dCache.org

dCache introduction

Paul Millar

on behalf of the dCache team.

EU dCache developer meeting

Fermilab NDGF Nordic DataGrid Facility DESY Helmholtz Association LSDMA

High-level Overview

dCache is...

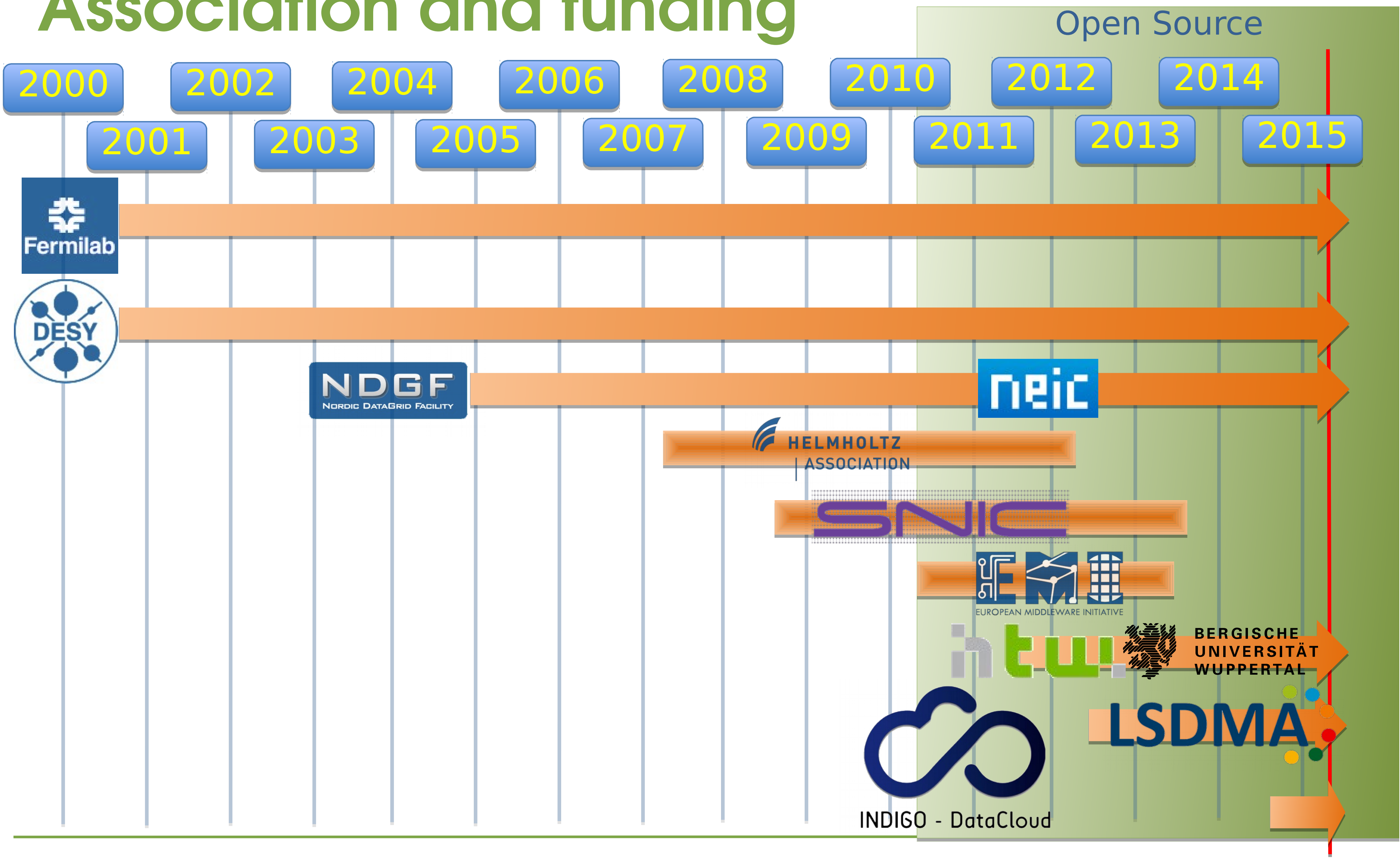
software for providing scalable, managed storage for huge amounts of data.

deployed at research institutes throughout the world and used by a diverse collection of user-communities.

supported through the dCache.org collaboration, which provides:

- regular feature releases that are maintained with subsequent bug-fix releases.
- Support and advice through a variety of channels.

Association and funding



dCache history

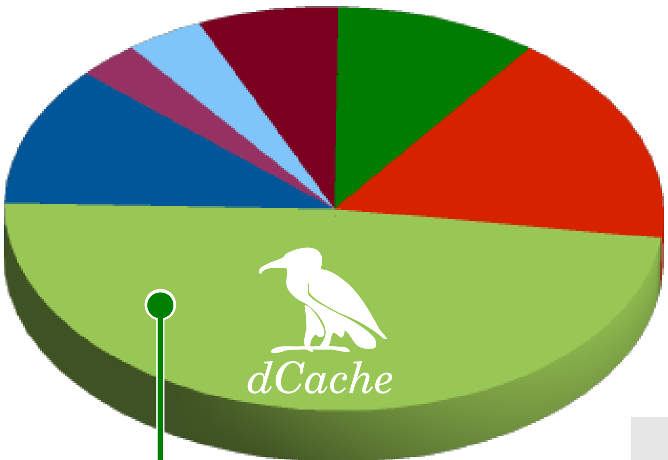
Era	Disk cache	Grid Storage	Generic Storage	Cloud Storage
Additional Communities	hermes, ZEUS, CDF, H1, HERA B	ALICE, ATLAS EXPERIMENT, CMS, IceCube, LHCb	Fermilab Intensity Frontier, European XFEL, Belle II, LOFAR, SNIC, Industry, CFEL SCIENCE, PETRA III	egi, LSDMA
Additional Authen-tication	Trusted host	X.509, Kerberos	Username+PW	SAML, OpenID, OAuth, Token, ...

dCache.org

open source

What is dCache today?

LHC data stored on each storage system



Source: BDII (2014-11-14)

Core team

DESY
8 FTEs

Fermilab
2 FTEs

neic
1 FTEs

dCache

Collaborations

egi

globus online

PHYSICS AT THE TERA SCALE
Helmholtz Alliance

LSDMA

EMI EUROPEAN MIDDLEWARE INITIATIVE

OpenGridForum

SNIC

Open Science Grid

Student mentor programme

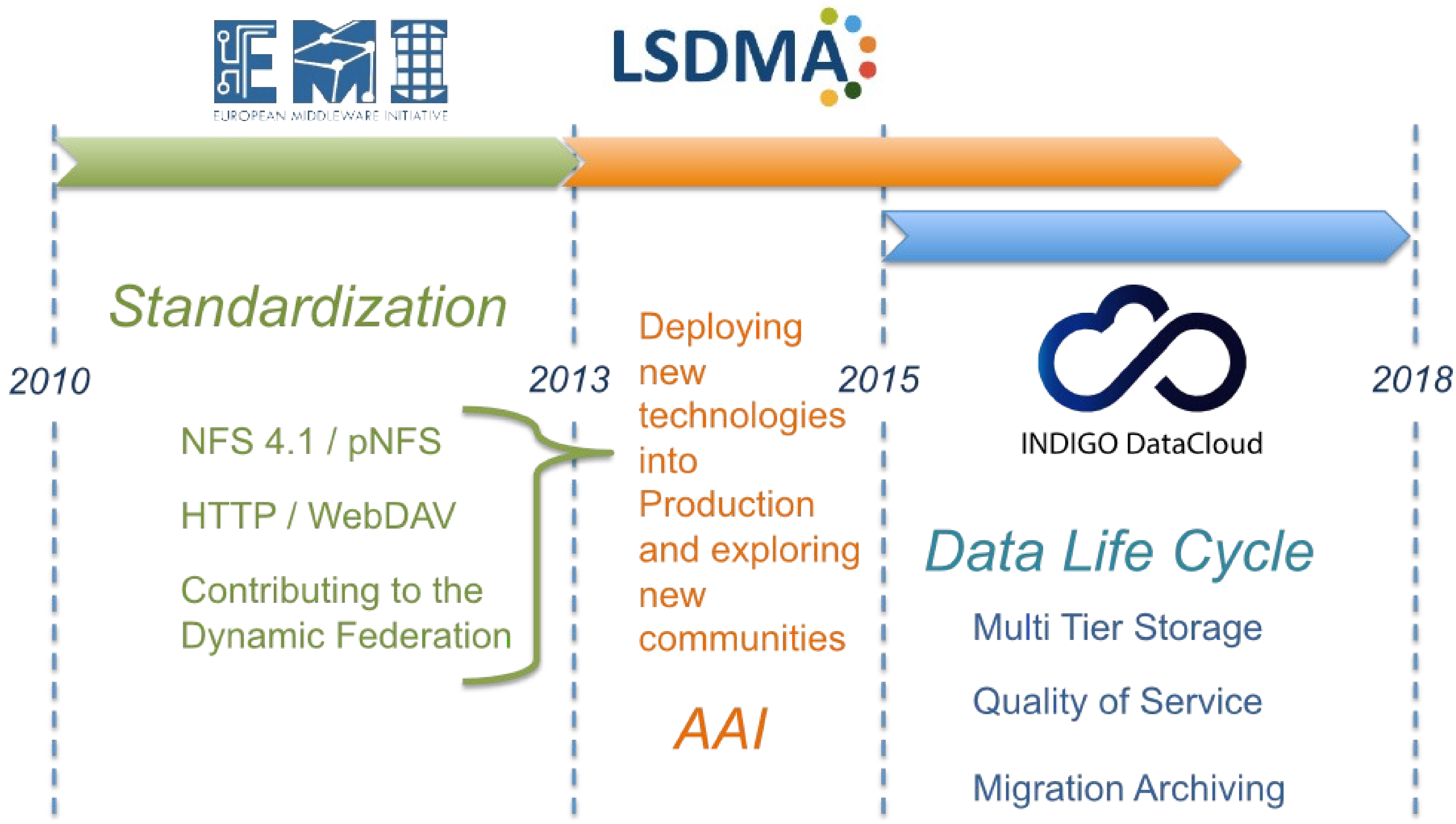
htw

Hochschule für Technik und Wirtschaft Berlin

BERGISCHE UNIVERSITÄT WUPPERTAL

3 students

Current and future project funding



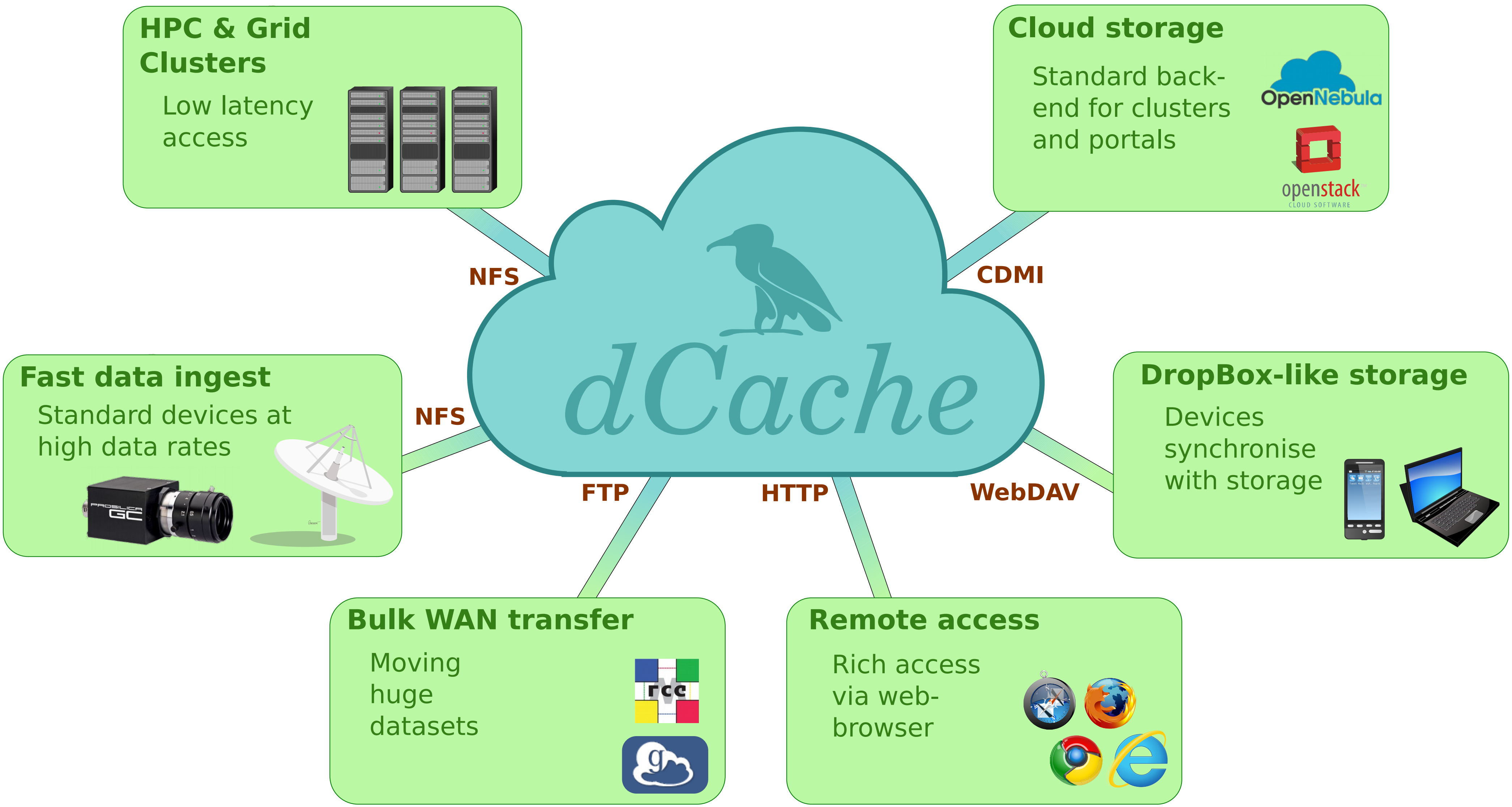
dCache key features include...

- Users see a single POSIX filesystem (hard- & soft-links, etc),
- Transparent support for tertiary (tape) storage,
- Scalable bandwidth,
- Steerable target when reading and writing,
- Space management,
- Resilience to storage node failure,
- Supports transparent storage device life-cycle,
- Hot-spot detection and mitigation,
- Differentiable quality of service,
- Pluggable authentication,

...

dCache.org

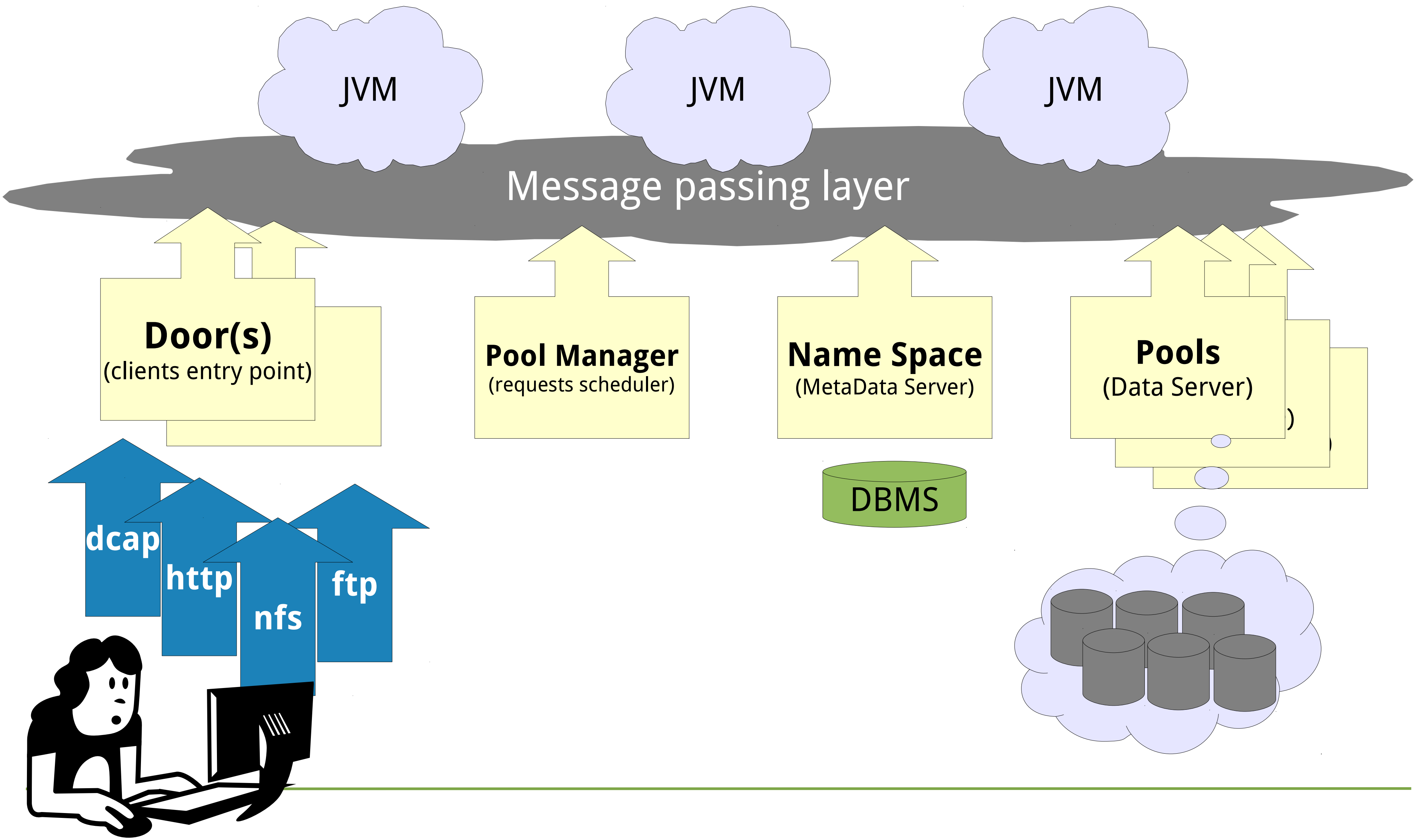
The scientific cloud vision



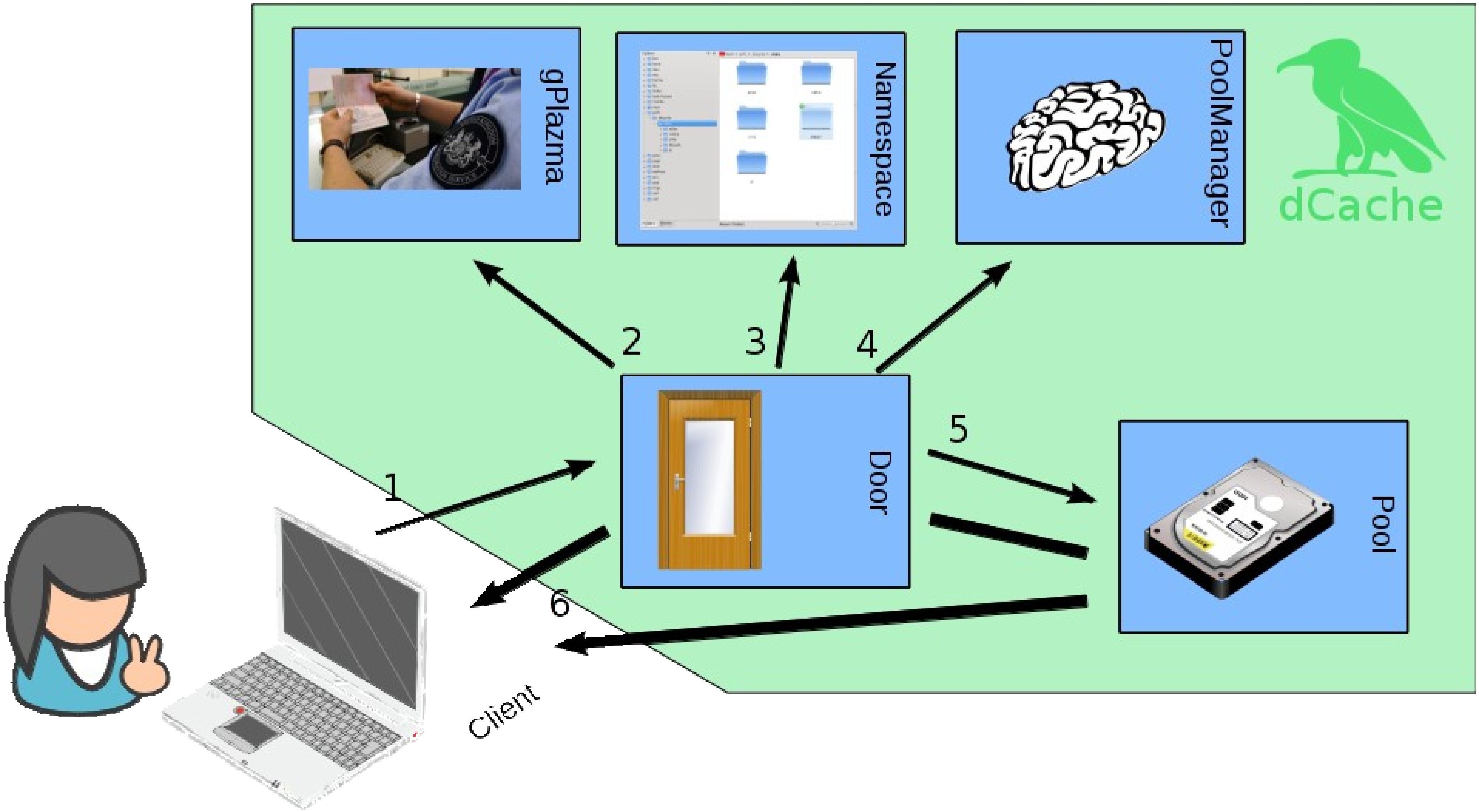
Some details* on how dCache operates...

* Some details are deliberately omitted to keep slides manageable.

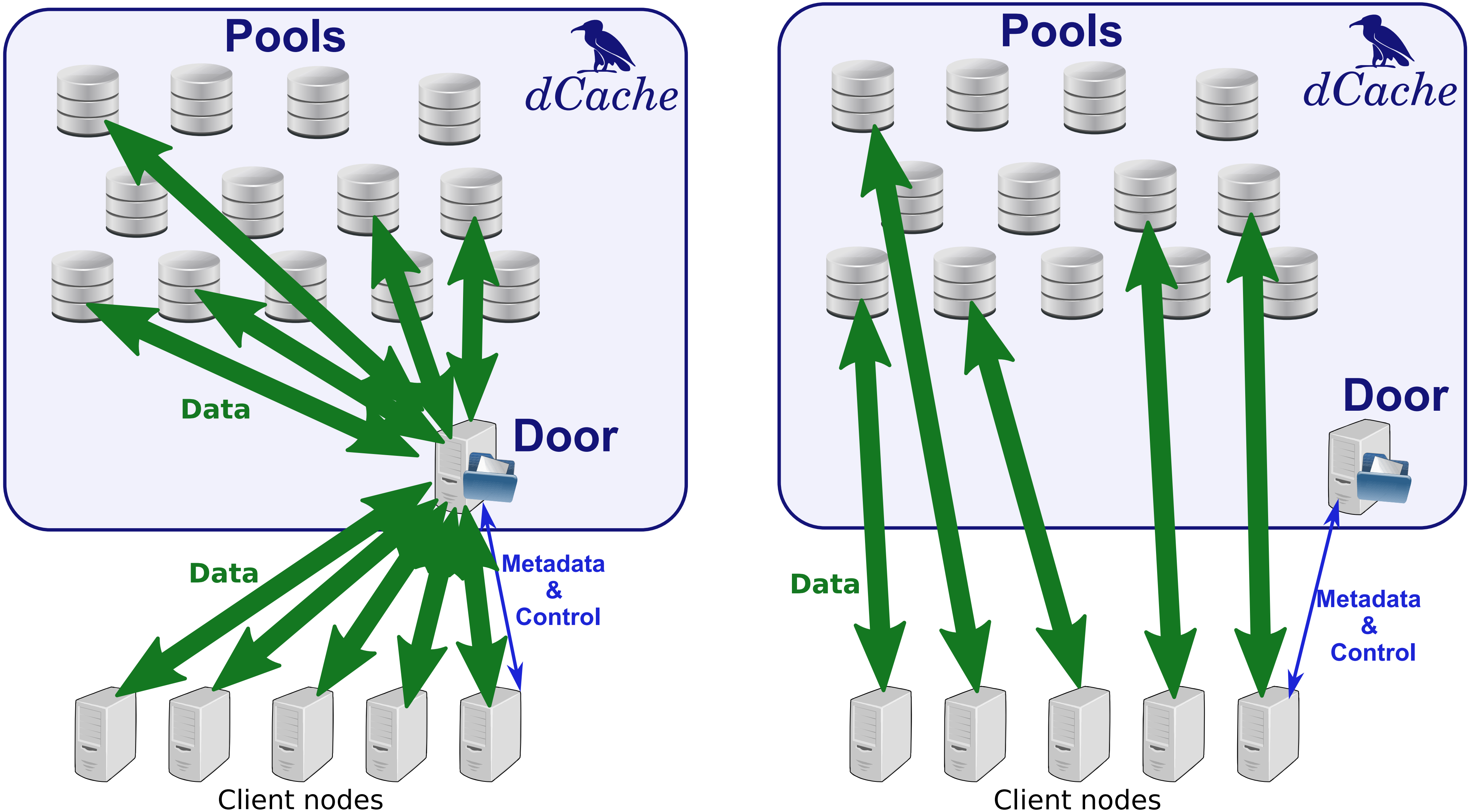
dCache – under the hood



Core components when transferring



Importance of redirection



Operational experience

Storage at DESY

- 6 dCache instances: Hera, CMS, ATLAS, Photon, "DESY" and Cloud:

  Hera is officially switched off,

  CMS & **ATLAS** for WLCG: CERN LHC experiments,

  Photon is for various photon user-communities,

  Cloud is for sync-and-share service,

  DESY is for the rest.

Comparative numbers

CMS	ATLAS	Photon	DESY	Cloud
~5x10^6 files	~1x10^7 files	~8x10^7 files *	~1x10^7 files	~2x10^6 files
~3 PiB	~3 PiB	~2.5 PiB *	~3 PiB	~10 TiB
~300 pool-nodes	~300 pool-nodes	~30 pool-nodes	~30 pool-nodes	~6 pool-nodes
~580 GiB/s ‡		~200 GiB/s ‡	~12 GiB/s ‡	~3 GiB/s ‡
~400 Hz (read)†		~180 Hz (write)†	~200 Hz (read)†	

* Photon instance accepts ~1 TiB per month as ~1x10^7 files.

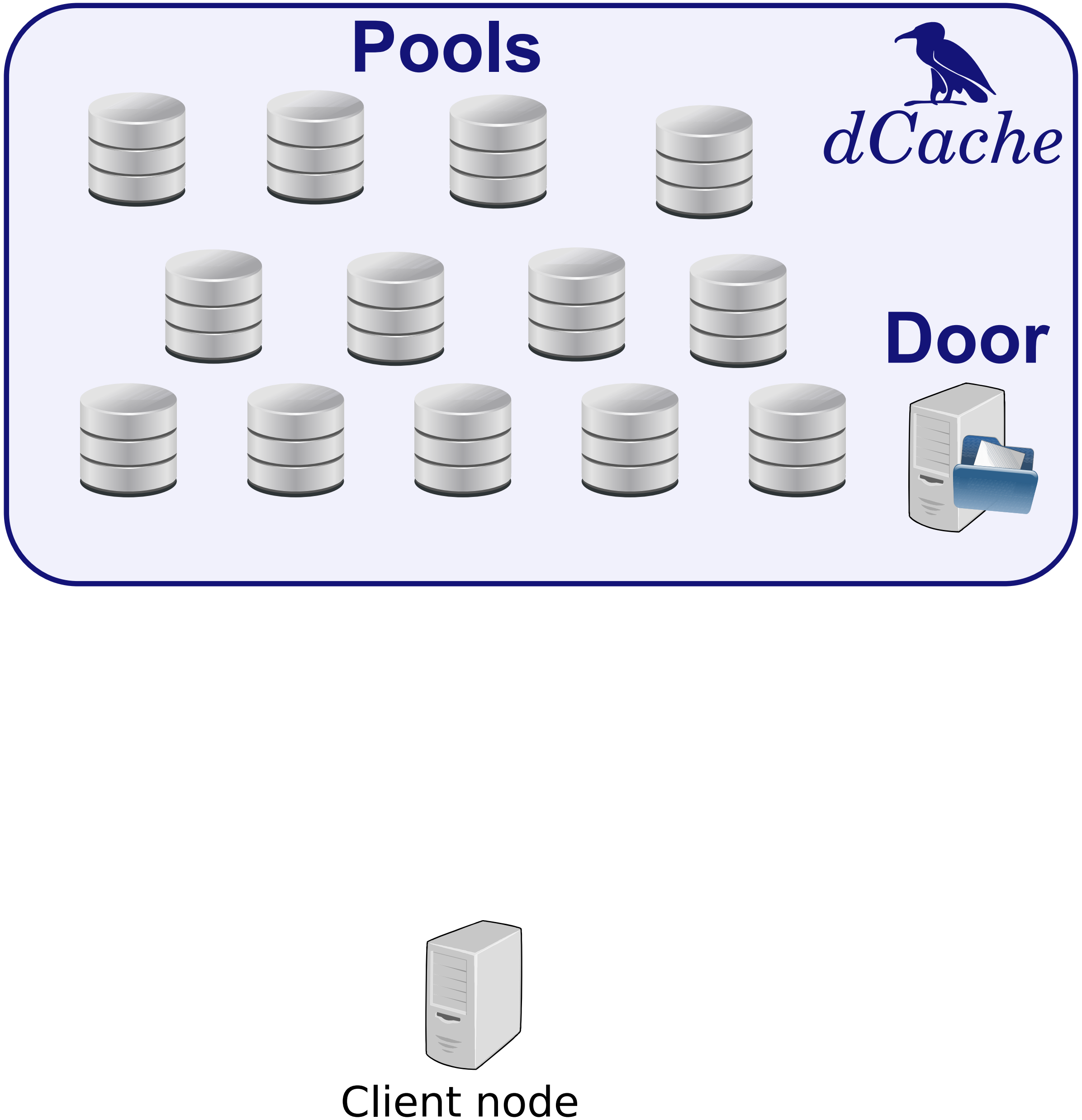
‡ Value is peak observed bandwidth aggregate over all clients within last 7 days.

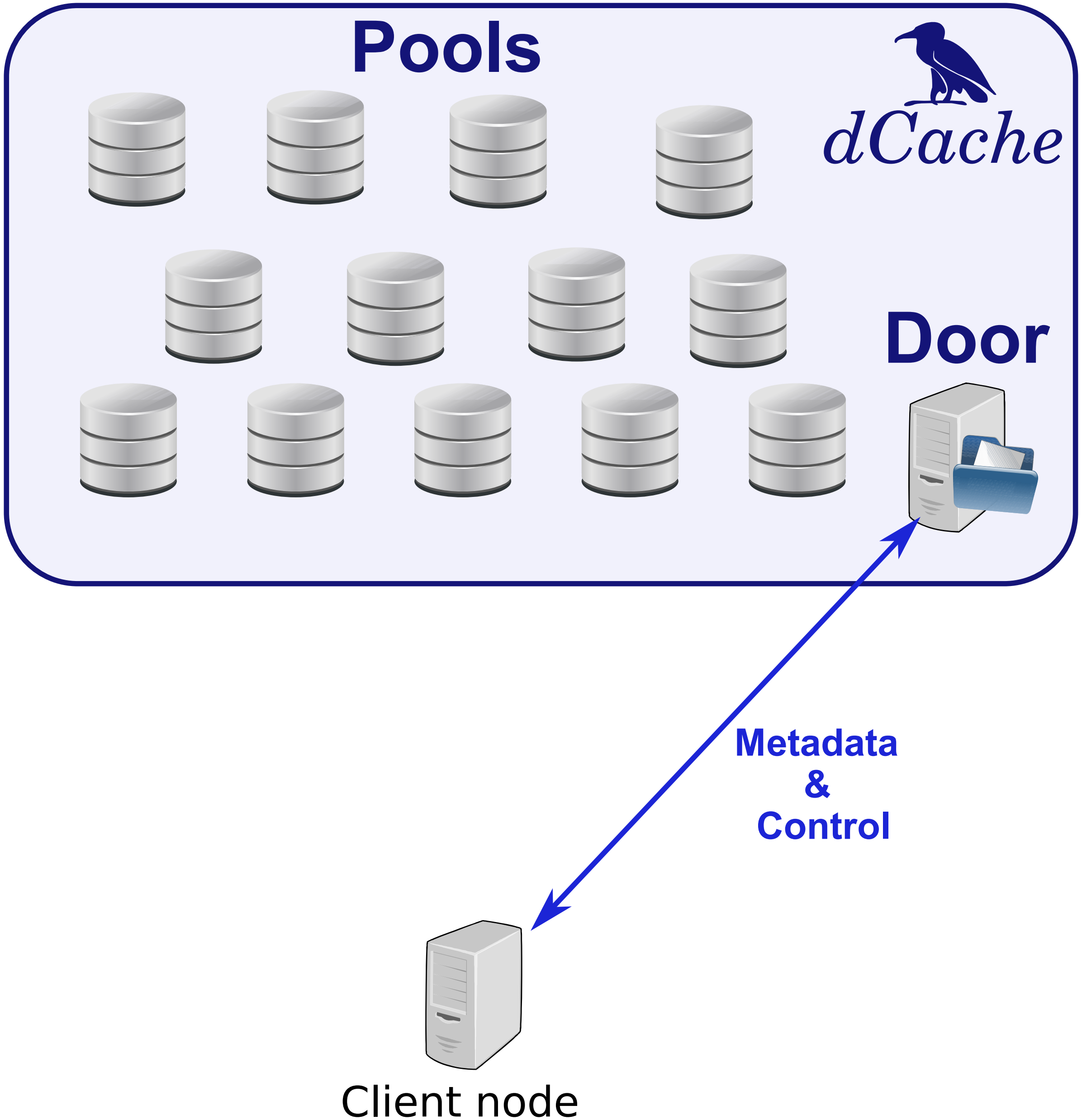
† Value is peak observed open rate (either read rate or write rate) observed within last 7 days.

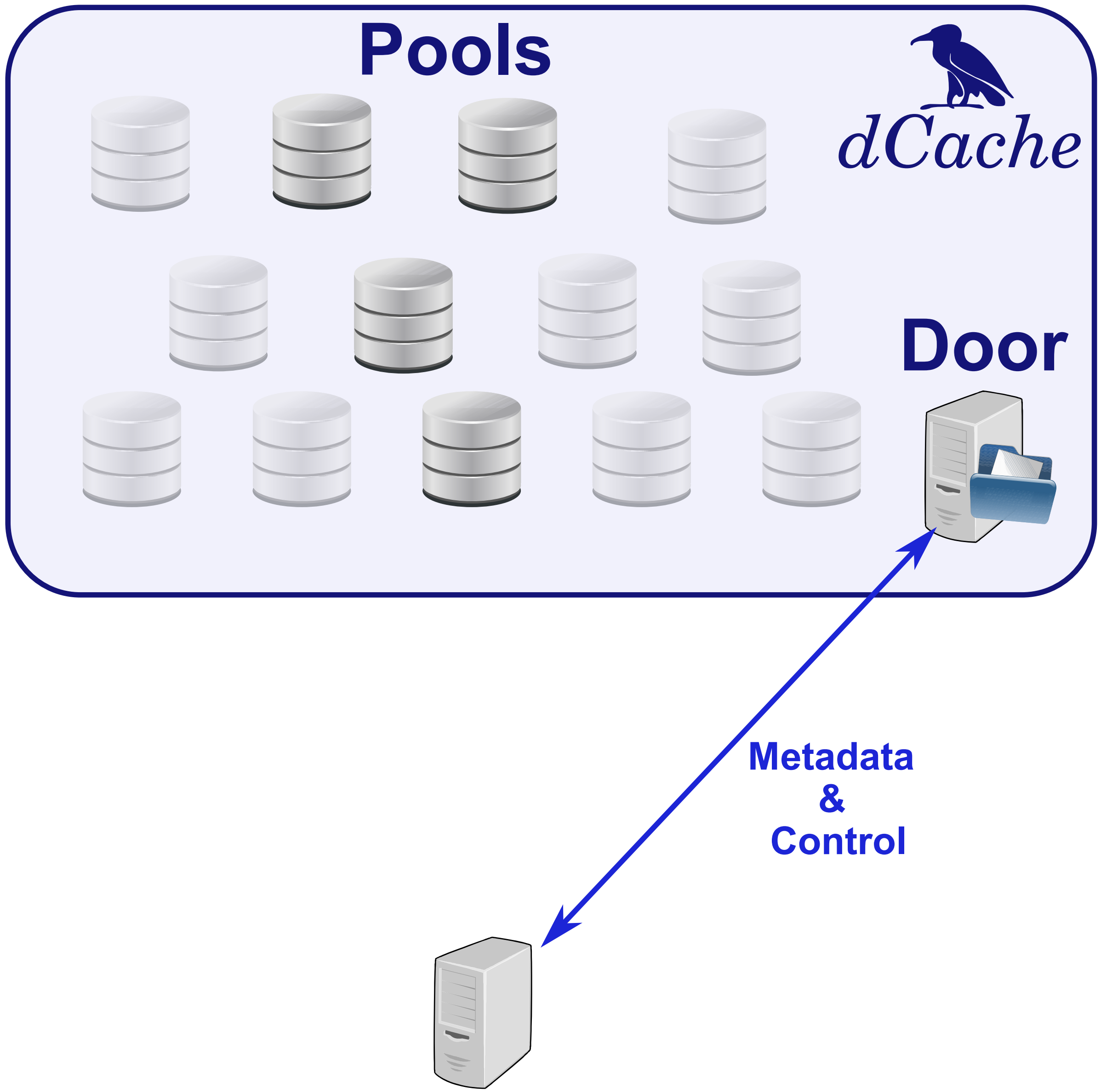
Other dCache instances

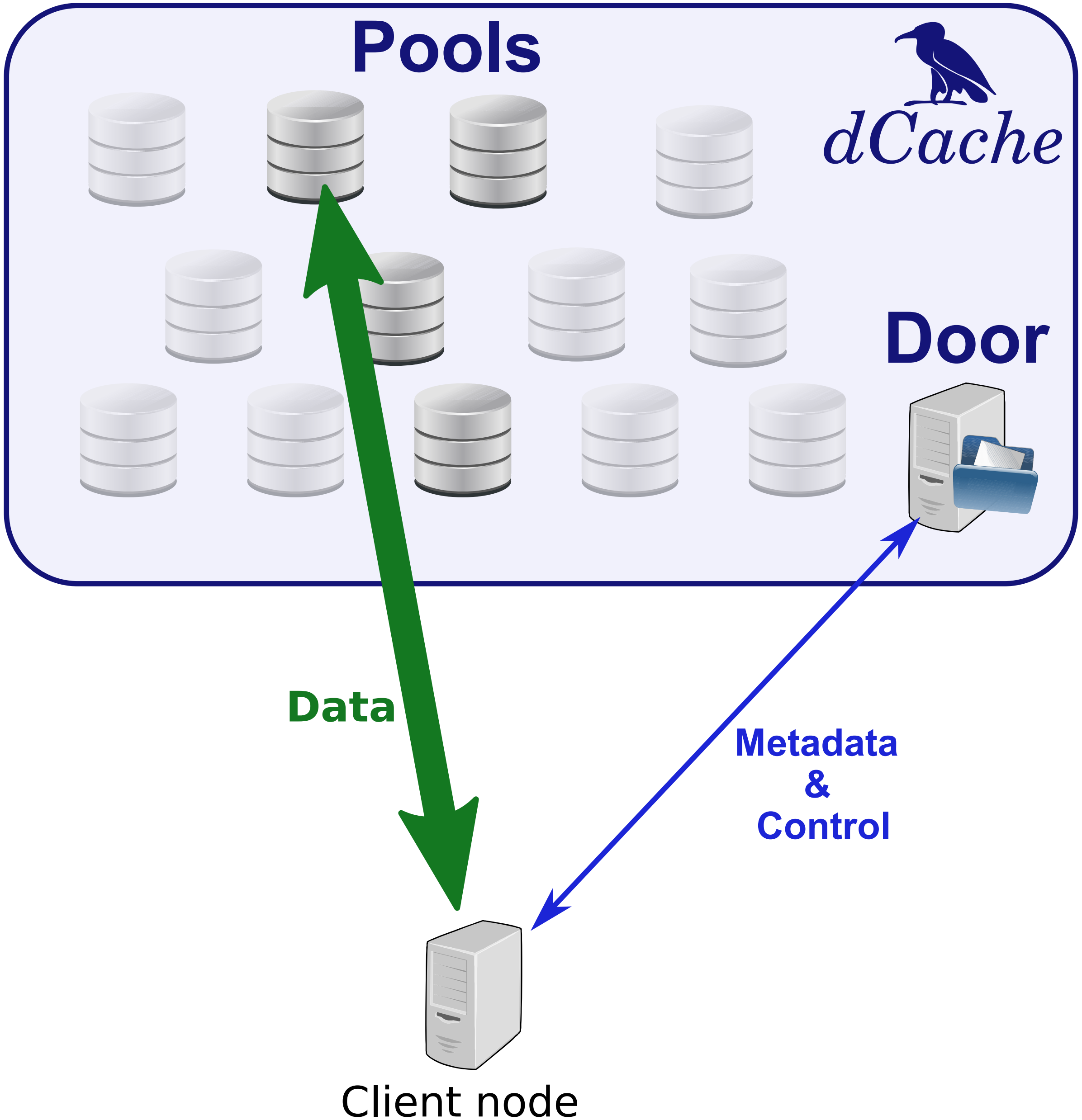
NT1	US-CMS T1	BNL	SARA
~5x10^7 files			
~6.3 PiB (2.1 PiB tape; 4.2 PiB disk)	~20 PiB (disk)	~15 PiB (disk)	~6.2 PiB (disk)

Backup slides

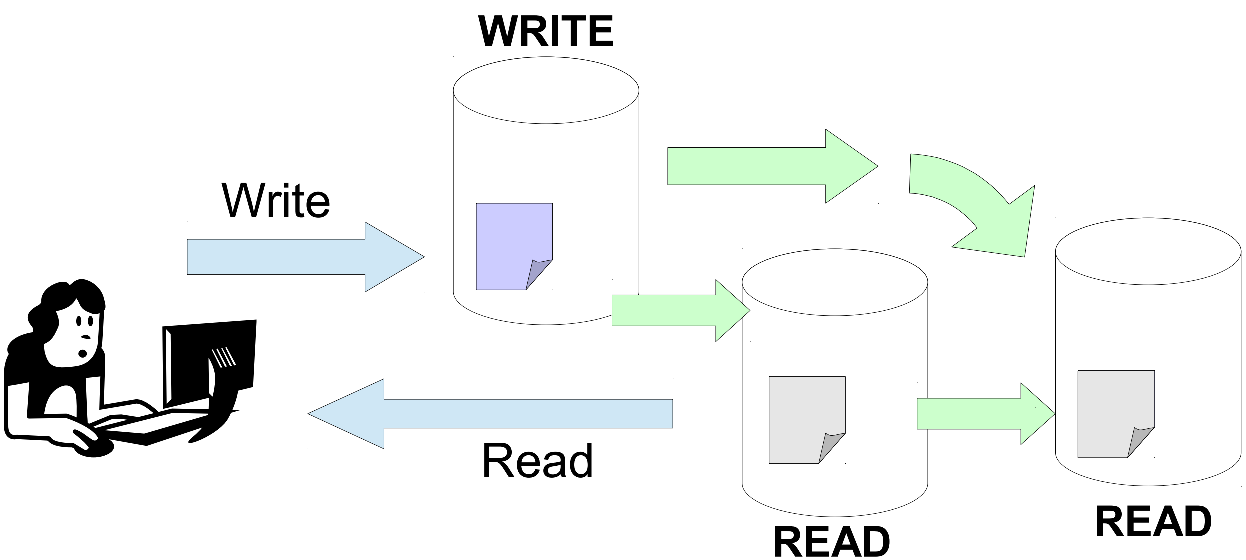

Pools
dCache
Door
Client node

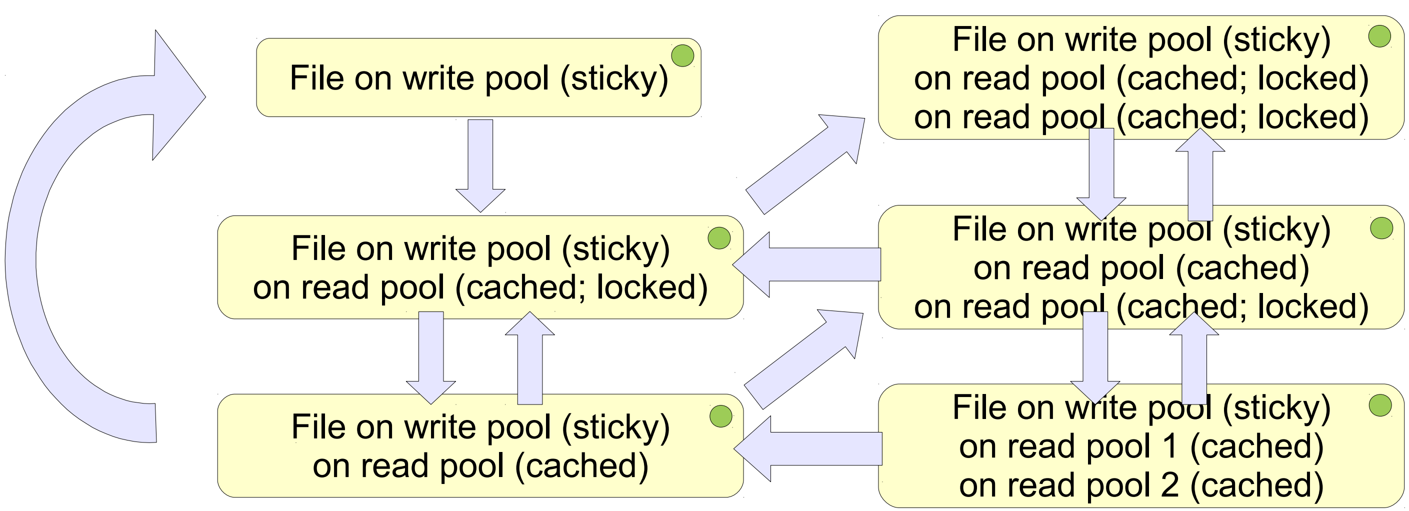

Pools
dCache
Door
Metadata
&
Control
Client node


Pools
dCache
Door
Metadata
&
Control
Client node


Pools
dCache
Door
Data
Metadata
&
Control
Client node

Guaranteeing QoS for write





Guaranteeing QoS for tape activity