

# GridPP

UK Computing for Particle Physics

## Ceph at RAL (The Tier 1 Awesome Plan)

James Adams, George Ryall, Shaun de Witt,  
Alastair Dewhurst, Andrew Lahiff



Science & Technology  
Facilities Council

- Same as at GridPP32
- No new SE going into Run 2
  - CASTOR disk layer
    - Ceph for tape buffer
    - Storage nodes becoming very large baskets
    - XRootD performance issues
  - Cloud
    - STFC Private IaaS cloud being built
    - Fast, resilient shared storage

- Working out how to deploy a production service
- This is hard
  - Monitoring
  - Day to day operation
  - Automated deployment of systems
    - Adding new nodes
    - Removing old nodes
    - Replacing nodes
- Learnt lots about running RADOS
- More to learn about CephFS
- Deployed three distinct clusters



- Benchmarking
- Tuning
- Testing new releases
- 6x Dell R520
  - 7x 3TB HDDs
  - 6x 1GbE NICs
  - One node with PCI-E SSD for journal testing

- Experiment in running Ceph on existing hardware
- Around 1PB raw capacity
- 30ish CASTOR grade storage nodes
  - Mostly 24x 2TB HDDs
  - Either single 1GbE or 10GbE NIC ← Problem!
    - Data traffic can starve out coherence and heartbeat traffic
- CephFS
- Testing Erasure Code
  - CephFS requires replicated pool - tiered pools to use EC

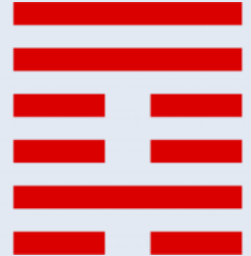
- Storage for STFC Private Cloud
- 30ish Dell R520
  - 8x 4TB HDDs
  - Dual 10GbE NICs
    - Separate client and cluster networks
- Replicated objects
- RADOS Block Device
  - Disks for KVM+Qemu VMs in OpenNebula
- Object Gateway
  - Amazon S3



- As shared FS on WNs and ARC CEs
  - Kernel updated to 3.11
  - CephFS as scratch space
  - Copied ~200TB CMS MC data from CASTOR and ran jobs against it
  - Real production jobs!
- GridFTP
  - Works well
  - Trivial to implement (even used Argus)
  - FTS worked too well
    - Drowned out heartbeats and mon traffic, cluster lost track of itself
- XRootD
  - Works
  - Not trivial



- Aquilon (new Quattor instance) from Day 1
  - No legacy baggage
- Ceph component for Quattor
  - Developed by University of Ghent HPC Group
  - Spent a month at RAL working with us
- Automate as much as possible
  - Make use of new service orchestration
    - Rules map nodes to clusters
  - CRUSH map built from Quattor's view of the infrastructure
    - Buildings, rooms, racks
    - OSD weight = drive size in TB







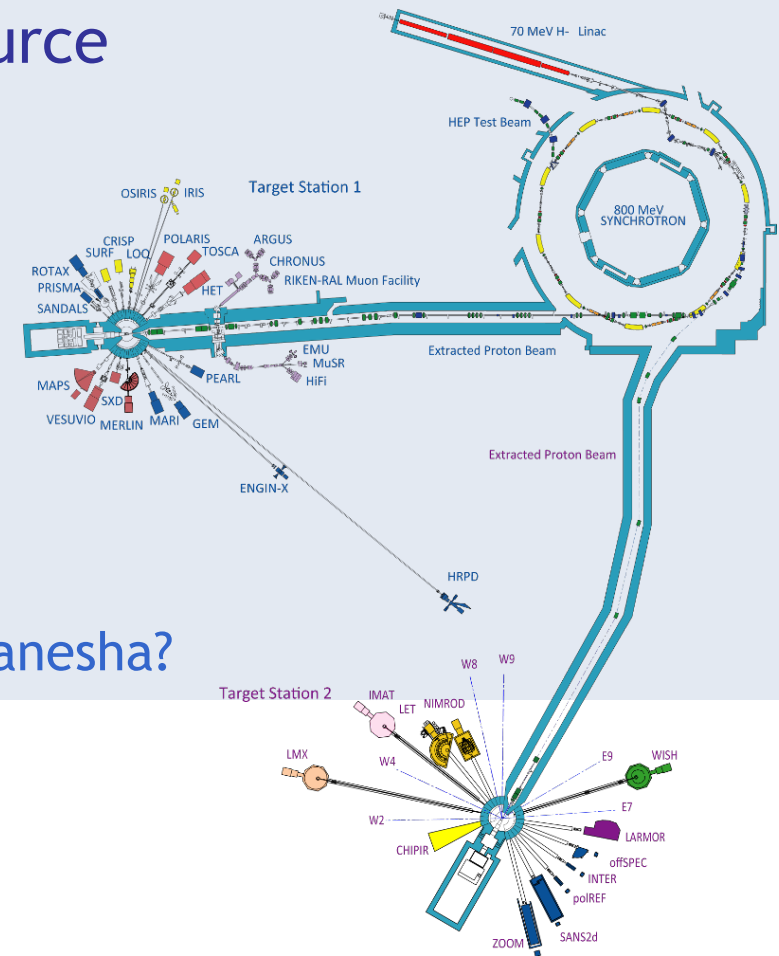
```
# id      weight  type name  up/down  reweight
-1       880.1   root default
-24      200     rack r89rack134
-8       40      host gdss498
-9       40      host gdss501
-10      40      host gdss505
-11      40      host gdss506
-12      40      host gdss507
-25      240     rack r89rack135
-13      40      host gdss508
-14      40      host gdss509
-15      40      host gdss510
-16      40      host gdss511
-17      40      host gdss514
-18      40      host gdss516
-26      160     rack r89rack136
-19      40      host gdss521
-20      40      host gdss522
-21      40      host gdss523
-22      40      host gdss524
```

Rack134 - 200TB

Rack135 - 240TB

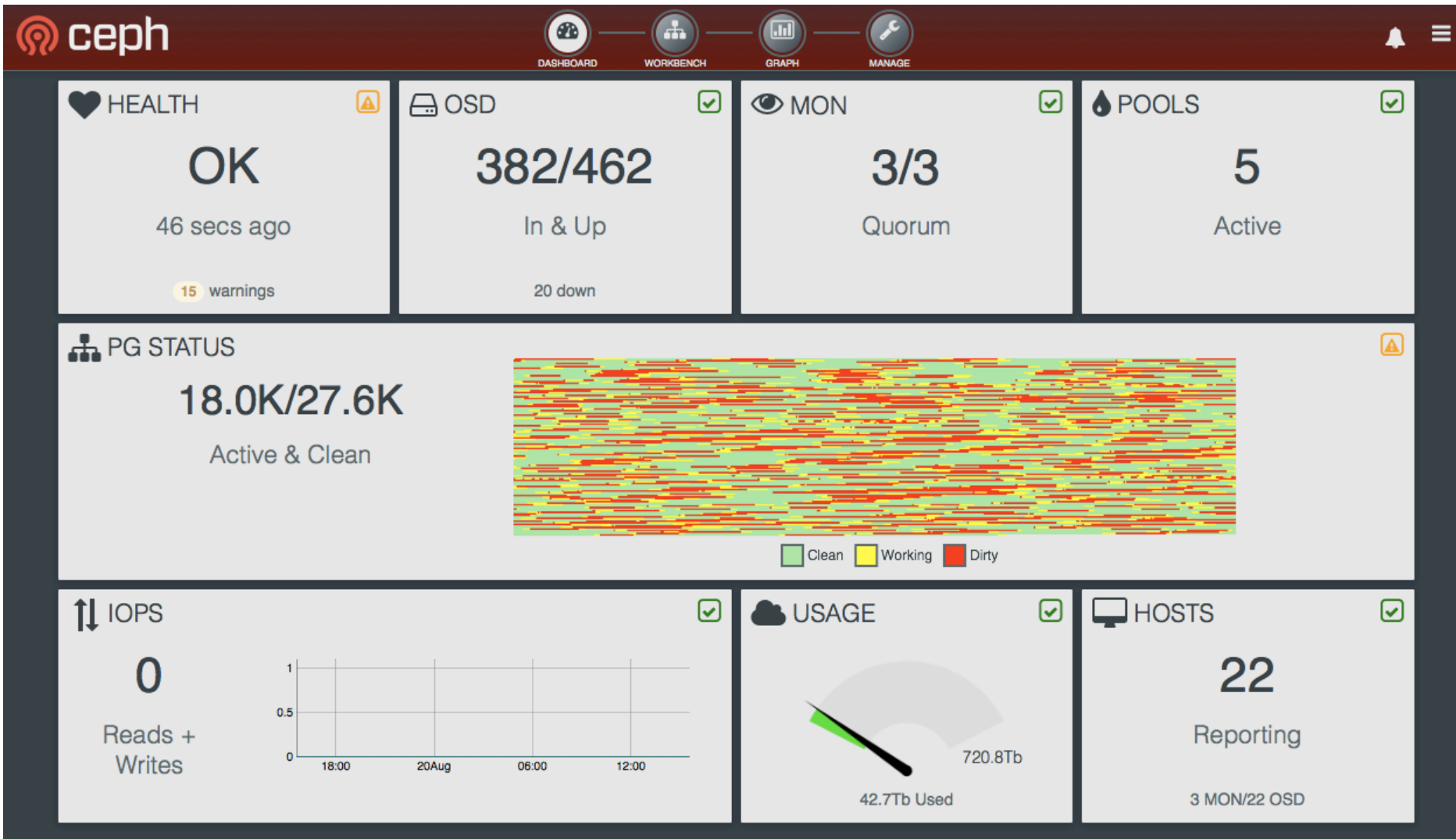
Rack136 - 160TB

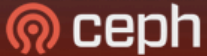
- **ISIS - Pulsed Neutron and Muon Source**
  - Taking data close to experiment
  - Analysing with SCD HPC resources
  - Bridge the gap with Ceph?
    - Tiered storage model?
    - Datacentre replication?
    - Federated clusters?
- **Exposing to other resources**
  - Re-exporting CephFS over NFS with Ganesha?





- Contacted team at KIT benchmarking Ceph
- RAL ERASMUS Student (Gabriela Fidyk)
  - Object Gateway
    - S3 + Kerberos + LDAP
  - Benchmarking
  - Tuning
  - Journal optimisations

- Web based management console
  - Dashboards!
  - Graphs!
  - Powerful buttons!
- Formerly Ceph Enterprise feature
  - RedHat purchase on InkTank opened it up for all
  - Installation still slightly arcane
- Screenshots
  - Show how badly FTS broke us






 HEALTH 

## OK


46 secs ago

15 warnings

 PG STATUS

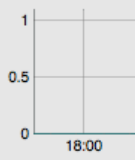
## 18.0K/27.6

Active & Clean

 IOPS

## 0

Reads + Writes







## Health Report

### SEVERITY DETAILS

- WARN 1 pgs backfilling
- WARN 1759 pgs degraded
- WARN 233 pgs down
- WARN 201 pgs incomplete
- WARN 5413 pgs peering
- WARN 66 pgs recovering
- WARN 5 pgs recovery\_wait
- WARN 678 pgs stale
- WARN 7156 pgs stuck inactive
- WARN 678 pgs stuck stale
- WARN 9078 pgs stuck unclean
- WARN 131 requests are blocked > 32 sec
- WARN recovery 389152/9576430 objects degraded (4.064%); 1060/3205750 unfound (0.033%)
- WARN mds cluster is degraded
- WARN nodown,noout,noscrub,nodeep-scrub flag(s) set

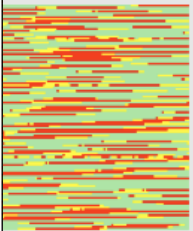
Close


 



 POOLS 

## 5

Active



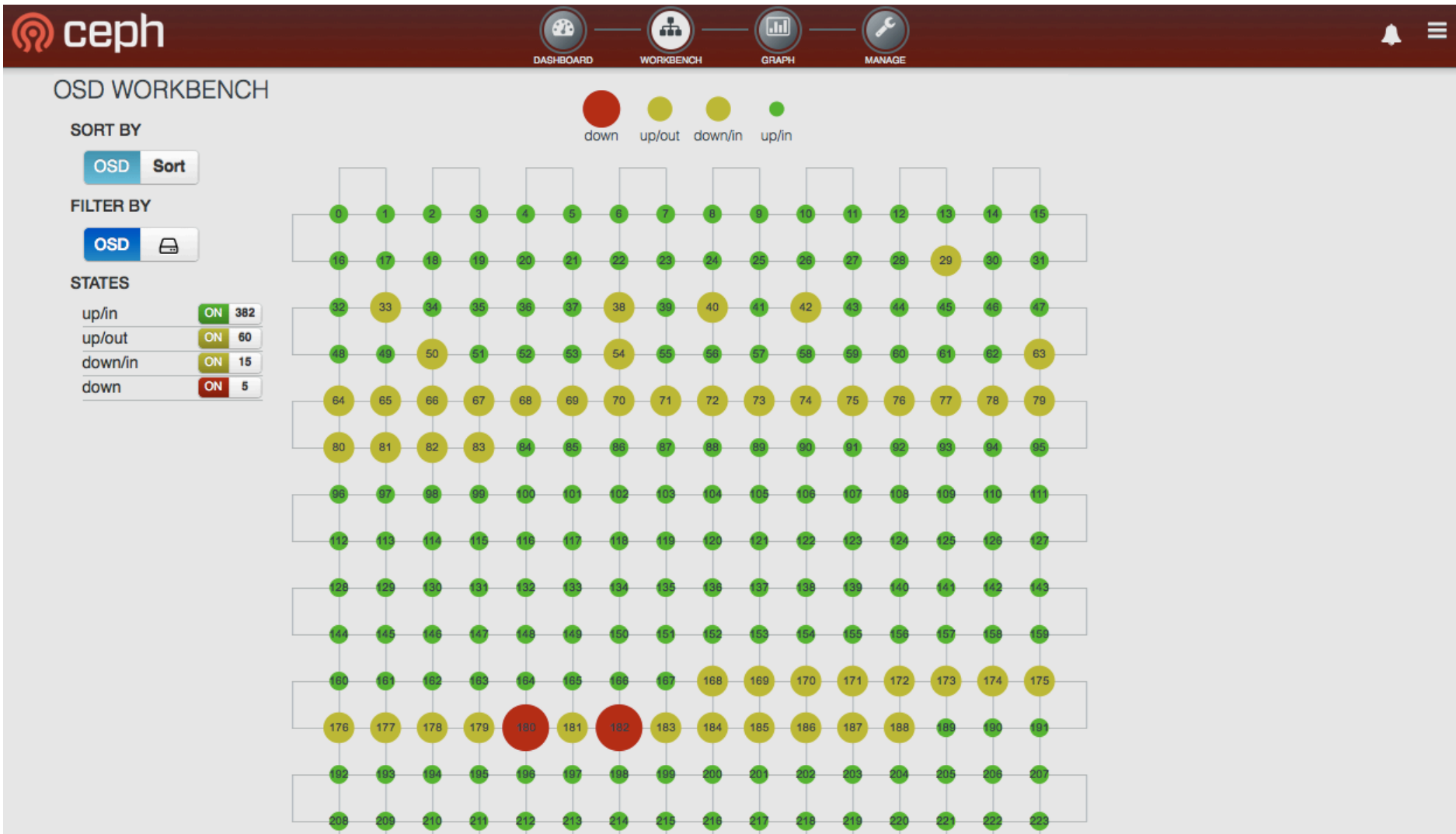


 HOSTS 

## 22

Reporting

3 MON/22 OSD



## CLUSTER SETTINGS

HOSTS CLUSTER SETTINGS CONFIG BROWSER

- No In
- No Out
- No Up
- No Down
- Pause
- No Scrub
- No Deep Scrub
- No Backfill
- No Recover

### Cluster Wide OSD Settings

Please be aware, settings on this screen affect the *entire* cluster. This *may* cause unexpected behavior, including data-loss, and prevent normal operation of service.

RESET

UPDATE



- RAM requirements per OSD double during recovery
  - Multiple OSDs fail in a node
  - Node runs out of RAM
  - Node leaves cluster (or starts flapping)
  - Failures cascade... admins start crying...
- Mons communicate with cluster on public network
  - VERY good idea to have dedicated systems (even VMs)
  - Sharing with OSD hosts can bring down cluster
  - Same badness as above
- Separate public and cluster networks
  - Recovery and replication traffic can interfere with cluster
  - Keeping this separate keeps admins sane

- Cloud cluster will enter production soon
  - RBD only at first
  - Pure RADOS and/or S3 later
- Extend testing of Grid cluster
  - Some re-architecting first
  - Working with trusted VO representatives
  - Trying different interfaces
- Start working on Ceph under CASTOR
- Tuning

## Questions?