IT at DESY: Strategic Report

90th Meeting of the Physics Research Committee – Open session

Presenter: Sophie Servan
Authors: Patrick Fuhrmann, Volker Gülzow, Thomas Hartmann, Yves Kemp, Birgit Lewendel, Christian Voß
5 Nov 2020

Content

01 Status of Existing Infrastructure
- Computing
- Cloud infrastructure
- Network
- Data management

02 Current and future challenges
- Growing Data volumes
- Distributed landscape
- Emerging needs for Data Analysis

03 Outlooks and Strategy
- Interdisciplinary Data Analysis Facility (IDAF)
- Integrating DESY into horizontal infrastructures
  - HEP and Astro data lakes
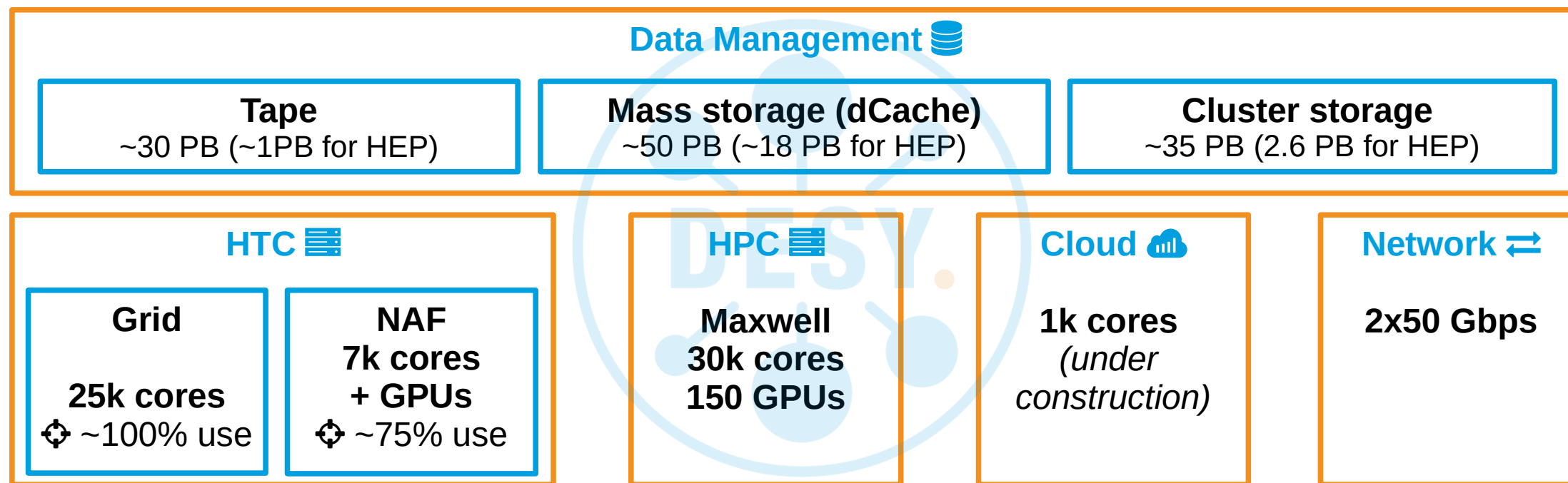 - Photon and Neutron federated cloud

04 DESY and COVID-19 compute
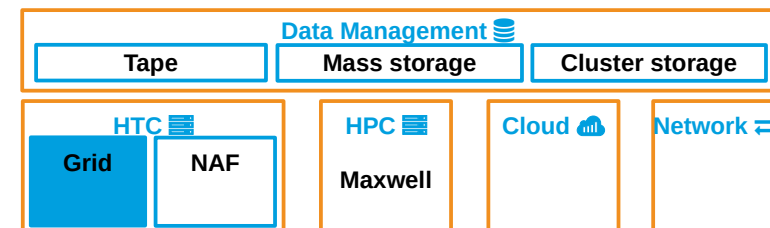
05 Conclusion

01. Status of existing Infrastructures

Status of existing Infrastructures

Overview

Data Management 🗄

Tape	Mass storage (dCache)	Cluster storage
~30 PB (~1PB for HEP)	~50 PB (~18 PB for HEP)	~35 PB (2.6 PB for HEP)

HTC

Grid

25k cores
✛ ~100% use

NAF
7k cores
+ GPUs
✛ ~75% use

HPC

Maxwell
30k cores
150 GPUs

Cloud ☁

1k cores
(under construction)

Network ⇄

2x50 Gbps

DESY Grid Computing

Resources for the supported HEP and Astro-Particle experiments

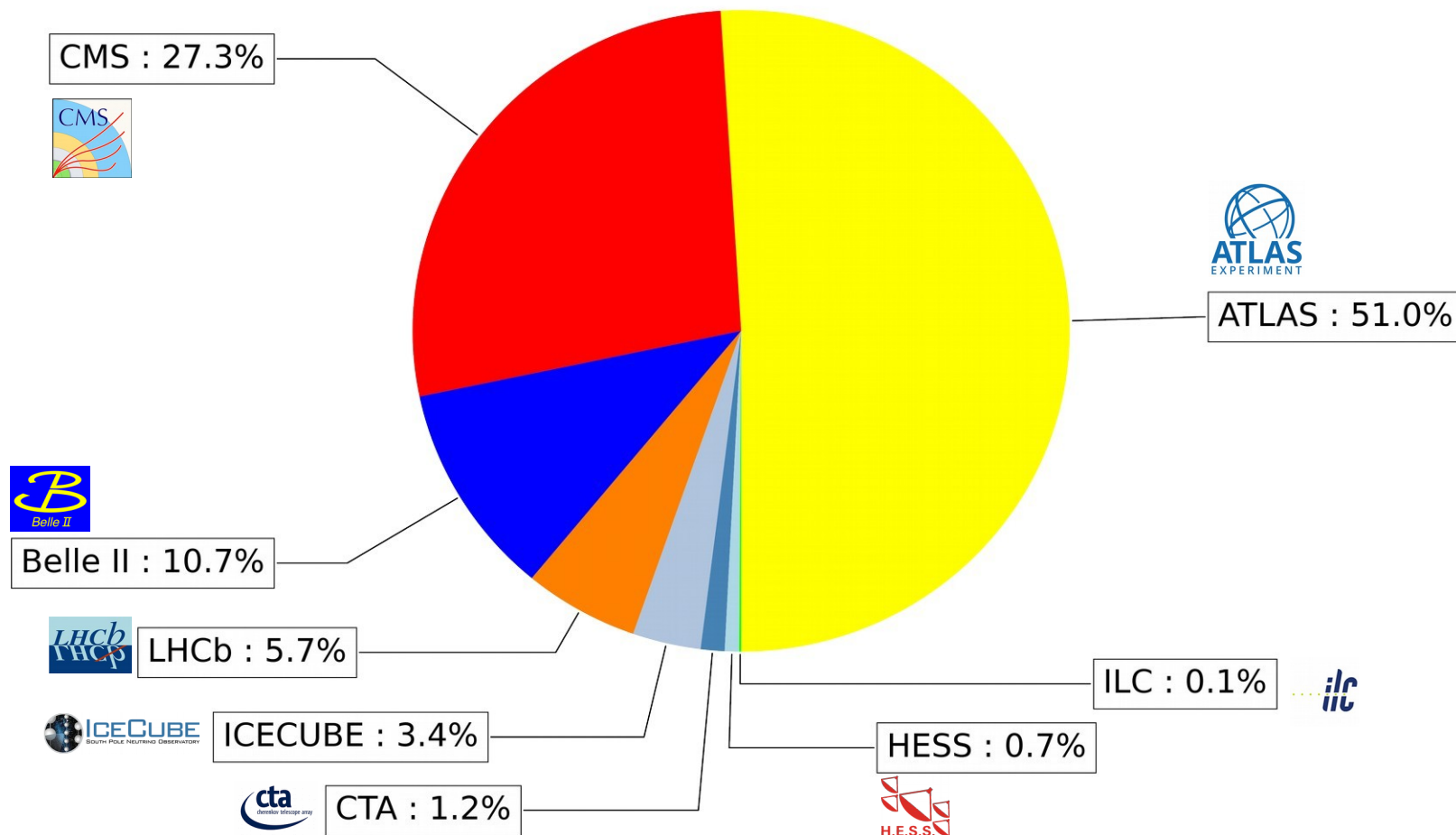**Tier-2 centre for WLCG**

- Secured participation in LHC Run3

Computing Centre for Belle II

- Preparations ongoing to establish ourselves as a raw data storage and recalibration computing centre for Belle II

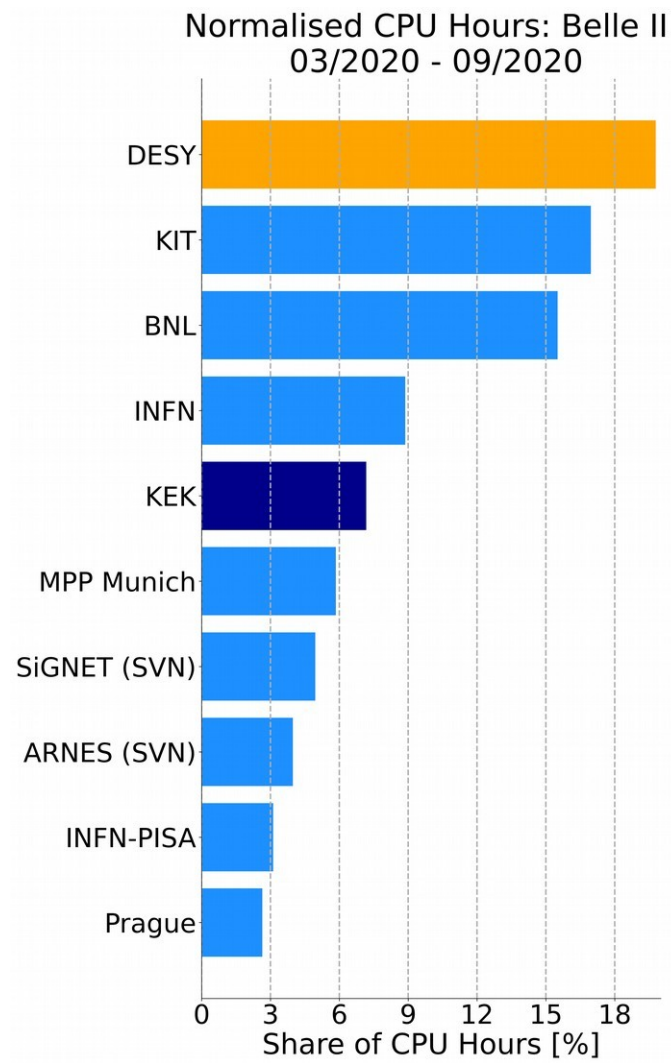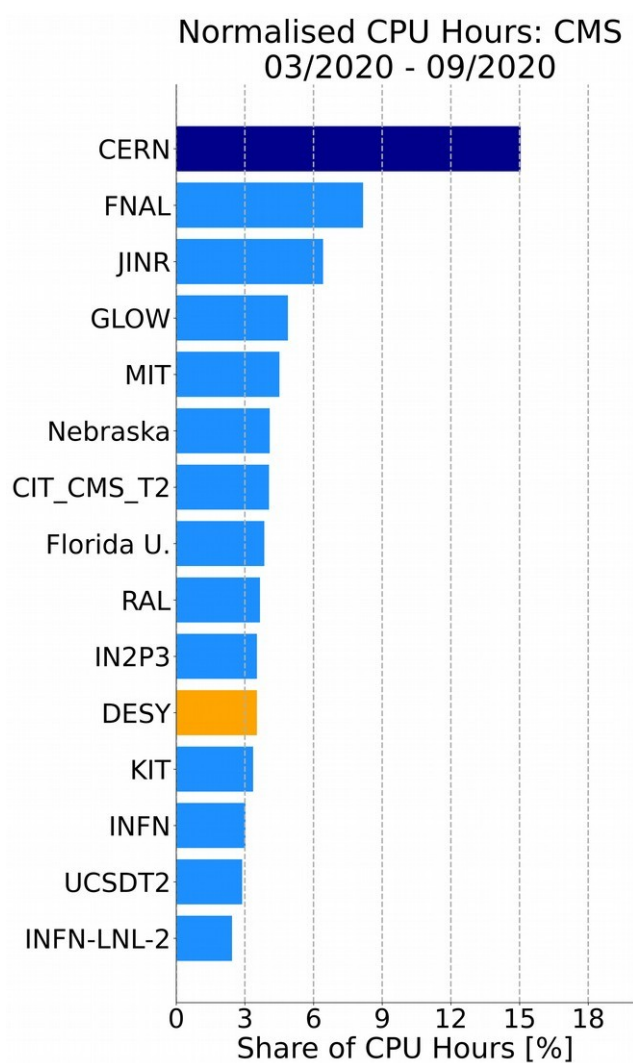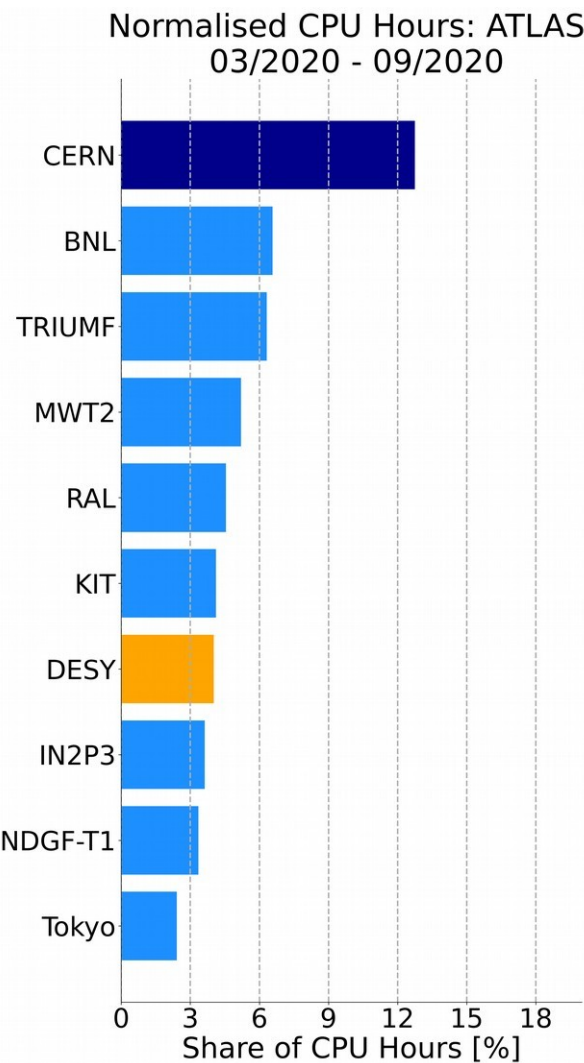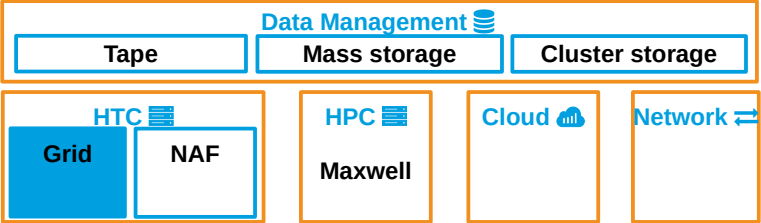**Fully migrated to HTCondor to consolidate our compute elements**

Relative Shares of CPU Hours 2020.03 to 2020.09
Total: ~1.5*e9 HS06 hours

- CMS : 27.3%
- ATLAS : 51.0%
- Belle II : 10.7%
- LHCb : 5.7%
- ICECUBE : 3.4%
- CTA : 1.2%
- HESS : 0.7%
- ILC : 0.1%

DESY Grid Computing

Contributions to major supported experiments (EGI and OSG Accounting only)

Data Management

| Tape | Mass storage | Cluster storage |

HTC

| Grid | NAF |

HPC

Maxwell

Cloud

Network

Normalised CPU Hours: ATLAS 03/2020 - 09/2020

Normalised CPU Hours: CMS 03/2020 - 09/2020

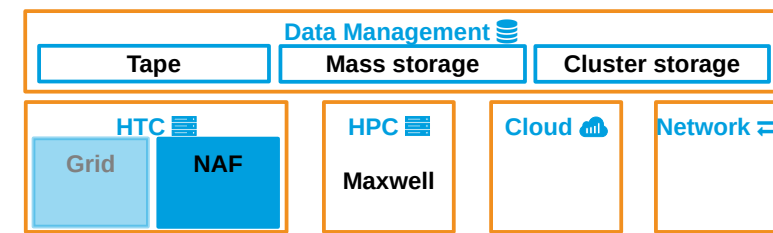
Normalised CPU Hours: Belle II 03/2020 - 09/2020

DESY maintaining a good ranking

- Next to KIT which is a Tier-1 center

- Lead center for Belle II

National Analysis Facility NAF

## A facility for interactive HEP analysis

NAF complementing the Grid:

- Easy access to data stored on DESY Grid dCache and additional project space

- Interactive machines for development

- Batch farm optimized for low latency and small jobs
  → fast turn around times

- Remote accessibility: ssh / FastX / Jupyter notebooks

- Documentation and direct support

Open to:
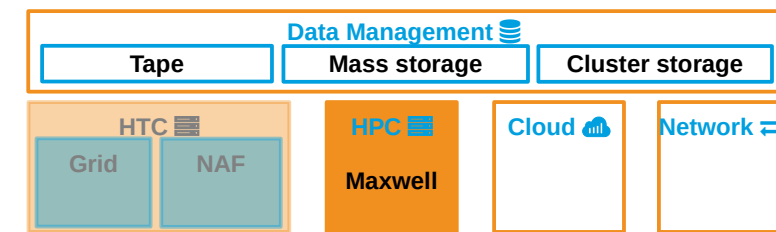
- German LHC institutes + ILC

- Belle II

NAF cluster utilization for the past six month
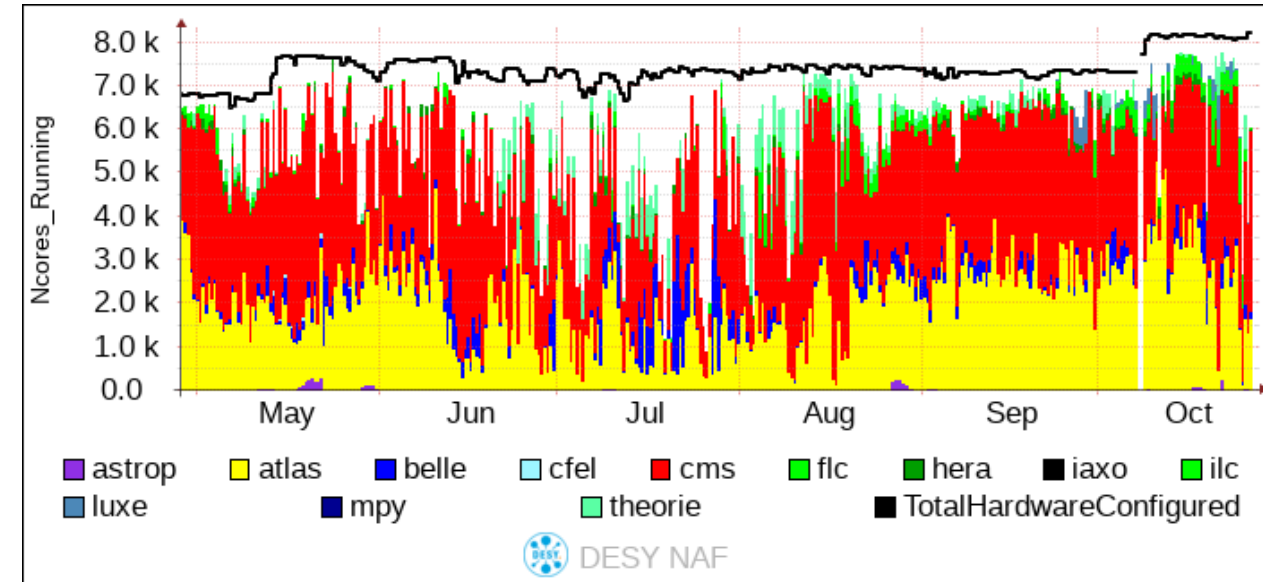

Aim for <75% utilization to allow for fast interactive use

- HERA legacy analyses

- New: LUXE

Maxwell: the DESY HPC system

Complementing the DESY computing landscape

Data Management
| Tape | Mass storage | Cluster storage |

HTC
Grid | NAF
HPC
Maxwell
Cloud
Network

The Maxwell set-up

- Cluster for parallel, high-memory computing (~550 nodes, ~30kcores, ~150GPUs) and low-latency network

- Communities: Machine department, photon science and recently HEP

- Applications: machine learning, simulations, data analysis

- Examples of CMS analysis doing ML: Supersymmetric Dark Matter searches and calorimeter simulation (Dirk Kruecker, DESY)
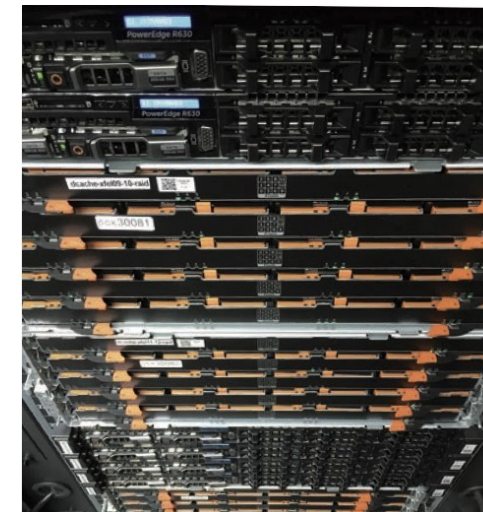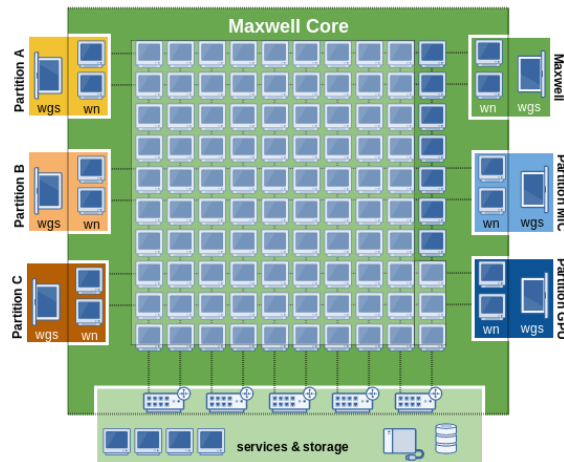
Racks in Maxwell HPC

DESY Cloud Computing

Using modern cloud technologies

DESY providing the stacks to use software in the cloud

- 1000 CPU cores Openstack and Kubernetes cluster under construction

- Jupyter available on HPC, HTC and cloud

- GitLab instance available to all users running in Openstack and Kubernetes

- DESY is a cloud provider in EGI Federated Cloud
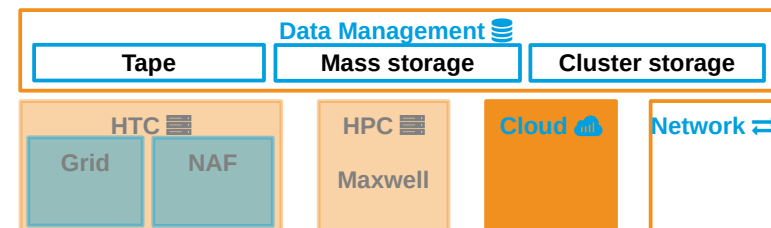
Software stack container orchestration

Software as a Service
Container-based environments
App deployments as code

Kubernetes as a Service
Container Orchestration
Kubernetes Package Manager

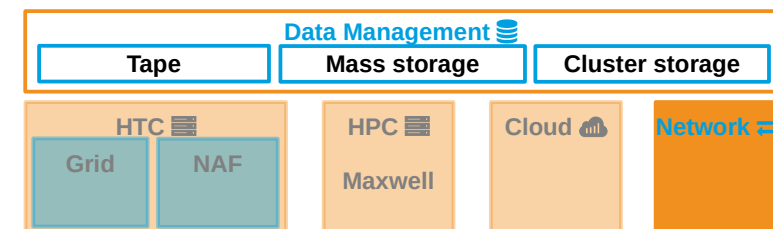**Container as a Service**
Cloud Native CI/CD
Docker registry

Infrastructure as a Service
Cloud Computing
Infrastructure as code

Slide from Michael Schuh, DESY IT

Connectivity

Bandwidth improvements proved useful

Data Management

| Tape | Mass storage | Cluster storage |

HTC
Grid   NAF

HPC
Maxwell

Cloud

Network

New 100 Gbps download capacity in Hamburg already pushed to its limit

- DESY received ATLAS traffic from BNL (see plots from 11th/12th October) which brought the firewall to its capacity limit

- Relief through adjustment of the traffic shaping - future upgrade of the firewall CPU

DESY ↔ WAN

2020-10-12 07:08:00
WAN -> DESY: **97.5 Gb/s**
DESY -> WAN: **6.0 Gb/s**
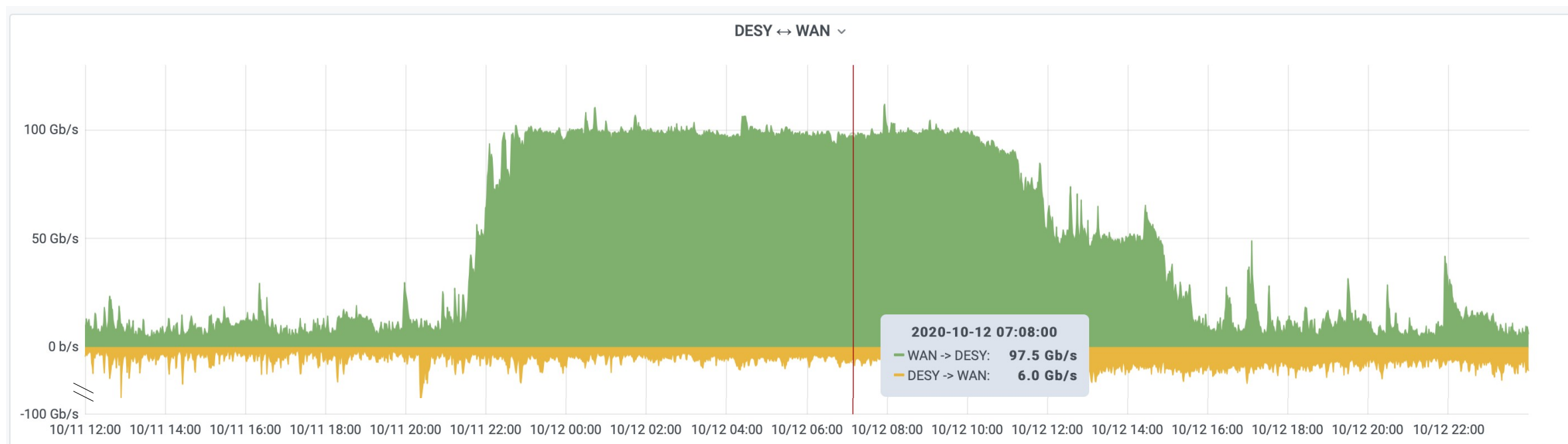
Data Management

dCache central in DESY data management

DESY-IT and data management

- Managing the whole data life cycle for in-house experiments (Photon Science, ALPS…)

- Data taking until archiving

- Providing long-term storage for the HEP experiments

dCache is a central system for both communities

- dCache is the strategic storage for long term data preservation at DESY

- dCache is a main component of the WLCG Data Lake prototype

- DESY is the headquarters for the dCache international organisation and development team

HEP storage over time (Hamburg)
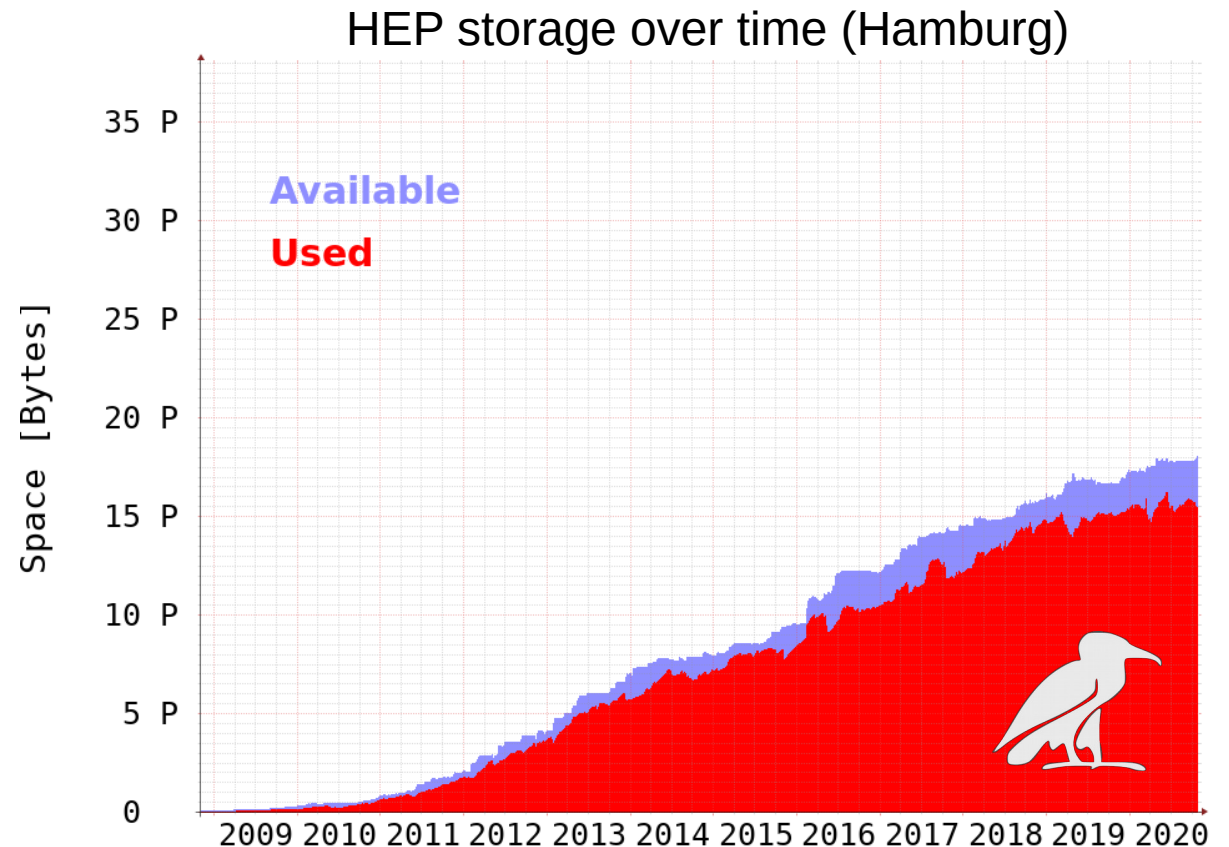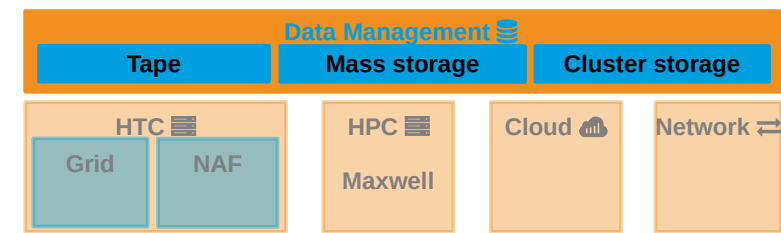
02. Current and future challenges

New orders of magnitude in data volumes

A challenge for both HEP and photon science

DESY Photon sources already produce PBs of data per year

- Eu.XFEL detectors produce 65 TB/hr of data (2500 frames/s) leading e.g to ~1 PB in 48h
- Synchrotrons: 10^4 - 10^5 x increased data rate foreseen in the next 10 years *(Source: Christian Schroer, PETRA IV)*
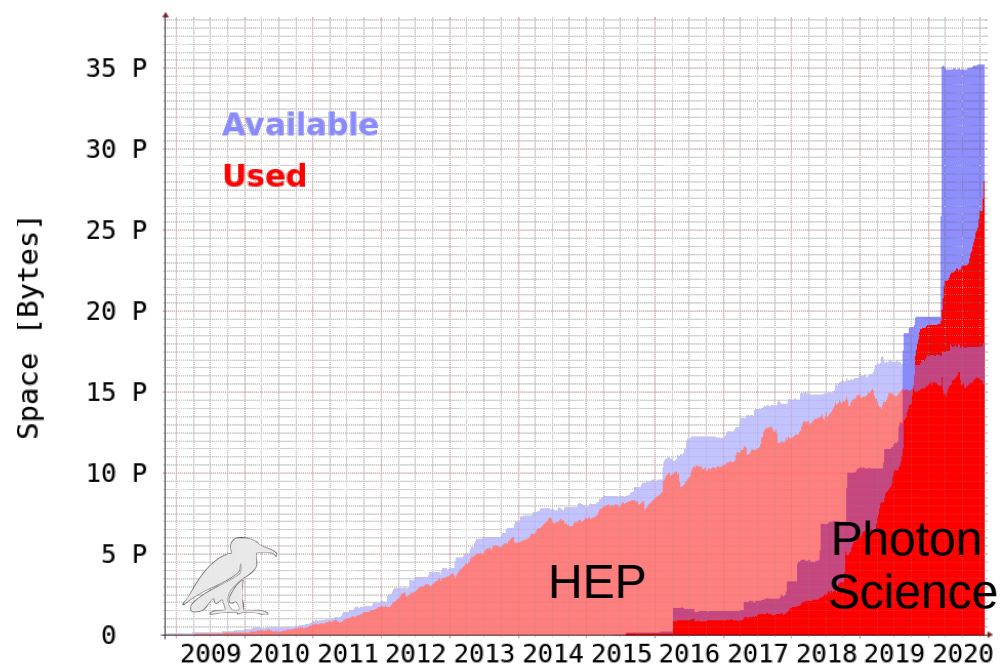
In HEP, next generation instruments will produce EBs of data per year

- High-Luminosity LHC starting in 2027
- Need for a new model
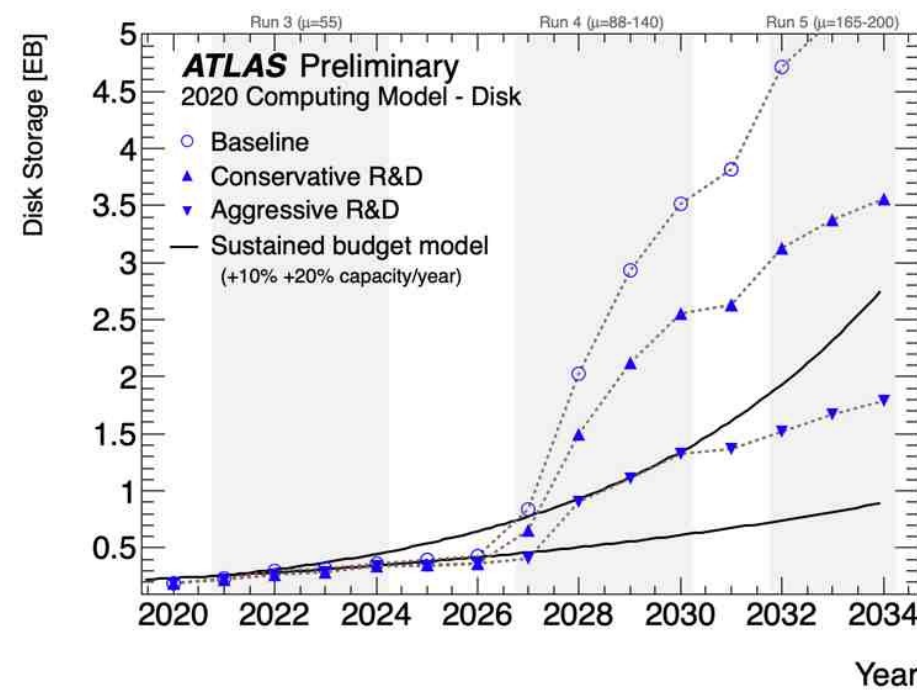
Projected disk storage requirements of ATLAS between 2020 and 2034 based on 2020 assessment.

A distributed landscape of facilities

DESY in a world-wide research community

The WLCG Collaboration - October 2020

66 MoU's
161 Sites
42 countries

Simone.Campana@cern.ch - WLCG RRB   27/10/2020   2

League of European Accelerator-based Photon Sources (LEAPS)

Emerging needs for data analysis at DESY
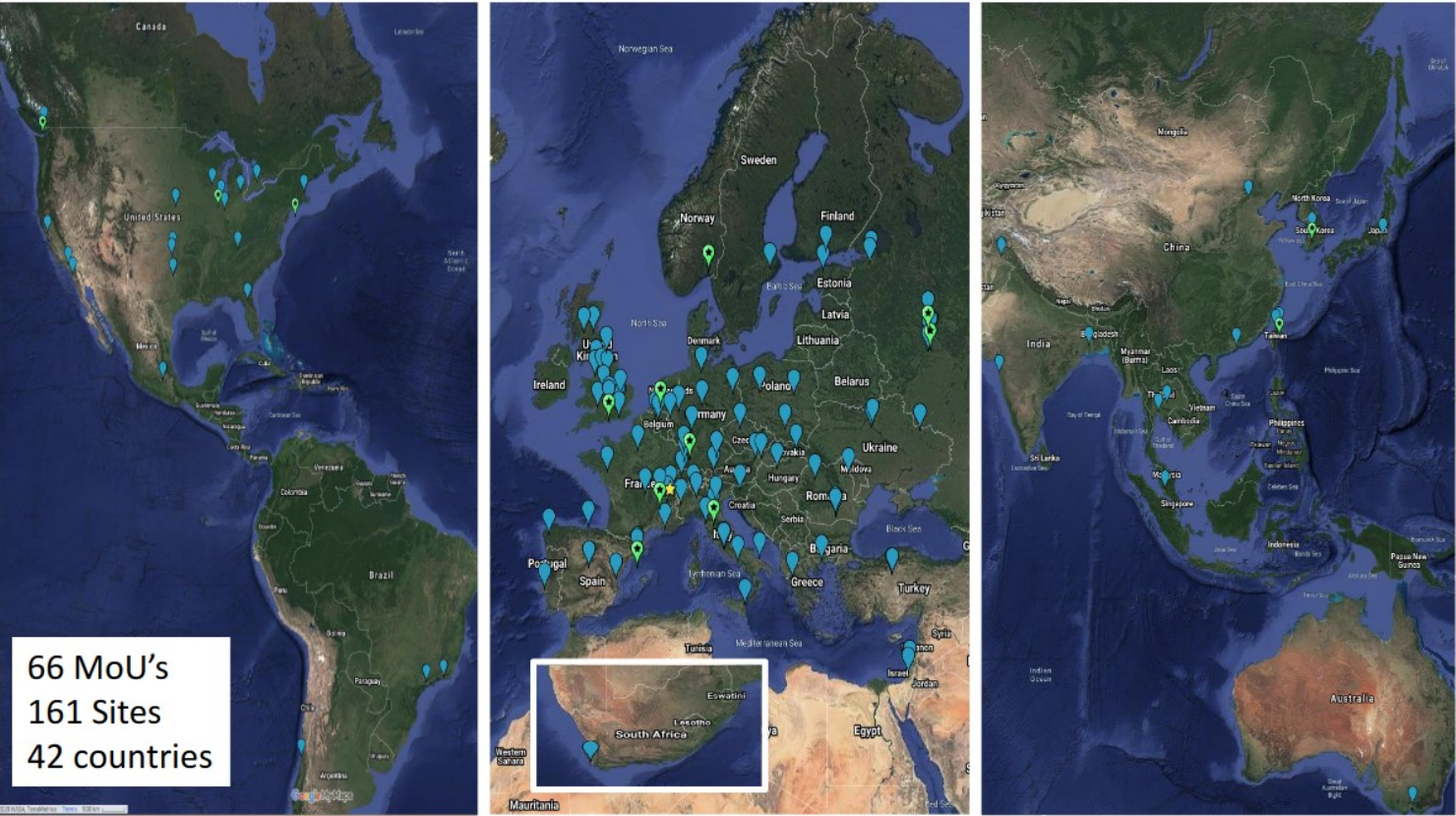
Real-time, HPC demanding and using modern cloud technologies

Real-time analysis

- PETRA III beam-times: Millions of images to be analysed in parallel, can no longer be made manually (benefiting from long triggering experience in HEP)

- Real Time data analyses to discover transient objects with the next generation of Astro-Particle Observatories

HPC-demanding analysis

- Simulations for next generation accelerators like plasma accelerators

- Start-to-end simulation ("Digital Twin")

- Artificial Intelligence for predictive maintenance, etc.

Need for easy deployment and scalability

- Data Analysis practices more and more moving towards modern cloud technologies using e.g. containers

Infrastructure is not enough

*From **Machine group** at DESY, simulation for new accelerators*

03. Outlook and strategy

Outlook and strategy

Federating resources for a seamless access to data and services

DESY committed to High-Luminosity LHC computing

- Participating in the design of the new data lake model

    – Developing the data lake prototype in ESCAPE

    – Working Group leader in WLCG-DOMA for Data Organisation, Management and Access

    – Playing a key-role in providing computing and storage

- Building the Interdisciplinary Data & Analysis Facility (IDAF) to answer the upcoming computing challenges

DESY integrating into horizontal infrastructures

- Working for federation and interoperability of resources at every level
- DESY is an observer in the newly formed EOSC Association
- DESY is chairing NGI-DE which will become member of EGI

Interdisciplinary Data & Analysis Facility (IDAF)

A vision for DESY resources to be made available to users

IDAF

- Combining HTC and HPC technologies for a seamless access to all necessary compute resources

- High performance access to scientific data lakes

- Providing resource monitoring and accounting to the users

- Continuing a strong support to the HEP community with an interdisciplinary consolidated infrastructure
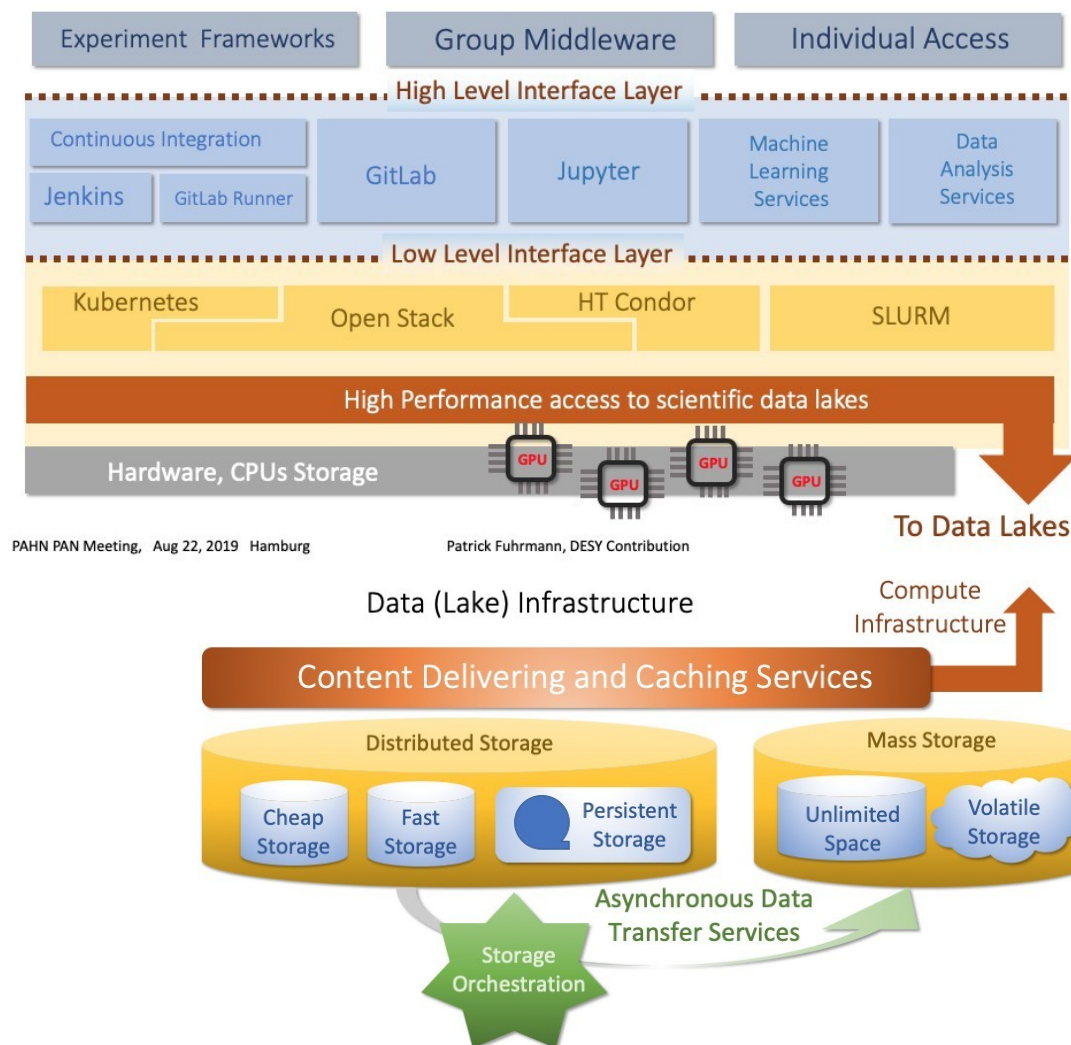
Slide from Patrick Fuhrmann, DESY IT

ESCAPE data lake

Smart data management

Data distribution, smart caching and quality of service for HEP and Astronomy

- Developing tools to achieve data lakes through DOMA (WLCG) and ESCAPE (EU)
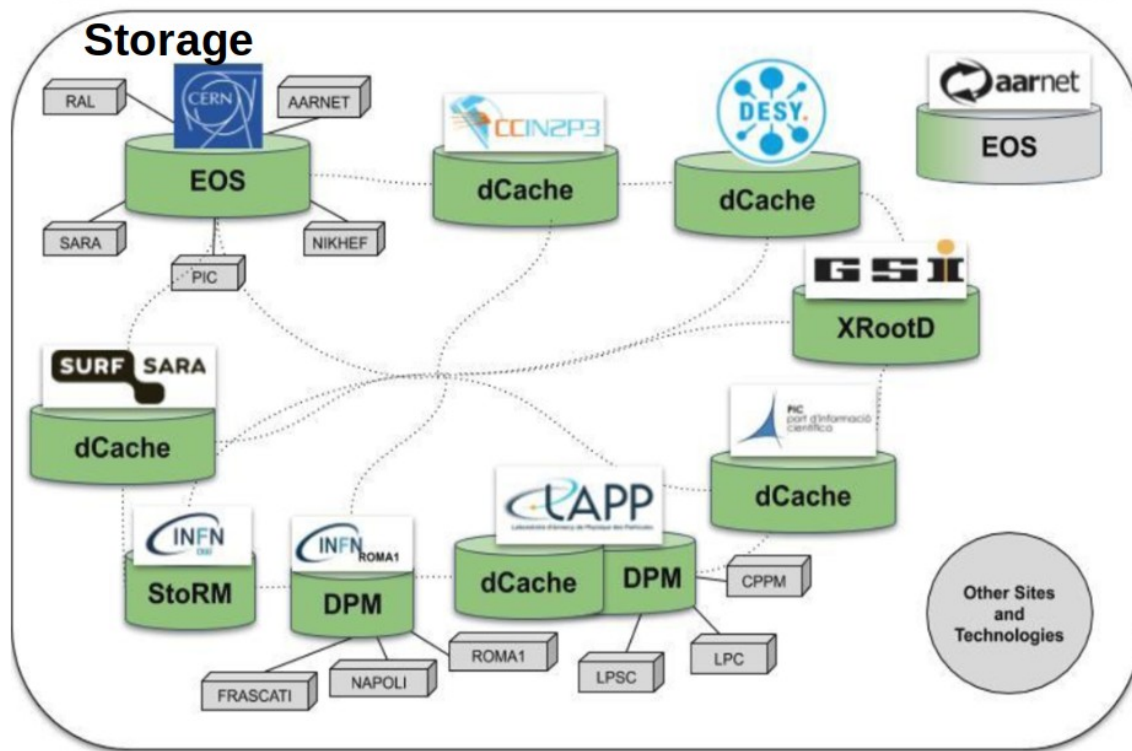
- DESY has historically a leading part in QoS thanks to dCache in EMI and Indigo-DataCloud

Current testbed for ESCAPE data lake from Aleem Sarwar, DESY IT

Integrating DESY in horizontal infrastructures

Working for federation at each level – through the ICT perspective

DESY continued participation in EU-projects for federation of cloud services at the European level

📅 2013 📅 2017 📅 2018 📅 2020 📅 2021 📅 2024

European Middleware Initiative — INDIGO - DataCloud — HELIX NEBULA THE SCIENCE CLOUD — eXtreme DataCloud — ARCHIVER — EGI-ACE — EOSC FUTURE

DESY involved at national level in the National Grid Infrastructure and thus part of EGI

NGI-DE   **ERUM-data**

DESY coordinating the HIFIS platform for a sustained federation of ICT infrastructures in Helmholtz

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES   HIFIS   HELMHOLTZ Data Federation | HDF

DESY building the IDAF to federate compute resources for our user communities

MT IDAF

Integrating DESY in horizontal infrastructures

Working for federation at each level – with scientific use-cases

Prototype-EOSCs for Photon and Neutron communities and for HEP and astronomy

📅 2014   📅 2021   📅 2022   📅 2023   📅 2023

DESY coordinating two projects proposed to be financed by NFDI to promote FAIR principles

DAPHNE   **PUNCH**

DESY supporting the HIP platform for a sustained federation of image processing know-how

DESY awarded by BMBF and Helmholtz AI with AI-projects for ML-based data analysis pipelines

MDMLA, UniSEF   **DeGeSim**

Integrating DESY in horizontal infrastructures

Learning to work with modern cloud technologies

Our multi-domain skill-set for industry-standard cloud technologies benefiting users

- PaN portal working on Kubernetes cluster at DESY as part of the EGI FedCloud

- HIP, HIFIS and HAI projects deploying their pipelines on GitLab and Kubernetes

- Learning to deploy HEP software on Kubernetes

- GitLab and Jupyter Notebooks already available in the DESY Cloud / NAF to HEP users

- DESY fully committed to open source

A federated Authentication and Authorisation Infrastructure using Keycloak

- Integrating our AAI with Keycloak to facilitate integration into Helmholtz AAI, EGI AAI and EduTeams

04. COVID-19

DESY and COVID-19 compute

A combined DESY / IT / Maxwell / ATLAS&CMS / WLCG / Hamburg / Zeuthen / … effort

Direct computing support

- Analysis of Petra-III experiments on Corona virus proteins: giving users prioritized access during data taking and analysis on Maxwell

Computing volunteer effort, directly and with WLCG

- Parts of the Hamburg and Zeuthen compute clusters have been dedicated to Rosetta @Home and Folding @Home

- In agreement with the DESY groups of ATLAS and CMS, the Grid clusters in Hamburg and Zeuthen have taken part in Folding @home activities organised by WLCG

- In a later step, the GPU systems for astro-particle computing in Zeuthen and the NAF in Hamburg have also contributed to this effort

Between April, 1st and September, 30th, ~6 % of the DESY central compute resources has been dedicated to Corona Compute

05. Conclusion

Conclusion

- **DESY is committed to continue playing a leading role in WLCG computing**
  - DESY funding is secured for LHC run 3 and not yet for run 4
  - DESY actively contributes to the development of the next computing model

- **DESY is a player in a federated environment and needs the European Projects**

- **DESY contributes to the development of common services, open source and FAIR principles**

- **It is hard to attract adequate personnel**

Thank you

Contact

DESY. Deutsches
Elektronen-Synchrotron

www.desy.de

Sophie SERVAN

Research and Innovation in Computing

sophie.servan@desy.de

+49 160 7455017