Storage Technologies

… or Managing Expectations

Paul Millar

GridKa School 2015 (2015-09-07)
Storage software: Free, Open-Source

https://github.com/dCache/dCache

mailto:dev-subscribe@dcache.org
Software that scales up to tens of PiB
… and down to a single Raspberry Pi
Software running throughout the world
Used to search for the Higgs boson
Feed data for HPC applications

HPC jobs on supercomputer

HPC jobs get access to dCache storage.
Research: pushing frontiers

Power supply

HGST Disk

Clip

Yves
dCache and INDIGO-DataCloud

THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION

HORIZON 2020
Developing two new features

Quality of Service

and

Data Life-Cycle
Quality of Service
Store data on disk or tape?
Now we have more media options
Replicating data

How many copies? Where are they located?
Motivation: budgets
How to make this a possibility
Bridging the gap

Concepts that users easily understand

Concepts that storage systems easily understand
Attributes and islands
Combining QoS attributes

Discrete  Independent  Continuous

“Fee selection of QoS”

Discrete  Dependent

“Islands of QoS”
Figure-of-merit: how users choose
Data Life-Cycle
DLC use-cases: the story of a file

- **Created**
  - Change QoS

- **Accept/Reject Deadline**
  - Change QoS

- **Main analysis complete**
  - Allow public access

- **Public embargo ends**
  - Anticipated end of interest
  - Change QoS

- **End of life**
  - Delete data

Time (not to scale)
Format for DLC rules

<trigger> <action>

(e.g., <after 6 months> <add public-access ACE>)
The plan

- **RDA**: Define abstract terms
- **SNIA**: Define network protocol
- **INDIGO - DataCloud**:
Thanks for listening