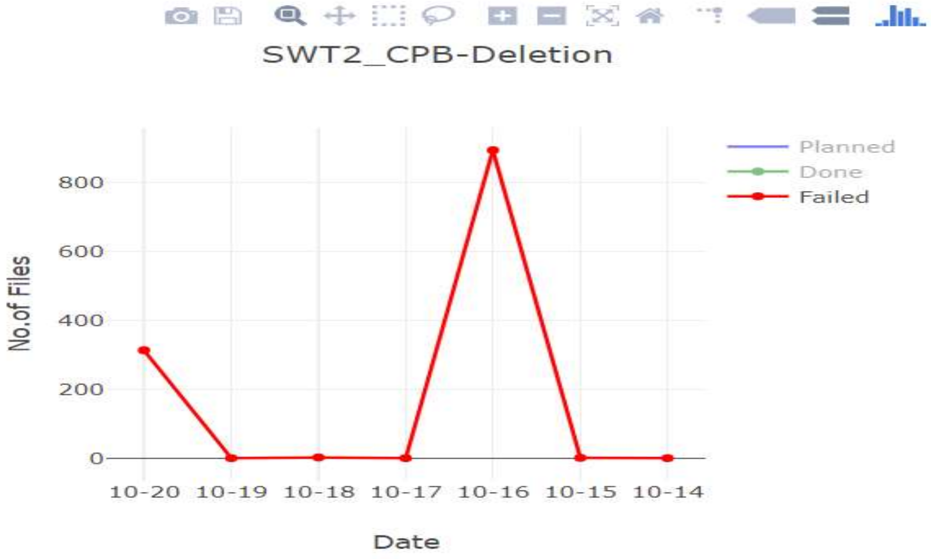


Figure 5-9 Number of deletions failed in SWT2_CPB

When clicking on the ‘Done legend, the chart shows only the files that are deleted successfully. See Figure 5-10 for number of successful file deletions in SWT2_CPB.

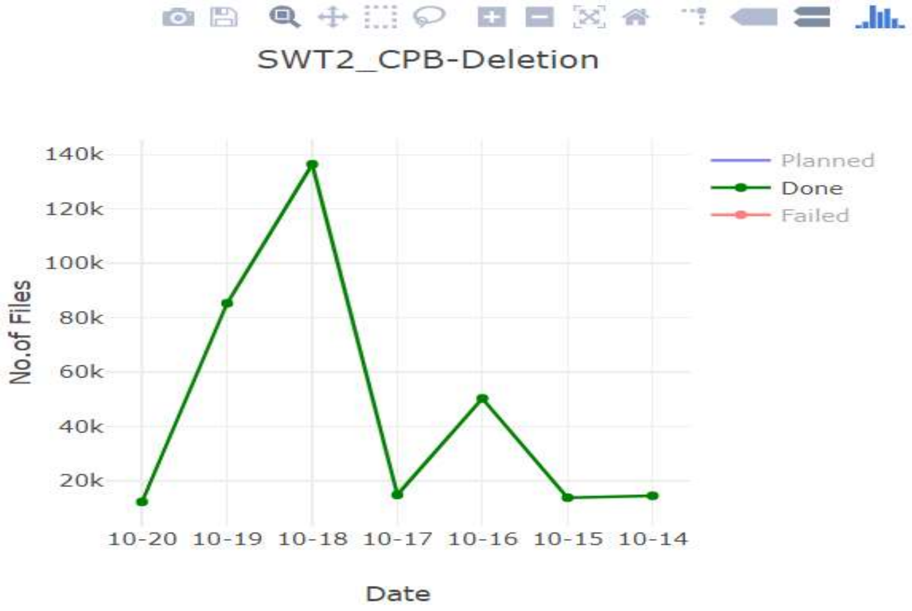


Figure 5-10 Number of successful file deletions in SWT2_CPB

As plotly library provides an interactive visualization, the administrators on hover of the chart, can view and compare the details on the number of files planned for deletion, number of files failed in deletion and number of files deleted successfully for the chosen day.

The graph can also be downloaded and saved in the local disk using camera icon in the chart. The user can also zoom in and zoom out the chart for an enlarged and detailed view. See Figure 5-11 to view the interactive visualization of deletion metrics in OU_OSCER_ATLAS site on 10/14.

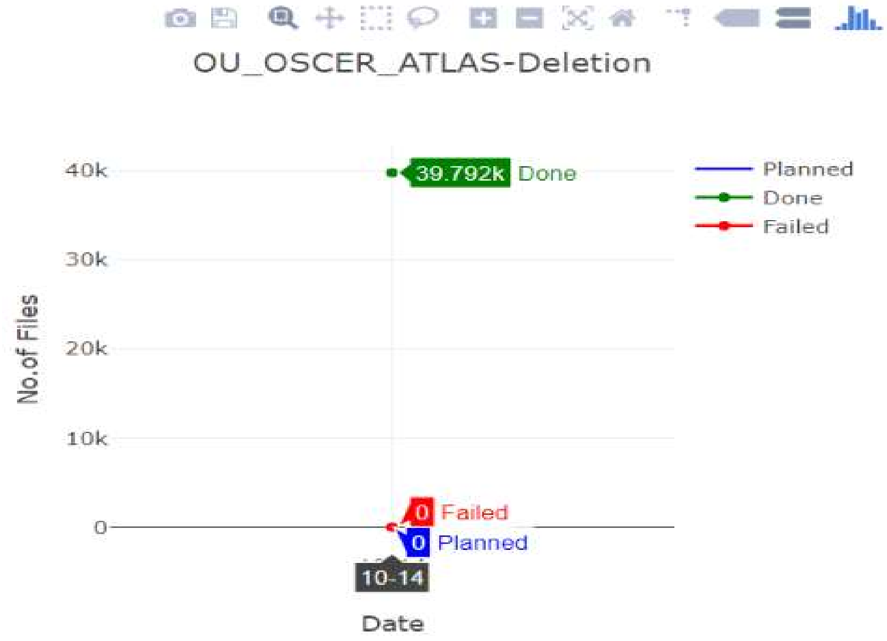


Figure 5-11 Deletion metrics in OU_OSCER_ATLAS

The deletion metrics can also be viewed for any day for the two-week time. See Figure 5-12 for overall cluster health in UTA_SWT2 for the past 14 days.

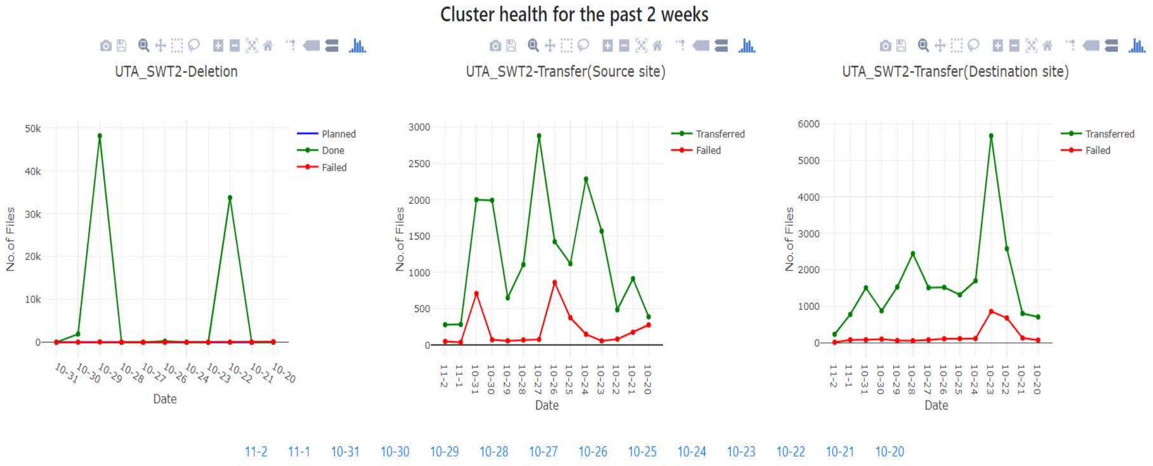


Figure 5-12 Overall cluster health in UTA_SWT2 for the past 2 weeks

The graph can be viewed for any chosen site. See Figure 5-13 for file deletion metrics in SWT2_CPB for the past 14 days.

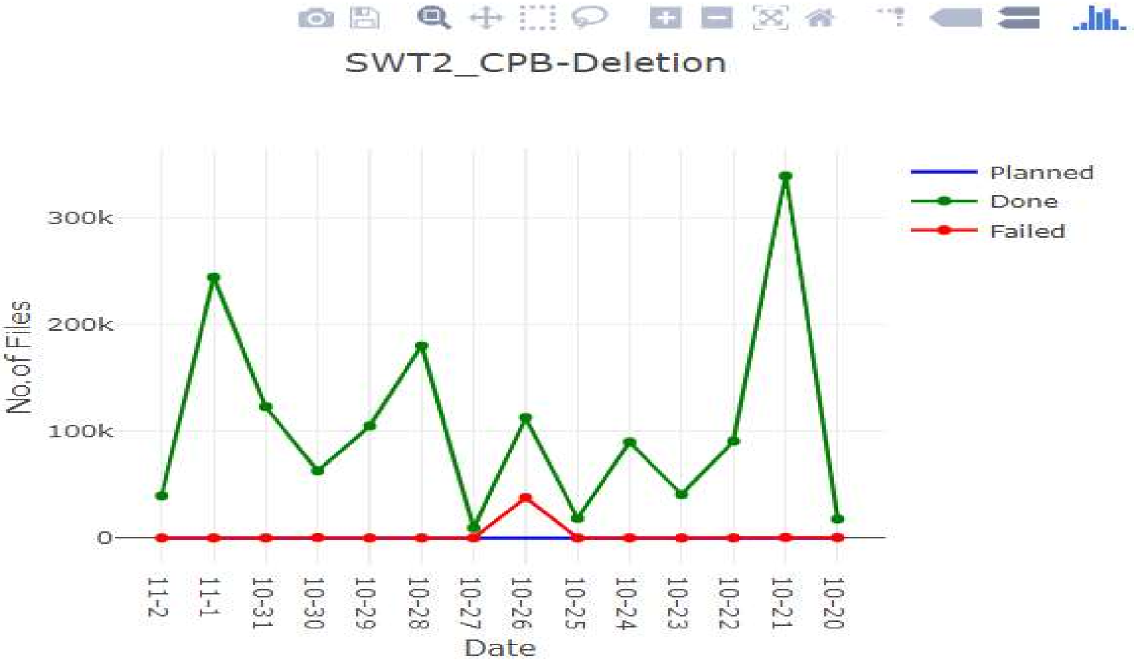


Figure 5-13 Deletion metrics in SWT2_CPB for the past 2 weeks

Starting with the current date as 11/2, the chart can be viewed for the following dates for the past two weeks.

11-2 11-1 10-31 10-30 10-29 10-28 10-27 10-26 10-25 10-24 10-23 10-22 10-21 10-20

For example, we click on the date 10-26, to view the deletion metrics for the 24-hour time starting from 12.00am of the day to 11.59pm. The x-axis scale value 1, shows the metrics of the first hour.ie. 12.00 am to 12.59am. See Figure 5-14 for file deletion metrics in SWT2_CPB for a chosen date.

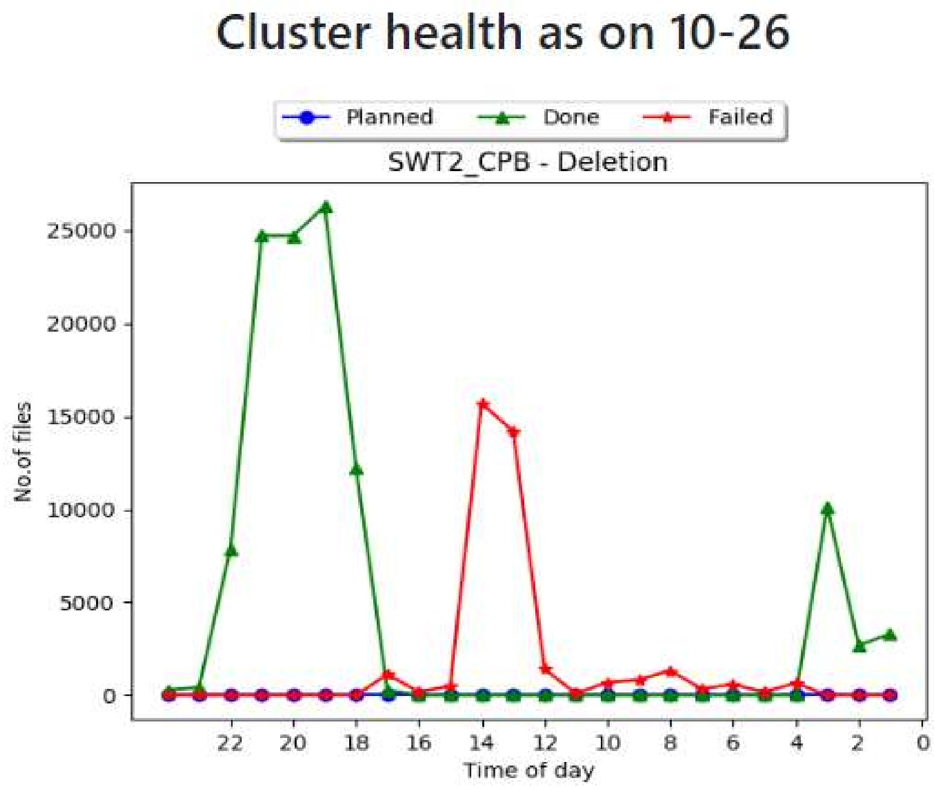


Figure 5-14 File deletion metrics in SWT2_CPB for the chosen date

With the above monitoring details, the administrators can view the deletion metrics for any site and any time just on the click of a button. They can take actions in prior, even before receiving alerts from the existing Rucio monitoring system.

5.2.2 File transfer metrics

The file transfer metrics shows the following details for each site:

a) No.of files transferred

b) No.of files failed

The 24-hour cluster health for file transfer shows how many files are transferred successfully and how many files have failed in transfer both at the source and destination for any time of the day. With file transfers, it is important for the administrators to know the details with respect to source as well as destination. See Figure 5-15 for the last 24-hour file transfers with UTA_SWT2 as source site.

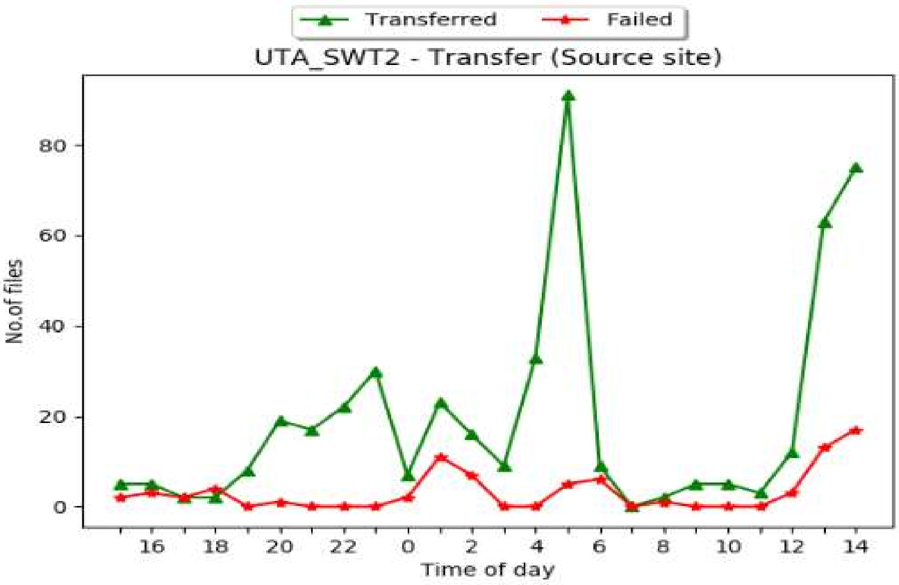


Figure 5-15 Last 24-hour file transfers with UTA_SWT2 as source

In the above graph, it is clear, that the graph is generated at 2.00pm of the day, and it shows the metrics till 3.00pm of the previous day. It is useful for the admin to know when the graph was generated and the file transfer metrics for any hour of the day. See Figure 5-16, for the last 24-hour file transfers with UTA_SWT2 as destination site.

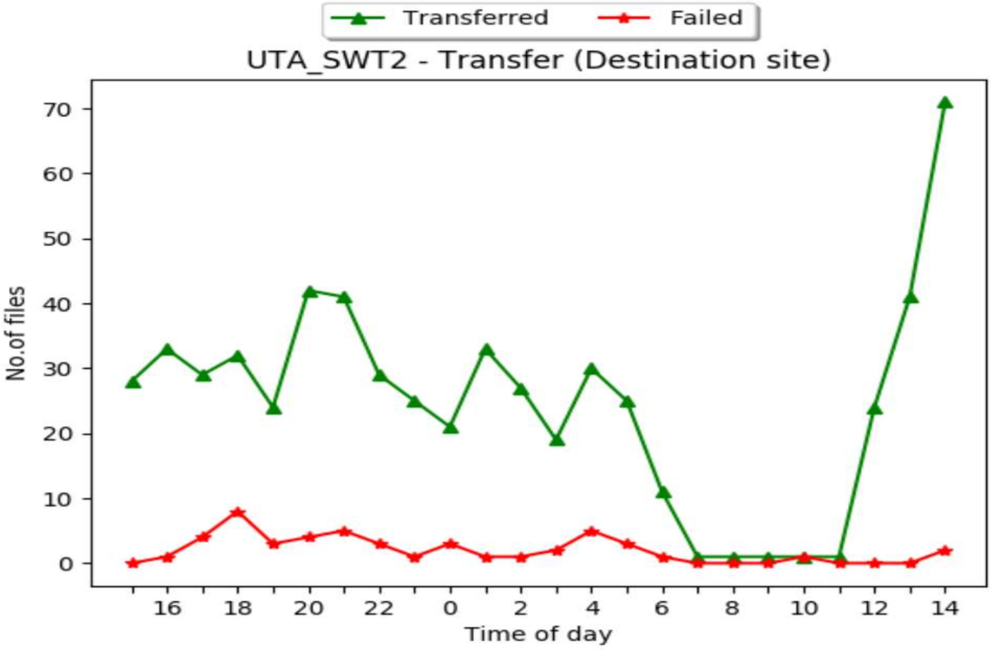


Figure 5-16 Last 24-hour file transfers with UTA_SWT2 as destination

See Figure 5-17, for the last 24-hour file transfers with SWT2_CPB as source site.

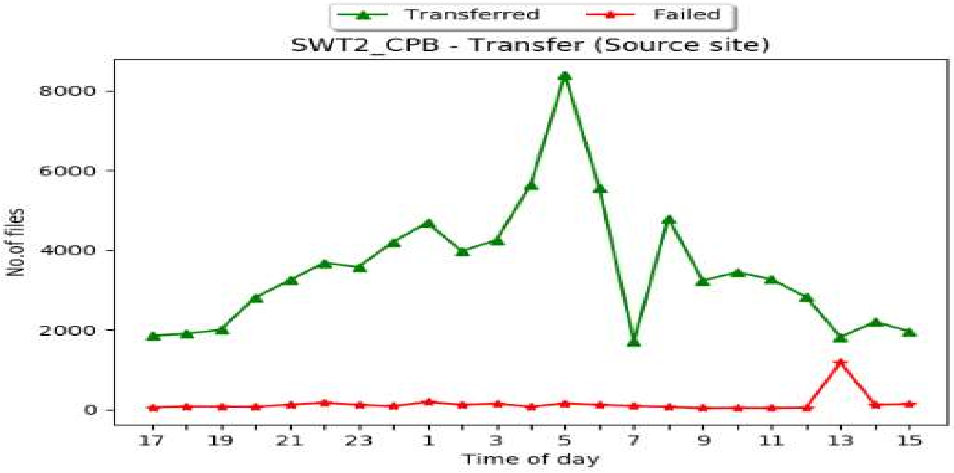


Figure 5-17 Last 24-hour file transfers with SWT2_CPB as source

See Figure 5-18, for the last 24-hour file transfers with OU_OSCER_ATLAS as destination site.

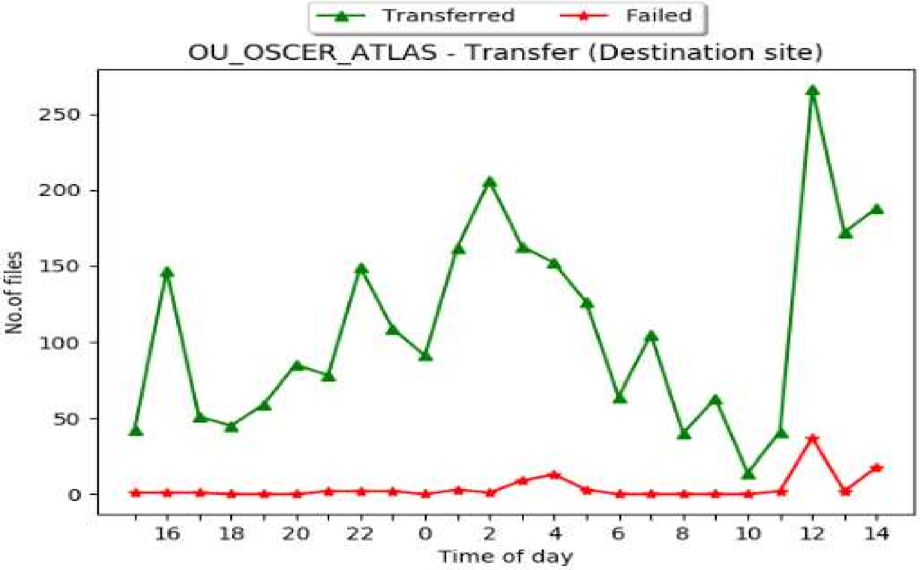
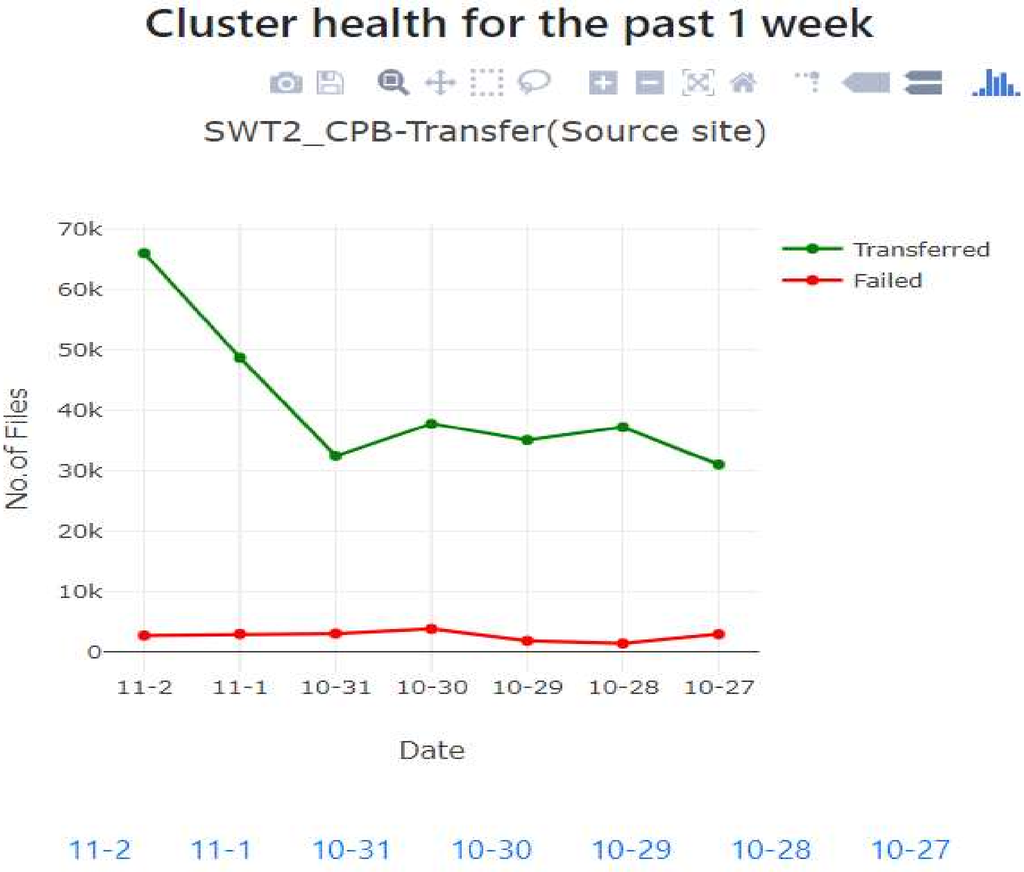


Figure 5-18 Last 24-hour file transfers with OU_OSCER_ATLAS as destination

The admin can also view the transfer metrics for the past one week and past two weeks for any site for both source and destination. See Figure 5-19 for the transfer details for the past one week in SWT2_CPB as source.



Cluster health for the past 1 week

Figure 5-19 SWT2_CPB as source transfers for the past one week

The administrator can click on any date to view the transfer metrics for the site with respect to source and destination. See Figure 5-20 for the transfer details for the past one week in SWT2_CPB as destination site.

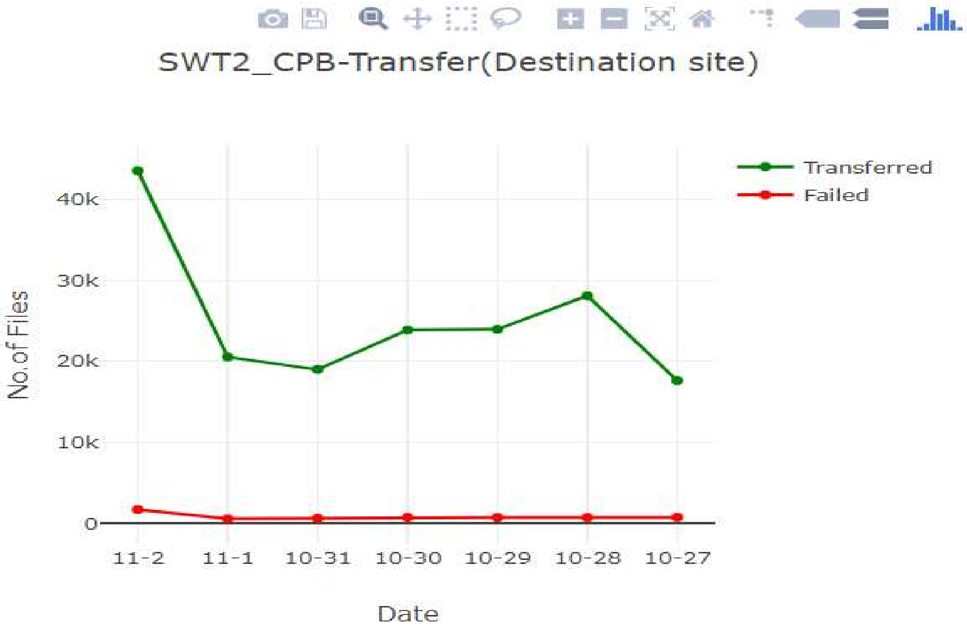


Figure 5-20 SWT2_CPB as destination transfers for the past one week

The admin can have a comparative view of transfers and failures like the above graph or can view the number of successful and failed transfers separately by clicking on the respective legends. See Figure 5-21 that shows only the successful transfers at UTA_SWT2 as the source site for the past one week. See Figure 5-22 that shows only the failed transfers at UTA_SWT2 as the source site for the past one week.

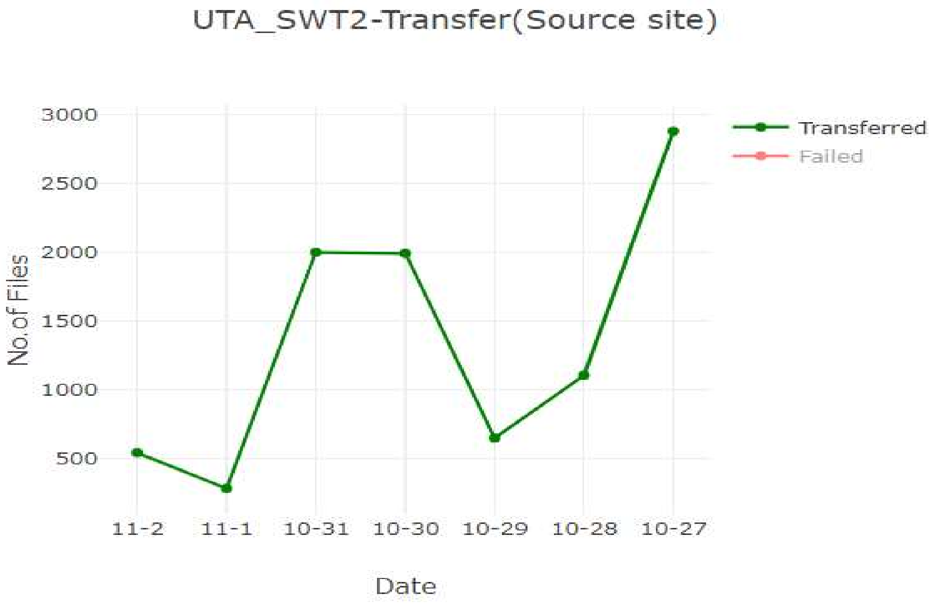


Figure 5-21 Successful transfers at UTA_SWT2 as source

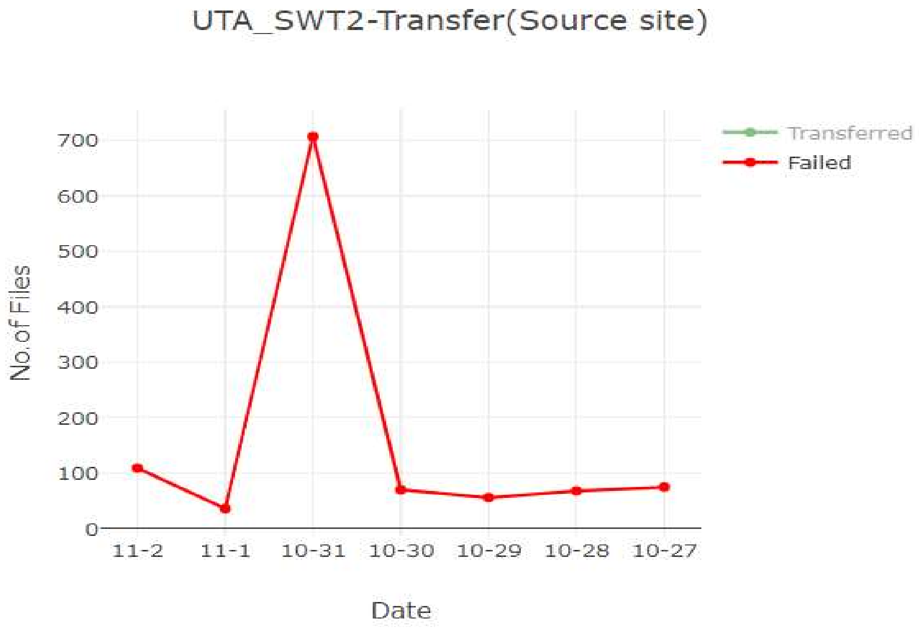


Figure 5-22 Failed transfers at UTA_SWT2 as source

The admin can also view the transfer details for the past two-week period. See Figure 5-23 for transfers at OU_OSCER_ATLAS as the source site for the past two weeks.

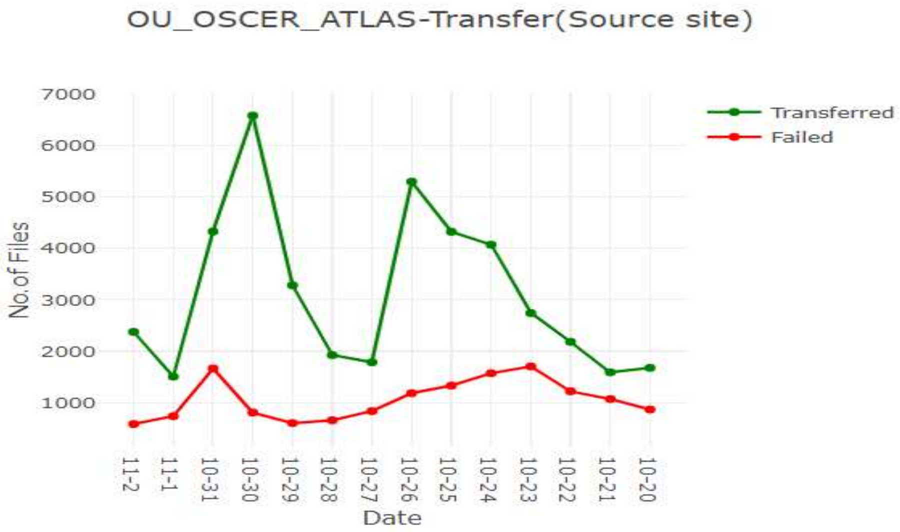


Figure 5-23 Transfers at OU_OSCER_ATLAS as source for the past two weeks

See Figure 5-24 for transfers at OU_OSCER_ATLAS as the destination site.

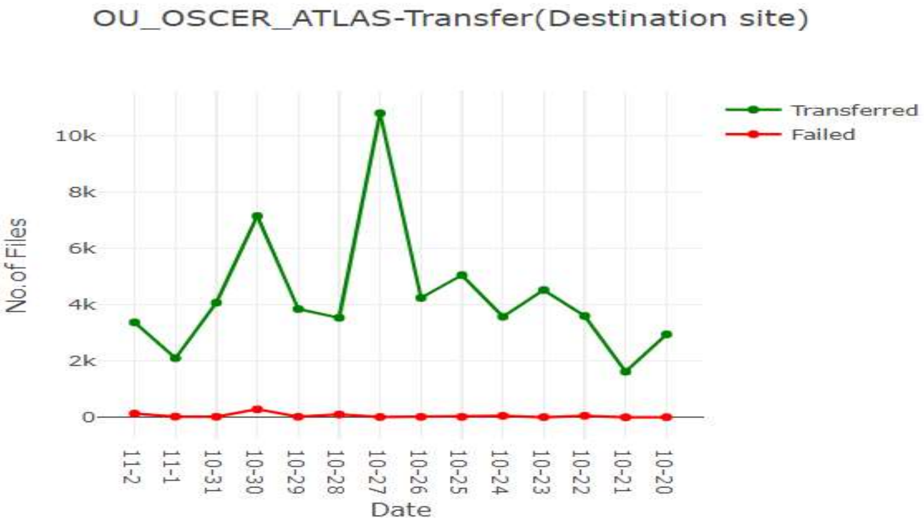


Figure 5-24 Transfers at OU_OSCER_ATLAS as destination for the past two weeks

The admin can view the transfer details of any site for any date under the two weeks period. See Figure 5-25 for transfers at SWT2_CPB as source for a chosen date.

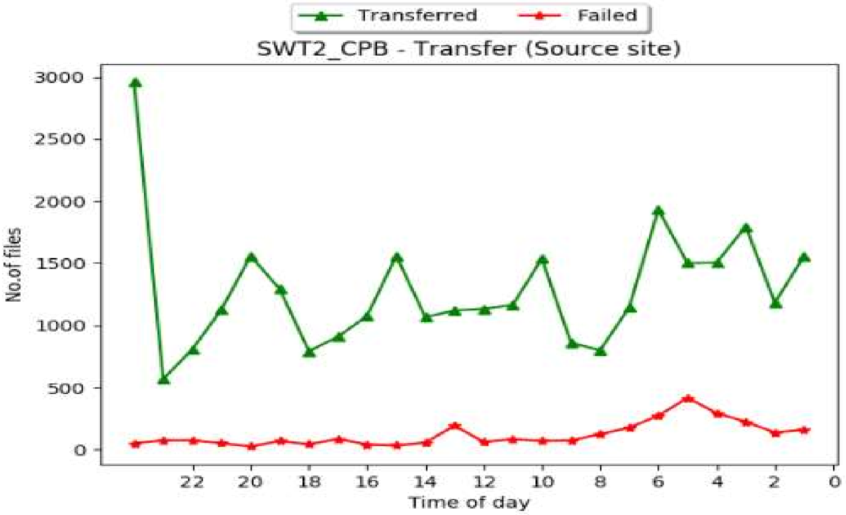


Figure 5-25 Transfers at SWT2_CPB as source for a chosen date

See Figure 5-26 for transfers at SWT2_CPB as destination for a chosen date.

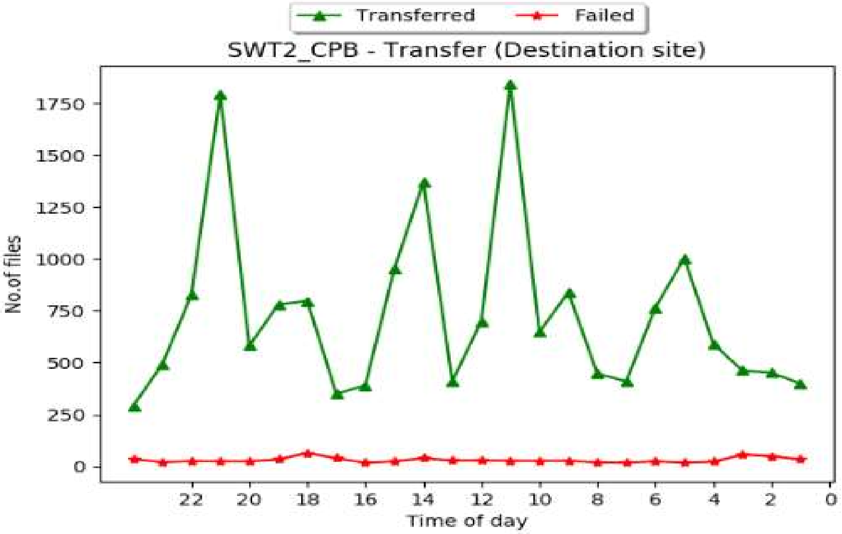


Figure 5-26 Transfers at SWT2_CPB as destination for a chosen date

The monitoring of file deletion and transfer metrics in this thesis provides administrators a more efficient solution than the existing Rucio monitoring system. This thesis provides real time metrics and required information for monitoring the health of clusters and the solution thus becomes more effective. This work does not stop in monitoring the cluster health, but also analyzes the failures in file transfers happening between two performing sites. The thesis explains how the failures are analyzed in the next chapter.

Chapter 6

Failure Analysis in File transfers

ATLAS experiments create a large volume of data. High computing systems [37] are used to analyze the data. The data is growing every year and the computation involves more input output operations per second (IOPS) [38]. To handle such enormous volume of data and perform computations, the applications need to be distributed over clusters [39].

6.1 Cluster Failures

The clusters must be protected against the three important types of outages [40] which are mentioned below:

a) Application /Service failure – This outage affects applications and services on the network.

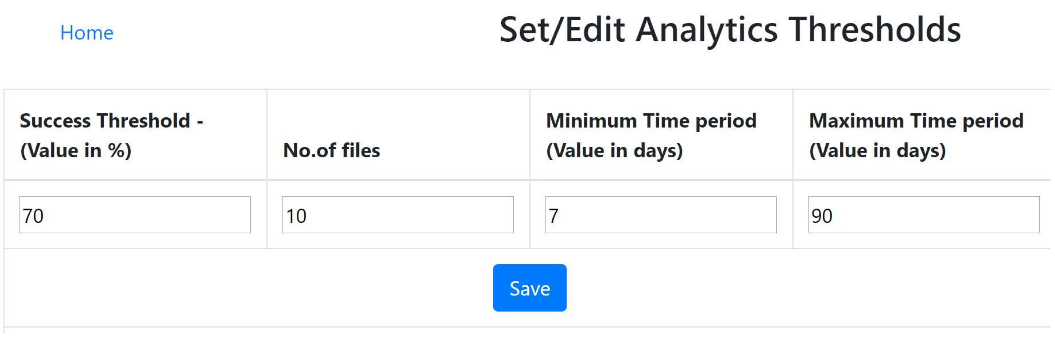
b) System / Hardware failure – This outage affects hardware components like CPU, drives, memory.

c) Site failure – These are generally caused by natural disasters.

The current problem is, as there are multiple sites within CERN, the failures may not be discovered for a long period of time. In this thesis, we are identifying the partner sites, which are having issues with the Tier 2 sites in US SW region, being source and destination. This thesis analyzes the success percentage in transfers and ranks the sites having highest difference in success rates between the maximum and minimum thresholds.

6.2 Define threshold for failure analysis

The thesis allows to configure the analytics thresholds within the web framework. See Figure 6-1 for analytics thresholds configuration for failure analysis.



Set/Edit Analytics Thresholds

Success Threshold - (Value in %)	No.of files	Minimum Time period (Value in days)	Maximum Time period (Value in days)
70	10	7	90

Figure 6-1 Analytics thresholds configuration for failure analysis

The following parameters are considered when analyzing failures during file transfers between sites.

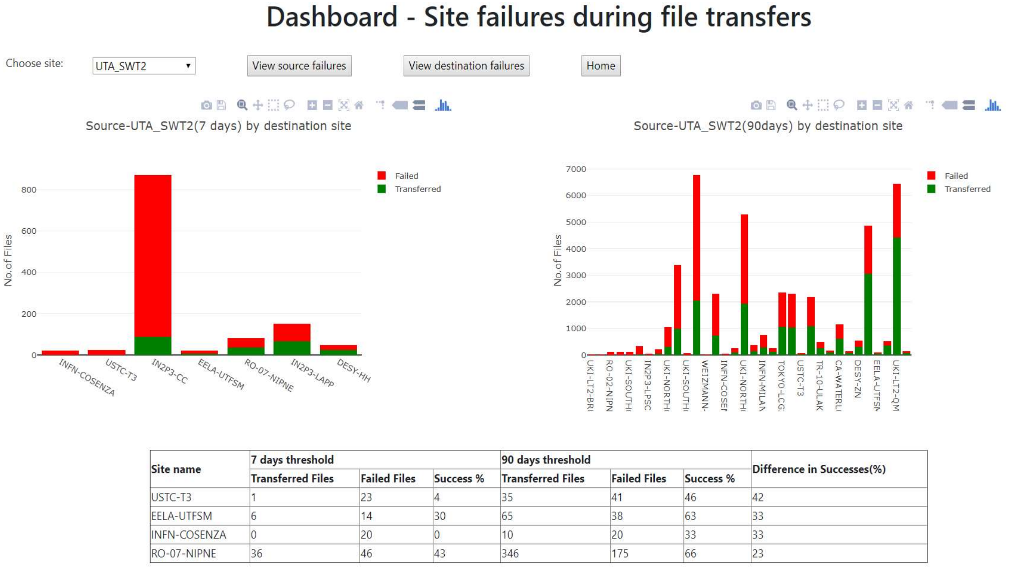
a) Success Threshold defines to analyze the sites having successful transfer rates less than 70%.

b) No.of files defines the total number of failed transfer files for any site. If the number of files failed is less than 10, we can ignore as it is not a huge failure to analyze.

c) The minimum and maximum time period are the days through which we compare the failures. We analyze the failure for a weeks' time, then we analyze the failures for 3 months' time and compare the failures between the sites that occurs in both the time periods and rank them based on their difference in success rate.

6.3 Failure analysis implementation

The thesis analyzes the failures in Tier 2 default sites being the source and destination with other sites. See Figure 6-2 for failure analysis dashboard.



Dashboard - Site failures during file transfers

Choose site: UTA_SWT2

Site name	7 days threshold			90 days threshold			Difference in Successes(%)
	Transferred Files	Failed Files	Success %	Transferred Files	Failed Files	Success %	
USTC-T3	1	23	4	35	41	46	42
EELA-UTFSM	6	14	30	65	38	63	33
INFN-COSENZA	0	20	0	10	20	33	33
RO-07-NIPNE	36	46	43	346	175	66	23

Figure 6-2 Failure analysis dashboard

See Figure 6-3 for failures by destination sites with SWT2_CPB as source for the past 7 days' time.

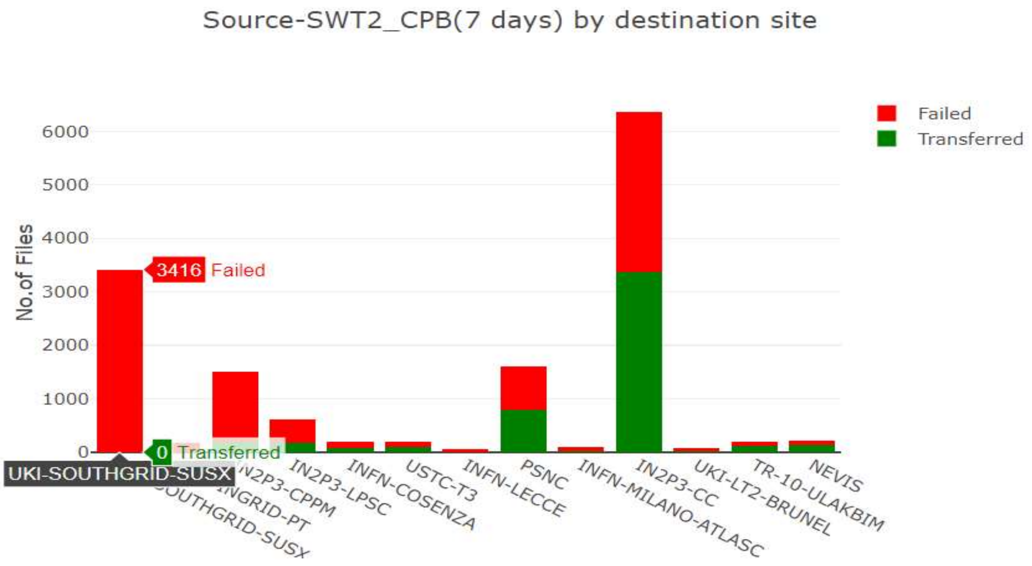


Figure 6-3 Failures by destination sites with SWT2_CPB as source for the past 7 days

From the above graph, we could see that the transfers within SWT2_CPB and UKI_SOUTHGRID_SUSX had failed and their success rate is 0%, as no files have transferred between them. The destination sites are arranged with increasing success rates. See Figure 6-4 for failures by destination sites with SWT2_CPB as source for the past 90 days' time.

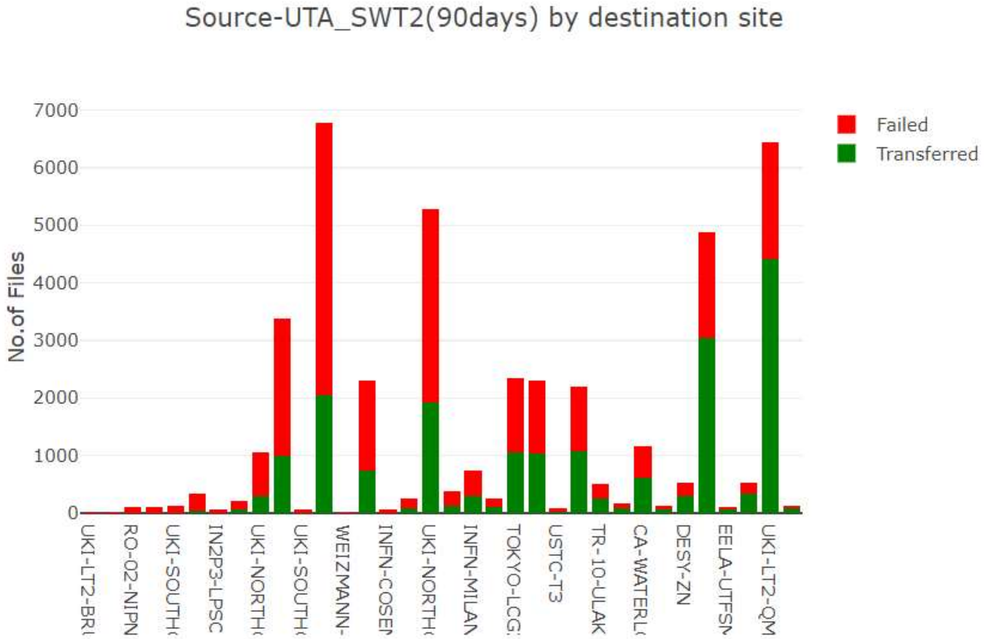
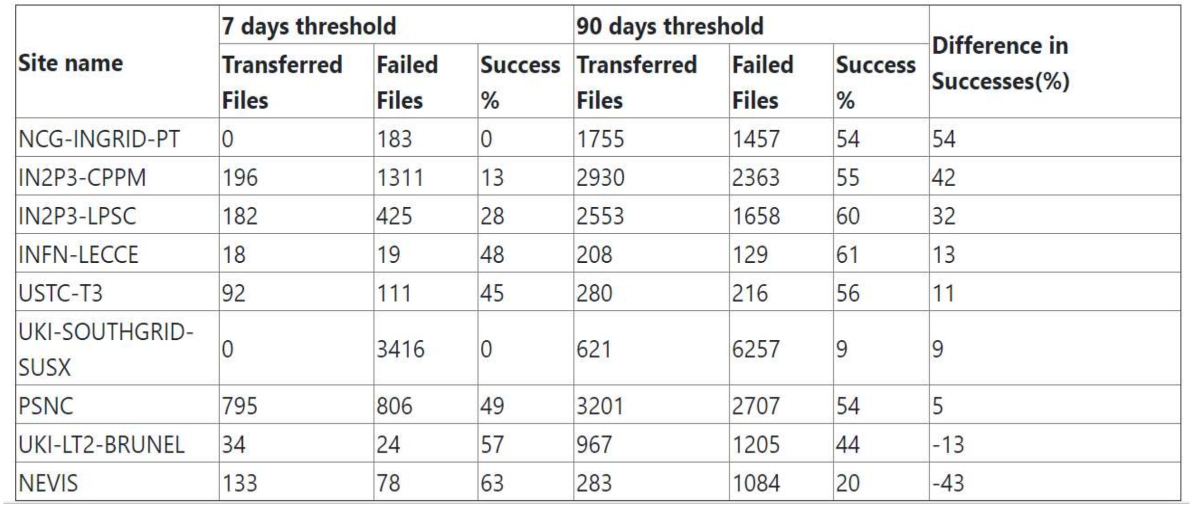


Figure 6-4 Failure by destination with SWT2_CPB as source for the past 90 days

Now we can compare the failure with SWT2_CPB as source to the destinations in the 7 days and 90 days' time and identify the sites which are having a problem as destination during transfer. This thesis then ranks the sites based on their success rate difference in the two time periods. See Figure 6-5 for the sites that failed as destination with SWT2_CPB as source.



Site name	7 days threshold			90 days threshold			Difference in Successes(%)
	Transferred Files	Failed Files	Success %	Transferred Files	Failed Files	Success %	
NCG-INGRID-PT	0	183	0	1755	1457	54	54
IN2P3-CPPM	196	1311	13	2930	2363	55	42
IN2P3-LPSC	182	425	28	2553	1658	60	32
INFN-LECCE	18	19	48	208	129	61	13
USTC-T3	92	111	45	280	216	56	11
UKI-SOUTHGRID-SUSX	0	3416	0	621	6257	9	9
PSNC	795	806	49	3201	2707	54	5
UKI-LT2-BRUNEL	34	24	57	967	1205	44	-13
NEVIS	133	78	63	283	1084	20	-43

Figure 6-5 Failed destination sites with SWT2_CPB as source and success rate

Similarly, we can identify the sites that are failing as source with Tier 2 default sites being the destination. See Figure 6-6 for failures by source sites with UTA_SWT2 as destination for the past 7 days' time.

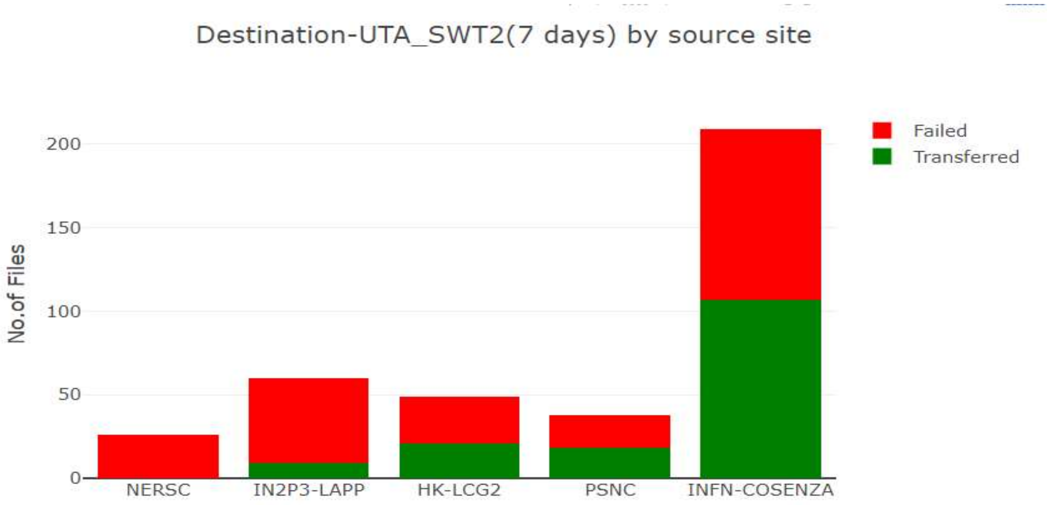


Figure 6-6 Failure by source with UTA_SWT2 as destination for the past 7 days

See Figure 6-7 for failures by source sites with UTA_SWT2 as destination for the past 90 days' time.

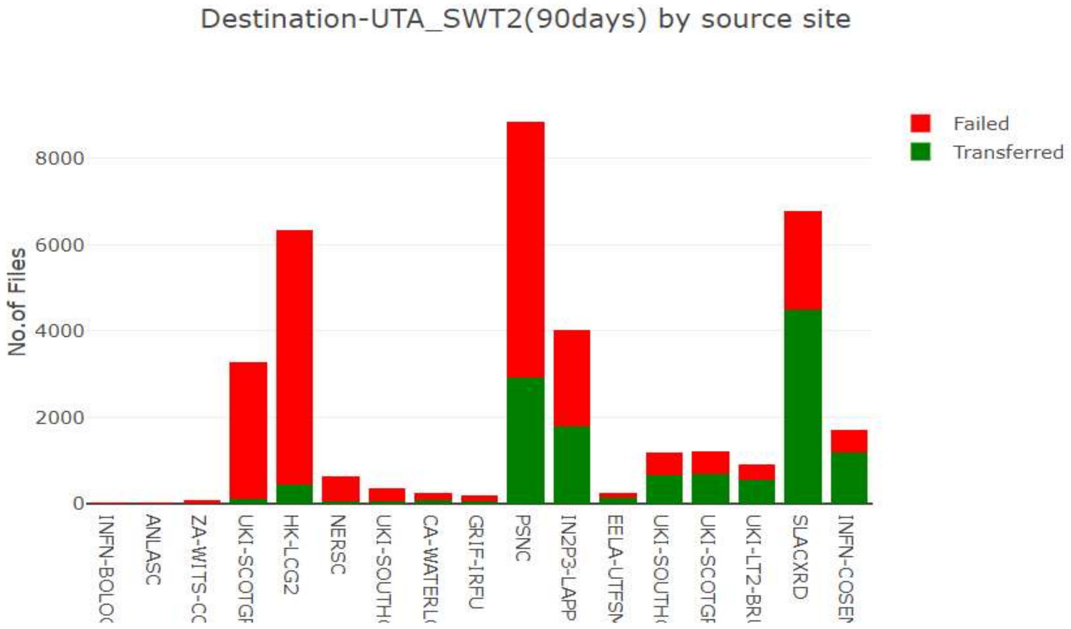
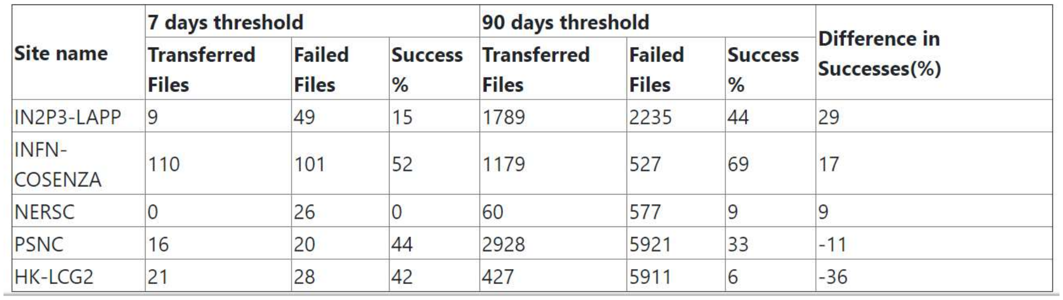


Figure 6-7 Failure by source with UTA_SWT2 as destination for the past 90 days

We can compare these failures with UTA_SWT2 as source and identify the sites failing during transfers. See Figure 6-8 for the sites that failed as source with UTA_SWT2 as destination.



Site name	7 days threshold			90 days threshold			Difference in Successes(%)
	Transferred Files	Failed Files	Success %	Transferred Files	Failed Files	Success %	
IN2P3-LAPP	9	49	15	1789	2235	44	29
INFN-COSENZA	110	101	52	1179	527	69	17
NERSC	0	26	0	60	577	9	9
PSNC	16	20	44	2928	5921	33	-11
HK-LCG2	21	28	42	427	5911	6	-36

Figure 6-8 Failed source sites with UTA_SWT2 as destination and success rate

For example, when analyzing the failure results, it is seen that the transfers to and from the site PSNC, is failing and obvious that the site has some problems. These failure sites may take a longer time to be detected by CERN and failure analysis in this thesis helps to discover the failed sites easily.

Chapter 7

Summary and Conclusion

This thesis started with a goal to monitor the health of clusters in an effective way, analyze the failures and use machine learning algorithms to predict the site failures in future. This research provides a web-based framework for monitoring the health of clusters in Tier 2 facility of US South west region. This work collects real time data every hour through cron and alerts are sent to the administrators when the success rates fall below the defined threshold. This thesis focuses on the deletion and transfer metrics, which are important in cluster health monitoring. The deletion metrics shows the number of files planned for deletion, number of files deleted successfully, and number of files failed in deletion process. The transfer metrics shows the number of files transferred successfully and the number of files failed to transfer both as source and destination.

The cluster health is monitored for various time intervals of the last 24 hours, past one week and past two weeks. This research monitors the day wise health for any day of the week. Site failures are also analyzed in this work for 7days period and 90 days period and the failures are compared and ranked based on their success rates. All the thresholds defined in this thesis are configurable, which enables to monitor the cluster health for any period.

In addition to monitoring and failure analysis, this thesis also predicts the sites which may fail due to diminishing deletions or transfers by machine learning algorithm. Cluster health monitoring and failure analysis is an important part of any data center and this solution is designed keeping in mind the problems in the current monitoring solution. This thesis provides a solution to monitor cluster health in an effective way , tracks the hourly deletion/ transfer metrics, analyzes failed sites and predicts site failures in future.

Chapter 8

Future Work

The Tier 2 data centers must have 99.74% uptime [41] and experiences only 22 hours of down time per year. The experiments at CERN involves enormous data flowing and the data centers must process petabytes of data every day. For such complex computation, the reliability of data centers is very important, and the sites must be performing the operations without failing. However, failure in data centers is unavoidable due to manual errors is network management, installations and maintenance works. The solution to know about the failures and the current cluster health is given in this thesis.

The failures in deletion/transfer between the sites gives us several logs about it. Many logs may be irrelevant for the failure. The administrator must go through the failure logs, understand the reason behind the failure and then provides solution. In future, if we can come up with any solution or algorithm to analyze the logs from the failed sites and provide the administrator with the exact reason for failure, a lot of time can be saved. A small increase in the uptime percentage in such enormous data flow can increase the reliability which could save the computing resources.

References

[1] Data Never Sleeps [Online]

https://www.iflscience.com/technology/how-much-data-does-the-world-generate-every-minute/

[2] What does High-Performance Computing (HPC) mean? [Online]

https://www.techopedia.com/definition/4595/high-performance-computing-hpc

[3] Dr. Vincent Garonne, Oslo University, “How does Atlas experiment manage petabytes of data?”, in Computing Techniques Seminar [Online]

https://indico.in2p3.fr/event/14429/attachments/15220/18709/How_Does_The_ATLAS_Experiment_Manage_Petabytes_of_Data_-.pdf [Accessed April 2018]

[4] 4 V’s of big data [Online]

https://www.ibmbigdatahub.com/infographic/four-vs-big-data

[5] CERN [Online] https://home.cern/about

[6] LHC [Online] https://home.cern/topics/large-hadron-collider

[7] WLCG [Online] https://home.cern/about/computing/worldwide-lhc-computing-grid

[8] ATLAS [Online] http://www.atlas-swt2.org/

[9] CERN Data Center [Online]

https://home.cern/about/updates/2017/07/cern-data-centre-passes-200-petabyte-milestone

[10] WLCG Mission [Online] http://wlcg.web.cern.ch/

[10] WLCG Tier Architecture [Online] http://wlcg-public.web.cern.ch/tier-centres

[11] SRM [Online] https://en.wikipedia.org/wiki/Storage_Resource_Manager

[12] Dr. Flavia Donno, “Storage Resource Manager Version 2.2: design, implementation, and testing experience” [Online] https://core.ac.uk/download/pdf/44187778.pdf [Accessed August 2018]

[13] Grid FTP [Online] http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/

[14] Fault tolerant implementation [Online] https://en.wikipedia.org/wiki/GridFTP

[15] Ganglia [Online] http://ganglia.sourceforge.net/

[16] RRD Tool [Online] https://en.wikipedia.org/wiki/RRDtool

[17] Time-series data in RRD tool [Online] https://www.caida.org/tools/utilities/rrdtool/

[18] Dr. Oleg Ivasenko, "Setup of a Ganglia Monitoring System for a Grid Computing Cluster" [Online] https://www.institut3b.physik.rwthaachen.de/global/show_document.asp?id=aaaaaaaaapwtye [Accessed September 2018]

[19] Rucio [Online] https://rucio.cern.ch/

[20] MLasnig, "Monitoring and controlling ATLAS data management: The Rucio web user interface", in Journal of Physics Conference, 2015 [Online] http://iopscience.iop.org/article/10.1088/1742-6596/664/6/062028/pdf [Accessed May 2018]

[21] ETL [Online] https://en.wikipedia.org/wiki/Extract,_transform,_load

[22] JSON [Online] http://www.json.org/

[23] CRON [Online] https://en.wikipedia.org/wiki/Cron

[24] Cluster Health Monitoring [Online] http://www.brightcomputing.com/blog/cluster-monitoring-vs.-health-checking-whats-the-difference

[25] Matplotlib [Online] https://matplotlib.org/

[26] Advantages of Matplotlib [Online]

https://jakevdp.github.io/PythonDataScienceHandbook/04.00-introduction-to-matplotlib.html

[27] Plotly [Online] https://github.com/plotly/plotly.py

[28] Visualization toolbox [Online]

https://plot.ly/python/user-guide/#what-is-plotly

[29] A picture is worth a thousand words [Online]

https://data-visualization.cioreview.com/cxoinsight/what-is-data-visualization-and-why-is-it-important-nid-11806-cid-163.html

[30] Power of Data Visualization [Online]

https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception

[31] Data Visualization [Online]

https://searchbusinessanalytics.techtarget.com/definition/data-visualization

[32] Human brain [Online]

https://dzone.com/articles/6-ways-data-visualization-can-change-your-company

[33] Cluster [Online]

http://www.developerscoding.com/258/what-is-the-differences-between-a-node,-a-cluster-and-a-datacenter-in-a-cassandra--database?

[34] Cluster health monitoring system [Online]

https://docs.oracle.com/database/121/CWADD/GUID-078AB1E1-22EE-4CAE-8158-48847B946DA7.htm#CWADD92242

[35] G.Bauer, “Health and Performance Monitoring of The Online Computer Cluster Of CMS”, in Journal of Physics Conference Series Volume 396, June 2012 [Online]

https://cds.cern.ch/record/1462972/files/CR2012_158.pdf

[36] Interactive visualization [Online]

https://www.gartner.com/it-glossary/interactive-visualization

[37] High Computing systems [Online] https://home.cern/about/experiments/atlas

[38] IOPS [Online]

http://superuser.openstack.org/articles/at-cern-storage-is-the-key-to-the-universe/

[39] Dr. Matei Zaharia, UC,Berkeley, “Discretized Streams: An Efficient and Fault-Tolerant Model for Stream Processing on Large Clusters”, in Technical Report No. UCB/EECS-2012-259, December 14, 2012[Online]

https://www.usenix.org/system/files/conference/hotcloud12/hotcloud12-final28.pdf

[Accessed September 2018]

[40] Types of outages [Online]

https://www.volico.com/understanding-clustering-servers-capabilities/

[41] Tier 2 Uptime [Online]

https://www.colocationamerica.com/data-center/tier-standards-overview.htm

Biographical Information

Meenakshi Balasubramanian joined the University of Texas at Arlington in Spring 2017. She has six years of experience as a Software Engineer in India. In the United States, she worked as a Graduate Research Assistant in the University of Texas at Arlington, working on Android mobile applications.

She also interned with Avaya, as DevOps Intern, working on Ansible and Security Compliances.

Her research interests are Software Development and Cloud computing.