Ceph Storage at RAL

Tom Byrne, Bruno Canning, Alastair Dewhurst,
George Ryall, George Vasilakakos

(Name.Surname@stfc.ac.uk)

Introduction

- A lot of interest in Ceph currently:
  - Talk on two most mature projects.
  - Some CephFS projects not discussed.
- RBD 'Cloud' cluster
  - Setup
  - Experience / Lessons learnt
- Object Store 'Grid' Cluster
  - Current setup and testing
  - Future plans

RBD Storage

Motivation

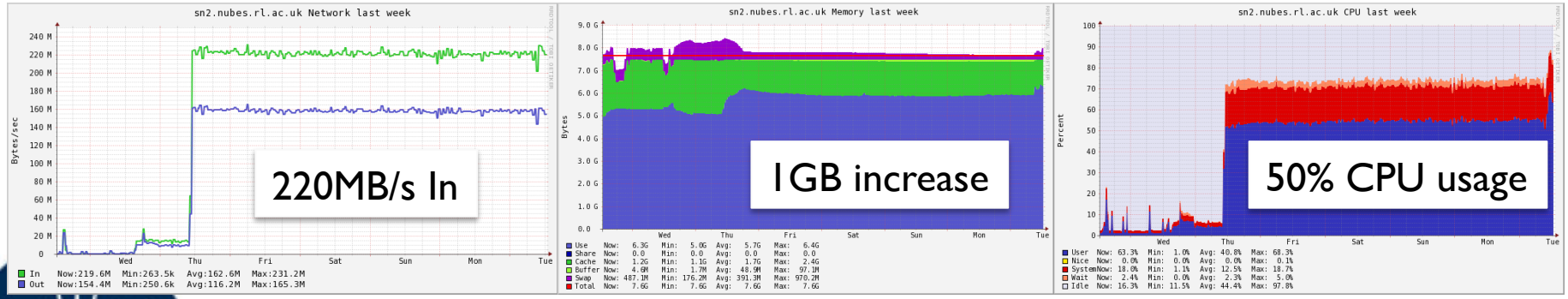
- Ceph RBD supporting Cloud infrastructure:
  - "Cloud @ RAL, an update", George Ryall, Thursday 14:00
  - "Ceph vs Local Disk For Virtual Machines", Alex Dibbo, Thursday 14:50
- Designed for low latency access.
- Using 3 replicas as small amount of space is required.
- Working since October 2014
  - Now being advertised internally as a pre-production service.

‘Cloud’ Cluster Setup

- 2 pairs of racks
  - 1 rack in each pair contains 14 Hypervisors while the other contains 15 storage nodes.
  - Each pair is connected to its own switch and these switches are inter-connected.
- 8 x 4TB drives in each Storage node
  - 1 drive for OS, 7 drives for storage.
  - 2 x 10GB/s links.
  - 8 GB memory (upgrade to 32GB on order).
  - 2 x Intel(R) Xeon(R) CPU E5-2403 v2 @ 1.80GHz

Performance

- Until recently, cloud cluster has been just one pair of racks.
  - Expect [almost] double performance now.
- Ran a ‘test’ where we had 50 VM randomly writing large amounts of data.
  - Rate we hit was 1044MB/s (8.2Gb/s).
  - 3 replicas means actual network activity was 24.6Gb/s.
- Limiting factor was storage node disks.



Ganglia metrics for one Cloud storage node

Upgrade to Giant

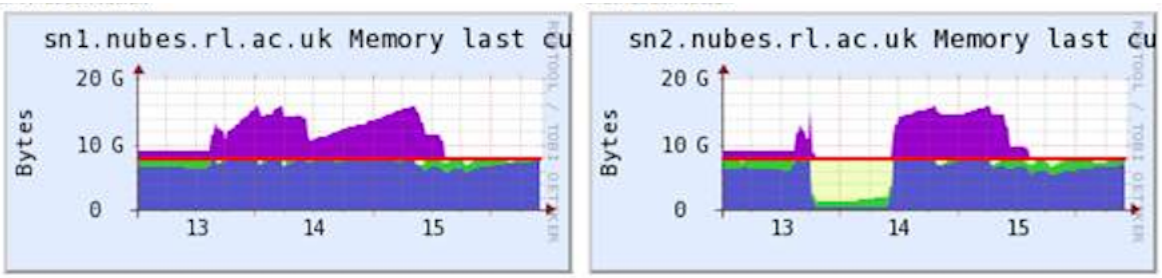
- In mid January we upgraded from Firefly to Giant.
- Clusters kept running (although were not loaded).
- Amazingly painless:
  - Clusters took between 30 mins and 3 hours.
- Cluster came back in Health Warn status because of 'requests are blocked > 32 sec' errors
  - Most problems were solved by restarting the OSD
  - A few OSD failed and were re-installed. Could be a problem as cluster grows?!
- For future upgrades will probably set 'no-out' and 'no-down' flags

Don't lose your Monitors!

- Early setup used VM monitors.
  - They were all using the same disk array…
  - Disk got corrupted…
  - Had to re-create cluster from scratch…
- Both clusters have 3 physical monitors
  - Not ruled out using VMs in future.
- Looking at ways of ensuring data reliability:
  - 5 monitors?
  - Monitors in different machines rooms?

Don't scrimp on memory!

- Cloud cluster nodes have 8GB memory.
  - Upgrade to 32GB on order
  - Grid Cluster nodes has 64GB.
- Cluster performance related to slowest node:
  - Memory failures have caused significant slow down in entire cluster.
- Rebuilding/balancing the cluster is what stresses Ceph.
- Also observed long term memory usage rise:
  - Restarting OSDs occasionally. Believe this will be fixed.



Rebalance started on 13th tried to put nodes back on 14th

Object Storage

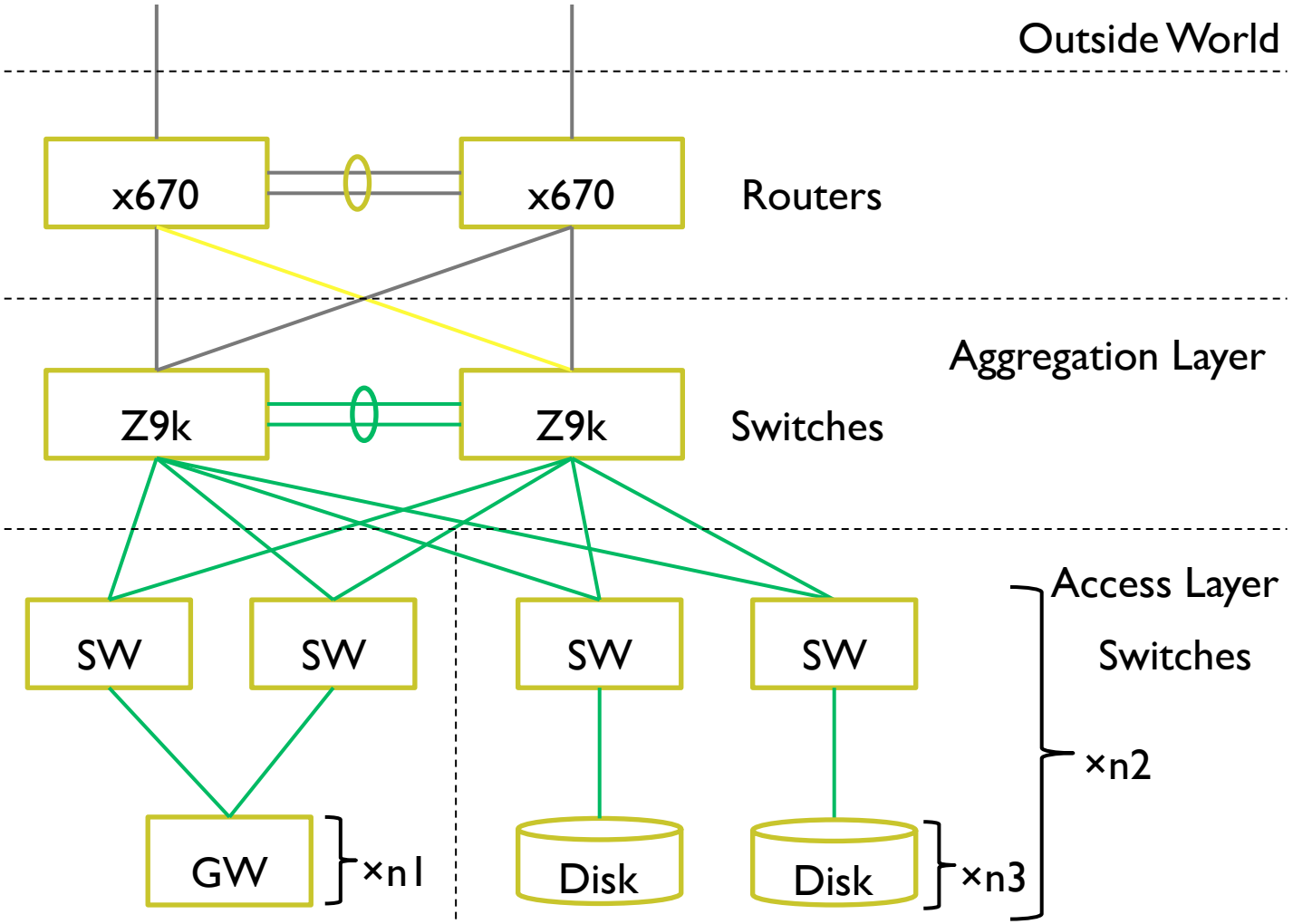
Motivation

- RAL Tier 1 provides ~10PB disk storage to WLCG VOs.
- RAL is the only site using Castor for disk only storage:
  - Difficult to operate, no future.
  - SRM is ~~crap~~ of limited use to other communities.
- Investigating Ceph as a large scale object store.
- Aim to provide thinnest layer possible on top of Ceph:
  - Xrootd, GridFTP and [in future] http protocols are required for LHC VOs to work.
  - The same data needs to be accessed through different protocols.

‘Grid’ Cluster Setup

- Grid Cluster working with old hardware:
  - 3 VM Monitors
  - 27 disk servers from 2009. Single 1GB/s links and only running limited number of OSD to match available CPU and memory.
- New hardware has been delivered and is undergoing acceptance testing:
  - 3 Physical Monitors (Dell R430 + SSD for levelDB)
  - 3 Gateways (Dell R430 + 2 x 10GB/s network links)
  - 21 x 120TB and 26 x 100TB storage nodes
    - 64GB memory, 2 x 6 x 2.4GHz CPU, 2 x 10GB/s network cards.
    - RAID Card – Allows nodes to be used in Castor.
    - Single SSD – Purchased before we understood journaling.

Network


Outside World
x670
x670
Routers
Aggregation Layer
Z9k
Z9k
Switches
SW
SW
SW
SW
Access Layer
Switches
×n2
GW
×n1
Disk
Disk
×n3

RADOS Gateway

- Our Grid Cluster has a RADOS Gateway supporting S3/ SWIFT.
  - Encouraging users to try it as extremely easy to support.
  - ATLAS Event Service.
  - Will take part in FTS3 S3 testing.
- RADOS Gateway creates several pools.
  - 'Data' and 'Logs' are the only ones that need a lot of placement groups.
  - Can backup all the others pools to keep user access credentials safe.

Xrootd / GridFTP

- Sebastien Ponce (CERN), has contributed the LibRadosStriper to Ceph mainline.
  - Working versions are in Hammer!
- He has also written xrootd and GridFTP plugins on top of this:
  - Xrootd will be in xrootd 4.2 and is being actively developed by CERN and xrootd developers.
  - RAL will continue to develop the GridFTP plugin.
- Created ceph-talk@cern.ch to discuss HEP specific development work.

Monitoring

- We have started to integrate our Ceph instances with our existing Nagios and Ganglia monitoring.
- Using Nagios plugins developed by Ceph community.
  - Only experience will tell us if genuinely useful.
- Ganglia monitoring of individual nodes.
  - Adding 'ceph status'
- Not yet looked at Calamari.

Host	Service	Status	Last Check	Duration	Attempt	
gceph-mon-1	Check CEPH Health	OK	15:50:04	3d 5h 53m 53s	1/3	HEALTH OK
	Check CEPH MON	OK	15:50:04	3d 5h 53m 53s	1/3	MON OK
gceph-mon-2	Check CEPH Health	OK	16:04:04	3d 5h 6m 57s	1/3	HEALTH OK
	Check CEPH MON	OK	15:46:04	3d 5h 21m 55s	1/3	MON OK
gceph-mon-3	Check CEPH Health	OK	16:04:04	3d 5h 6m 57s	1/3	HEALTH OK
	Check CEPH MON	OK	16:04:04	3d 5h 6m 57s	1/3	MON OK
gdss489	Check CEPH OSD	OK	15:46:04	3d 7h 2m 23s	1/3	OSD OK
gdss490	Check CEPH OSD	OK	15:44:05	3d 7h 8m 11s	1/3	OSD OK

Dashboard


Ceph Cluster Overall Status
Cluster Status:HEALTH_OK
Ceph Cluster Monitor Status
Monitor GCEPH-MON-1: HEALTH_OK
Monitor GCEPH-MON-2: HEALTH_OK
Monitor GCEPH-MON-3: HEALTH_OK
Ceph Cluster OSD Status
Total
204
In
204
Up
204
Unhealthy
0
Ceph Cluster Placement Group Status
Storage
0 %
20 %
40 %
60 %
80 %
100 %
Cluster storage utilization
2.13 TB / 370.36 TB (1%)
Write / second
0 bytes
Read / second
0 bytes
Operations / second
0
PG Status
23296 placementgroups in cluster

Erasure Coding

- How are we planning on storing data?
  - 3 replicas is too expensive
  - Have to use erasure coding (EC)
- EC breaks data into ‘k’ chunks and creates ‘m’ parity chunks.
  - Can lose any ‘m’ OSDs without losing data.
- Use case is such that EC should work well.
  - LHC VOs write objects once and read them a few times.
- EC does not support partial writes.
  - Might need to use Cache Tier

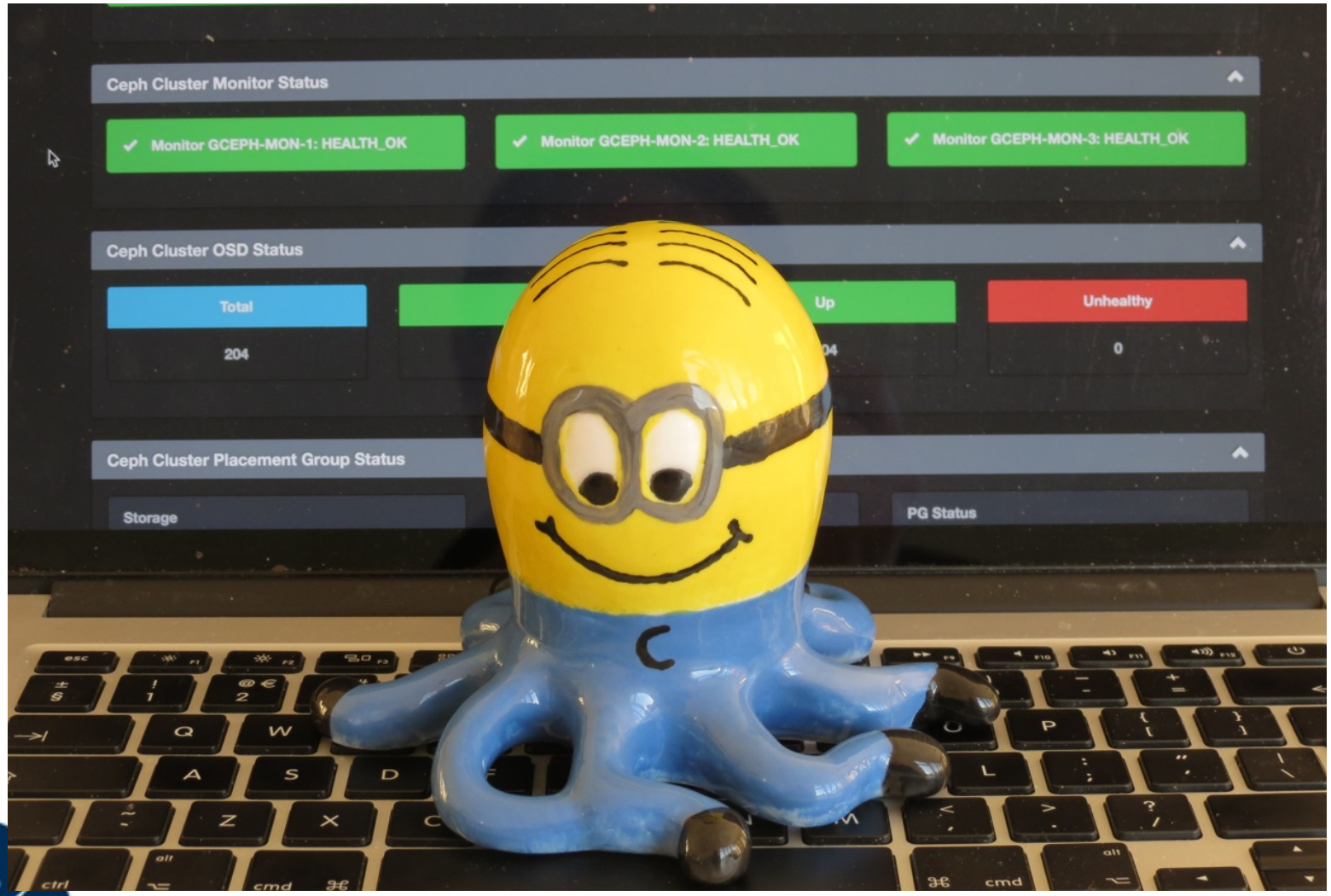
Erasure Coding(2)

- To be cost effective we need <30% overhead.
- ‘m’ = 2
  - EC algorithms fast as original data unchanged.
  - Heavily optimized as equivalent to RAID6.
- ‘k’ = 16? 8? Something else?
  - The greater k is the more load there will be during a rebuild.
- First thing to test with new hardware!
- Note: libRadosStriper and EC both ‘chunk’ the file.
  - Object size = [stripe size] / k

Summary

- Pre-production RBD Ceph Cluster.
- Active development on large scale object store.
- Greatly benefitted from experience of others.
  - Hope we have shared some of ours

Questions?


Ceph Cluster Monitor Status
Monitor GCEPH-MON-1: HEALTH_OK
Monitor GCEPH-MON-2: HEALTH_OK
Monitor GCEPH-MON-3: HEALTH_OK
Ceph Cluster OSD Status
Total
204
Up
Unhealthy
0
Ceph Cluster Placement Group Status
Storage
PG Status