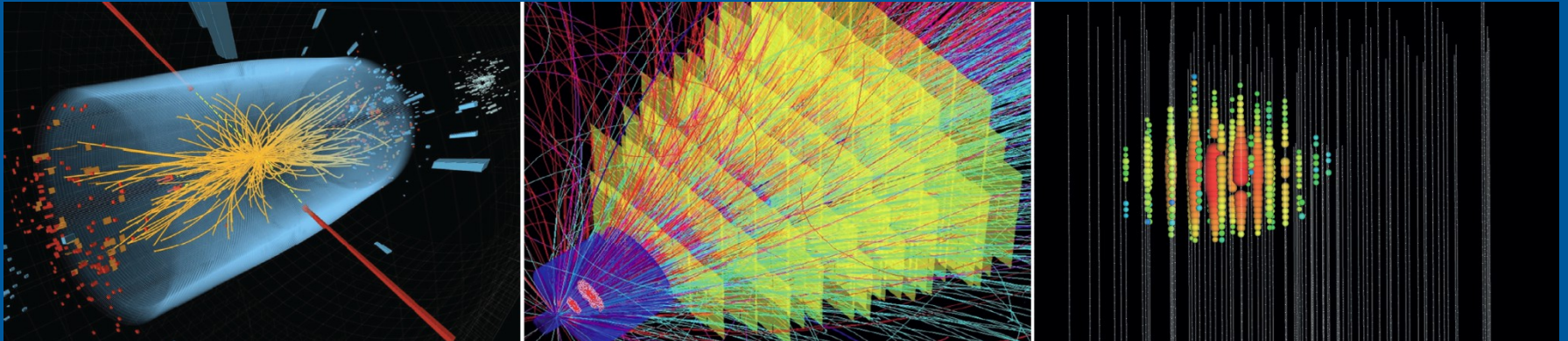# IT: development

Paul Millar – DESY

HELMHOLTZ | ASSOCIATION

dCache: Big Data throughout the world

1. Asia 0.886
2. South America 1.0
3. North America 43.
4. Europe 74.1

Source: BDII

1. Others 0.79
2. Italy 1.2
3. Belgium 1.3
4. Switzerland 2.4
5. Sweden 2.9
6. UK 3.7
7. Russia 4.2
8. The Netherlands 5
9. Scandinavia 6
10. Spain 7.3
11. France 9.2
12. Germany 30

Source: BDII
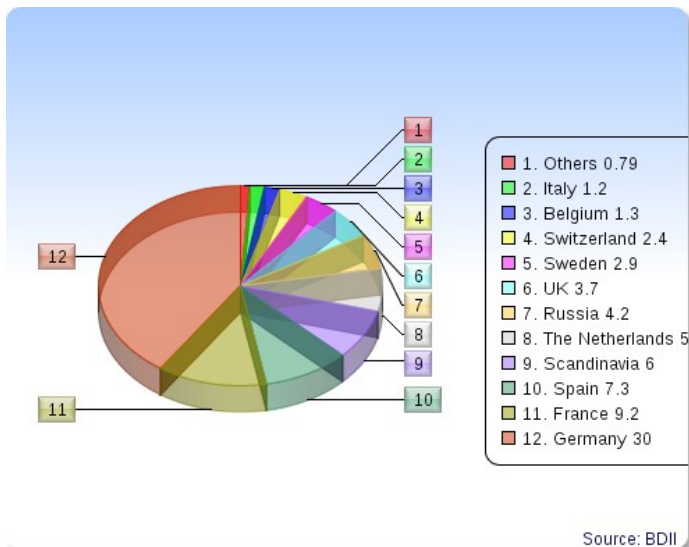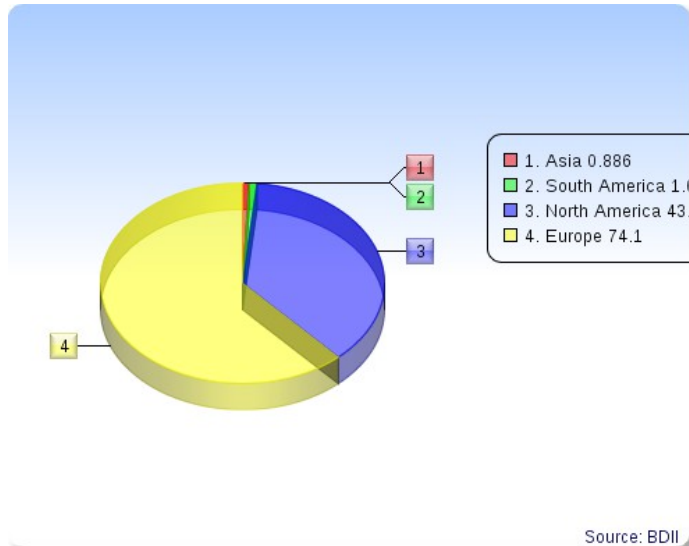
Collaboration and funding for dCache

Open Source

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

Fermilab

DESY

NDGF
Nordic DataGrid Facility

HELMHOLTZ
ASSOCIATION

SNIC

EMI
EUROPEAN MIDDLEWARE INITIATIVE

ntw.

LSDMA

Paul Millar   IT: Development
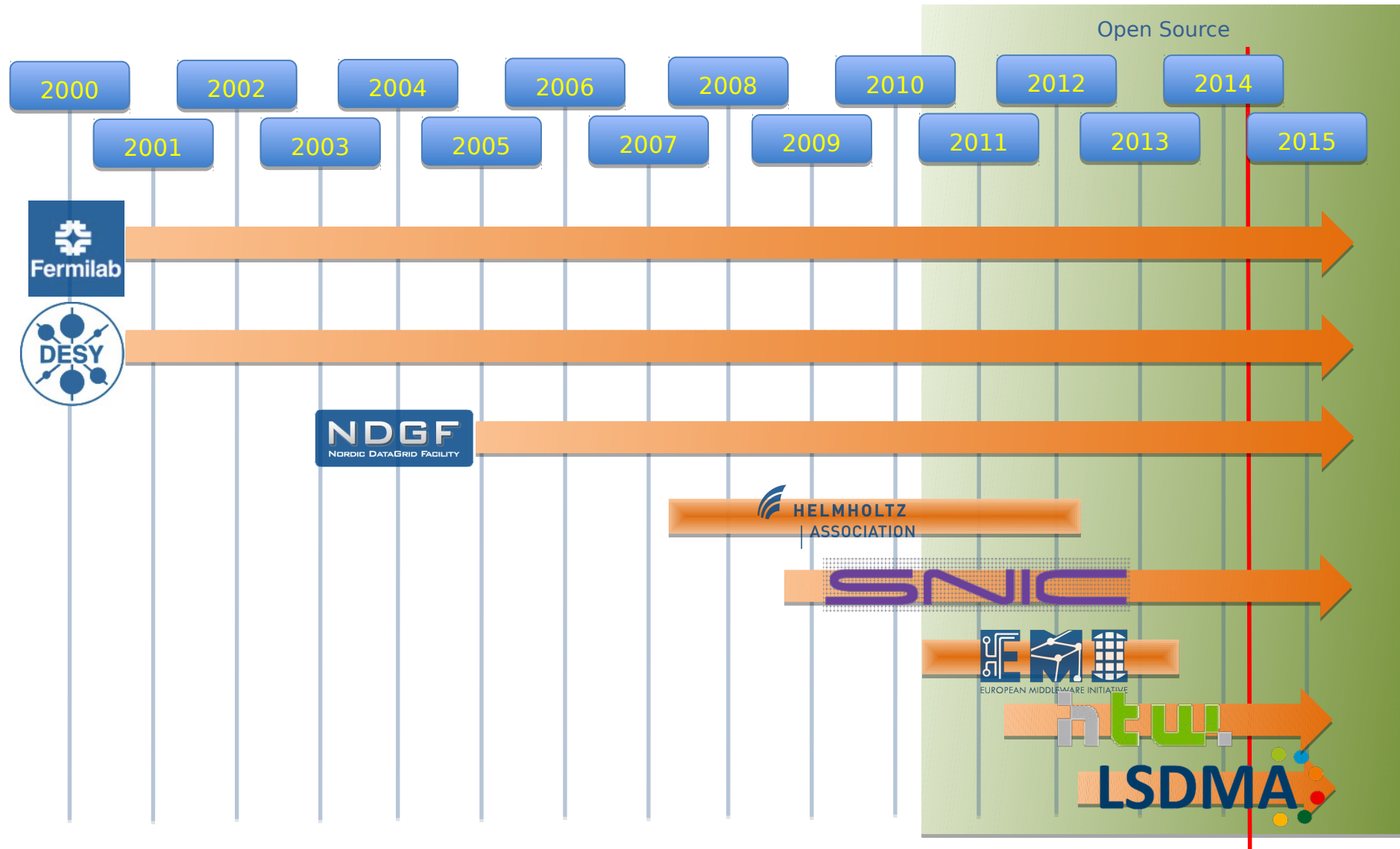
dCache evolution for Big Data and Cloud

Era			
Additional Communities			
Authentication			
Trusted host	Trusted host, **X.509, Kerberos**	Trusted host, X.509, Kerberos, **Username+PW**	Trusted host, X.509, Kerberos, Username+PW, **SAML, OpenID, OAuth, Token, ...**

Big Data Strategy

HPC & Grid Clusters
Low latency access

Cloud storage
Standard back-end for clusters and portals
OpenNebula
openstack
CLOUD SOFTWARE

NFS

CDMI/S3

FTP

HTTP

WebDAV

Bulk WAN transfer
Moving huge datasets
LCG

Remote access
Rich access via web-browser

DropBox-like storage
Devices synchronise with storage

Paul Millar   IT: Development

Advances in networking

Paul Millar   IT: Development
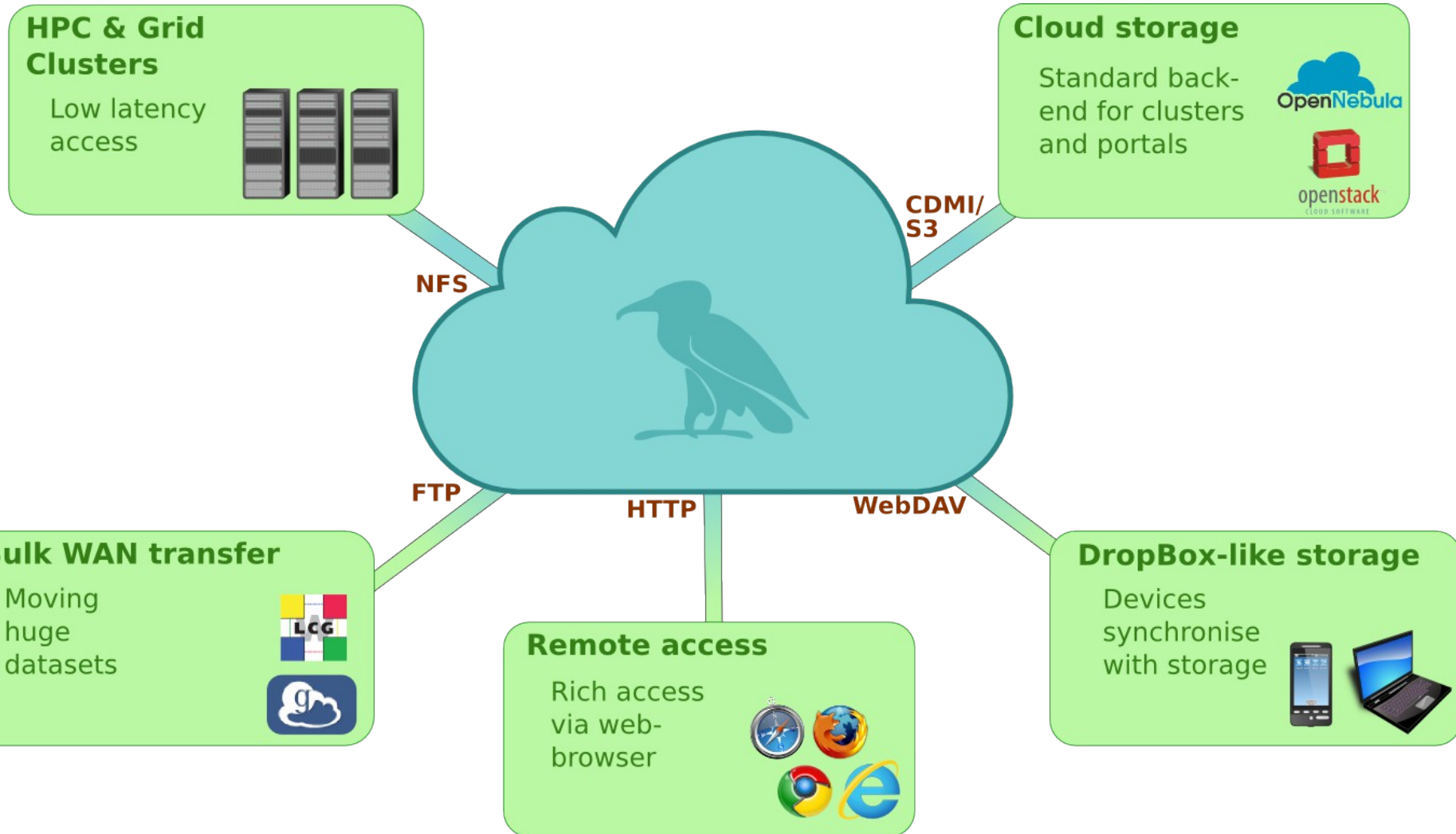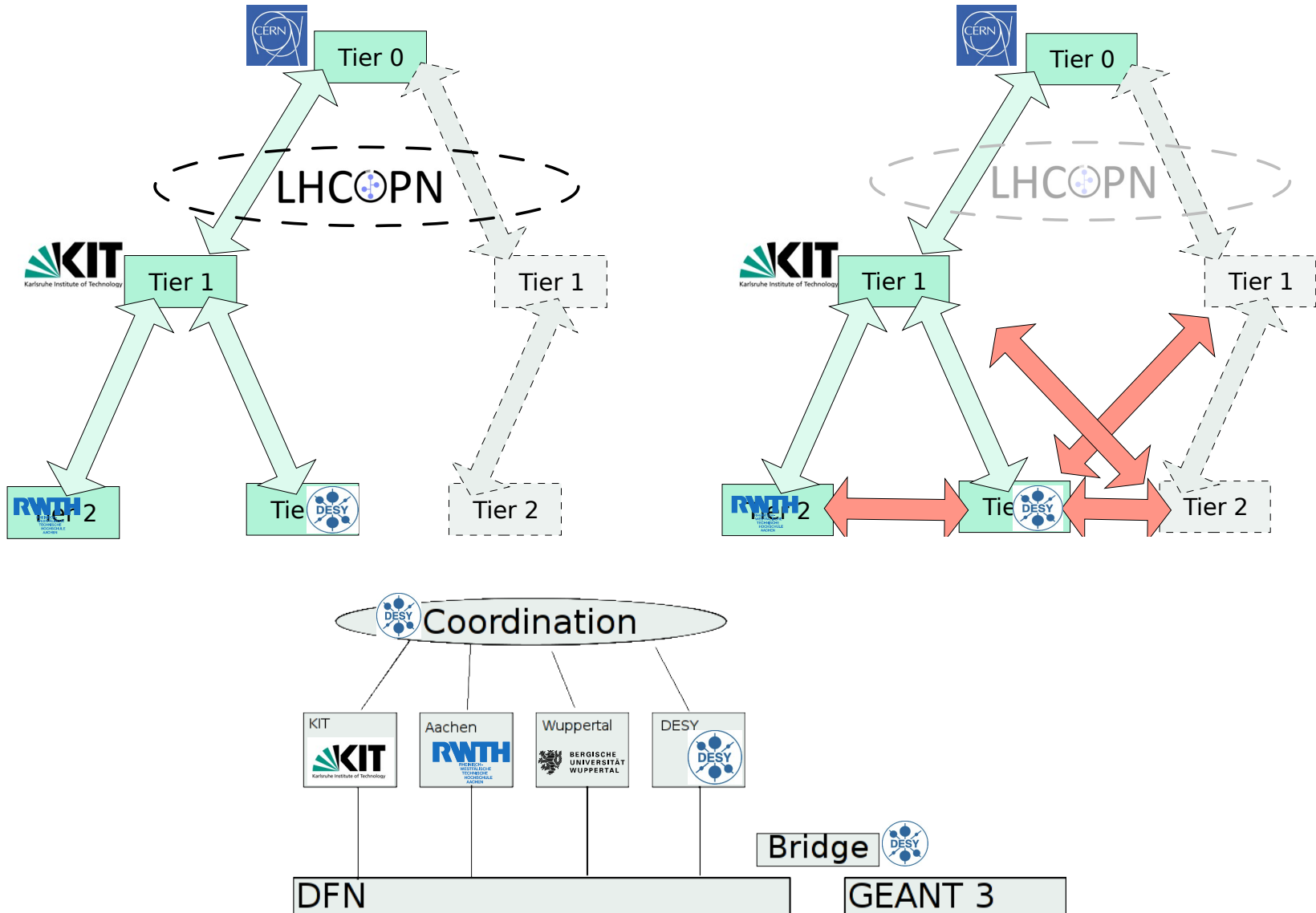
Supporting Heavy HPC users:
Lattice QCD, Simulation for Photon Science, Plasma, ...

Masters and PhD students

Testing new ideas and methods

Post-processing

HPC usage:
Jülich, Munich, Berlin

Paul Millar   IT: Development
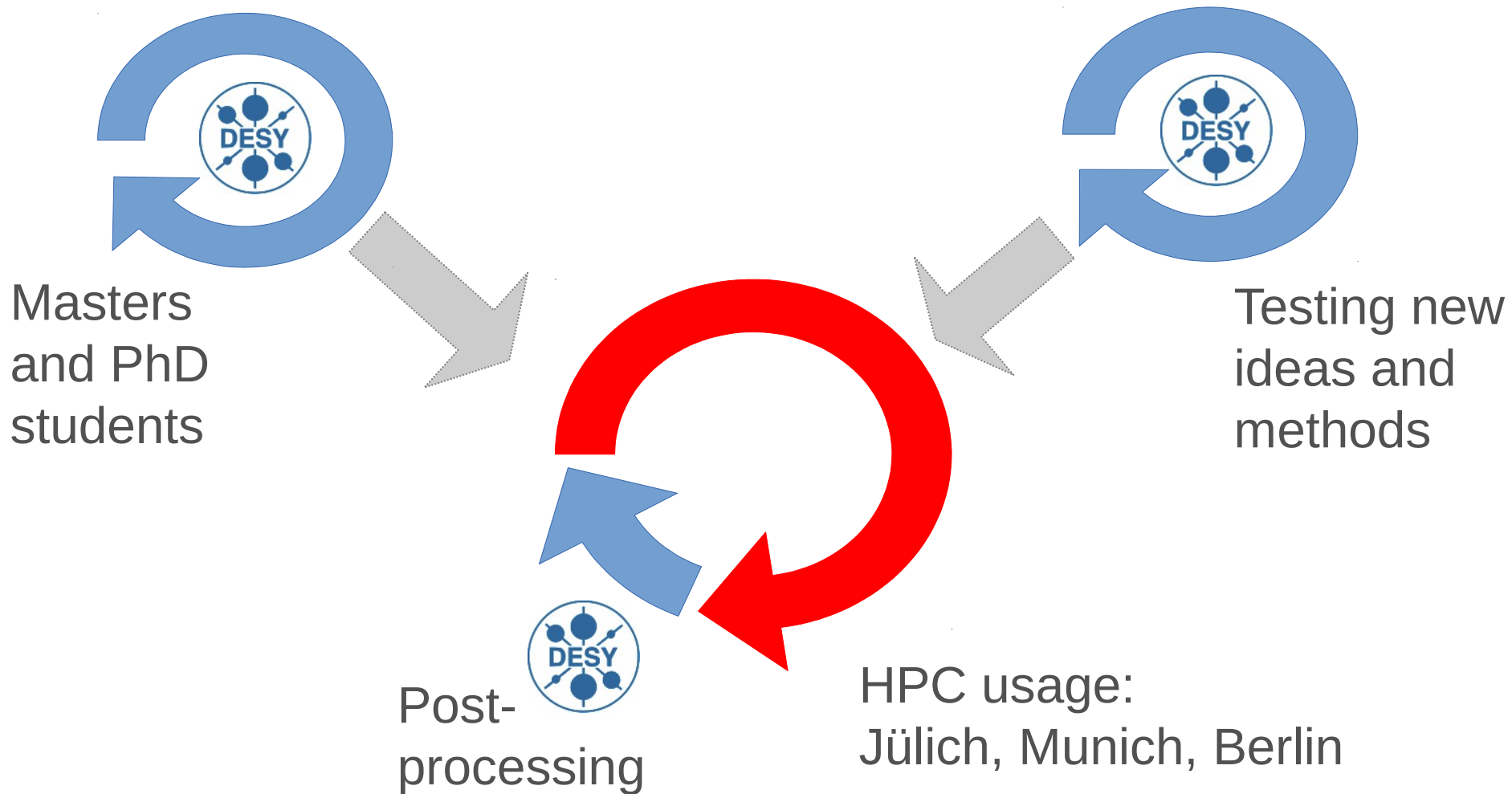
Handling the complete data life-cycle

Paul Millar   IT: Development
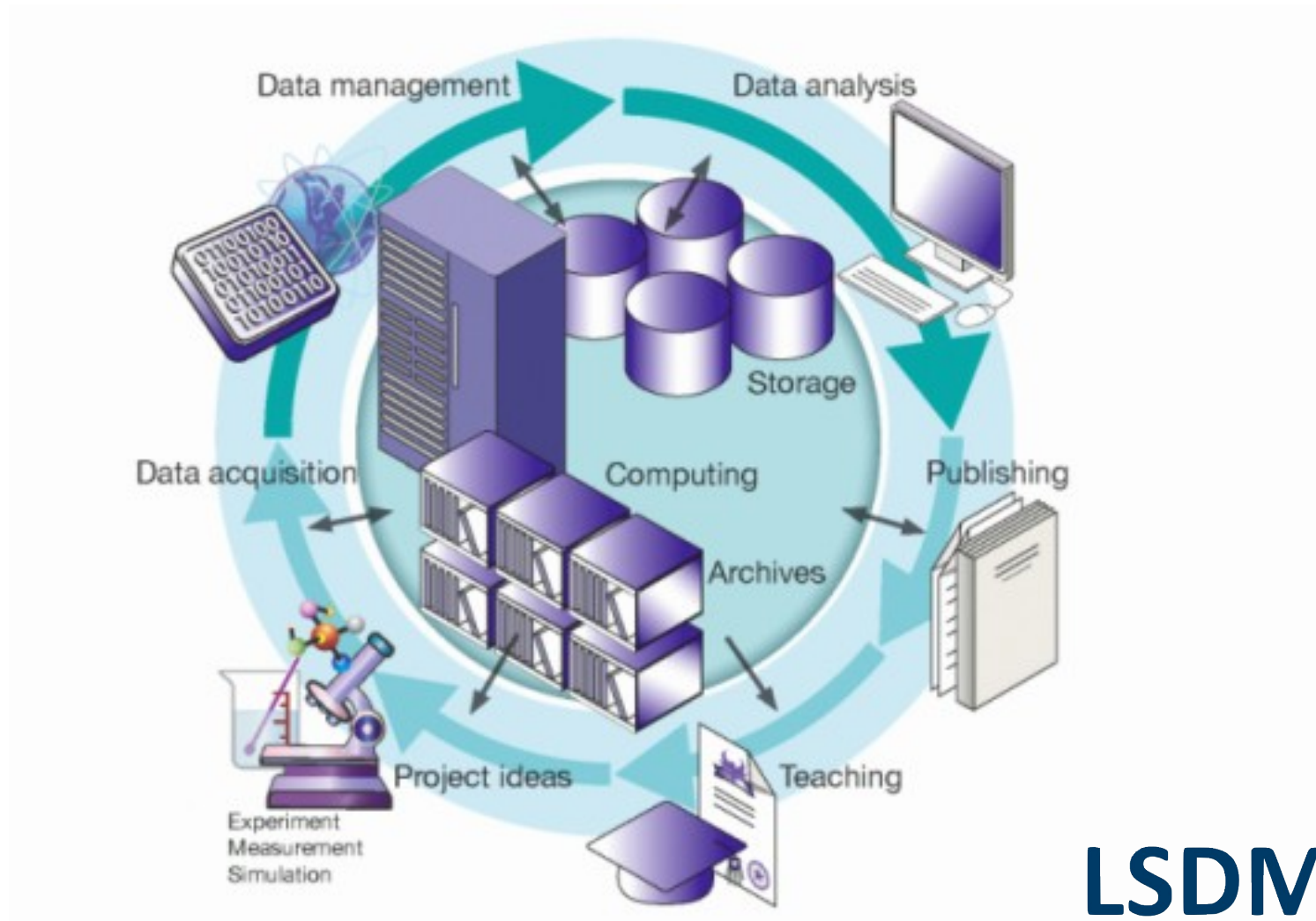
Responsible curation of unique Hera data

CERN joins collaboration

INFN and IN2P3 expected to join collaboration

Report publish

RDA IG formed

2011   2012   2013   2014   2015

DESY initiates study group

DP-HEP formed, CERN appoints manager

HEPix group

DESY to join collaboration

RDA and DP-HEP to apply for H2020 funding

Paul Millar   IT: Development
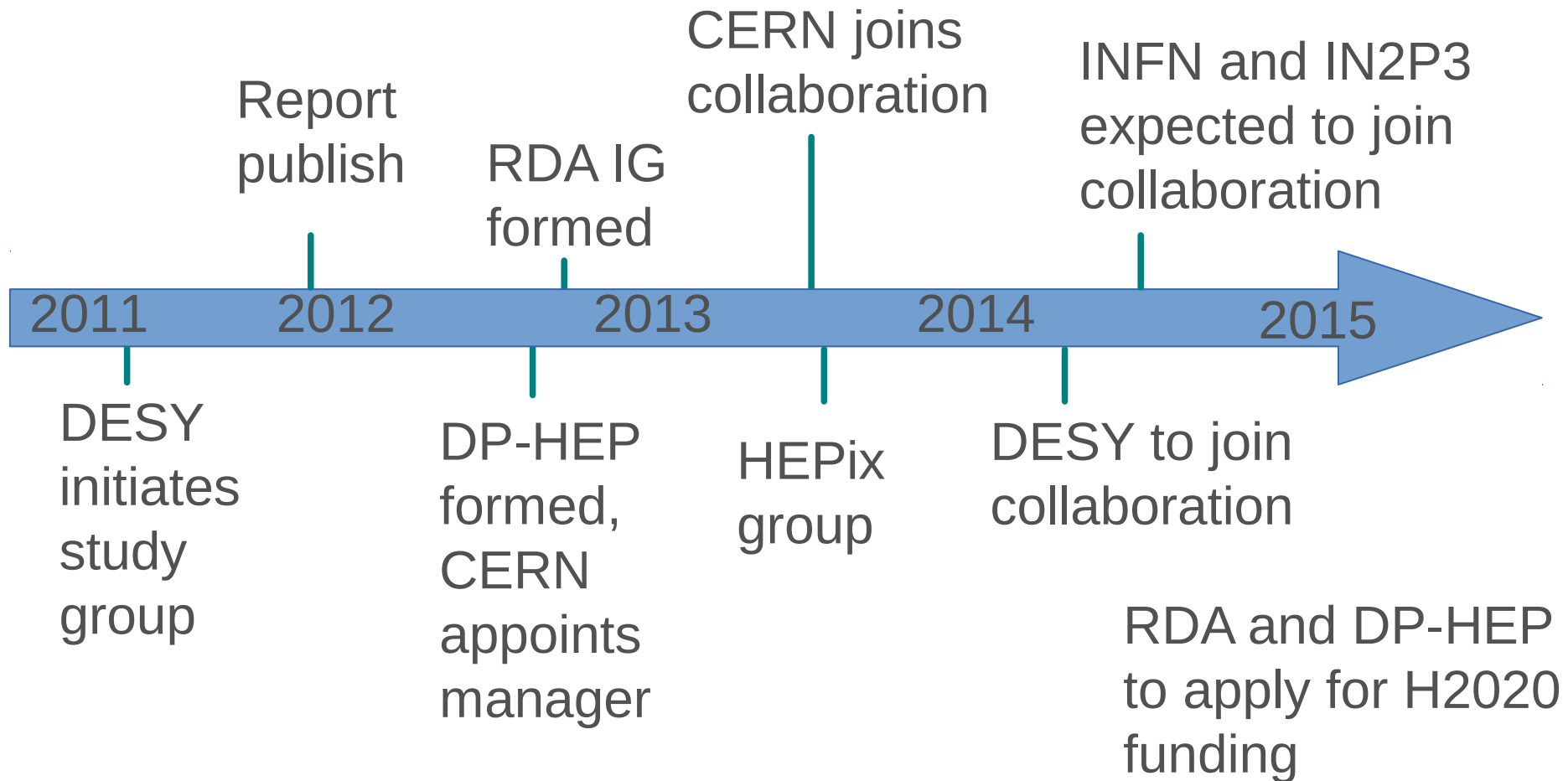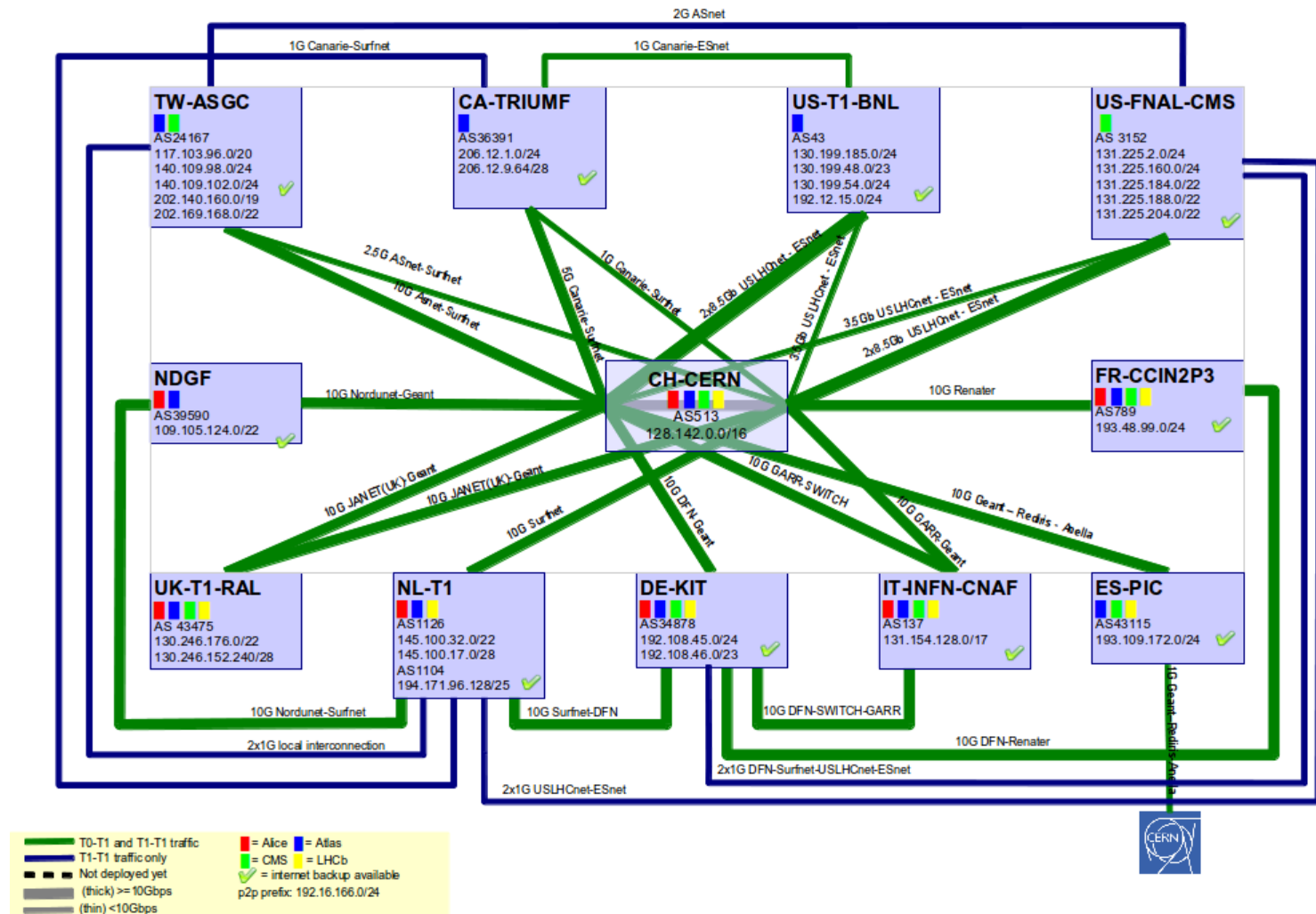
Strategic outlook

DESY is advancing IT by

- developing Big Data solutions,

- pushing network connections,

- exploring bit preservation,

- providing systems to handle complete lifecycle

- providing online/DAC systems to support fast data ingest

Paul Millar   IT: Development

Backup slides

Paul Millar   IT: Development

Advances in networking

Paul Millar   IT: Development

Advances in networking

LHCONE: A global infrastructure for the LHC Tier1 Data Center – Tier 2 Analysis Center Connectivity

Paul Millar   IT: Development