

EIC PD Computing and Software Services I: Control and Monitoring Services

De Stefano, *et al*,
on behalf of the Scientific Computing & Data Facility (SCDF)
and other collaborators

EIC Computing PD at BNL AHM: 5 Dec 2025



@BrookhavenLab

Disclaimer

This work is the sum of many efforts in and without our organization; only the details of those directly related to this PD, and who have reported their efforts here, are included in detail. Apologies to those whose efforts are not mentioned.

This work is evaluatory in nature and still in development. As such, much of the technology is not yet available outside the working group or for wider testing.


Monitoring: Status & Achievements




Monitoring technology: Prometheus

Evaluated and built Prometheus instances for development and production (Pelosi)

- Initial deployments of node exporter and other service exporters (e.g., PostgreSQL exporter)

Prometheus: Targets

 Prometheus Alerts Graph Status ▾ Help



Targets

All scrape pools ▾ All Unhealthy Expand All

☐ Unknown ☐ Unhealthy ☒ Healthy

federate (1/1 up) [show more](#)

gpfs_test (0/0 up) [show more](#)

gs-opensearch (4/4 up) [show more](#)

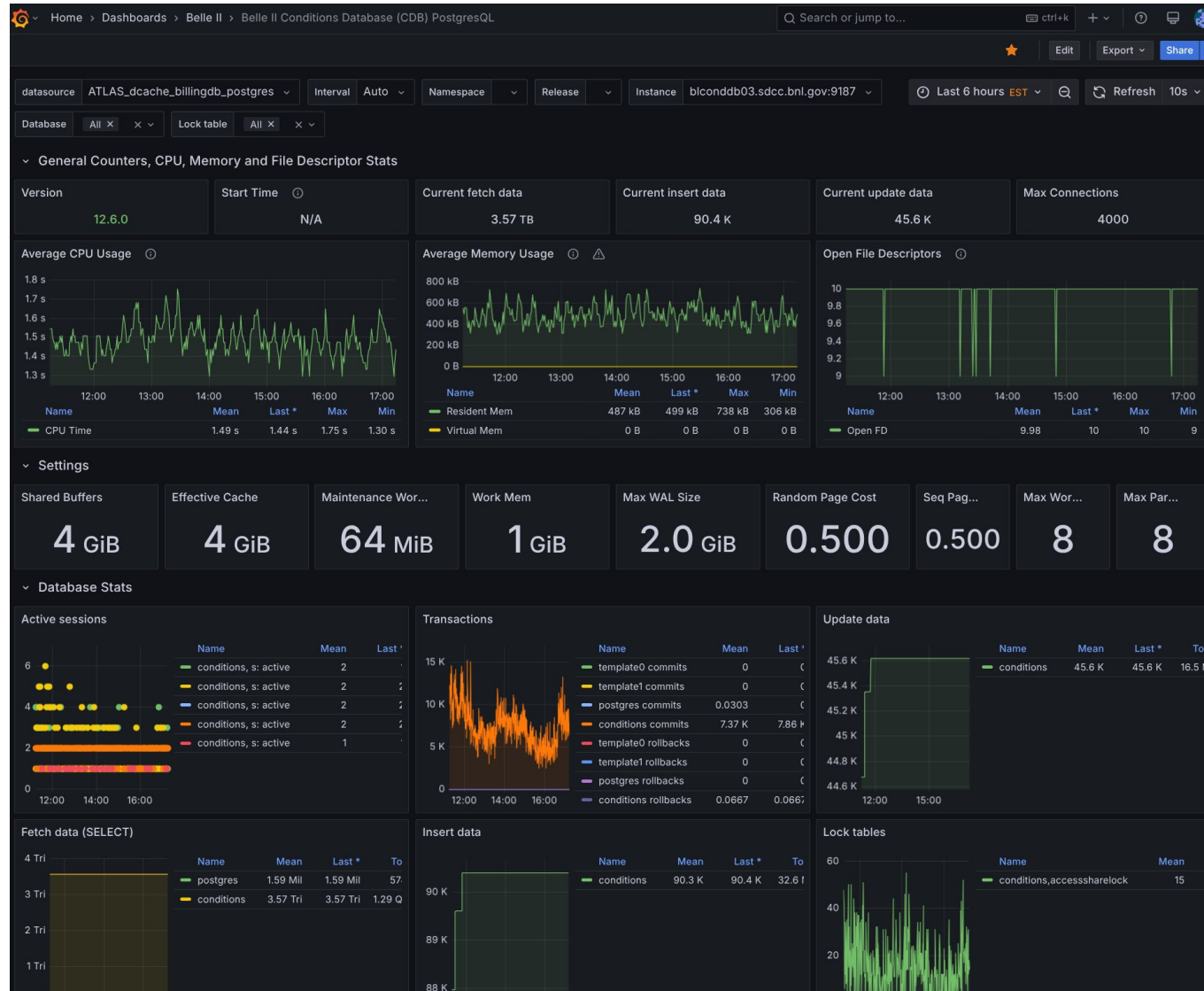
linuxfarm (2079/2079 up) [show more](#)

loki (1/1 up) [show more](#)

node_exporter (16/16 up) [show less](#)

| Endpoint | State | Labels | Last Scrape | Scrape Duration | Error |
|---|-------|---|-------------|-----------------|-------|
| http://gs-opensearch04.sdcc.bnl.local:9100/metrics | UP | instance="gs-opensearch04.sdcc.bnl.local:9100" job="node_exporter" ▾ | 43.851s ago | 16.751ms | |
| http://gs-osdashboards.sdcc.bnl.local:9100/metrics | UP | instance="gs-osdashboards.sdcc.bnl.local:9100" job="node_exporter" ▾ | 48.116s ago | 18.253ms | |
| http://gs-opensearch01.sdcc.bnl.local:9100/metrics | UP | instance="gs-opensearch01.sdcc.bnl.local:9100" job="node_exporter" ▾ | 1m 50s ago | 22.335ms | |
| http://gs-opensearch03.sdcc.bnl.local:9100/metrics | UP | instance="gs-opensearch03.sdcc.bnl.local:9100" job="node_exporter" ▾ | 1m 15s ago | 15.804ms | |
| http://blcondnb03.sdcc.bnl.gov:9100/metrics | UP | instance="blcondnb03.sdcc.bnl.gov:9100" job="node_exporter" ▾ | 1m 20s ago | 28.204ms | |

Prometheus: PostgresQL exporter



Monitoring technology: OpenSearch

Rebuilt legacy Elasticsearch nodes as new Opensearch cluster (Pelosi)

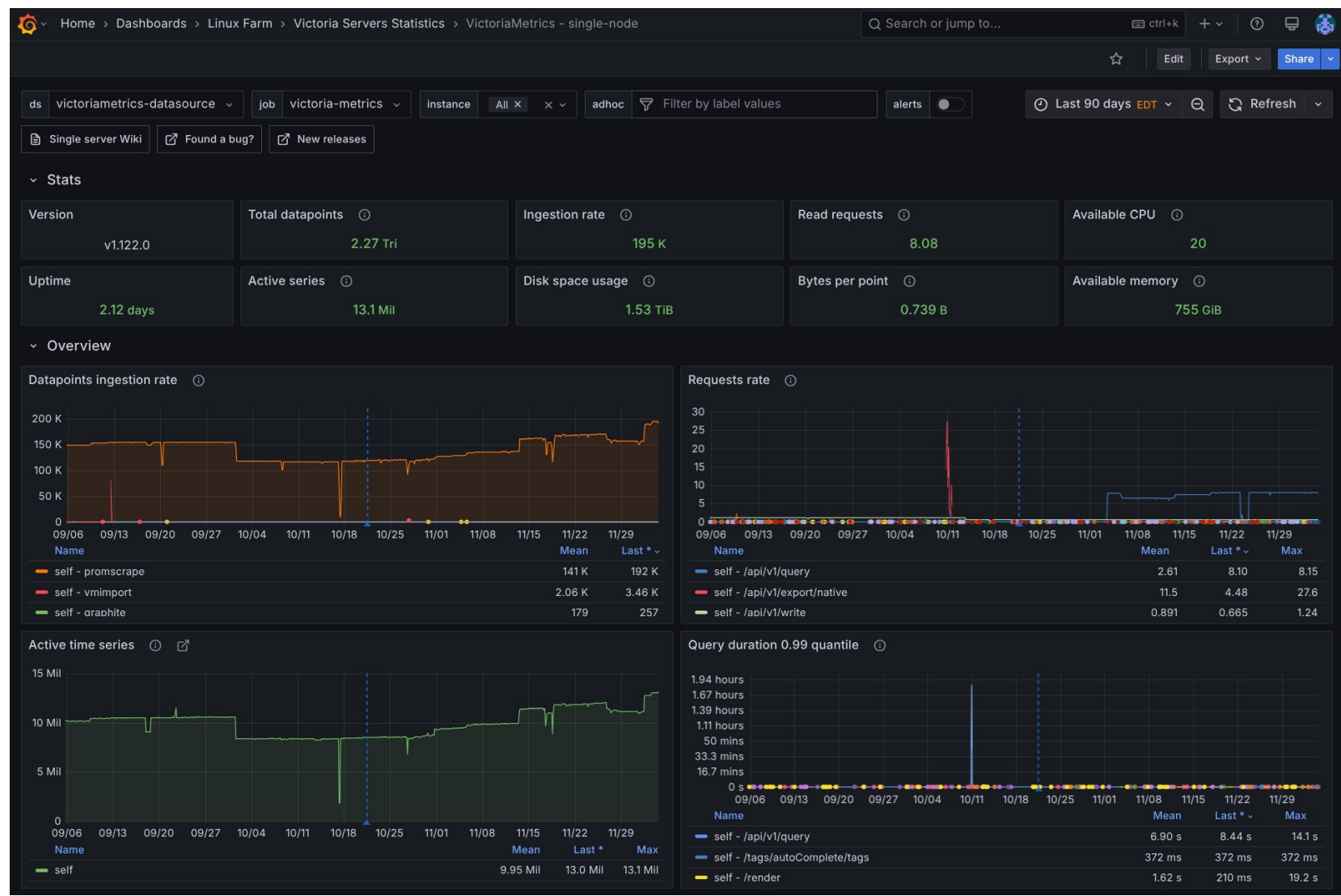
- Rebuilt legacy Logstash as development platform for OpenSearch
- AdStash tested on OpenSearch v1 (T. Smith)
- Kafka, Zookeeper nodes integrated with Logstash, OpenSearch pipelines (Garonne)

Monitoring technology: Victoria Metrics

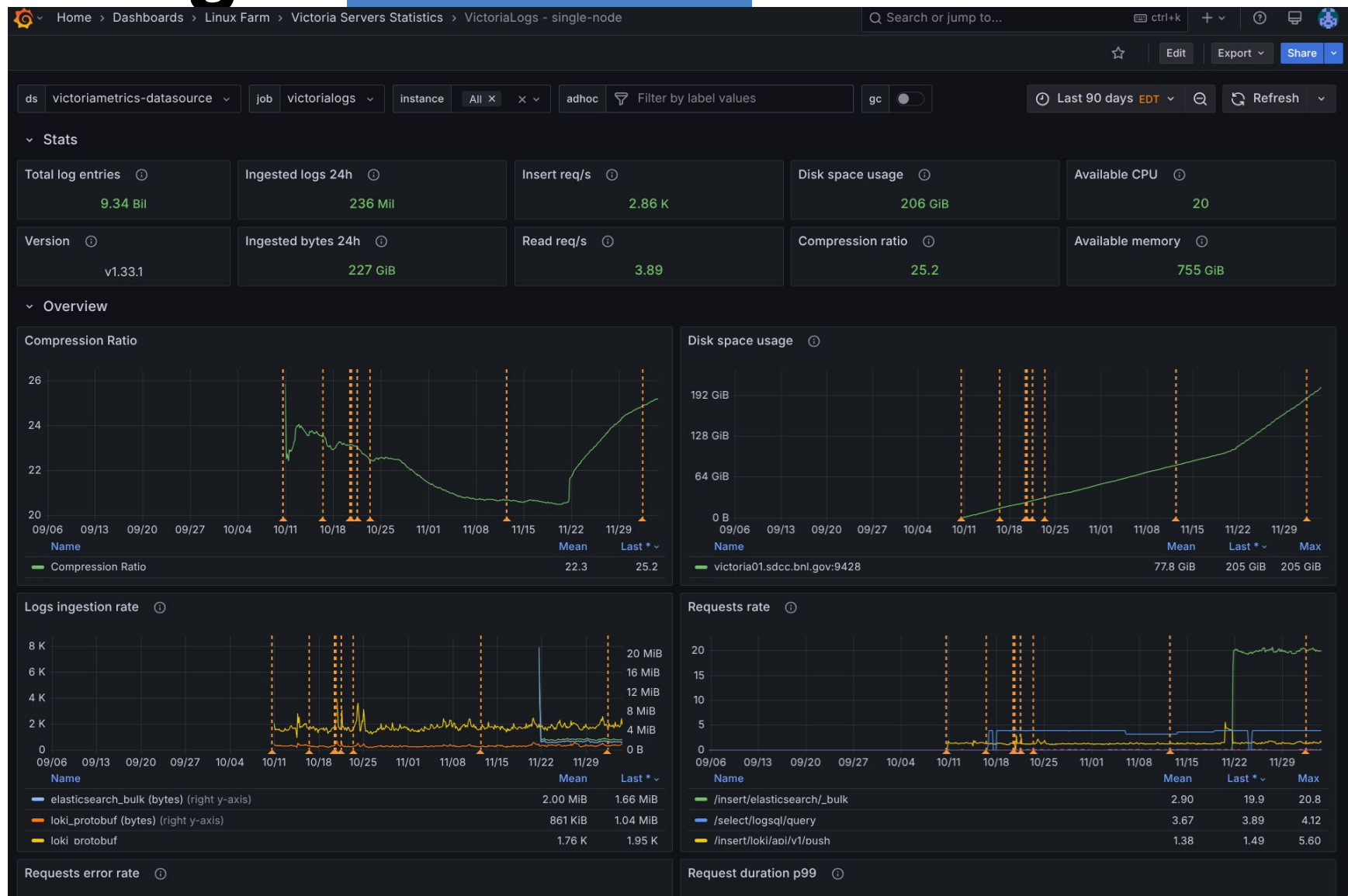
Evaluated Victoria Metrics and Victoria Logs functionality with ELK and Prometheus (Caramarcu)

- Built and evaluated services, moved to production for worker node farm (and beyond)
- Replaced collectd with node exporter and other custom exporters (e.g., CVMFS client exporter)
- Migrating Graphite/Carbon clusters to Victoria Metrics
- Replaced, created Grafana dashboards with Victoria Metrics data source
 - Includes new metrics monitoring, Grafana dashboards for NetApp storage (Frith)

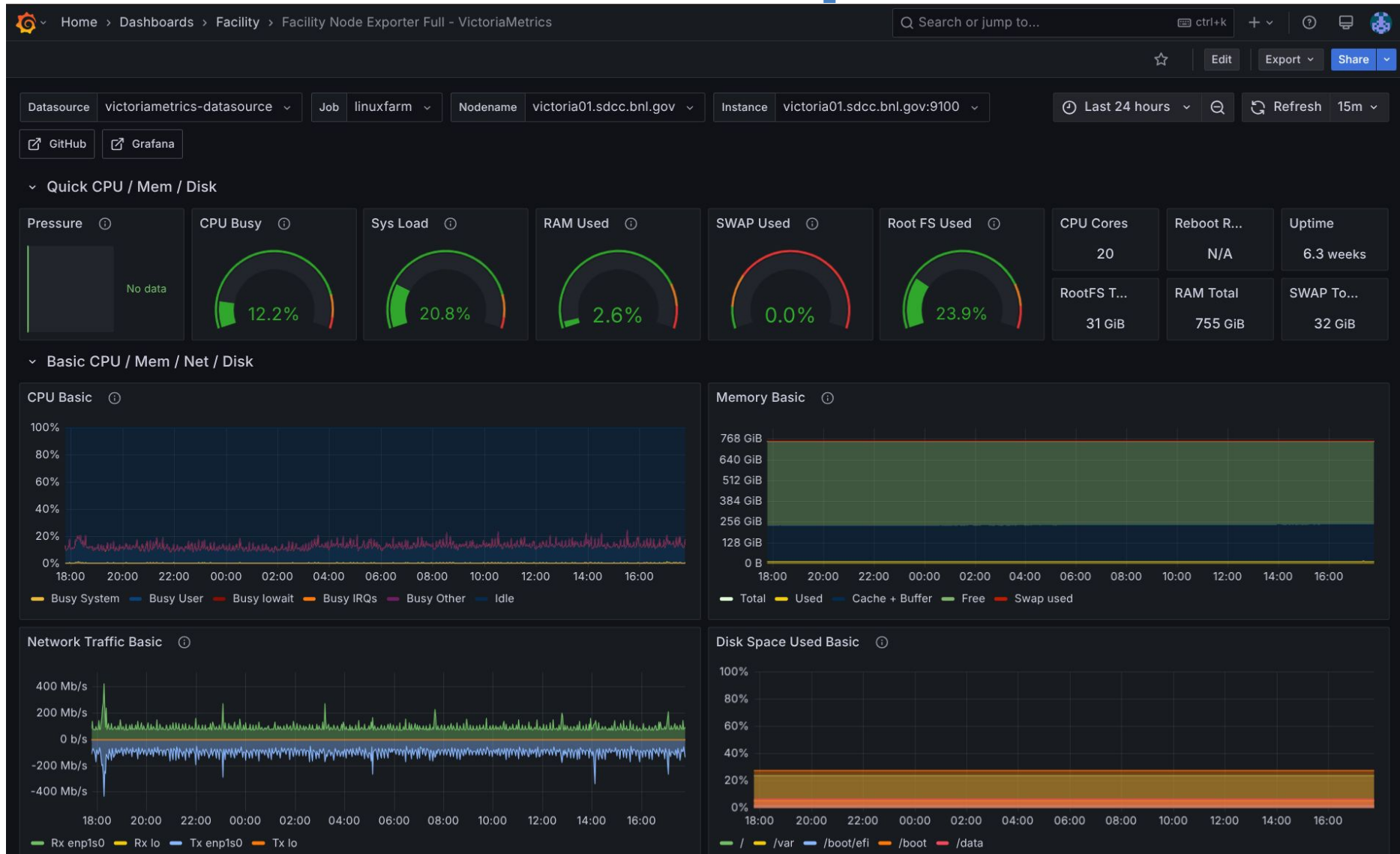
Victoria Metrics: Statistics



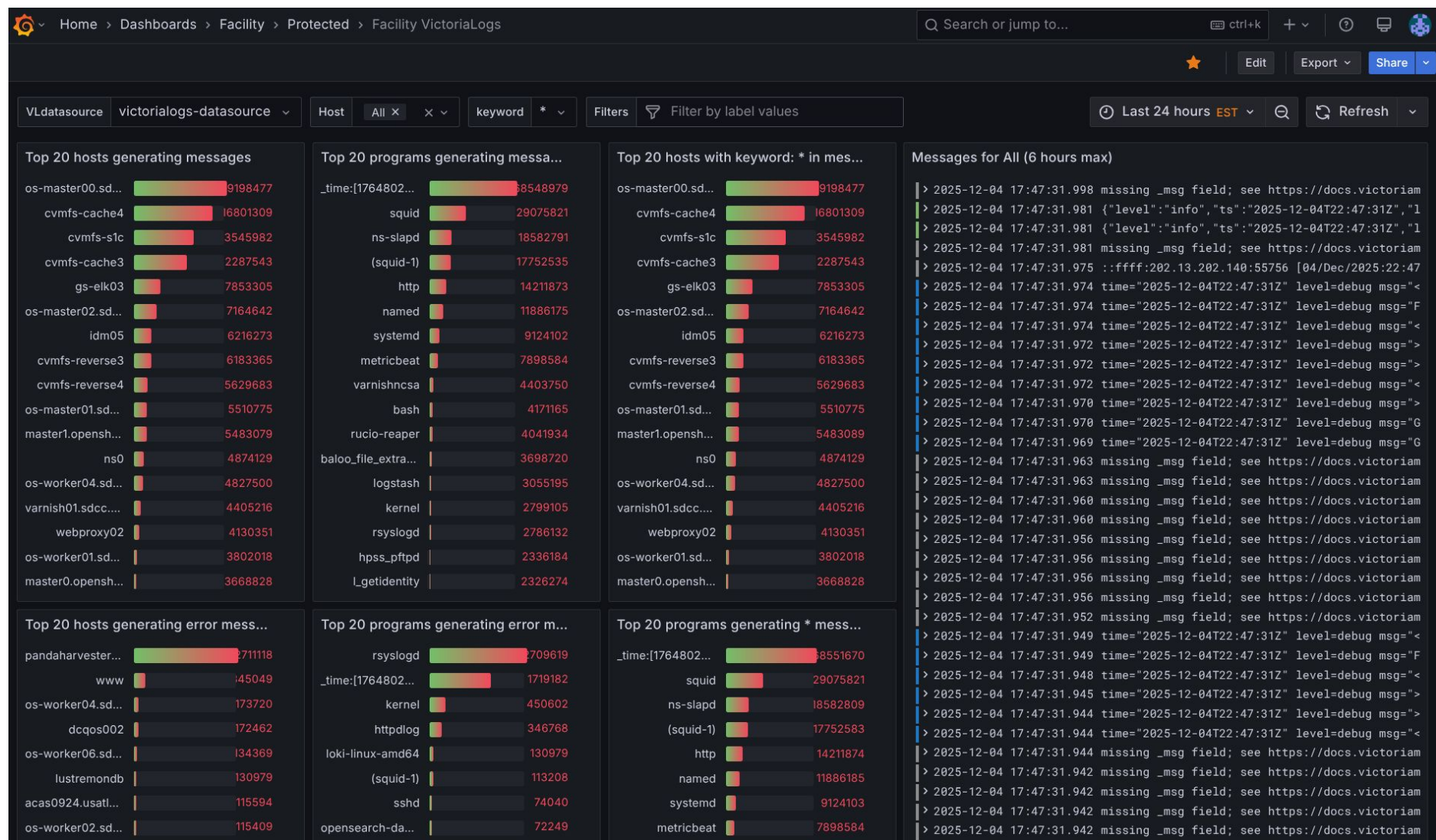
Victoria Logs: Statistics



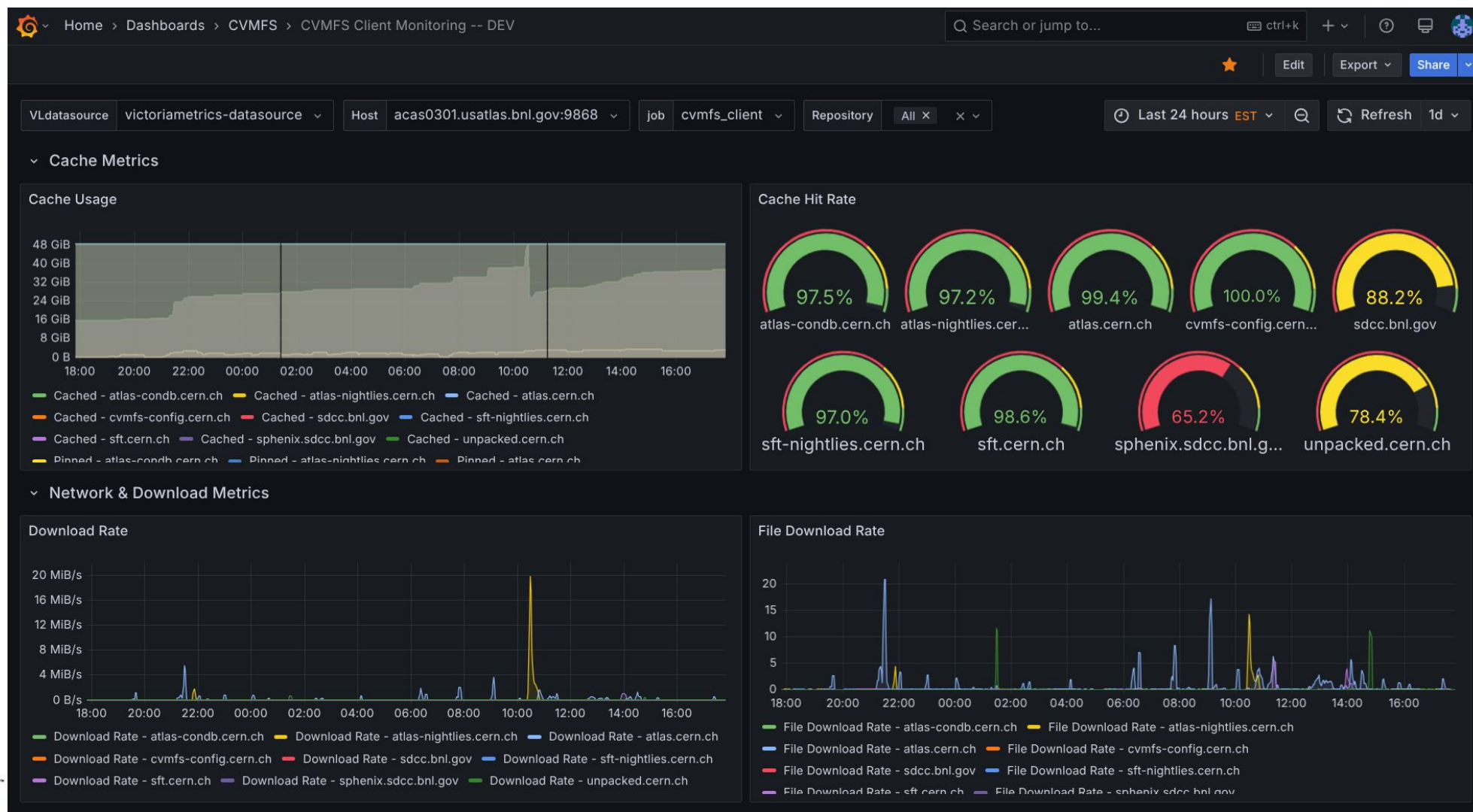
Victoria Metrics: Node exporter



Victoria Logs: Dashboard




Victoria Metrics: CVMFS client exporter



Evaluation: Storage

VAST storage (Hancock):

- DF-3015
- 338 TB (200 TiB licensed, 260 usable)
- Added new workflows for network block storage (NVMe over TCP)



Cluster and Rack Details - Fully Scaled

VAST Data Platform

Cluster Type

2x1 Ice Lake + Ceres 338 Cluster

Capacity

| | |
|-----------------|--------|
| Usable capacity | 0.2 PB |
| Logical @ : 1 | 0.5 PB |

Performance

| | |
|------------------|-----------|
| Read throughput | 40.0 GB/s |
| Write throughput | 10.0 GB/s |
| Read IOPS | 650k IOPS |
| Write IOPS | 93k IOPS |

*Key performance for selected protocol and options

Cluster Details

| | |
|----------------------------------|---------|
| 2x1 Ice Lake + Ceres 338 Cluster | |
| Max files / objects | 2.0 bn |
| Min usable capacity per DBox | 246 TiB |

Connection to customer network

| | |
|-----------------------------|--|
| CNodes to customer switches | |
| 16x 100Gb/s ports | |



Evaluation: CPU

Dell servers (Hancock):

- PowerEdge R7525 R1 (x2)
- Integrated into OpenShift cluster as worker nodes dedicated to monitoring hosts and workflows

PowerEdge R7525 Server

SAS/SATA Backplane

PowerEdge R7525 Server

No Trusted Platform Module

16X 2.5 SAS/SATA with XGMI supports 6X T4 cards

AMD EPYC 7543 2.8GHz, 32C/64T, 256M Cache (225W) DDR4-3200

AMD EPYC 7543 2.8GHz, 32C/64T, 256M Cache (225W) DDR4-3200

Heatsink for 2 CPU configuration (CPU greater than or equal to 180W)

Performance Optimized

3200MT/s RDIMMs

24x 64GB RDIMM, 3200MT/s, Dual Rank, 16Gb

C7, Unconfigured RAID for HDDs or SSDs (Mixed Drive Types Allowed)

PERC H755 Adapter FH

2x 960GB SSD SATA Read Intensive 6Gbps 512e 2.5in Hot-plug AG Drive, 1 DWPD

3x 3.84TB SSD SATA Read Intensive 6Gbps 512e 2.5in Hot-plug AG Drive, 1 DWPD

Monitoring technology: HPSS

Modernization of HPSS web infrastructure (Undrus)

- New RHEL 8 host baseline
- Migration from PHP 5 to PHP 8
- Code refactoring & cleanup
 - Current code size: 211,000 PHP and 344,000 HTML
 - Remove deprecated APIs, modernize language constructs, and introduce stricter typing, error handling
- Security tuning to align configuration with best practices

Monitoring technology: HPSS

Preparing the next-gen HPSS Web Platform

This testbed is dedicated to the preparation and deployment of the **Upgrade & Modernize HPSS Web Infrastructure** project, including migration to RHEL 8, migration from PHP 5 to PHP 8, and associated code refactoring and hardening.

- RHEL 8 migration
- PHP 8 upgrade
- Legacy PHP 5 decommissioning
- Code refactoring & testing

⚡ Why PHP 8 vs PHP 5?

🚀 Open New HPSS Web

ENVIRONMENT

- Isolated testbed – safe for experimentation

SCOPE

RHEL 8 adoption, PHP 8 runtime, dependency refresh, and systematic refactoring of the HPSS web codebase to remove PHP 5 patterns and legacy assumptions.

UPGRADE WORKSTREAMS

RHEL 8 · PHP 8 · Refactor

- **RHEL 8 baseline**
New OS image, updated libraries, modern TLS/OpenSSL defaults, and hardened defaults for HPSS web workloads.
- **PHP 8 migration**
Replace PHP 5 with PHP 8, align extensions, and validate compatibility with existing HPSS web modules.
- **Code refactoring & cleanup**
Remove deprecated APIs, modernize language constructs, and introduce stricter typing, error handling, and test coverage.
- **Performance & security tuning**
Benchmark critical HPSS paths, enable PHP 8 Just-In-Time (JIT) compiler where appropriate, and align configuration with security best practices.

🔧 Safe space for breaking changes & rollbacks

🔒 Security-driven modernization

📊 Measured performance impact

TARGET STACK

Modernized HPSS Web Platform

Operating system

RHEL 8 ← RHEL 6 (legacy)

PHP runtime

PHP 8.x ← PHP 5.x (current HPSS Web)

Web tier

Apache (tuned for PHP-FPM)

Application code

Refactored HPSS web modules with reduced legacy & improved testability

All experiments here are isolated from production.

Monitoring: Plans & Goals

To do:

- Continue deployment and configuration of key services (Victoria Metrics, Prometheus, OpenSearch)
 - Rebuild OpenSearch, LogStash instances on VAST storage
 - Prometheus: complement Victoria Metrics, or merge?
- Evaluate VAST storage:
 - compare performance with other NAS solutions (NetApp)
 - compare storage types (e.g., block vs. image storage)
- Evaluate potential AI use cases and solutions:
 - Anomaly detection (e.g., [vmanomaly](#)), predictive analysis, observability, natural language
- Continue modernization of HPSS monitoring
 - Update web services, consolidate and update background processes, upgrade backend database
- Convert Loki log output to Victoria Logs (Liu)
- (Potential) Dashboard reorganization and clean-up, data source consolidation (Yang)