Performance and Service Monitoring for GPFS

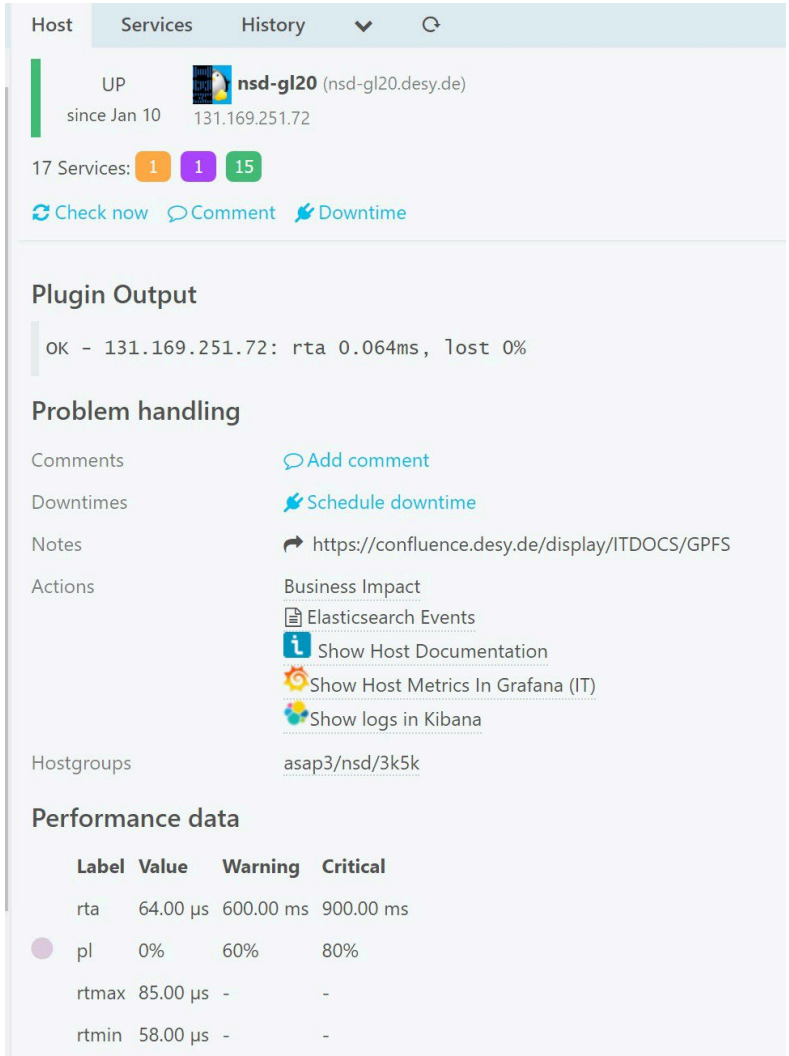
Stefan Dietrich, Martin Gasthuber, Jürgen Hannappel
Hamburg, 2022-07-14

HELMHOLTZ

DESY.

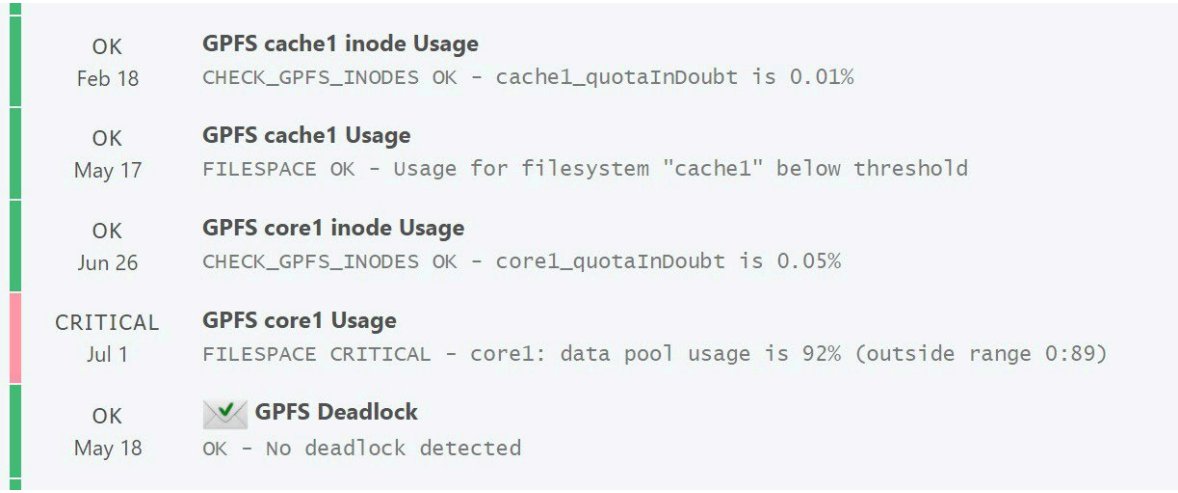
Service Monitoring: Icinga

- Central Icinga 2 Installation
  - Provided by IT-Operating group
  - Monitoring for all hosts in datacenter
  - Handles hardware- and service monitoring
  - Alerting via mail or pager for on-call shift
- Hardware issues handled by Operating staff
  - Service issues partially handled by Operating
- Standard RHEL 7/8 on ESS/DSS-G
  - Majority of already available Linux checks for free
  - IBM POWER slightly cumbersome



GPFS Service Checks

- Several GPFS specific service checks
  - Written in Python 3, partially with nagiosplugin https://github.com/mpounsett/nagiosplugin
  - Checks for pool & inode utilization, cluster export services (CES), deadlock, GPFS native RAID etc.
- Checks call mm* commands, parse output
  - Check script running on GPFS node
  - Icinga triggers check through remote agent
- Performance output of checks available in Grafana
- Checks currently not public, could be shared with KIT

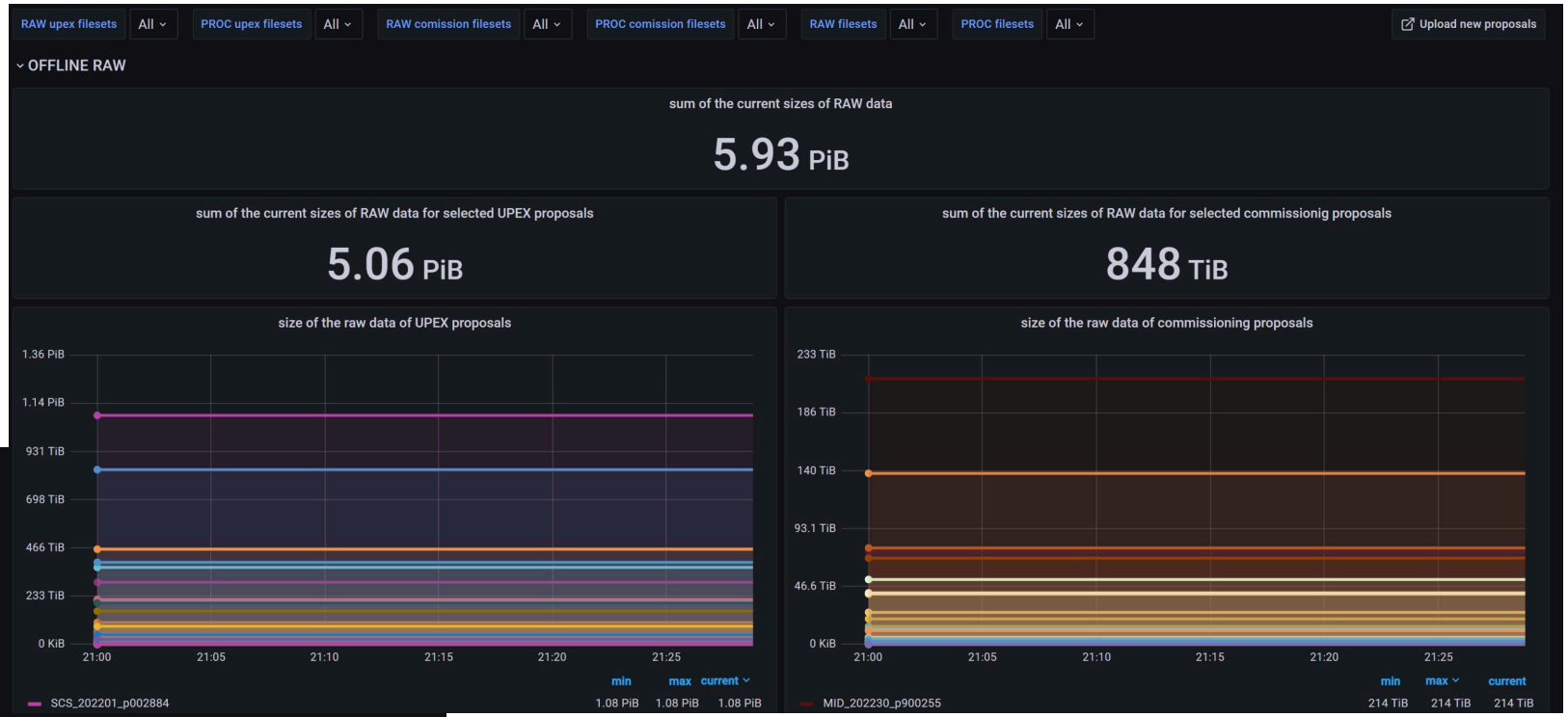


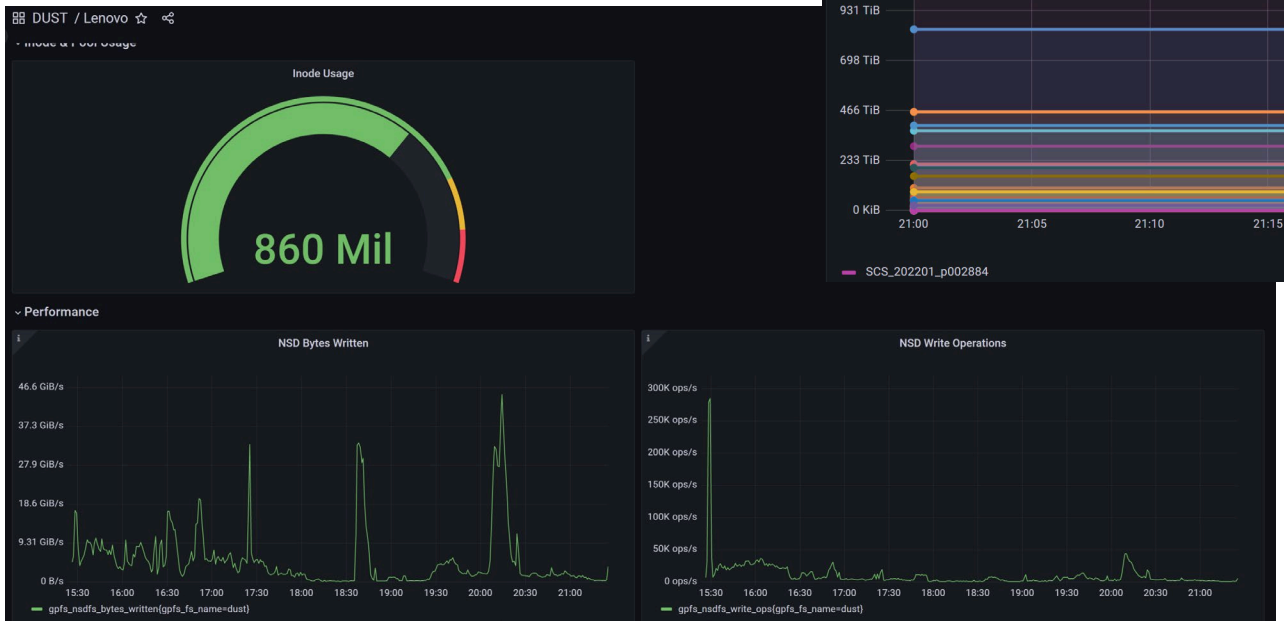
Performance Monitoring

- Several metrics from different data sources
  - Basic Linux metrics
    -> Icinga Service Checks
  - GPFS specific metrics
    -> Spectrum Scale Performance Monitoring Tool aka ZIMon
  - Custom metrics, e.g. Open FDs of Ganesha
    -> Script pusing metrics to Carbon/Graphite
  - InfiniBand (ibqueryerrors)
    -> InfiniBand Radar and Prometheus
- Visualization of metrics with Grafana
  - …Alerting for important metrics not handled by Icinga
- Several dashboards, primary for admins

- Grafana Bridge for ZIMon
  https://github.com/IBM/ibm-spectrum-scale-bridge-for-grafana
  - Grafana for visualization instead of GPFS GUI
  - Works mediocre, e.g. wrong query can easily block bridge or ZIMon collector
  - IBM provides periodically updates for bridge
  - Created script to dump metrics from ZIMon and push to InfluxDB
    -> currently not in use

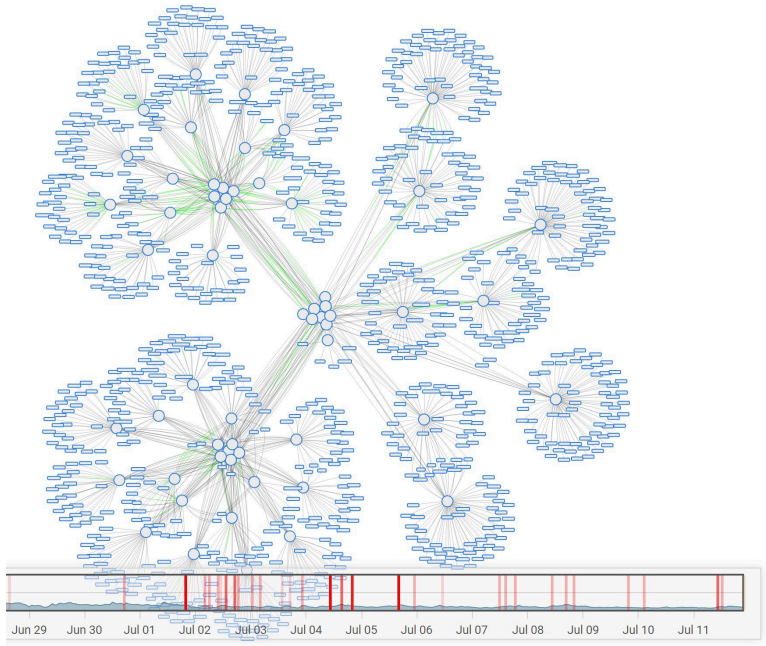
Example Grafana Dashboards



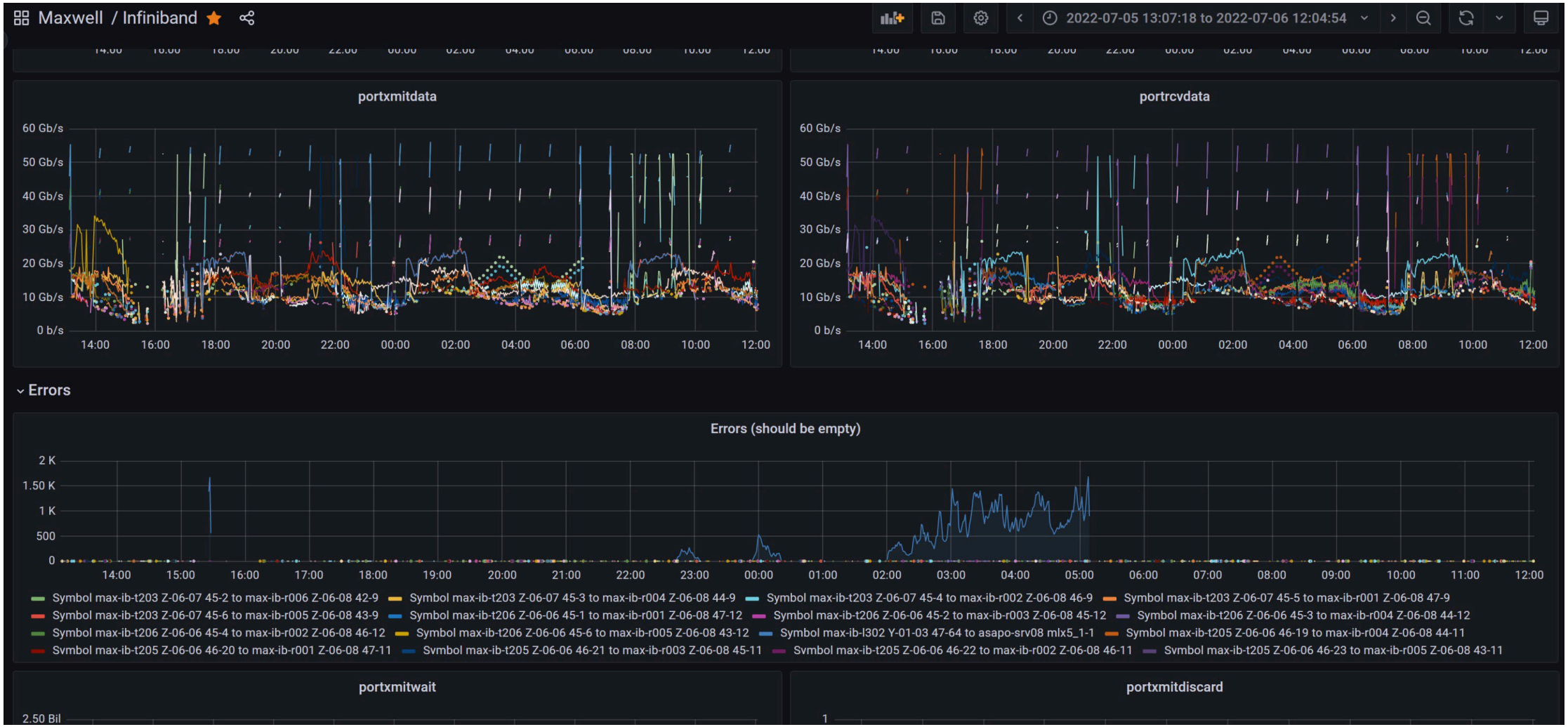


InfiniBand Monitoring

- InfiniBand Radar
  - https://github.com/infiniband-radar/infiniband-radar-web
  - Developed by former Bachelor Student
- InfiniBand Prometheus Exporter
  - https://github.com/guilbaults/infiniband-exporter
- InfiniBand has lots and lots of metrics…
  - …documentation on metrics is sparse
  - Mellanox: „Just buy UFM“
  - Still trying to identify the key metrics for InfiniBand, e.g. congestion/overload or dropped packets



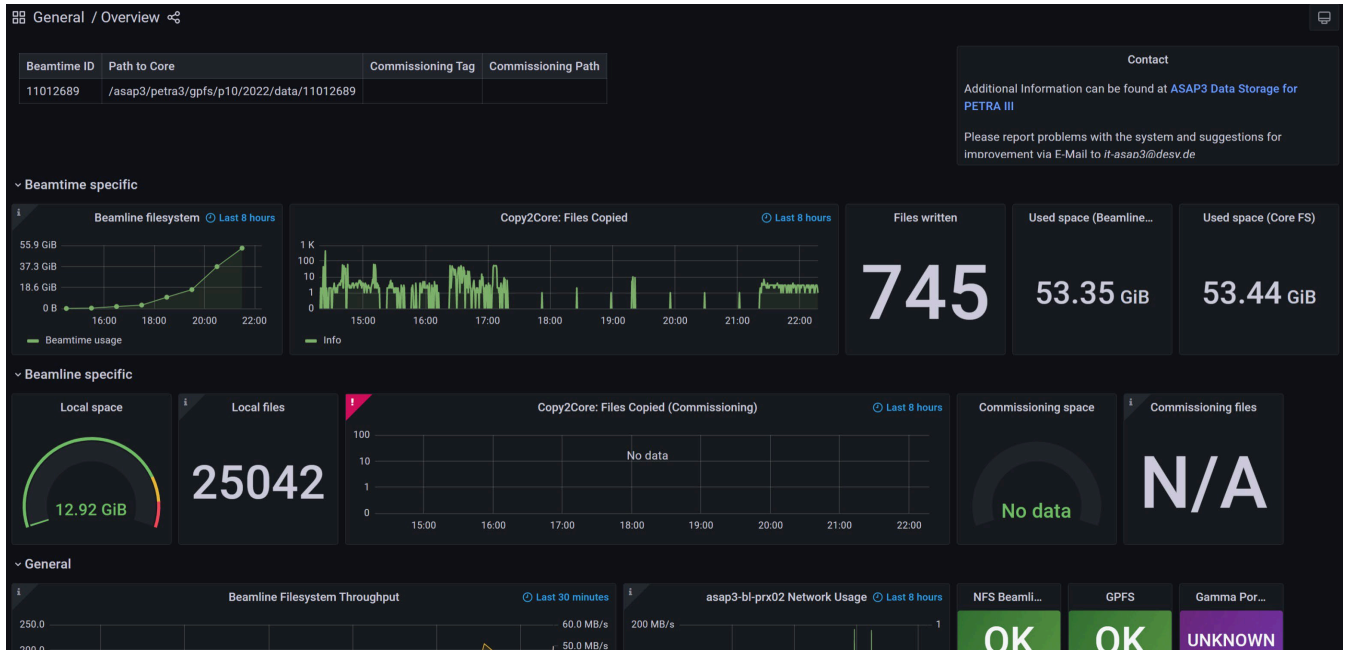
InfiniBand Prometheus Exporter



Monitoring for end-users

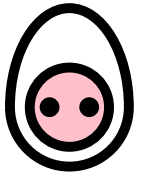
…won‘t somebody please think of the users!

- Specific to ASAP3 storage system
- Admins have all of the important metrics and service states, but end-users?
- Provide simple view to the beamline
  - Only important metrics, aggregate service states
- Dashboard uses ZIMon Bridge for GPFS metrics
- InfluxDB for additional data
  - Aggregate service states from Icinga (traffic light)
  - Text data, e.g. active beamtime
- Elasticsearch for log entries



Copy Tools - ewmscp

https://gitlab.desy.de/ewmscp



- ewmscp: Fast and versatile copy program
- Written in C++, partially with GPFS support
  - Support for multiple threads
  - Batteries included, like chown or setting XATTR
  - fixGpfsAcls: Recursively change NFS4 ACLs!
- ASAP3: Used with inotify to copy data beamline -> core filesystem

Thank you