

DOMA General Meeting

QoS Update

Oliver Keeble on behalf of the group
22/05/2019

QoS Site Survey

- Site Survey has been published
 - Intention is to identify current QoS-oriented activity and document solutions under evaluation
 - Sent out on 14th May.
 - Merged with the planned WLCG Ops survey
 - Now two “chapters”, storage and compute
- The WG will compile the results
 - Likely to lead to identification of a small number of themes around which future efforts can be focused.

QoS: Survey Overview

- <https://forms.gle/mhWPrDfq8n2bDGES9>
- Deadline 15th June.
- Describe your current system
 - Media, configurations, storage system, effort
- Users and use-cases
 - Constraints from use-cases, Caching systems, non-WLCG users
- Future
 - What have you tried?
 - Did it work?
 - What would you like to try?
 - What cost-saving opportunities do you see?
 - How could the VOs help you save?
 - e.g. read behaviour, job scheduling, deletion, reliability/availability expectations...

CERN QoS Directions (survey response preview)

- Hardware
 - Diverse types of drives
 - In particular, consumer and SMR (shingled) have been tried
 - Inconclusive cost/benefit figures
 - Different reliability/complexity/performance characteristics
 - We have significant bulk discounts on PMR (“normal”) which are difficult to get for the others
 - Overhead reductions
 - New systems have 192 12TB disks per node
- Erasure Coding
 - Demonstrated in the alice daq instance of EOS
 - RS(12,10) : the encoding of 2.4 (4.8 raw) PB freed 2 PB of storage in the pool
- Ceph
 - Smaller instances with various QoS
 - SSD, SSD/HDD, HDD, Erasure Coding, 3 replicas, Regions, mirroring...
- We can report more fully on any of the above if desired.

Familiar QoS concepts

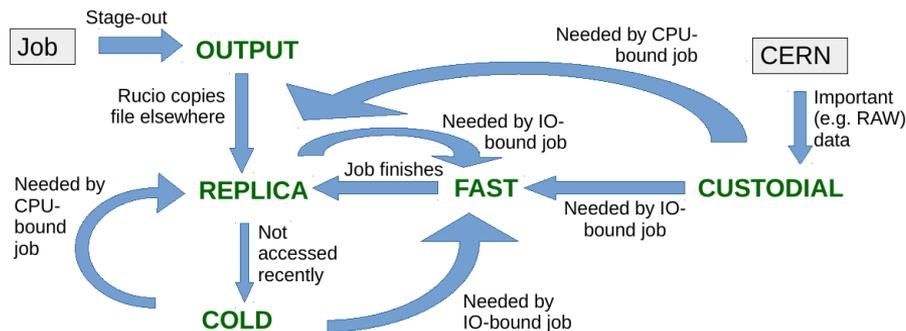
- Disk
 - Huge QoS variations possible under this category
 - All relevant workflows mapped onto this
 - For a particular workflow, can be overspecified in some ways (e.g. reliability) and underspecified in others (e.g. concurrent clients)

- Tape
 - Covers both “durable” and “low-cost”

- These terms should become: “Disk”, “Tape”

- Example additional storage QoS possibilities:

- Enterprise HDD as RAID: OUTPUT
- Tape-backed disk cache : CUSTODIAL
- Consumer HDD as JBOD: REPLICAS
- (public) cloud storage: COLD
- SSD as JBOD: FAST
- Internal replicas existing on multiple server nodes: FAST



Rucio discussions

- There has been a dedicated discussion on QoS in Rucio
 - How should Rucio interact with storage?
 - What interface/abstraction is seen at the experiment level?
 - Storage level
 - Latencies, bandwidths, availabilities
 - Creates a large combinatorial phase-space
 - Experiment level
 - A small number of meaningful classes
 - Many questions...
 - Where is cost exposed?
 - How are the classes defined?
 - How is fulfillment of a QoS level validated?
 - Are classes shared between experiments
 - Would be necessary if this reaches “pledge” level
- How will we make progress?
 - These issues will be summarised in the upcoming White Paper
 - This will form the basis for dedicated community discussion on the most important points



Quality of Service

Objectives

1. Find the matching RSE based on the requested QoS of the rule
 - E.g. put 1 copy of file.001 on an RSE in country=uk with qos archive
2. Manage QoS over time
 - 1 copy of file.001 on an RSE in country=uk with qos latency<50 for 30 d after that change qos to latency<1000 for ∞ d
 - Existing rule concept very well suited to express data placement over time

CHEP Abstract

- Title: Quality of Service (QoS) for cost-effective storage and improved performance
- Track: 4: Data Organisation, Management and Access
- Authors: A. Paul Millar, Martin Barisits, Mario Lassnig, Xavier Espinal, Oliver Keeble, Edward Karavakis, Patrick Fuhrmann, Markus Schulz
- Abstract:

The anticipated increase in storage requirements for the forthcoming HL-LHC data rates is not matched by a corresponding increase in budget. This results in a short-fall in available resources if the computing models remain unchanged. Therefore, effort is being invested in looking for new and innovative ways to optimise the current infrastructure, so minimising the impact of this shortfall.

In this paper, we describe an R&D effort targeting "Quality of Service" (QoS), as a working group within the WLCG DOMA activity. The QoS approach aims to reduce the impact of the shortfalls, and involves developing a mechanism that both allows sites to reduce the cost of their storage hardware, with a corresponding increase in storage capacity, while also supporting innovative deployments with radically reduced cost or improved performance.

We describe the strategy this group is developing to support these innovations, along with the current status and plans for the future.