

Construction of real-time monitoring system for Grid services based on log analysis at the Tokyo Tier-2 center

T. Kishimoto^{*1}, T. Mashimo¹, N. Matsui¹, T. Nakamura², H. Sakamoto¹

¹ *International Center for Elementary Particle Physics, The University of Tokyo*

² *Computing Research Center, KEK*

E-mail: tomoe@icepp.s.u-tokyo.ac.jp

The Tokyo Tier-2 center, which is located in the International Center for Elementary Particle Physics at the University of Tokyo, is providing computer resources for the ATLAS experiment in the Worldwide LHC Computing Grid. Logs produced by the Grid services provide useful information to determine whether the services are working properly. Therefore, a new real-time monitoring system based on log analysis has been constructed using the ELK stack. This paper reports the configuration of the new monitoring system at the Tokyo Tier-2 center, and discusses improvements in terms of stability and flexibility of the site operation by introducing the new monitoring system.

*International Symposium on Grids and Clouds (ISGC) 2018 in conjunction with Frontiers in Computational Drug Discovery
16-23 March 2018
Academia Sinica, Taipei, Taiwan*

*Speaker.

1. Introduction

The Tokyo Tier-2 center, which is located in the International Center for Elementary Particle Physics (ICEPP) [1] at the University of Tokyo, is providing computer resources for the ATLAS experiment [2] in the Worldwide LHC Computing Grid (WLCG) [3]. The official site operation in the WLCG was launched in 2007 after several years of development. The site has been achieving a stable and reliable operation since then.

The Grid middleware is a software stack, which provides services for users to access the distributed computing resources of the Grid. A stable operation of the Grid services is a key role of the sites. Logs produced by the Grid services provide useful information to determine whether the services are working properly. Therefore, a new real-time monitoring system based on log analysis has been constructed. This paper reports the configuration of the new monitoring system, and discusses improvements in terms of stability and flexibility of the site operation by introducing the new monitoring system. The paper is organized as follows, Section 2 provides the configuration and status of the Tokyo Tier-2 center. Section 3 describes the construction of the new monitoring system. Section 4 gives a summary.

2. Tokyo Tier2 center

In the Tokyo Tier-2 center, almost all hardware devices are leased, and are replaced in every three years in order to satisfy the requirement of the ATLAS experiment. The latest hardware upgrade was done in January 2016, and the new system (so-called 4th system) is stably running. In the 4th system, 6144 CPU cores (256 worker nodes) and 7392 TB disk storages are reserved for the ATLAS experiment. Table 1 summarizes the WLCG pledge and actual deployed resources. It can be confirmed that the WLCG pledges are well satisfied with the 4th system. The new system will be provided for year 2019.

		CPU [HS06]	Disk [TB]	LOCALGROUPDISK [TB]
2017	Pledge	34,000	4,000	-
	Deployed	111,268 (6144 CPU cores)	4,000	1,000
2018	Pledge	40,000	4,800	-
	Deployed	111,268 (6144 CPU cores)	4,800	1,000

Table 1: The WLCG pledge and deployed resources at the Tokyo Tier2 center. All deployed resources are reserved for the ATLAS experiment. LOCALGROUPDISK is dedicated disks for the ATLAS-Japan group.

2.1 Status in ATLAS experiment

Figure 1 shows the slot occupancy at the Tokyo Tier-2 center (top) and the proportion of number of completed jobs of Tier-2 sites in the ATLAS experiment (bottom). The figures show the results for year 2017. It can be seen that the slots were well filled by various types of jobs. The Tokyo Tier-2 site completed 5.5% of the jobs in Tier-2s. Also a high site availability, which is greater than 99%, has been achieved using the 4th system.

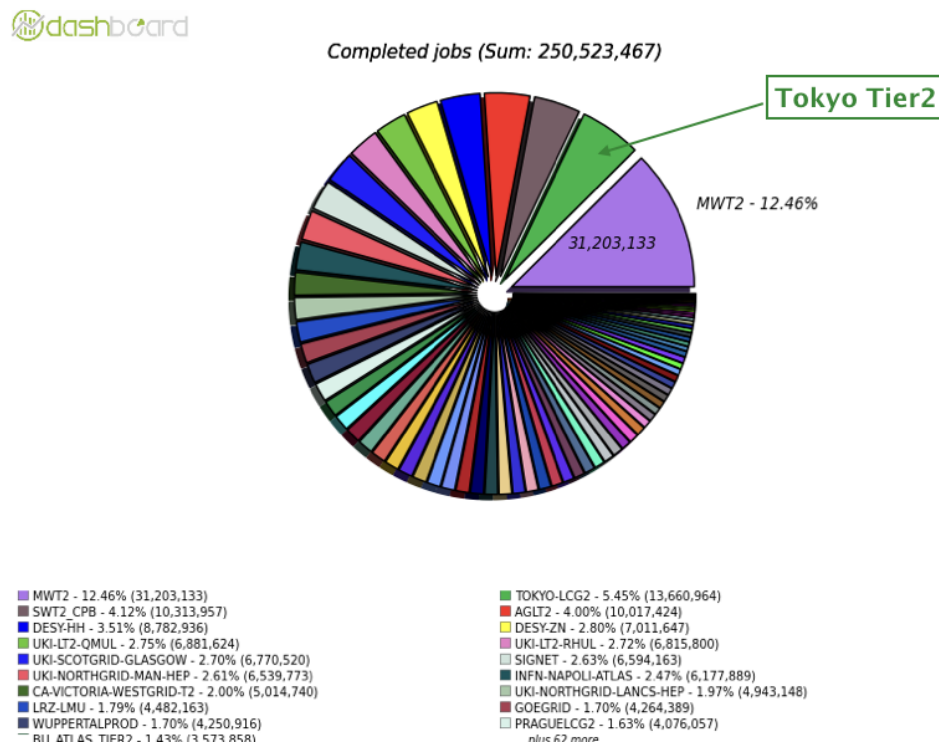
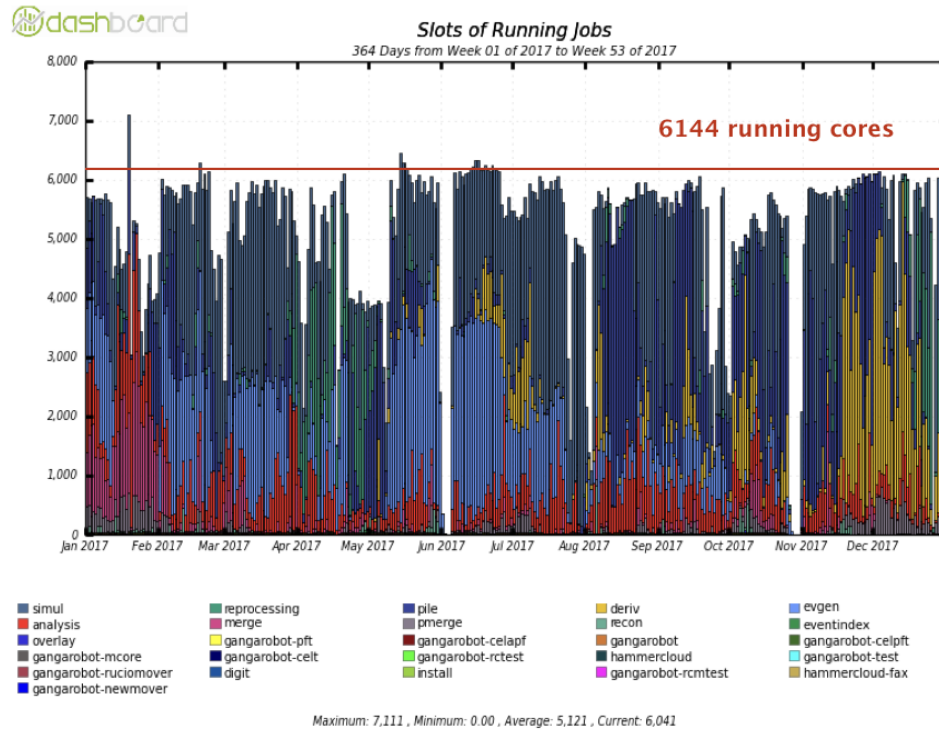


Figure 1: The slot occupancy at the Tokyo Tier-2 center (top) and the proportion of number of completed jobs of Tier-2 site in the ATLAS experiment (bottom).

2.2 Grid middlewares

The ARC-CE [4] is deployed as the computing element in front of the HTCondor batch job scheduler [5]. The configuration of the computing element is shown in Figure 2. Two ARC-CEs are installed for redundancy. The high availability configuration of the central manager of the HTCondor is also implemented.

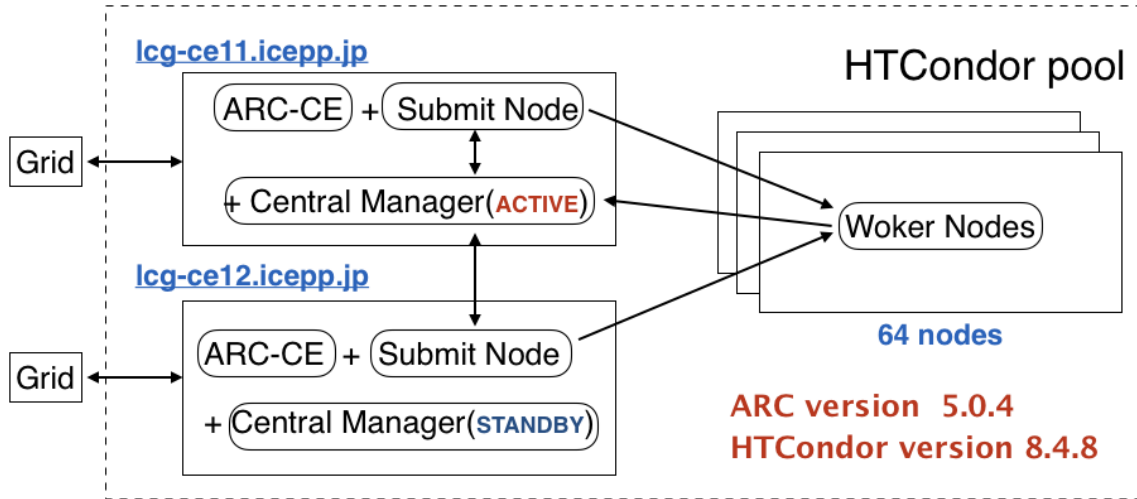


Figure 2: The configuration of the computing element at the Tokyo Tier-2 center.

The disk storage consists of 48 sets of a disk array and a file server, which are managed by the Disk Pool Manager (DPM) [6]. The configuration of the storage element is shown in Figure 3. The DPM supports several protocols for data transfers, such as GridFTP [7], XRootD [8] and https. The MySQL [9] is used as the database of the DPM, and the auto replication of the database is available in master-slave configuration.

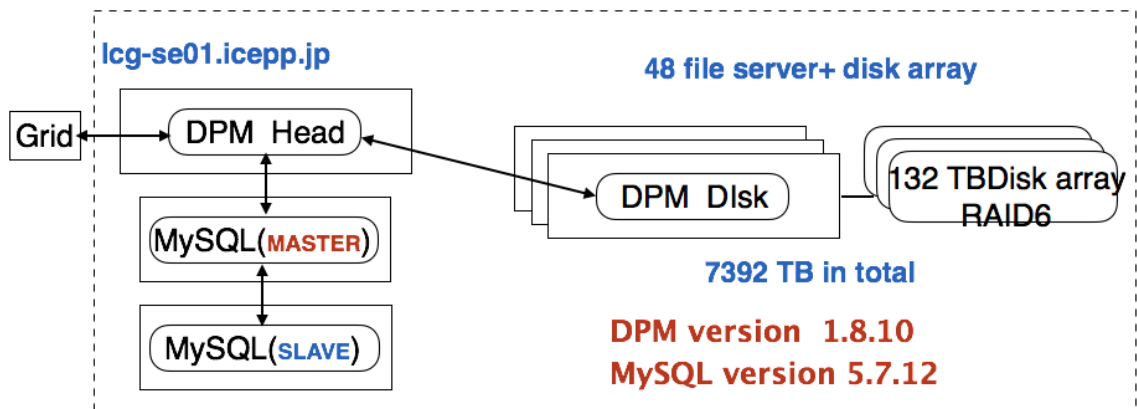


Figure 3: The configuration of the storage element at the Tokyo Tier-2 center.

3. New monitoring system for Grid services

Logs produced by the Grid services provide useful information to check health of the services.

Figure 4 shows an example of text logs produced by the ARC-CE. Time stamp, severity of message, job global ID and status of job transition are shown in this example. By checking the log, we can judge that the job was processed at the ARC-CE without errors.

Time stamp	Severity	Job global ID	State transition
[2018-02-05 04:05:23]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: ACCEPTED: parsing job description
[2018-02-05 04:05:23]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: ACCEPTED: moving to PREPARING
[2018-02-05 04:05:23]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: PREPARING from ACCEPTED
[2018-02-05 04:05:31]	[Arc.DataStaging.DTR]	[INFO] [6734/7]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: All downloads finished successfully
[2018-02-05 04:05:31]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: SUBMIT from PREPARING
[2018-02-05 04:05:31]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: state SUBMIT: starting child: /usr/share/arc/submit-condor-job
[2018-02-05 04:05:32]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: state SUBMIT: child exited with code 0
[2018-02-05 04:05:32]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: INLRMS from SUBMIT
[2018-02-05 04:22:01]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: Job finished
[2018-02-05 04:22:01]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: FINISHING from INLRMS
[2018-02-05 04:22:03]	[Arc] [INFO]	[6734/2]	roaNdm7Cn2rnG457in1NHNUqABFKDmABFKDmVi1XdmABFKDmpGQQXn: State: FINISHED from FINISHING

Figure 4: An example of text logs produced by the ARC-CE.

Therefore, a new real-time monitoring system based on log analysis has been constructed using the ELK stack [10] in order to detect a problem of the Grid services. The ELK stack consists of Elasticsearch, Logstash, Kibana and Beat, and provides an efficient way of log processing, storing, query and visualization. The function of each component of the ELK stack is summarized in Table 2. Logs produced by the ARC-CE, HTCondor and DPM are processed and monitored in the new monitoring system. A dedicated Logstash configuration for each service has been developed to parse the Grid service logs. Generally, 1 line in the log file is processed and stored as 1 record in the Elasticsearch database.

Beat 5.6	data shipper
Logstash 5.6	log processing, transformation
Elasticsearch 5.6	data storing, search
Kibana 5.6	visualization

Table 2: The function of each component of the ELK stack.

3.1 Resource usage of ELK host

The Elasticsearch, Logstash and Kibana are installed to a physical machine, which has 24 CPU cores and 32GB memory (Dell Power Edge M630). The total capacity of local disks is 1 TB. The logs are collected from the service hosts to the ELK host using *rsync* command. Table 3 shows the log frequency (number of log lines) of the services. The DPM is the dominant source of logs to be processed and stored. Figure 5 shows the resource usage of the ELK host. It can be confirmed that an enough computing resource is available in the ELK host since the observed average CPU usage is still $\sim 200\%$ (full usage of 24 CPU cores = 2400%). 12 GB memory is assigned to the JVM heap of the Elasticsearch and Logstash. The size of the Elasticsearch database increases by ~ 20 GB per day. Therefore, old records, which passed 1 month, are deleted from the database.

3.2 Examples of the new monitoring

Figure 6 shows selected figures in the new monitoring system for the ARC-CE. The left figure shows the number of job states, such as SUBMIT, ACCEPTED, PREPARING, at the computing

ARC-CE	HTCondor	DPM
~300/min	~5,000/min	~50,000/min

Table 3: The log frequency (number of log lines) of the ARC-CE, HTCondor, and DPM.

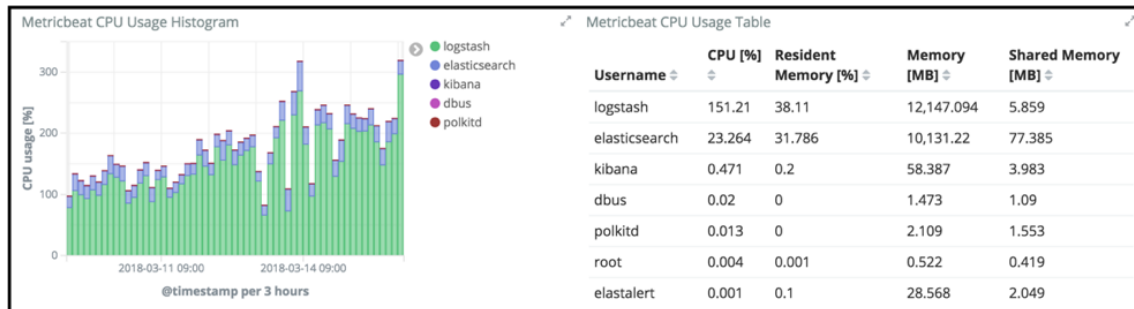


Figure 5: The resource usage of the ELK host, measured by the Metricbeat [11].

element. The stacked job state due to some problems in the computing element can be detected from this figure. The right figure shows the number of logs for each message severity. It can be seen that the error and warning counts are small in the figure.

Figure 7 also shows selected figures for the HTCondor. The left figure shows the number of completed jobs at the HTCondor. There is a good correlation between the number of FINISHED state at the ARC-CE and the number of completed jobs at the HTCondor. The right figure also shows the number of completed jobs in terms of the time it takes to complete. There is an upper limit of the job time in the HTCondor in order to kill runaway jobs. It can be monitored from this figure that the jobs are completed within the time limit.

Figure 8 shows the number of errors (left) and warning (right) for each protocol for the DPM. A bump of the warning counts can be observed in the right figure. There was a hardware problem in this period, and file deletions were failing using the httpd protocol. This experience shows a good example that the new monitoring system works well to detect a problem of the services. An alerting function is implemented using elastalert [12], which is a third party software. E-mail is sent to administrators when the error or warning counts exceed a threshold.



Figure 6: Examples of the ARC-CE monitoring.

3.3 Development of iftopbeat

The DPM supports several protocols for the data transfer. In order to measure and monitor the

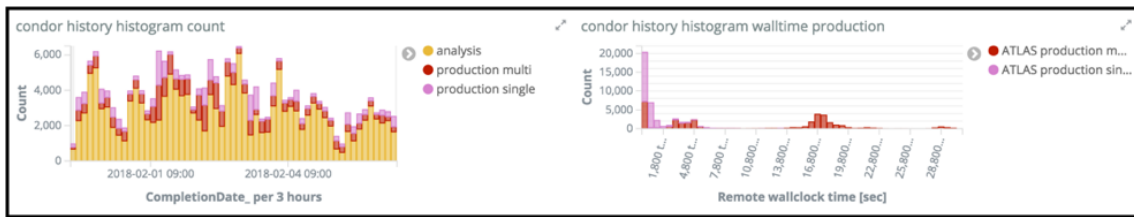


Figure 7: Examples of the HTCondor monitoring.

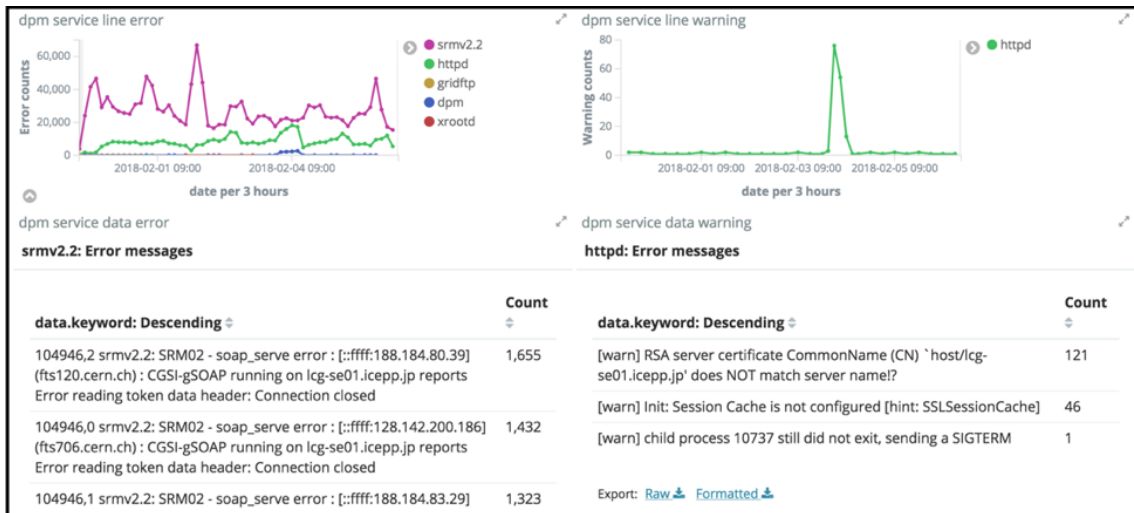


Figure 8: Examples of the DPM monitoring.

bandwidth usage of each protocol, a software (so-called iftopbeat) has been developed. Iftopbeat collects the information using *iftop* command, which provides the bandwidth usage, the direction and source of transfers and its ports. Then iftopbeat processes the information and directly sends the data to the Elasticsearch using the Beat libraries. Figure 9 shows the bandwidth usage of the file servers in the production system measured by iftopbeat. The protocol and transfer direction/source are well identified and visualized.

4. Summary

The Tokyo Tier-2 center is providing sufficient computing resources to the ATLAS experiment in the WLCG, and is keeping the high site availability. In order to operate the site more stably and flexibly, the new real-time monitoring system based log analysis for the Grid services has been constructed using the ELK stack. It has been confirmed that the ELK stack can handle the logs produced by the Grid services at the Tokyo Tier-2 center, and can visualize the logs well. Since the logs of the ARC-CE, HTCondor and DPM are integrated into the new monitoring system, it is easy to check a correlation between the services when a problem arises. Iftopbeat has been developed in order to monitor the bandwidth usage of each protocol, and is working well in the production system. The monitoring of the bandwidth usage for each protocol and direction/source will help to separate a problem into small parts.

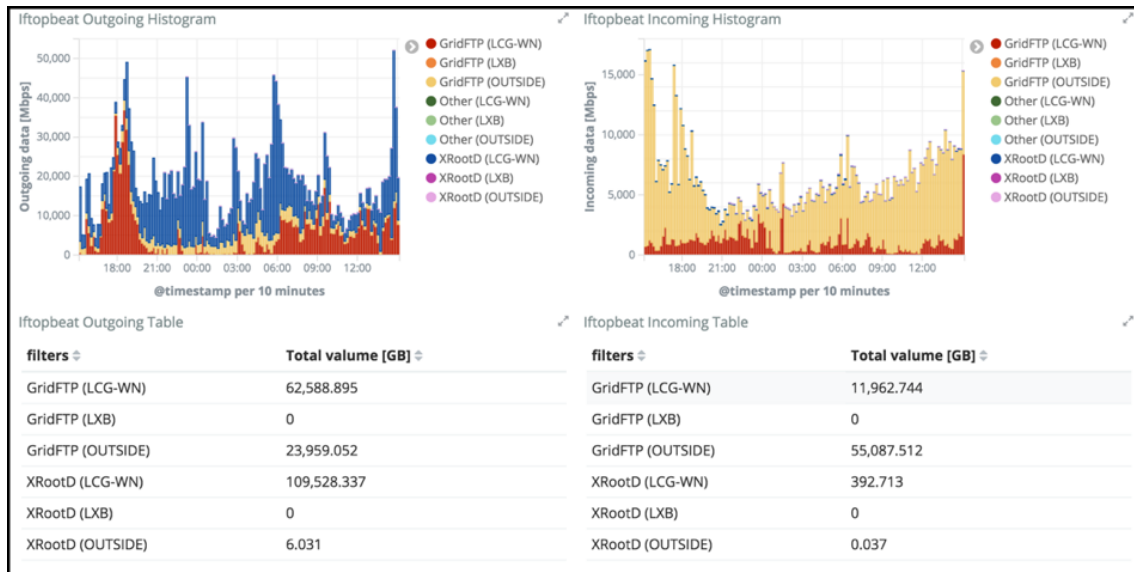


Figure 9: The bandwidth usage of the file servers in the production system, measured by iftopbeat. The left (right) figure shows outgoing (incoming) data.

References

- [1] ICEPP web page, <https://www.icepp.s.u-tokyo.ac.jp/en/index.html>
- [2] ATLAS experiment web page, <http://atlas.cern/>
- [3] WLCG web page, <http://wlcg.web.cern.ch/>
- [4] ARC-CE web page, <http://www.nordugrid.org/arc/ce/>
- [5] HTCondor web page, <https://research.cs.wisc.edu/htcondor/>
- [6] DPM web page, <http://lcgdm.web.cern.ch/dpm>
- [7] GridFTP web page, <http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/>
- [8] XRootD web page, <http://xrootd.org/>
- [9] MySQL web page, <https://www.mysql.com/jp/>
- [10] ELK stack web page, <https://www.elastic.co/jp/elk-stack>
- [11] Metricbeat web page, <https://www.elastic.co/jp/products/beats/metricbeat>
- [12] Eelastalert github, <https://github.com/Yelp/elastalert>