dCache.org

dCache and Storage Evolution

or

An update from the workshop

Paul Millar
on behalf of the dCache team.

GDB 2015-06-10

Fermilab NDGF Nordic DataGrid Facility DESY HELMHOLTZ ASSOCIATION LSDMA

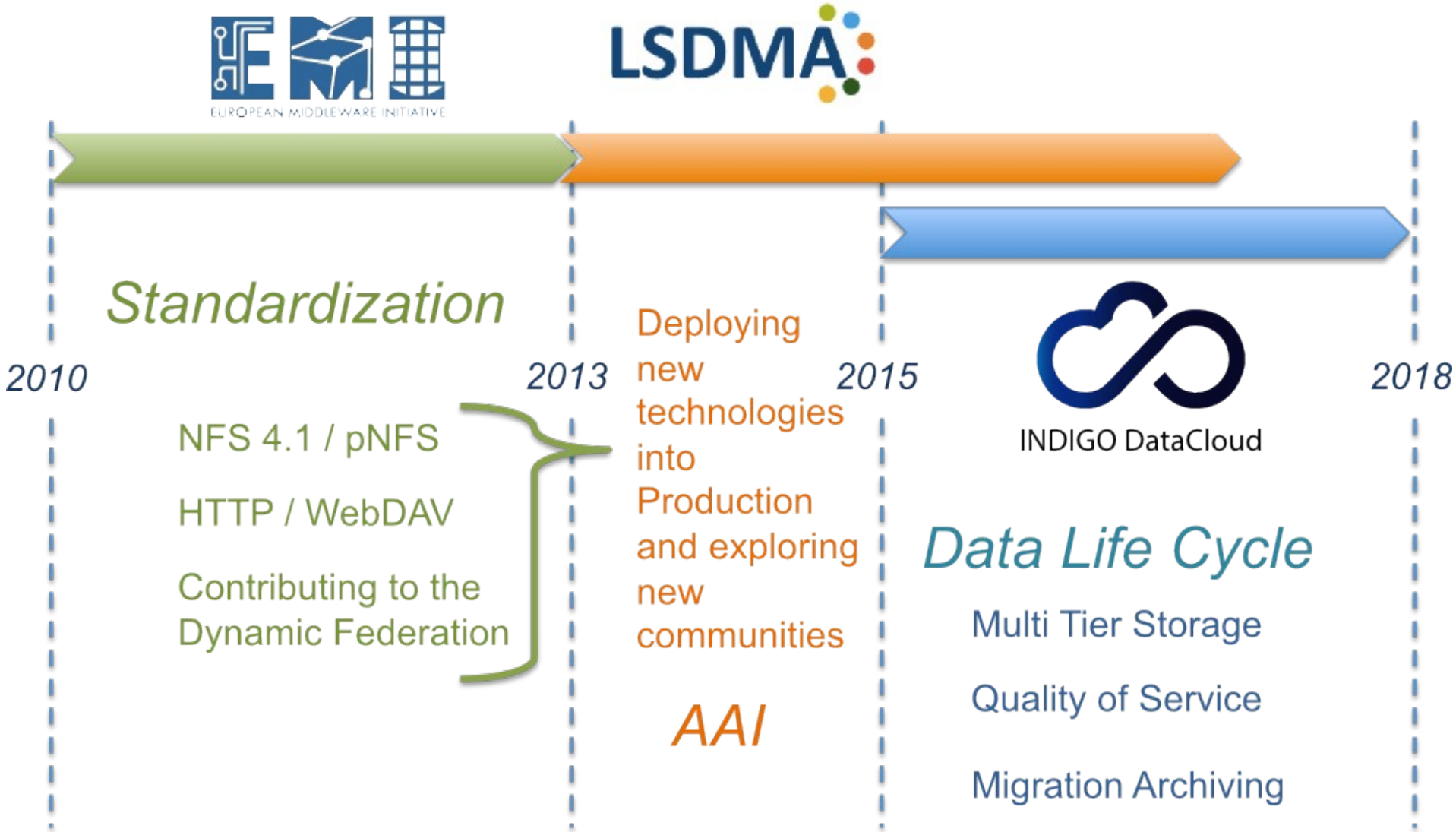
9th International dCache workshop

35 registrants from 13 countries world-wide.

Workshop covered 2 ½ days

Contributions from: **development team**, our collaborators at **CERN**, dCache **admins**, storage hardware **companies**. Several interactive sessions: hands-on tutorials, fish-bowl, genius bar.

Funding


EUROPEAN MIDDLEWARE INITIATIVE
LSDMA
2010
2013
2015
2018
Standardization
NFS 4.1 / pNFS
HTTP / WebDAV
Contributing to the Dynamic Federation
Deploying new technologies into Production and exploring new communities
AAI
INDIGO DataCloud
Data Life Cycle
Multi Tier Storage
Quality of Service
Migration Archiving

Indico DataCloud: Cheat Sheet

- H2020 project, started **April 2015** for **30 months.**
- INFN-Italy project lead; DESY **WP4-lead** & involved with **WP5**.
- Budget of 11.1 M€ (**0.8 M€ for dCache**).
- 26 partners from 11 European countries
- The project aims for an Open-Source Data and Computing platform targeted at scientific communities, deployable on multiple hardware, and provisioned over private and public e-infrastructures.

dCache focus is **Data Quality of Service** and **data-lifecycle management**.

SRM

- Very much **supported** by dCache team:

  e.g., recent GFAL/FTS bug

- ***If*** SRM goes away, **functionality remains**:
  - SRM is an interface, not the functionality itself
  - Will provide equivalent, RESTful interface (e.g., CDMI), through Indigo DataCloud.
  - **Take input** from experiments in shaping such an interface.

WLCG HTTP deployment WG

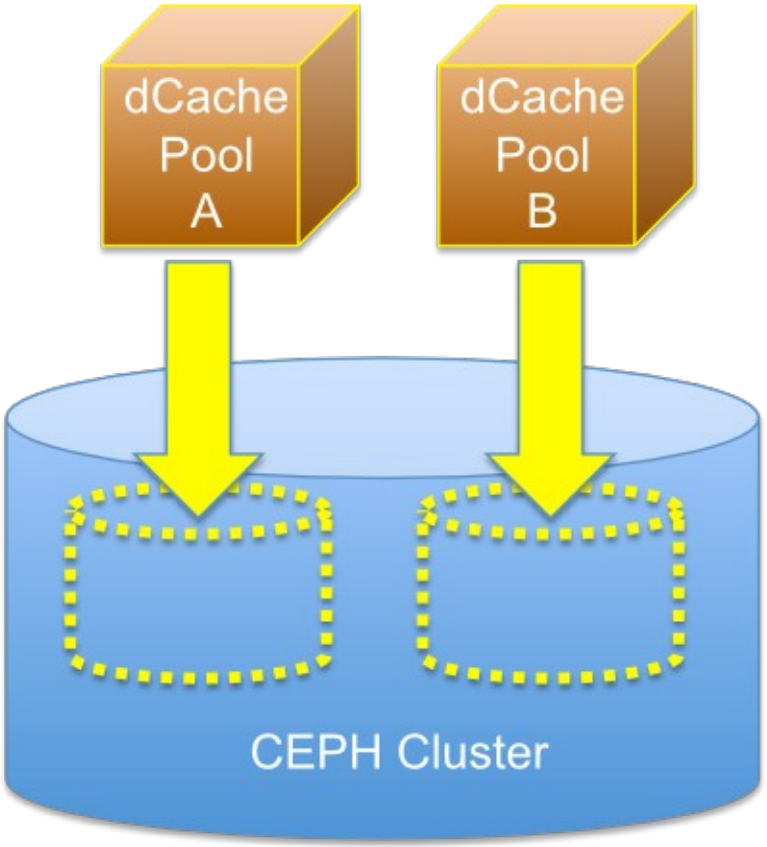
- dCache **active participant**, involved from beginning.
- Oliver is doing a **great job**! Strong agreement with the direction of WG
- dCache **supports** the HTTP dynamic federation

  … not in coding, but deployment and finding user-cases.
- Plan to use within Indigo DataCloud:

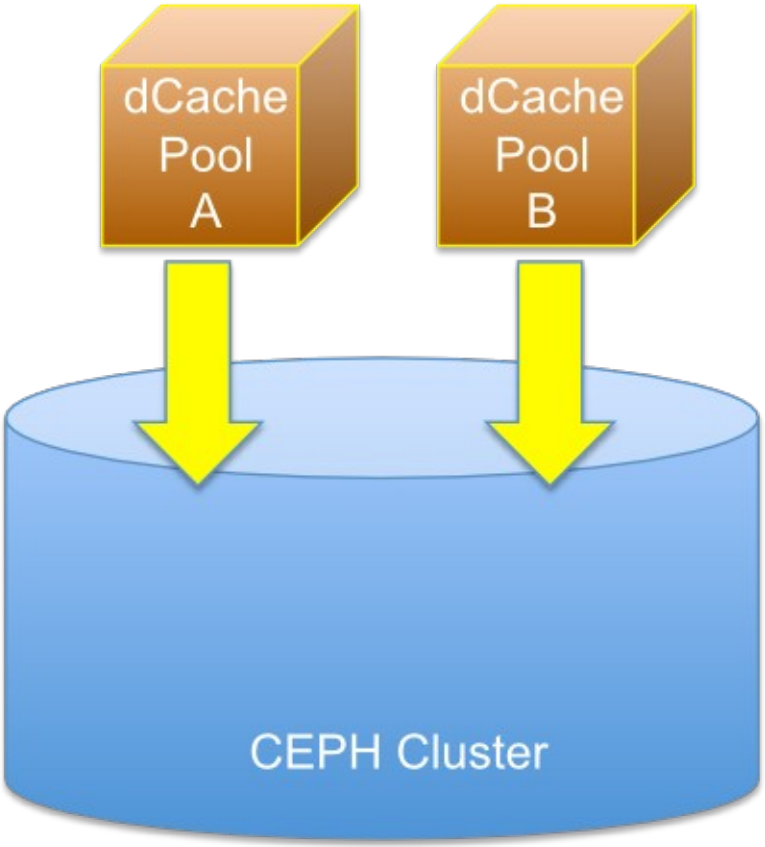
  Deploying FTS & WebFTS at DESY so we can further investigate

CEPH integration

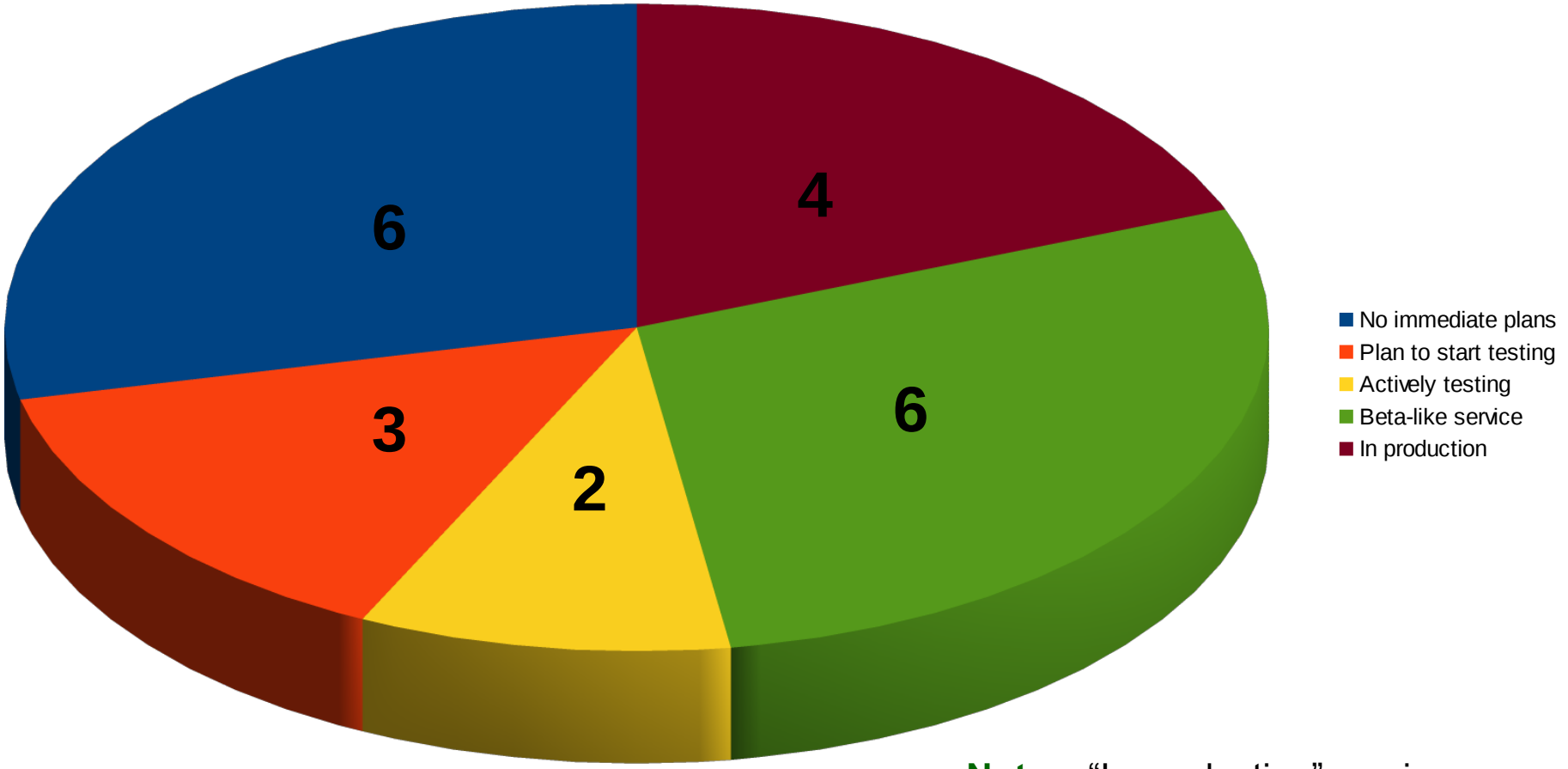
Each dCache pool still only 'sees' his own private repository.



dCache pools can use shared repositories. Requires new pool semantics. (Focus on protocol engine)



NFS v4.1/pNFS usage by dCache sites

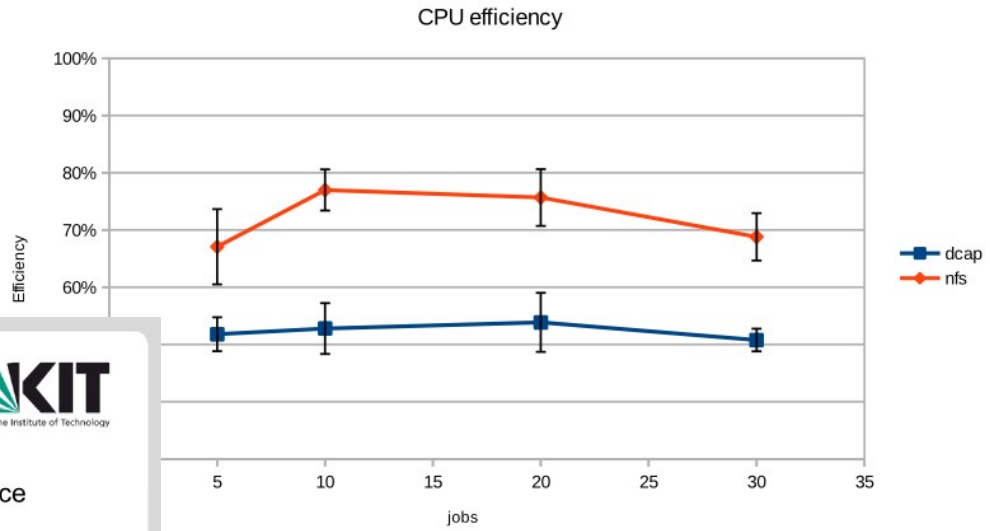


Notes: “In production” services currently do not include direct worker-node access.
Data source: dCache user-forum survey 2015-06-08

KIT evaluation of NFS v4.1/pNFS

Slides courtesy of **Preslav Konstantinov**, (SCC, KIT) presentation at GridKa TAB meeting 2015-05-06

CPU efficiency

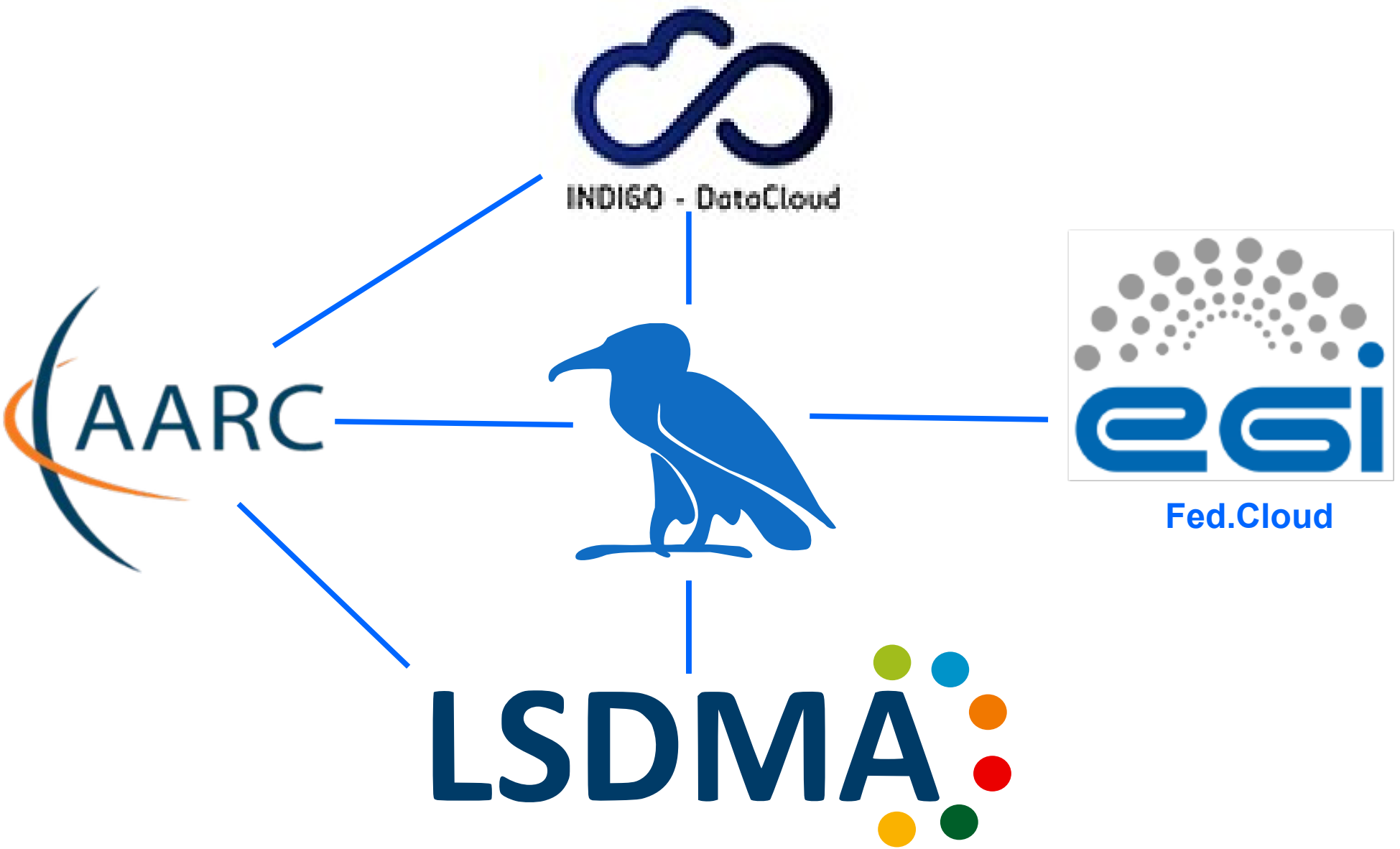


Eff = (User + System) / Wall
Roughly 35% better with NFS

Conclusions

- In addition to convenience of usage NFSv4.1 in dCache shows some performance improvements over dcap.
- I think it is worth pursuing further testing with production workflows and growing workload in order to achieve stability.

Federated AAI



Other highlights from the workshop

Future directions in storage tech.

Manfred Berger, **HGST**

Embedding dCache in DDN storage controllers

Simon Liu, **TRIUMF**

dCache and ELK

Johan Guldmyr, **CSC-IT**

The HTTP ecosystem

Fabrizio Furano, **CERN**

… plus many other presentations

Thanks for listening … any questions?

Backup slides