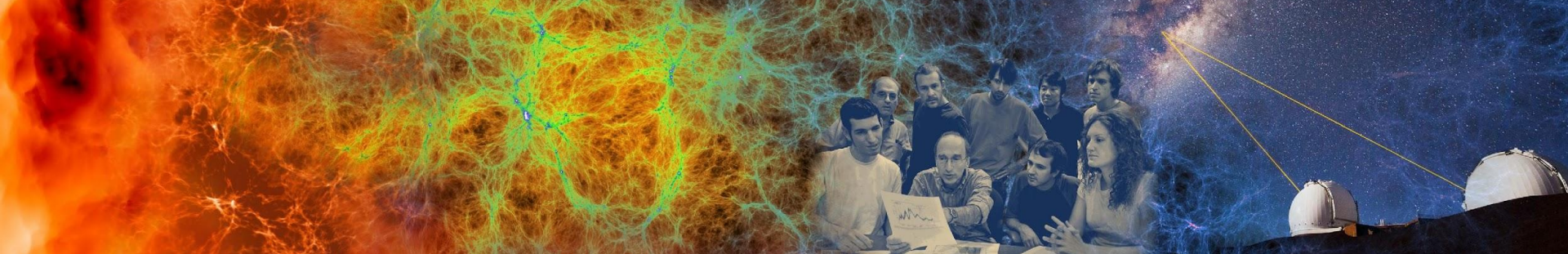


# Shasta Monitoring Improvements



John Stile [jstile@lbl.gov](mailto:jstile@lbl.gov)  
Computational Systems Group  
2023-05-08



# John Stile

CSG Group at NERSC

Supporting and Monitoring Shasta



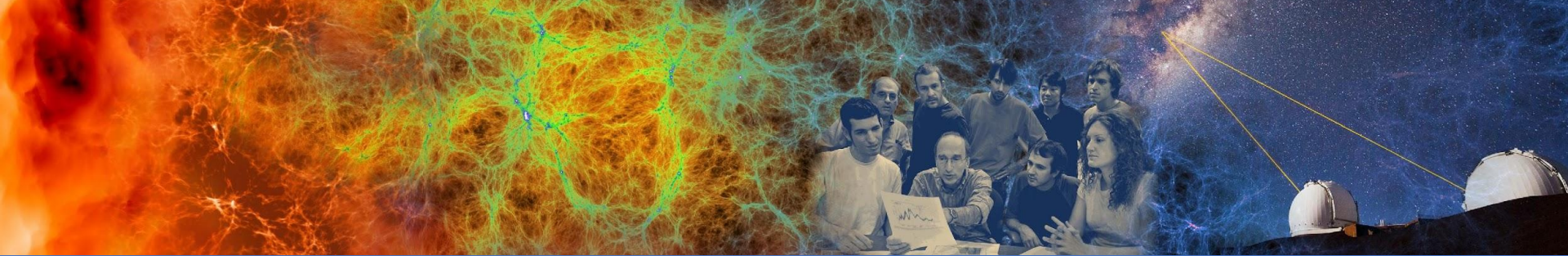
BERKELEY LAB



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science





“If you can’t monitor a service,  
you don’t know what’s happening,  
and if you’re blind to what’s happening,  
you can’t be reliable.”



~Betsy Beyer, Site Reliability Engineering: How Google Runs Production Systems

# Log Flow

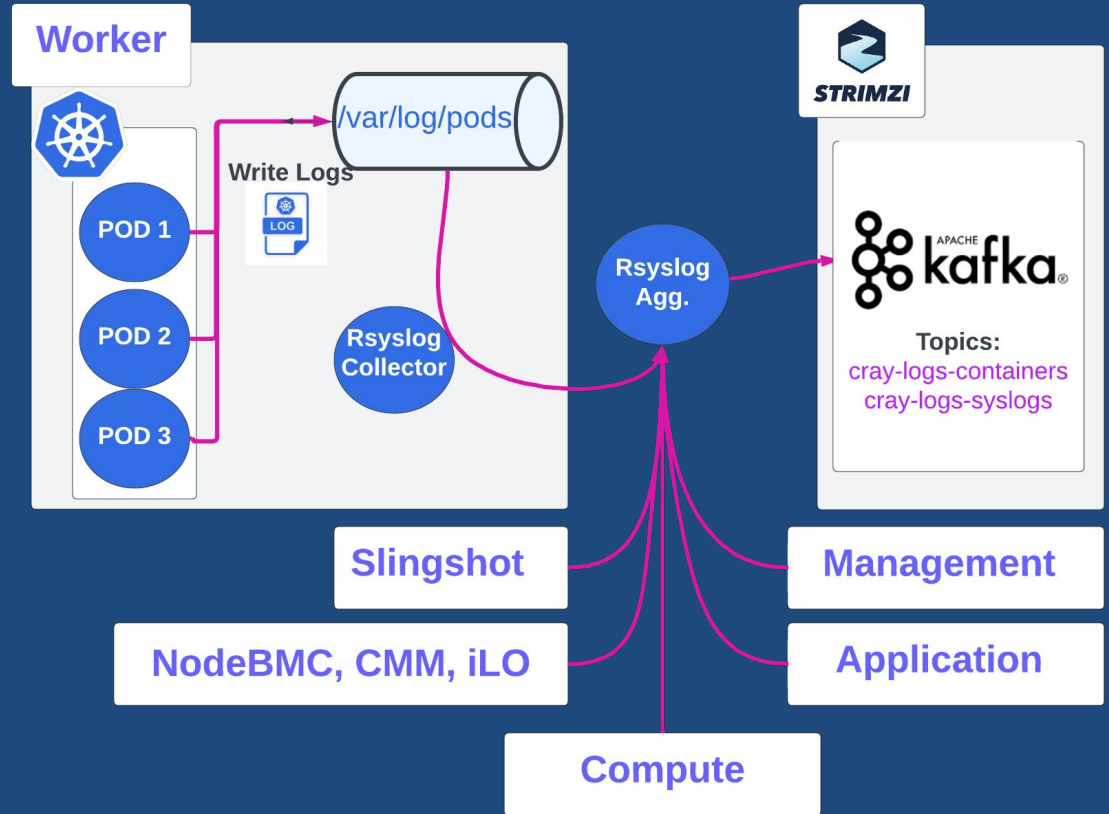
## Current configuration:

Logs flow to Rsyslog  
(exception: Aruba and E1000)

Rsyslog flow to Kafka topics:

- cray-logs-containers
- cray-logs-syslog

NERSC OMNI and OTG group consume  
Kafka for monitoring and alerting



# Rsyslog Monitor

Most logs lead to Shasta's Rsyslog

- No monitoring built in.

Our solution:

Modules: impstats & omprog  
python posts to Omni's VictoriaMetrics  
grafana dashboards

Add script to rsyslog container image

Add Vars to rsyslog charts  
(pod, node, cluster)

Shows rsyslog queue statistics with  
annotations (vertical red lines)

On the right, scaling ES pods effected  
logs all the way back to collector.k



# Rsyslog Monitor

Patch sma rsyslog-collector configmap

- Configure impstats and omprog
- Name, filter, and tune omfwd rule

```
module(load="omprog")
module(
  load="impstats"
  interval="300"           # post every 300 seconds
  format="json"           # Format json
  resetCounters="on"      # don't reset counters, do math on VM
  ruleset="process_stats" # Call external program
)
ruleset(name="process_stats") {
  action(
    name="stats_to_vmagent"
    type="omprog"
    binary="/rsyslog_stats_to_vmagent.py"
    forceSingleInstance="on"
  )
}
```

```
ruleset(name="collected") {
  if $msg contains 'ERROR: Unknown telemetry sourc' then
    stop
  else if $msg contains 'Error with Get_Metric_Type' then
    stop
  else if $msg contains 'Unexpected Integer key: gpu_id and value' then
    stop
  else {
    action(type="mmnormalize"
           rulebase="/etc/kubernetes.rb"
    )
    action(
      name="collector_fwd"
      type="omfwd"
      target="rsyslog-aggregator"
      port="514"
      protocol="tcp"
      action.resumeRetryCount="900"
      queue.type="linkedList"
      queue.size="10000"
      queue.maxFileSize="256m"
      queue.maxDiskSpace="40g"
      queue.filename="forwarding-queue"
      queue.spoolDirectory="/var/log/cray"
      queue.workerThreads="16"
      template="json_data"
    )
  }
}
```

# Rsyslog Monitor

## Patch sma rsyslog-aggregator configmap

- Configure impstats and omprog (see previous slide)
- Remove elasticsearch action queues
- Name and tune kafka action queues

```
action(  
  name="agg_kafka_containers"  
  type="omkafka"  
  broker= "cluster-kafka-bootstrap.sma.svc.cluster.local:9092"  
  topic="cray-logs-containers"  
  template="my-data"  
  Partitions.Auto="on"  
  action.resumeRetryCount="4"  
  queue.type="LinkedList"  
  queue.size="50000"  
  queue.maxFileSize="256m"  
  queue.workerThreads="8"  
  queue.maxDiskSpace="2g"  
  queue.filename="kafka5"  
  queue.spoolDirectory="/var/opt/cray/bufferdir"  
  #queue.minDequeueBatchSize="64"  
  #queue.minDequeueBatchSize.timeout="500"  
)  
}
```

```
action(  
  name="agg_kafka_syslog"  
  type="omkafka"  
  broker= "cluster-kafka-bootstrap.sma.svc.cluster.local:9092"  
  topic="cray-logs-syslog"  
  template="json_data"  
  Partitions.Auto="on"  
  queue.type="LinkedList"  
  queue.size="50000"  
  queue.dequeuebatchsize="500"  
  action.resumetrycount="4"  
  queue.maxFileSize="256m"  
  queue.workerThreads="8"  
  queue.maxDiskSpace="2g"  
  queue.filename="kafka5"  
  queue.spoolDirectory="/var/opt/cray/bufferdir"  
  #queue.minDequeueBatchSize="64"  
  #queue.minDequeueBatchSize.timeout="500"  
)  
}
```



# Rsyslog Monitor

Today's Grafana dashboard: more stuff

## Collector Resources(3 panels)

- CPU, Memory, Net

## Collector impstats(3 panels)

- rsyslogd process
- Main Q
- omfwd

## Aggregator Resources(3 panels)

- CPU, Memory, Net

## Aggregator impstats(5 panels)

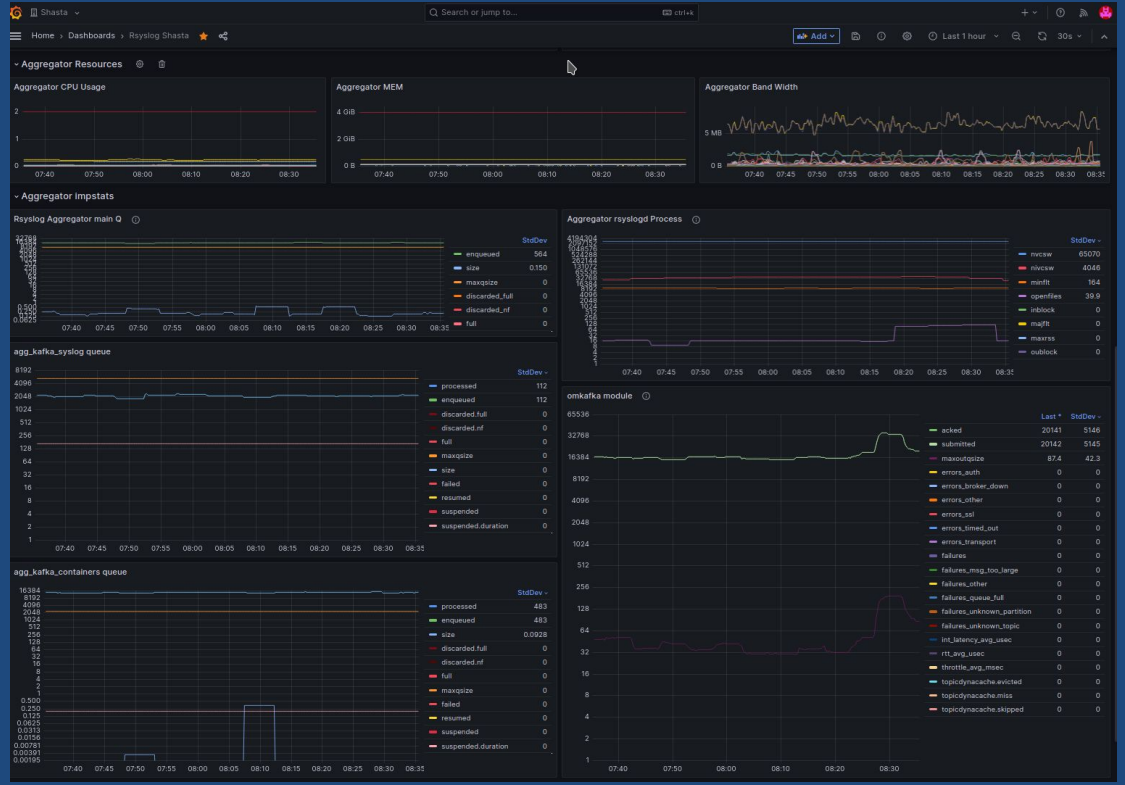
- rsyslogd process
- Main Q
- syslog queue
- container queue
- omkafka

## UDP Aggregator Resources(3 panels)

- CPU, Memory, Net

## UDP Aggregators impstats(2 panels)

- similar to above



To the right: Aggregator Resources and impstats





# Rsyslog Monitor

## Original configuration:

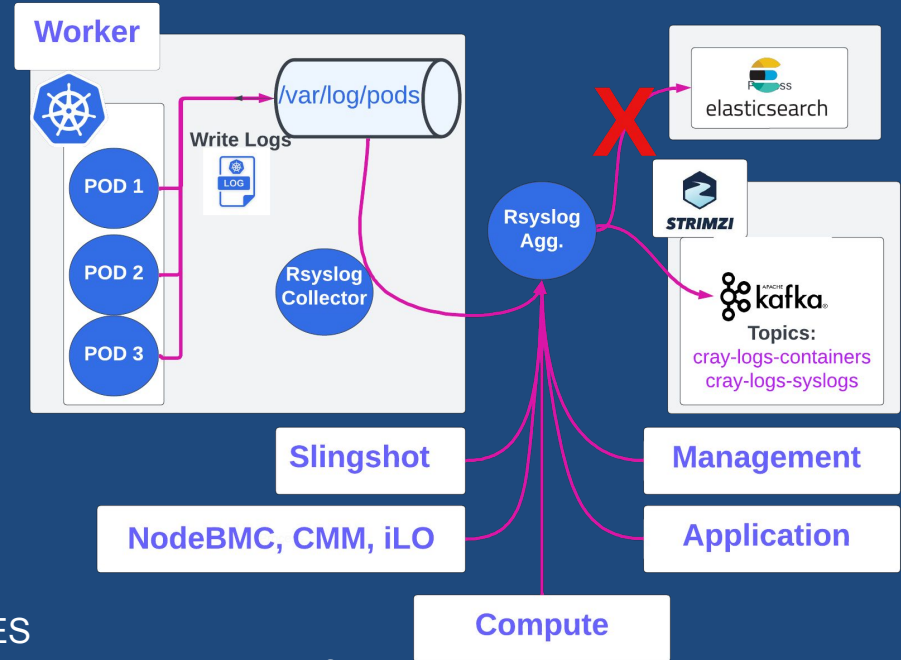
- Effects entire rsyslog infrastructure negatively

## Tried:

- K8s scaling CPU, Mem, replicas
- ES config: Shard counts
- Ceph SMF optimization
- Rsyslog omelasticsearch rule tuning

## Currently:

- Taken out of all rsyslog configs
- Broken index script in curator
- No good way to pause data during tuning or sick ES
- We see a sick ES effects rsyslog-aggregator, rsyslog-collectors, and kafka.



# Kafka Monitor

Shasta's Rsyslog feeds into Kafka

- No kafka monitoring enabled.

Our solution:

- Enable JMX included with Strimzi Kafka
- Enable jvm GC Log
- Enable metrics in kafka and zookeeper
- OTG provided Kafka Exporter
- Add PodMonitor to ns:sysmgmt-health
- Grafana dashboards from strimzi kafka github

On the right is one of 3 dashboards:  
Strimzi Kafka  
Strimzi Kafka Exporter  
Strimzi Operators

Missing 4 more...



# Kafka Monitor

Shasta's Rsyslog feeds into Kafka

- No kafka monitoring enabled.

Our solution:

- Enable JMX included with Strimzi Kafka
- Enable jvm GC Log
- Enable metrics in kafka and zookeeper
- OTG provided Kafka Exporter
- Add PodMonitor to ns:sysmgmt-health
- Grafana dashboards from strimzi kafka github

On the right is one of 3 dashboards:  
Strimzi Kafka  
Strimzi Kafka Exporter  
Strimzi Operators

Missing 4 more...



# Fabric Dashboard

## HPE Fabric Dashboard

- No dashboard, Until Now

## Our Contribution:

Create container Image

Create/Assabile helm charts

Deploy as k8s CronJobs

Integrate with keycloak

Running on Alvarez and Muller

<https://fabric-grafana.cmn.<system>.nersc.gov>

## Current Issue:

Perlmutter p-to-p file too big for ConfigMap

- Does not show log rates
- Does not show FM log events
- Does not show Switch log events

On the right Fabric Health Summary



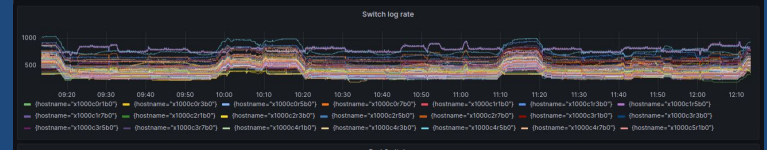


# Fabric Dashboard Needs

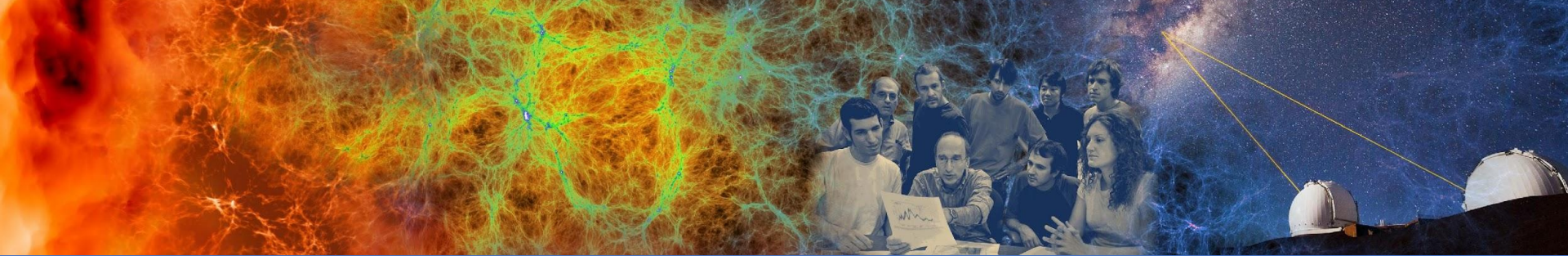
Point-to-point file does not fit in a configmap (need s3 bucket)

The NERSC created Fabric Manager Monitor offers a few things missing from the HPE Fabric Dashboard:

- Switch Log rates
  - High log volume has indicated an issue with a switch
- FMN messages
  - Fabric Manager tells you what it is doing
- Custom message alerts
  - We have add health checks as new bugs arise



```
Fabric Manager Logs
> 2023-04-12 12:18:34 ('switch': 'x3123cd44b8', 'port': '57', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
> 2023-04-12 12:18:34 ('switch': 'x3123cd44b8', 'port': '56', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
> 2023-04-12 12:18:34 ('switch': 'x3123cd44b8', 'port': '43', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
> 2023-04-12 12:18:23 ('critical' switch_line Wed Apr 12 15:47:57 UTC 2023, name: template-policy, description: Recommended: Fabric Admin should reduce telemetry rate limit | Cause: Disk usage 5.4
> 2023-04-12 12:18:23 [info] description:(Runtime): HEALTH, Configuration: HEALTH, Traffic: HEALTH, Security: HEALTH, APZ: HEALTH, resourceLink: http://1191qphst-fabri
> 2023-04-12 12:18:23 [info] description:(Runtime): HEALTH, Configuration: CRITICAL, Traffic: HEALTH, Security: HEALTH, APZ: HEALTH, resourceLink: http://1191qphst-fabri
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 4,fff', 'switch': 'x1182cd4780b', 'port': '51', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:06:22 ('action': 'taking link down', 'reason': 'b3/ifaill: b1: pcc 1k 0/ff, a8 f1 lfi hdb, serdes 0,fff', 'switch': 'x1182cd4780b', 'port': '58', 'krcal': 'x1182cd4780b151', 'nod
> 2023-04-12 12:05:34 ('switch': 'x3123cd44b8', 'port': '57', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
> 2023-04-12 12:05:34 ('switch': 'x3123cd44b8', 'port': '56', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
> 2023-04-12 12:05:34 ('switch': 'x3123cd44b8', 'port': '43', 'msg': 'pause: global/802.3x, tx on, rx on (active)', 'severity': 'critical', 'problem': 'fm_incorrect_link_flow_settings')
```



We are working with HPE to bring these improvements to Shasta

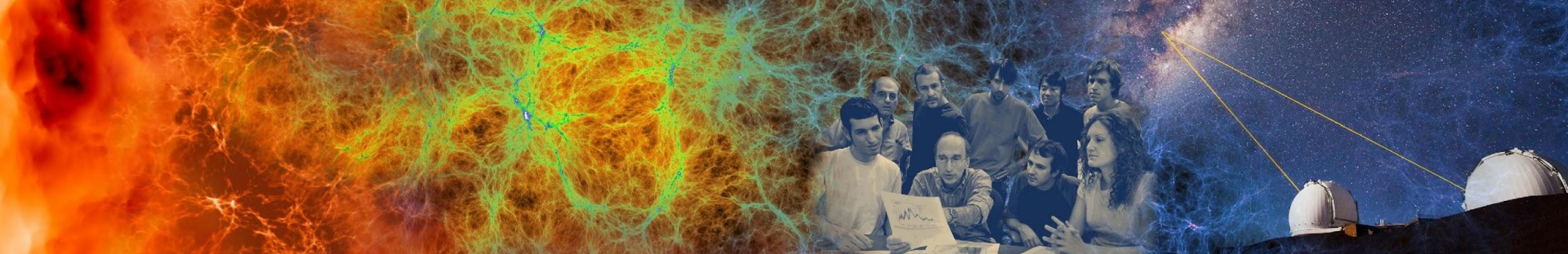


BERKELEY LAB



U.S. DEPARTMENT OF ENERGY

Office of Science



## Thank You:

### HPE:

Jeff Hanson	: Coordinating
Andrew Barry	: Rsyslog debugging, tuning, Kafka monitoring
Zakaria Bani	: Elasticsearch
Matt Silvia	: CSM integration
Craig DeLatte	: Ceph storage and Kafka turning
Nilakantan Mahadevan	: Fabric Dashboard
Forest Godfrey	: Fabric Dashboard
Rambabu Bolla	: Kafka Monitoring in Prometheus

### NERSC:

Siqi Deng	: Entire log pipeline and visualization
David Fox	: Promql and logql incantations and visualization
Brian Friesen	: Ingress and keycloak magic
Doug Jacobsen	: Investing resources to improve Shasta



BERKELEY LAB



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science



# Questions?

