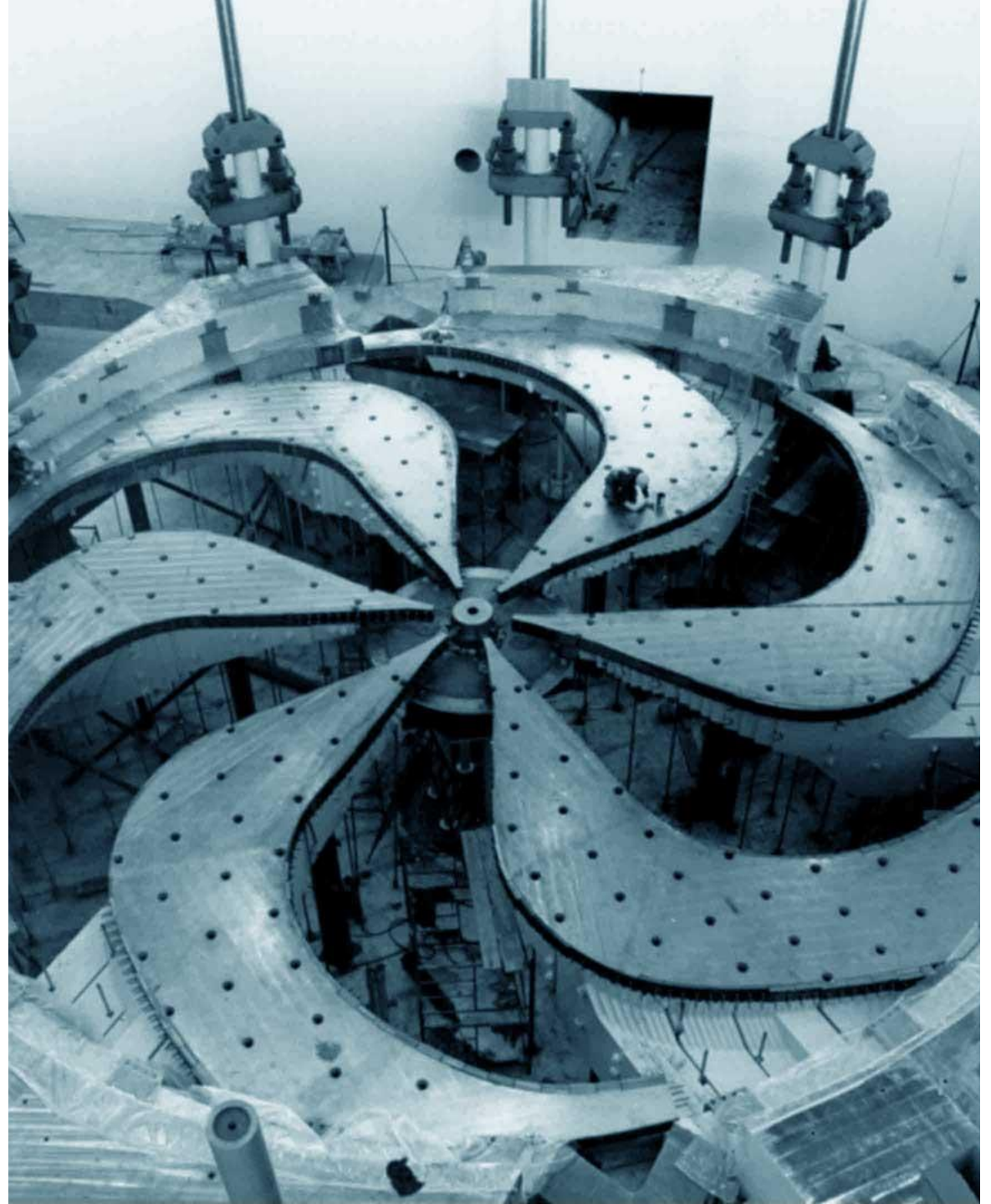


# Infrastructure and Analytics Summary (Tier-1)

Fernando Fernandez Galindo  
TRIUMF Scientific Computing Department

US ATLAS Computing Facilities Face-to-Face at SLAC  
December 1st, 2022



# INFRASTRUCTURE SUMMARY

- Compute:

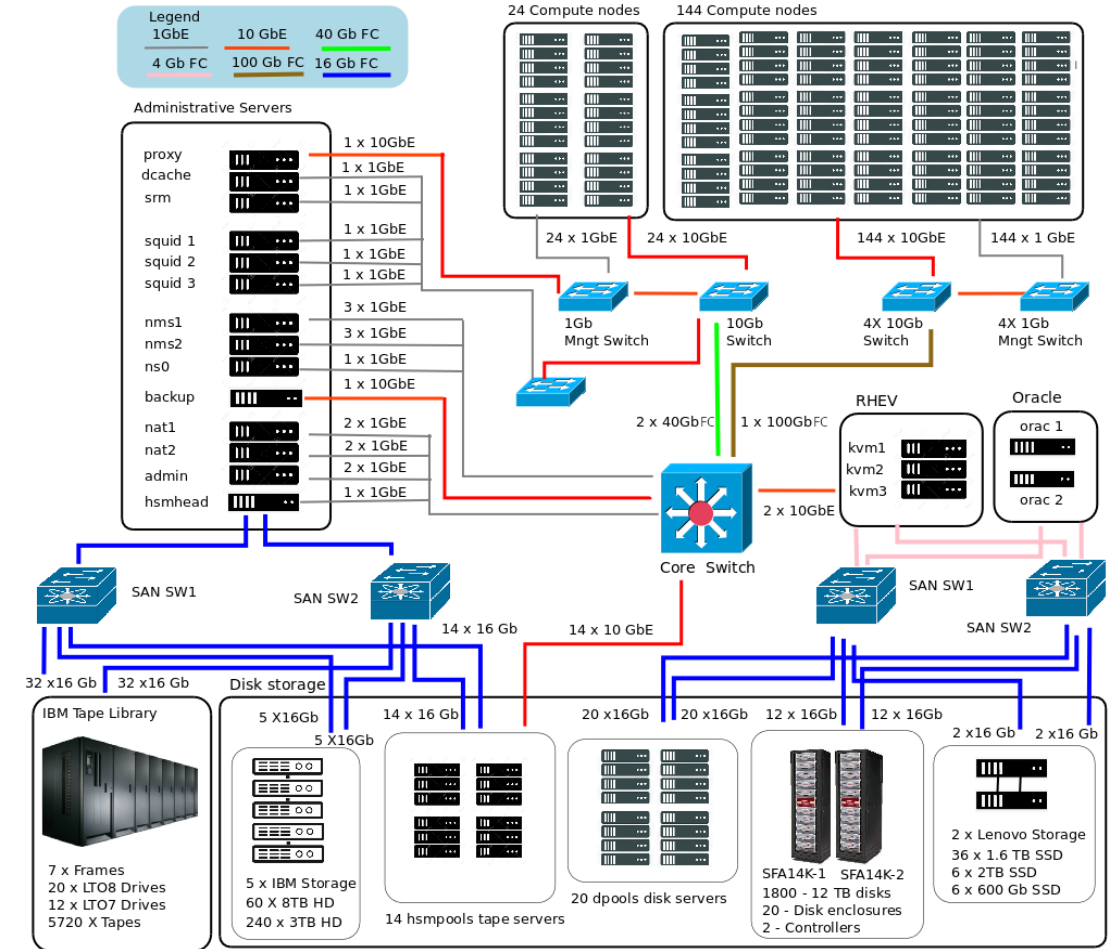
- ARC CE 6.15.1
- HTCondor 9.0.11
- 1,600 cores at TRIUMF.
- 7,820 at Simon Fraser University.

- Storage:

- dCache (6.2.39 -> 7.2.26 soon)
  - 30 dpool serving 17PB of usable disk.
  - 13 hsm pool nodes 36PB of usable tape.

- OS:

- Mostly running on SL 7.9
- RedHat has given us 1,000 licenses of RHEL 9 to which we will be upgrading to in the near future.



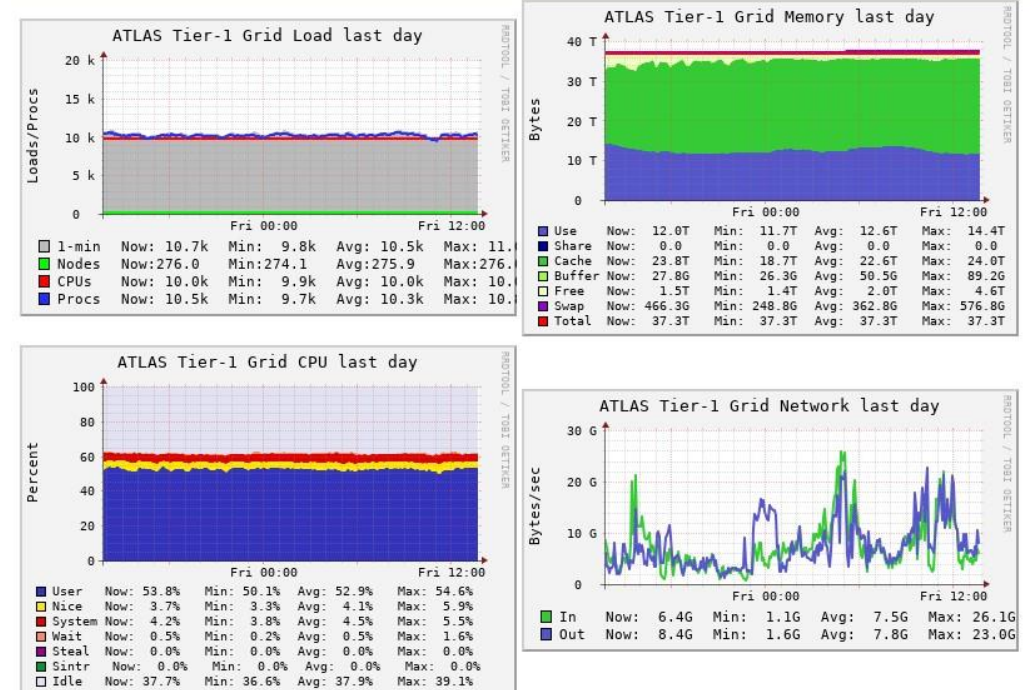
# ANALYTICS PROJECT MOTIVATION

- With growing infrastructure an ever-increasing collection of heterogeneous monitoring data is produced.
- Our existing monitoring and alerting implementation is robust and stable but rather static and isolated.
- Early 2020 we started the analytics project on hardware that was deprecated.
- The project's main objectives are:
  - Create a framework where the different datasets can be brought in together for analysis and monitoring.
  - Experiment with tools and techniques like ML, to assist in finding correlations between different areas of our infrastructure, detect anomalies, inefficiencies and maybe even predict issues.

## Nagios

Host Group	Host Status Summary	Service Status Summary
Admin Nodes (admin_nodes)	11 UP	63 OK 1 CRITICAL : 1 Disabled
Analytics Nodes (analytics_nodes)	2 UP	10 OK
Compute 1 Nodes (compute01_nodes)	24 UP	120 OK
Compute 2 Nodes (compute02_nodes)	24 UP	120 OK
Compute 3 Nodes (compute03_nodes)	24 UP	120 OK
Compute 4 Nodes (compute04_nodes)	24 UP	120 OK
Compute 5 Nodes (compute05_nodes)	24 UP	120 OK
Compute 6 Nodes (compute06_nodes)	24 UP	120 OK
Compute 7 Nodes (compute07_nodes)	24 UP	120 OK
Compute 8 Nodes (compute08_nodes)	3 UP	15 OK

## Ganglia



# NEW HARDWARE (FALL 2022)

## Frontend (Elasticsearch and clients)

- 1x PowerEdge R650
- **CPU:**  
2x Xeon 6336Y – 24 cores with Scikit-learn extensions.
- **GPU:**  
Nvidia Telsa T4
- **Memory:**  
256GB
- **Network:**  
Nvidia Mellanox ConnectX-5
- **Storage:**  
2x 480GB SSD (OS)  
2x 3.84TB NVMe

## Backend (Elasticsearch and Logstash)

- 2x PowerEdge R650
- **CPU:**  
Xeon 6326 – 16 Cores
- **Memory:**  
256GB
- **Network:**  
Nvidia Mellanox ConnectX-5
- **Storage:**  
2x 480GB SSD (OS)  
2x 3.84TB NVMe (Hot Data)  
2x 7.68TB SSD (Warm Data)
- KVM will be used to create 4 ES nodes: master, hot data, warm data, transform (with logstash).



# SOFTWARE OVERVIEW



## COLLECTION

- Custom Scripts
- Elastic beats
- Gmond
- Nagios
- SNMP traps
- Syslog
- Telegraph



## ENRICHMENT

- ES pipelines
- Logstash



## STORAGE

- MariaDB
- Elasticsearch
- influxDB
- RRD



## VISUALIZATION

- Ganglia
- Grafana



## ALERTING

- Email
- Grafana
- Nagios
- Pager (24/7)

\*Purple denotes additions from the analytics project.



# ELASTIC SUITE

- It is the workhorse of the analytics platform.
- Originally interested in their built-in machine learning tools and while promising, licensing is expensive. The free 'basic' license provides all our needs so far.
- Currently using version 7.17.6, upgrading to 8.3.3 soon.
- **Elasticsearch:**
  - Flexible database, can hold heterogeneous data.
  - Easy to grow horizontally as demands increase.
  - Many tools to aggregate and transform data.
- **Logstash:**
  - Many filters to parse and enrich data.
  - Multiple instances and pipelines to balance the load.
  - Many input and output protocols.
  - Persistent and 'dead-letter' queues.
- **Beats:**
  - 'Smart' collectors that monitor log files (Filebeat), service metrics (Metricbeat) and network ports (Packetbeat).
  - Balance loads to multiple outputs and queues data if they are unavailable.

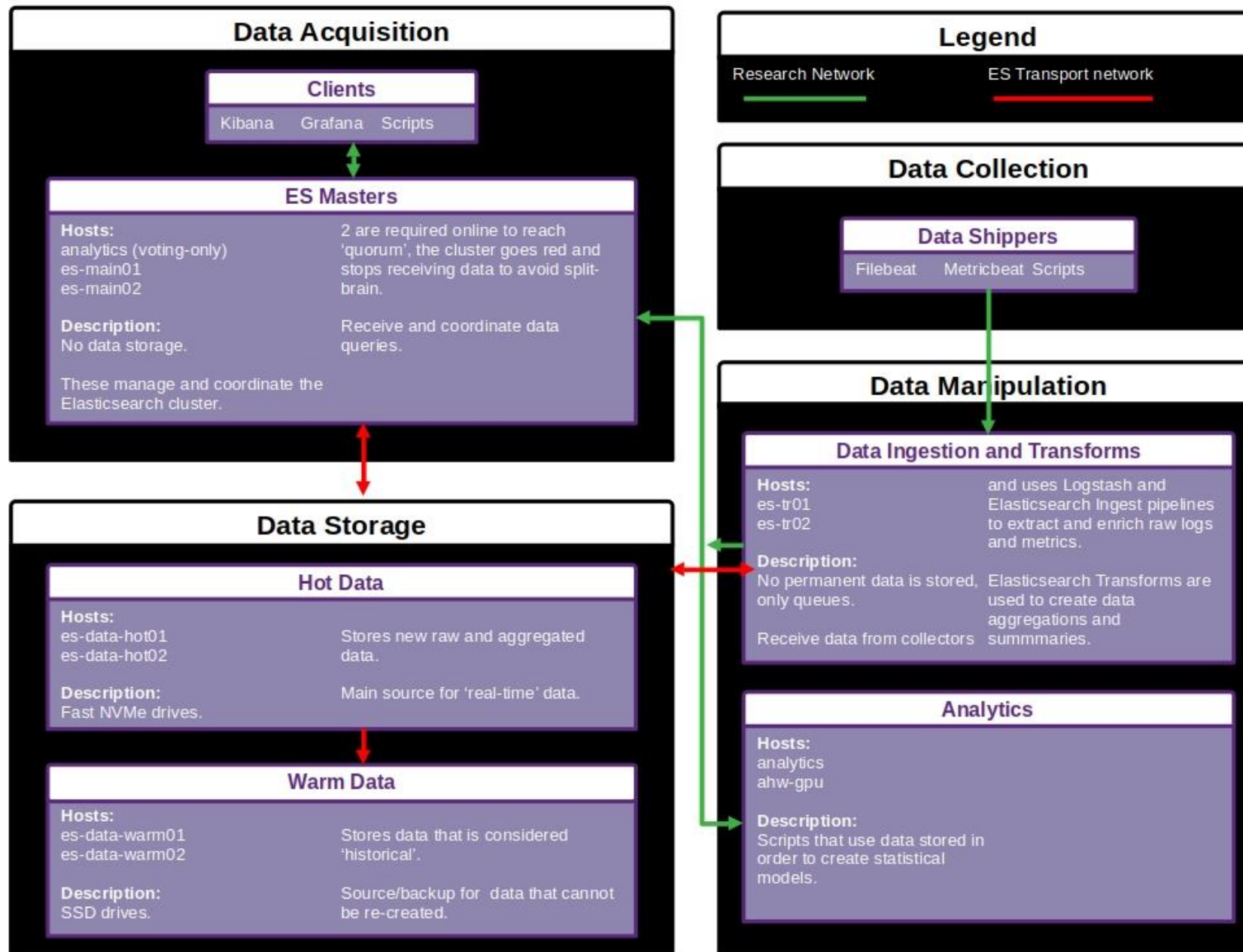


# GRAFANA

- Currently using version 9.2.1
- Our main visualization software for the following reasons:
  - It can use a large variety of different data sources.
  - Has many options for creating nice looking dashboards easily.
  - Powerful templating of panels.
  - It can further transform data on the fly.
  - It can generate alerts data query-based alerts.



# ELASTICSEARCH PIPELINE





# DATASETS

## Logs

Service	Size (GB)	Storage DB
dCache (billing, ftp srm webdav xrootd access)	4,500	Elasticsearch
network (router)	2	Elasticsearch
SNMP (traps)	15	MariaDB
system (auth, iptables, kernel)	400	Elasticsearch

## Metrics

Service	Size (GB)	Storage
dCache (queues, movers)	25	Elasticsearch
dCache (netflows)	1,250	Elasticsearch
HTCondor (job history, status)	10	Elasticsearch
infrastructure (DDN and inlet temps, SSD TBW)	50	Elasticsearch
infrastructure (humidity, PDU, temps, etc)	2*	RRD
mySQL (status)	1	Elasticsearch
network (router sflow)	15	influxDB
postgreSQL (activity, bgwriter, database)	200	Elasticsearch
system (cpu, mem, net, etc)	2*	RRD
tape library stats (device, volume)	2	influxDB
tape library (consumption, performance, staging, etc)	2*	RRD

\* 2GB for all datasets in RRD



# DCACHE

- Filebeat monitors and ships the contents of dCache's logs.
  - Billing logs contain transaction information within the scope of dCache.
  - Access logs contain transaction information pertaining to the different door protocols (FTP, SRM, WebDAV, XRootD).
- Logstash parses these logs into fields to create Elasticsearch documents, and enrich them as necessary (DNS resolution, GeoIP, tags).
- Packetbeat monitors the door protocol ports to obtain network flows and TLS handshake response times.
- A custom script parses dCache's pool queue table and sends it to Elasticsearch.



# DCACHE OVERVIEW DASHBOARD



# TAPE LIBRARY (HSM)

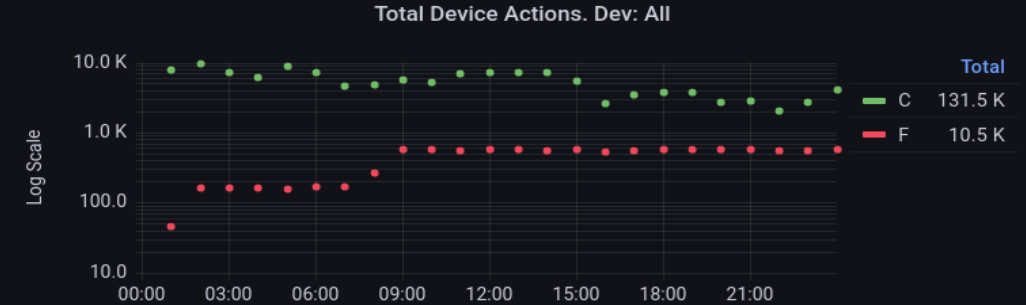
- The library produces SNMP traps when there are failures which are stored in MariaDB instance from where Grafana queries the information directly.
- Another MariaDB instance that records the tape library's devices and volumes activities.
- A custom script extracts the data and sends it to influxDB where we can manipulate it for later visualization.
- Here we chose influxDB to test and see how it compares against Elasticsearch for metrics data. Main benefit so far is the smaller storage footprint. Still investigating if there are other benefits.



# TAPE LIBRARY DASHBOARD

## Overview

Completed...	Bytes Writt...	Busiest Drives		Failed Actions by Drive	
		devname	actions	devname	failures
131.46 K	1.63 TiB	changer1	30584	LTO8F6C4R3	10514
		LTO8F2C3R3	23802	LTO8F6C2R3	2
		LTO8F6C4R3	21394		
		LTO8F6C2R3	9116		
		-----	----		
Failed Acti...	Bytes Read	Busiest Volumes		Failed Actions by Volume	
		volume	actions	volume	failures
10.52 K	2.26 TiB	S02428L8	22836	S01235L8	10512
		S01235L8	21378	S02296L8	2
		S02081L8	9028	S02081L8	2
		S01887L8	8964		
		-----	----		
Critical Tra...	Warning Tr...				
4	8				



### SNMP trap description

Time	Hosts	Total traps	Severity	Trap message
2022-11-24 10:42:30	ts4500-icc1	2	WARNING	Trap for drive TapeAlert 003. Flag: Hard error. Type: W Cause: The drive had an unrecoverable read, write, or positioni
2022-11-24 10:42:31	ts4500-icc2	2	WARNING	Trap for drive TapeAlert 003. Flag: Hard error. Type: W Cause: The drive had an unrecoverable read, write, or positioni
2022-11-24 10:42:32	ts4500-icc1	1	CRITICAL	Trap for drive TapeAlert 005. Flag: Read failure. Type: C Cause: The drive can not determine if an unrecoverable read f
2022-11-24 10:42:33	ts4500-icc1	2	CRITICAL	Trap for drive TapeAlert 005. Flag: Read failure. Type: C Cause: The drive can not determine if an unrecoverable read f
2022-11-24 10:42:33	ts4500-icc2	1	CRITICAL	Trap for drive TapeAlert 005. Flag: Read failure. Type: C Cause: The drive can not determine if an unrecoverable read f
2022-11-24 10:42:34	ts4500-icc2	2	CRITICAL	Trap for drive TapeAlert 005. Flag: Read failure. Type: C Cause: The drive can not determine if an unrecoverable read f

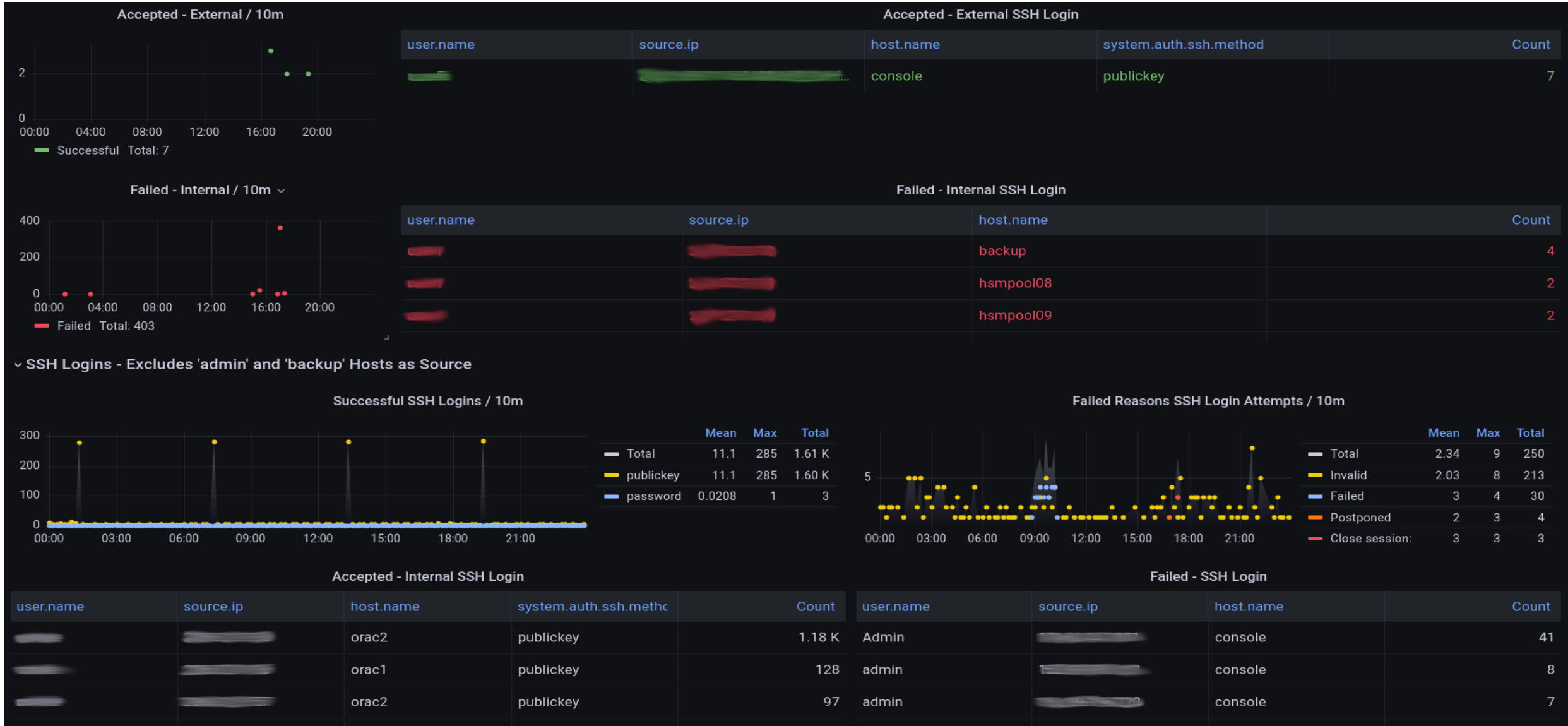


# SYSTEM SYSLOG

- All our hosts' kernel, auth and iptables logs are sent centralized via Syslog to one location and file.
- Filebeat monitors and ships the data.
- Logstash separates the three datasets, parsing and enriching as necessary.
- Our goal is to monitor and detect hardware issues, unauthorized logins and network traffic rejections.
- This is one of the datasets we would like to apply machine learning anomaly detection.



# LOGINS OVERVIEW DASHBOARD



# CURRENT AND FUTURE WORK

- Migrating Elasticsearch to the new hardware.
- Cleanup of existing datasets and re-processing in some instances.
- Creating 'events' database for overlay on Grafana and classification.
- Creating of time aggregated datasets (e.g. 1hour bins) to both reduce storage usage and normalization.
- Creating Grafana alerts from existing dashboards.
- “Tokenizing” logs. (e.g. 1.1.1.1 -> <IPADDRESS>).
- Creating tools for testing and implementing machine learning tools like anomaly detection, classification, correlation, prediction.
- Identification of datasets that can be brought together to create "vectors" for correlation analysis.
- Investigate the use of GPUs for this type of work.

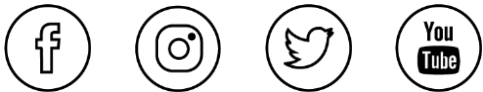




Thank you  
Merci

[www.triumf.ca](http://www.triumf.ca)

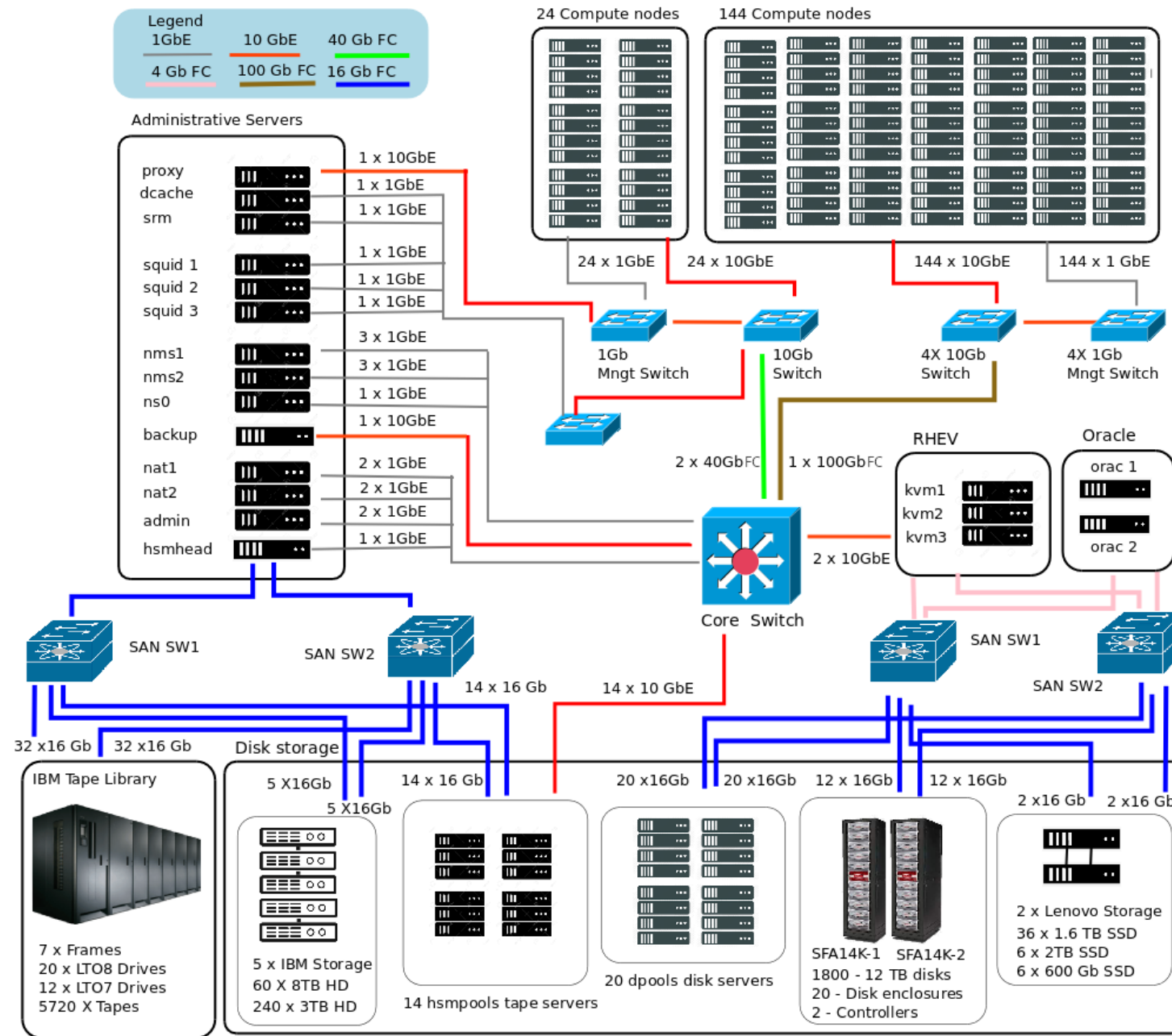
Follow us @TRIUMFLab



# ADDITIONAL MATERIAL



# TIER-1 INFRASTRUCTURE DIAGRAM



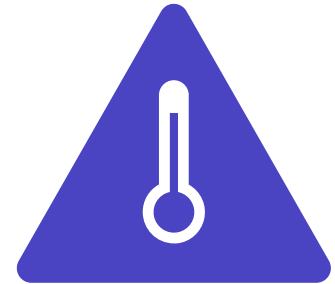
# CORE ROUTER SFLOW

- Telegraf receives sflow data from our Juniper core router. Only a percentage sample of all data is captured due to its large magnitude.
- Data is stored on influxDB.
- One idea is to implement Snort/Suricata as an intrusion detection system.
- Logs would be sent to Elasticsearch



# WORKER NODES INLET TEMPS

- A custom script queries all worker nodes' iDrac interfaces to obtain current temperature.
- It writes all information to a logfile.
- Filebeat monitors and ships the data.
- Logstash parses these logs into fields to create Elasticsearch documents, enriching it with infrastructure data.



# WORKER NODES INLET TEMPS

## Overview

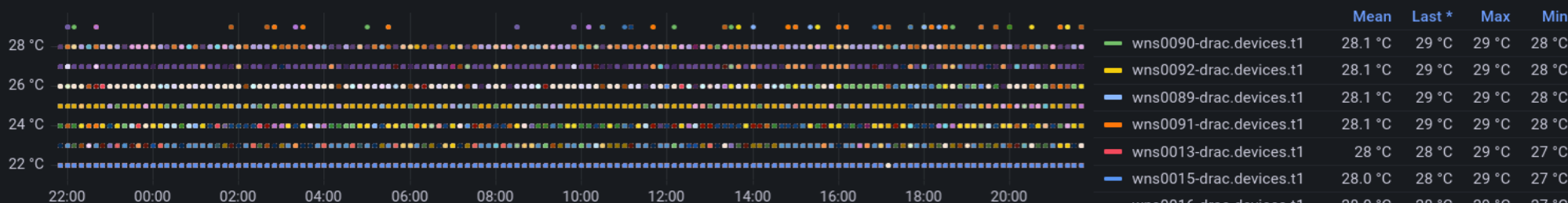
Max Device Inlet Temperature History - Per Rack / 1h

	22:00	23:00	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00
2	29	28	28	29	29	29	28	28	28	28	29	29	29	29	28	29	28	29	29	29	28	29	29	29
6	26	25	25	25	25	25	25	25	26	25	25	26	25	25	25	26	26	26	26	26	26	26	26	26
7	29	28	28	28	29	29	28	29	28	28	28	28	28	29	29	29	29	29	29	29	29	29	29	29
8	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28

Average Max Temperature for i...

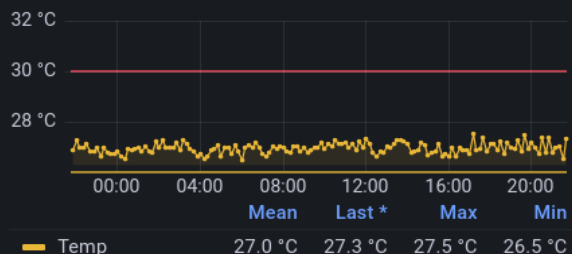


Average Inlet Temperature / 10m

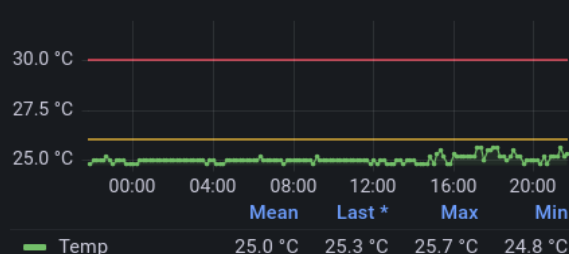


## Rack Temperatures

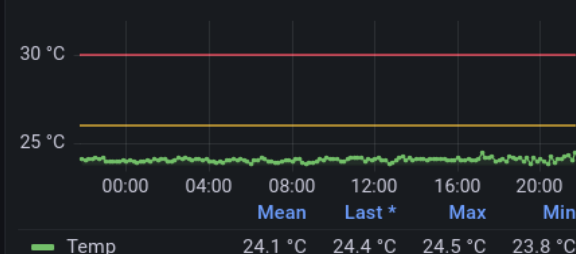
Rack: 2 - Average Inlet Temperature in / 10m



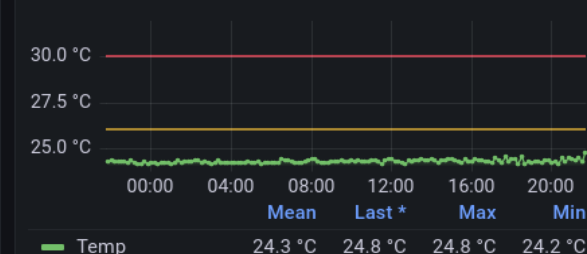
Rack: 6 - Average Inlet Temperature in / 10m



Rack: 7 - Average Inlet Temperature in / 10m



Rack: 8 - Average Inlet Temperature in / 10m



# HTCONDOR

- Two custom scripts query the HTCondor:
  - Every 15 minutes to obtain current jobs status.
  - Every 1 hour to obtain job history.
- Both write all information to two different logfiles.
- Filebeat monitors and ships the data.
- Logstash parses these logs into fields to create Elasticsearch documents.



# HTCONDOR JOBS' STATUS

Current Total Jobs

5765

Current Running Jobs

4998

Current Idle Jobs

763

Current Held Jobs

4

Current Removed Jo...

0

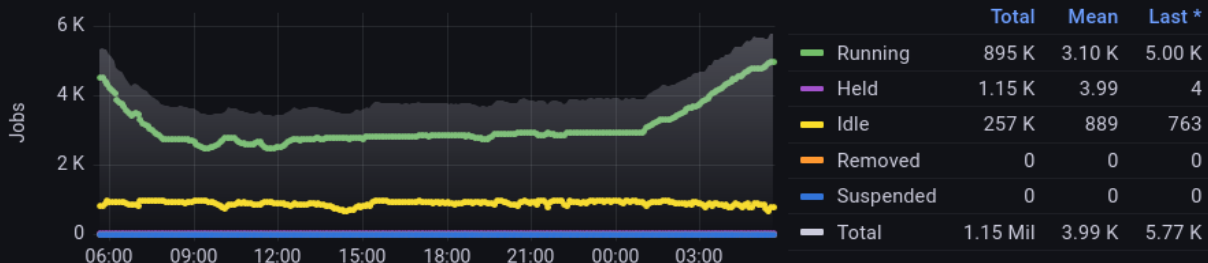
Current Suspended J...

0

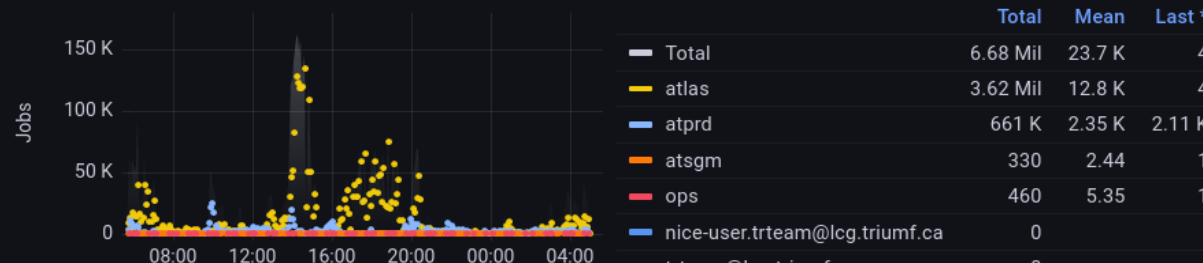
CPU Slots Provisioned

572

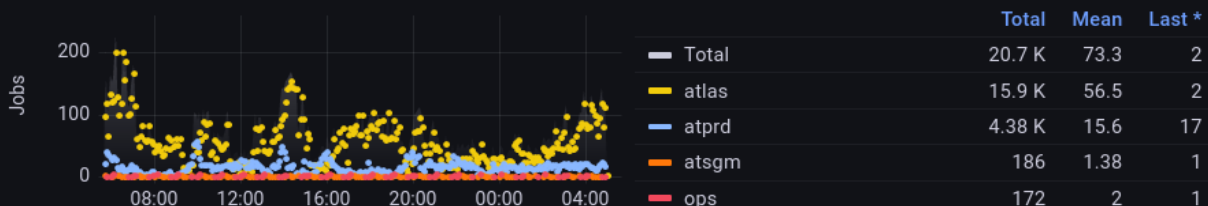
Jobs Status (All) / 5m



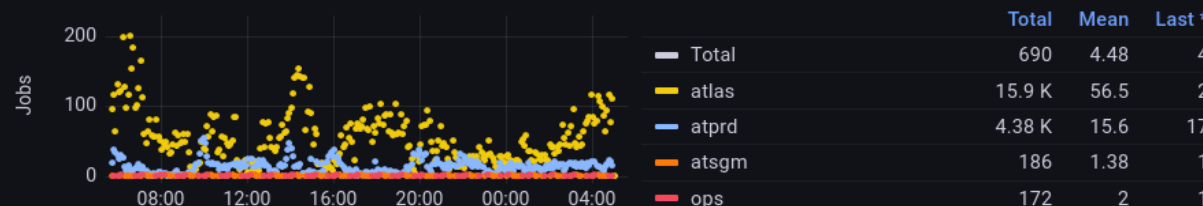
Completed Job Counts - CPU Normalized (All) / 5m



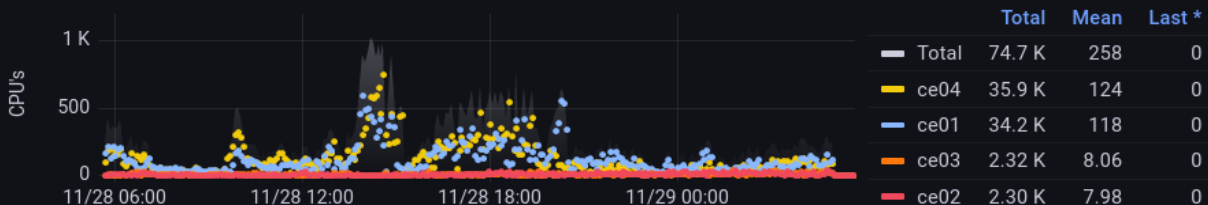
Completed Job Counts (All) / 5m



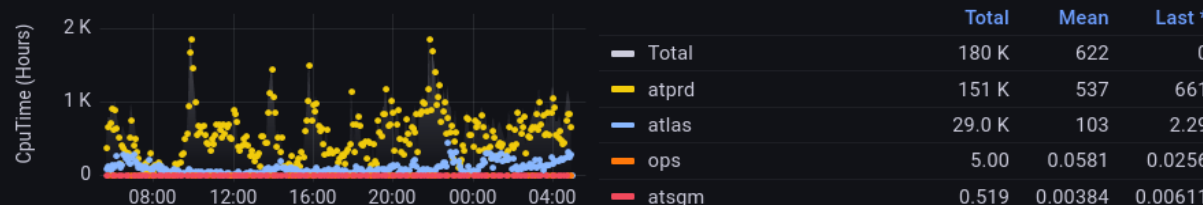
Removed Job Counts (All) / 5m



CPU's Provisioned (All) / 5m



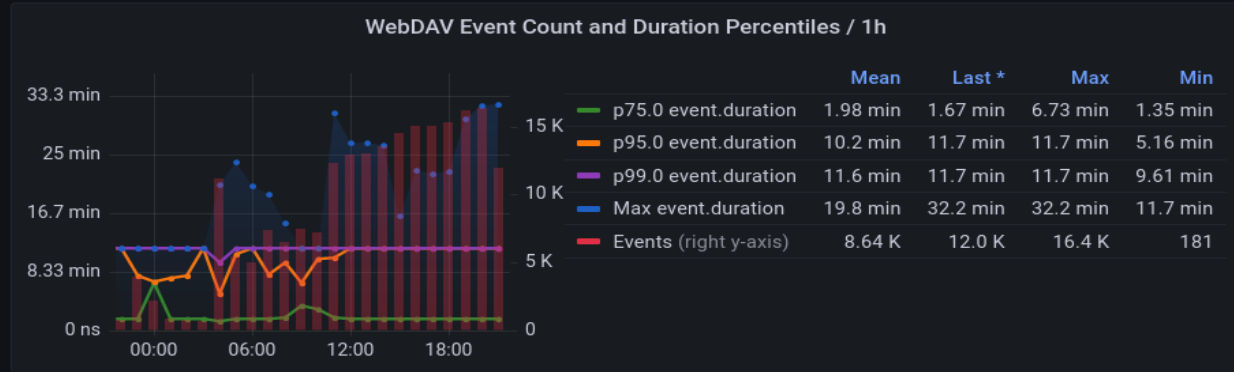
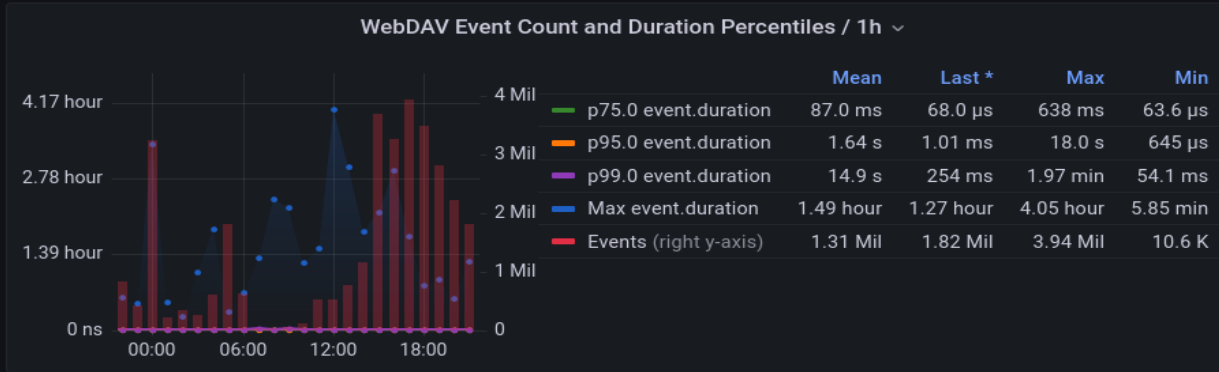
CPU Time (All) / 5m





# DCACHE PROTOCOLS (PACKETBEAT)

## Packetbeat Flow - Event Count and Duration Percentiles by Protocol - Plots



## Packetbeat Flow - Event Count and Duration Percentiles by Host - Plots

