IT-Report.

Status of Scientific Computing

Christian Voß – DESY-IT

PRC 88 12/13 Nov. 2019

November 12, 2019

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

DESY.

Content

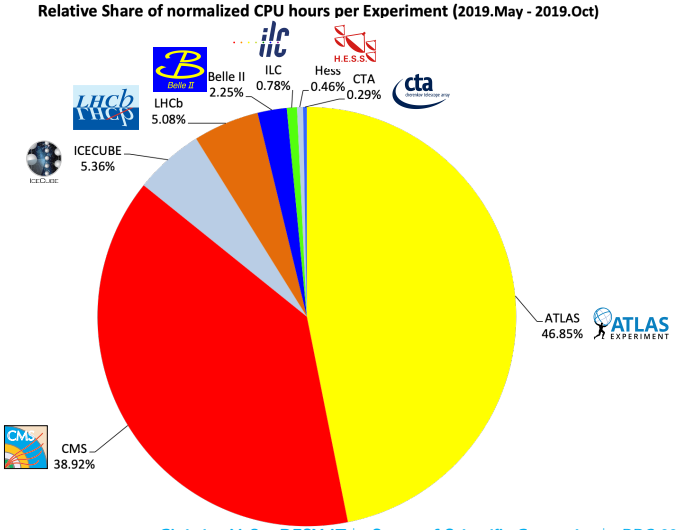
1. Status of Existing Infrastructure
   - Grid: Large Scale Production and Analysis
   - NAF: User Analysis
   - Real-Time Astroparticle Data-Analysis
   - Future Challenges: Belle II Data Taking and LHC Run-III and IV
2. Development to Improve User Experience
   - Support for Modern Data Analysis Methods
   - Remaining Challenges to Improve Usability
3. Future Plans for Computing at DESY
   - Interdisciplinary Data Analysis Facility – PoF IV
   - National, European and Wordwide Projects

Content

1. Status of Existing Infrastructure
   - Grid: Large Scale Production and Analysis
   - NAF: User Analysis
   - Real-Time Astroparticle Data-Analysis
   - Future Challenges: Belle II Data Taking and LHC Run-III and IV
2. Development to Improve User Experience
   - Support for Modern Data Analysis Methods
   - Remaining Challenges to Improve Usability
3. Future Plans for Computing at DESY
   - Interdisciplinary Data Analysis Facility – PoF IV
   - National, European and Wordwide Projects

Overview: DESY Grid Computing Hamburg and Zeuthen

Resources for the supported HEP and Astro-Particle Experiments



Supported in Hamburg

- ATLAS
- Belle-II
- CMS
- ILC
- LHCb

Supported in Zeuthen

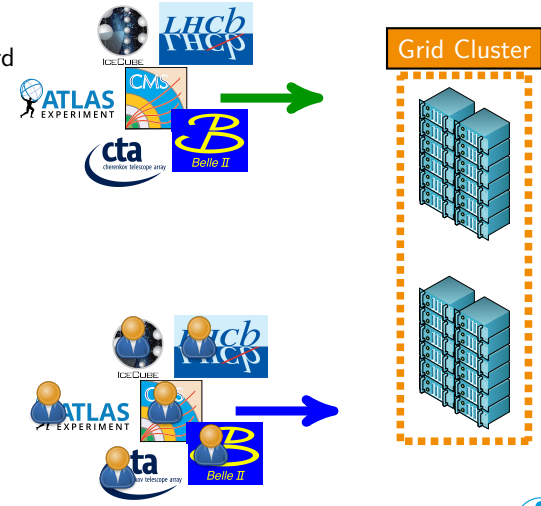
- ATLAS
- CTA
- HESS
- ICECUBE

Overview: Interactive and Batch Computing

High Throughput Computing

> Support different usage modes with the same hard and software

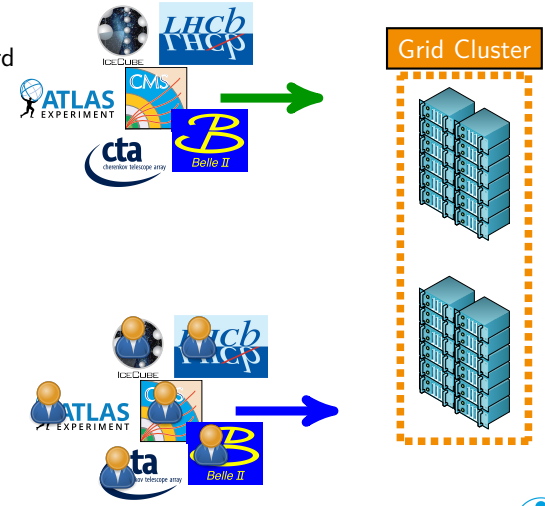
1. Global production and simulation
2. Interactive user analysis



Overview: Interactive and Batch Computing

High Throughput Computing

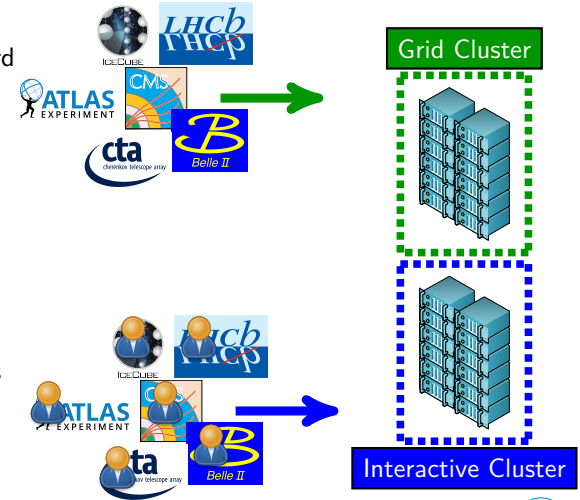
- Support different usage modes with the same hard and software
  1. Global production and simulation
  2. Interactive user analysis
- User analysis wants
  1. Fast turn-around of jobs
  2. Lower job latency
  3. Fast scratch space



Overview: Interactive and Batch Computing

High Throughput Computing

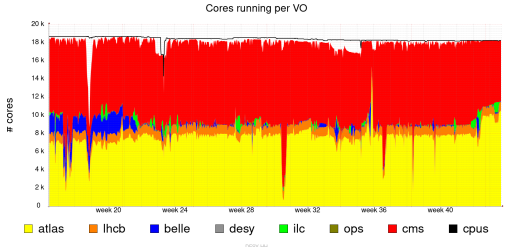
- Support different usage modes with the same hard and software
  1. Global production and simulation
  2. Interactive user analysis
- User analysis wants
  1. Fast turn-around of jobs
  2. Lower job latency
  3. Fast scratch space
- Split the HTC cluster in two independent entities
  1. Grid Cluster
  2. Interactive Cluster



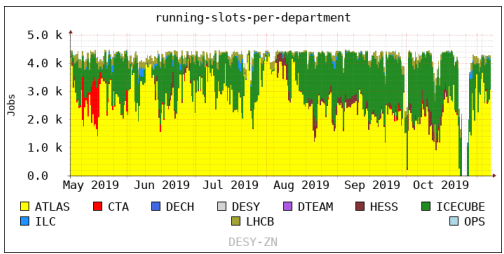
DESY Grid Computing

Global Production and Simulation

Grid Cluster – CPU Utilisation



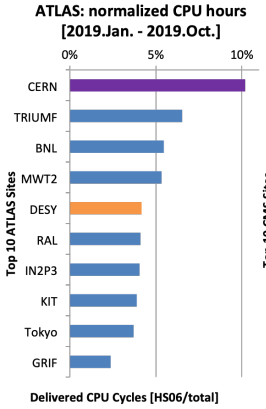
DESY-Hamburg

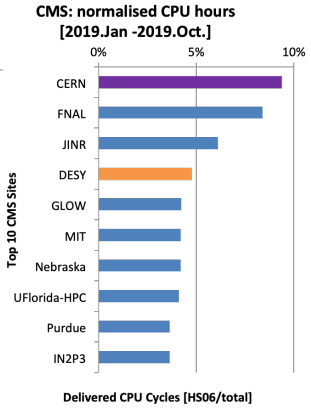


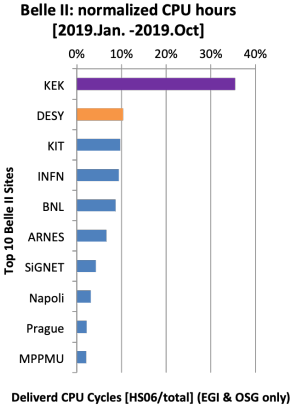
DESY-Zeuthen

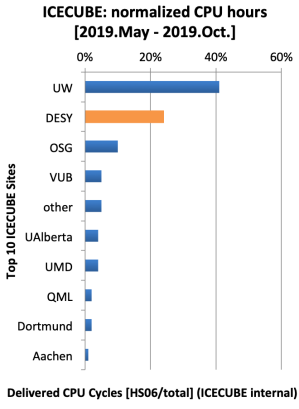
DESY Grid Computing

Computing Contributions to Major Supported Experiments (EGI and OSG Accounting only)









National Analysis Facility

High Throughput Computing for Interactive User Analyses

Purpose of the NAF

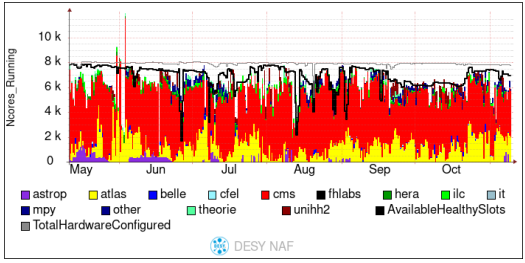
> Platform for user jobs

> Lower latency and smaller jobs

> Faster turn around time

→ Aim for less than 75% utilisation

NAF Cluster Utilisation – October 2019



National Analysis Facility

High Throughput Computing for Interactive User Analyses

Purpose of the NAF

> Platform for user jobs

> Lower latency and smaller jobs

> Faster turn around time

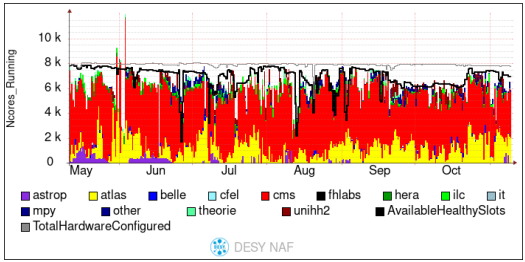
→ Aim for less than 75% utilisation

NAF CPU Usage above 75%

> Affecting User Experience

→ Increase number of cores

NAF Cluster Utilisation – October 2019



National Analysis Facility

High Throughput Computing for Interactive User Analyses

Purpose of the NAF

> Platform for user jobs

> Lower latency and smaller jobs

> Faster turn around time

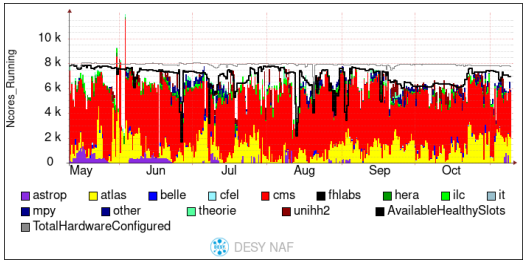
→ Aim for less than 75% utilisation

NAF CPU Usage above 75%

> Affecting User Experience

→ Increase number of cores

NAF Cluster Utilisation – October 2019



GPU Resources within the NAF

> Interactive development servers

> Batch nodes for production runs

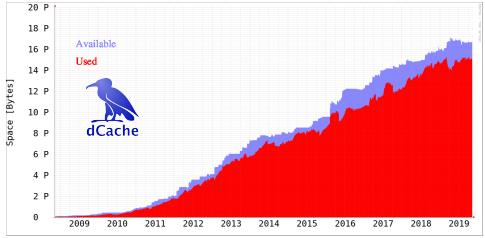
Long Term Storage

HEP dCache installations at DESY

DESY Storage

- DESY a major EU Grid Storage site
- Platform: dCache – co-developed by DESY
- Capacity: Hamburg 16PB, Zeuthen 6.3PB

HEP storage over time (Hamburg)



Network

Recent and Future Improvements

DESY Network Improvements

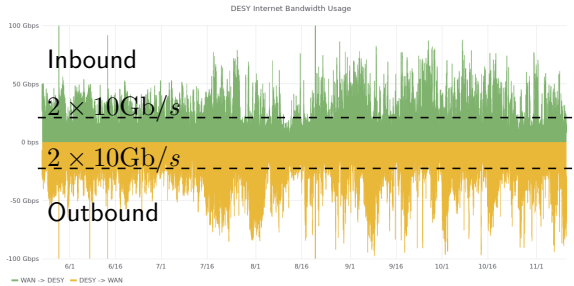
- WAN connection (PRC87)
- Connection Hamburg – Zeuthen
  - Bandwidth in Hamburg: $2 \times 50\mathrm{Gb}/s$
  - Zeuthen – Hamburg: $4 \times 10\mathrm{Gb}/s$

Global Improvements – BELLA

EU funded Project for $100\mathrm{Gb}/s$:

- Direct connection Europe – South America
- Connect EU and SA research communities
- Connectivity for observatories in the Andes

Network-Bandwidth Usage (Hamburg)



AMPEL Framework for Zwicky Transient Facility

Alert Management, Photometry, and Evaluation of Light curves

General Idea: to automate the search for new cosmic light sources

> Message-streaming based analysis framework:
  1. Observatory produces a message containing time, direction, intensity, meta-data
  2. Messages can be consumed/analysed at any AMPEL centre

Live AMPEL instance at DESY-Zeuthen

> Alerts in place since June 2018

> During first half year
  1. $64\mathrm{M}$ alerts
  2. Latency from shutter close to ingestion at DESY $\sim 10\,\mathrm{min}$
  3. Real-Time AMPEL analysis accounts for $1/3$ of Supernovae discovered in the northern sky since summer $'18$

> Technical details follow a bit later

Belle II Resource Demands

CPU and Storage Resources pledged the Belle II Collaboration

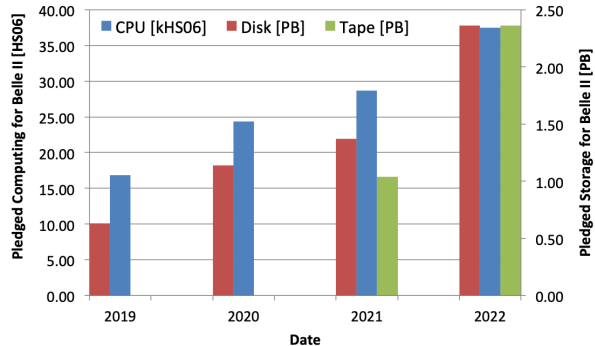
DESY will become a Raw Data Centre for Belle II

CPUs: Pledged Resources

- By 2022 similar pledges to CMS and ATLAS

Storage: Pledged Resources

- By 2022 Belle II surpasses ATLAS
- Very limited group disk requirements
- Tape resources from 2021 onwards
- Tape access needs to be balanced with on-site Photon Sources



Towards LHC Runs III and IV

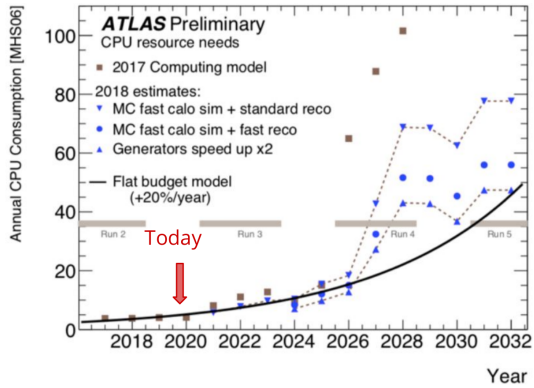
Upcoming Challenges in HEP Computing

Towards HL-LHC

- Run-III → increased Pile-Up per events
- Run-IV → steep increase in luminosity

HL-LHC Computing

- Increase in luminosity ∼ increase in compute resource demands
- Bridge expected funding gap
  1. Optimise Code/new Event models
  2. Opportunistic CPU cycles (MC prod.)
- How to optimise storage?
  1. Smaller data containers
  2. Turbo Stream Model (LHCb)



Content

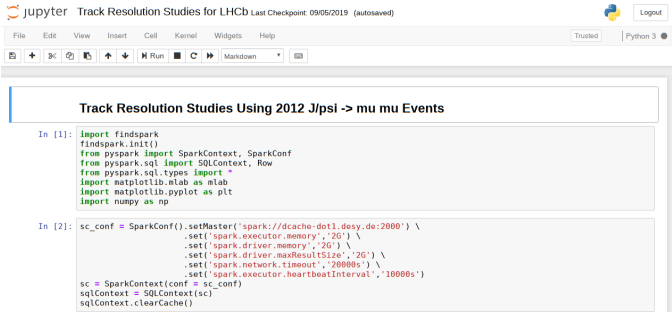
1. Status of Existing Infrastructure
   - Grid: Large Scale Production and Analysis
   - NAF: User Analysis
   - Real-Time Astroparticle Data-Analysis
   - Future Challenges: Belle II Data Taking and LHC Run-III and IV
2. Development to Improve User Experience
   - Support for Modern Data Analysis Methods
   - Remaining Challenges to Improve Usability
3. Future Plans for Computing at DESY
   - Interdisciplinary Data Analysis Facility – PoF IV
   - National, European and Wordwide Projects

Support for Notebooks

New Focal Point for Data Analysis

What are Jupyter Notebooks – Data Analysis in Your Browser

- Python based interpreter for `Python`,`C++`/`ROOT`,...
- Heavy data lifting happens on WGS/WN
- Access via Web-Browser



Support for Notebooks

New Focal Point for Data Analysis

What are Jupyter Notebooks – Data Analysis in Your Browser

- Python based interpreter for `Python`,`C++/ROOT`,...
- Heavy data lifting happens on WGS/WN
- Access via Web-Browser

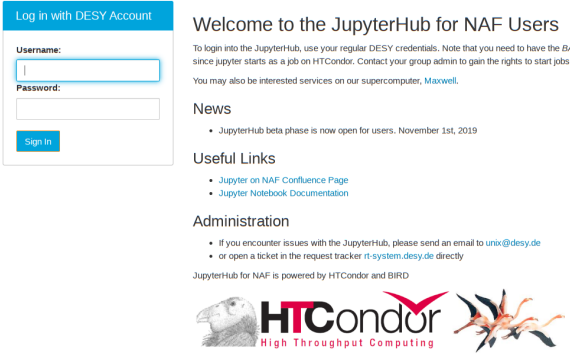
Support for `Jupyter` Notebooks requested by Belle II for some time

- Existing `Jupyter`-Hub on Maxwell but reserving whole nodes
- Running `Jupyter` on dynamic batch resources a difficult endeavour
  → sparked major interest in the `HTCondor` community
- Work done in close collaboration with HTCondor development team – (Thanks)
- Support for custom user-specific Kernels (Thanks to M. Ritter@LMU)

NAF Jupyter-Hub

Ready for Access

Deutsches Elektronen-Synchrotron DESY
A Research Centre of the Helmholtz Association



Access to Jupyter-Hub

- Officially announced
- Documentation for setups
- Feel free to test
- Feel free to send us feedback

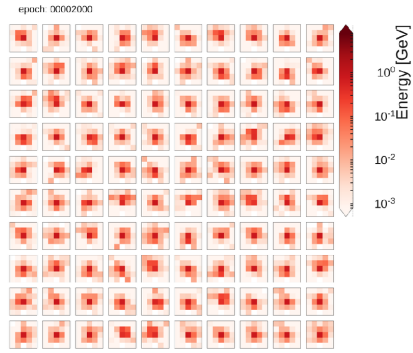
Machine Learning Infrastructure

Improving on Existing Data Analysis Methods

- HEP scientists probe the Standard Model with ever increasing precision
- Need for improvements in Signal-to-Background ratios to find ever rarer signatures

Machine Learning as Tool of Choice

- Observe user interest to improve
  1. Trigger sensitivity → higher rates/improved results
  2. b-Tagging, tracking, User analysis,...
- Neural Networks highly parallelizable → Using GPUs
- DESY-IT hosts dedicated ML-seminar well attended across divisions
- DESY-IT organises ML round table (next 29/11/19)
- DESY-IT active in ML projects with Photon Science and HEP – HAF, AmaleA



Calorimeter Hit simulation with NNs (T.Ferber/Belle II)

Some Open Challenges to Improve Usability

Some Possible Road-Blocks for Scientific Users

Parallelism and Scaling

> Batch systems require manual data splitting and merging
Possible Solution: Apache Spark – collaboration with Belle II

> Batch resources on Demand
Possible Solution: Cloud Batch nodes on demand – Collaboration with University of Victoria

Analysis Software Development and Deployment

> Good Practises in Software Engineering not well established, especially user context
Possible Solution: Dev. pipeline – Prototype on EGI Cloud/European Open Science Cloud

Automated Workflows

> Workflows usually depend on personal communication (bound to specific scientists)
Possible Solution: Message Systems and Function as a Service – AMPEL Cluster and EOSC

Analysis on Arrival of Data

Experimental Events, Storage Events and Function as a Service

Messages: Producer – Broker – Consumer Model

- Any Frame from observatory delivers a notification to a message system
- Message arrival triggers further actions at consumer, *e.g.* filter or restructure
- Arrival at AMPEL cluster triggers final analysis and selection
- Result produces new message → send and consumed by scientific portals

Analysis on Arrival of Data

Experimental Events, Storage Events and Function as a Service

Messages: Producer – Broker – Consumer Model

- Any Frame from observatory delivers a notification to a message system
- Message arrival triggers further actions at consumer, *e.g.* filter or restructure
- Arrival at AMPEL cluster triggers final analysis and selection
- Result produces new message → send and consumed by scientific portals

Storage Events – dCache as Workflow-Engine

- Idea driven by DESY `dCache` development team
- Any transactions delivers a notification to a message system
- Message arrival triggers further actions → Batch jobs

Content

1. Status of Existing Infrastructure
   - Grid: Large Scale Production and Analysis
   - NAF: User Analysis
   - Real-Time Astroparticle Data-Analysis
   - Future Challenges: Belle II Data Taking and LHC Run-III and IV
2. Development to Improve User Experience
   - Support for Modern Data Analysis Methods
   - Remaining Challenges to Improve Usability
3. Future Plans for Computing at DESY
   - Interdisciplinary Data Analysis Facility – PoF IV
   - National, European and Wordwide Projects

Scientific Computing and PoF-IV

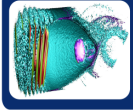
Consolidating Interdisciplinary Efforts into Matter And Technology

Changes with the Upcoming PoF-IV Funding Period

> Creation of a new topic within *Matter And Technology* – Data Management and Analysis

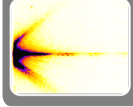
ST1: The Matter Information Fabric

- IT Solutions (Hard+Software) for Large Facilities
- Automated Data Lifecycle Management (LK II)
- Solution for Community Specific Demands



ST2: The Digital Scientific Method

- Method Research in Data Analysis & Simulation
- Machine Learning, Visual Analytics, Scientific Workflows
- Heterogeneous HPC, HTC, I/O
- New technologies, e.g. Quantum Computing



ST3: The Digital Experiment and Machine

- Start-to-End Simulations (Machine/Physics/Detector)
- Fast Feedback („Human in the Loop")
- Determination of Data Quality, Handling of Metadata
- Control Systems

Scientific Computing and PoF-IV

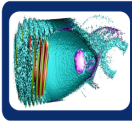
Consolidating Interdisciplinary Efforts into Matter And Technology

Changes with the Upcoming PoF-IV Funding Period

> Creation of a new topic within *Matter And Technology* – Data Management and Analysis

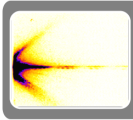
ST1: The Matter Information Fabric

- IT Solutions (Hard+Software) for Large Facilities
- Automated Data Lifecycle Management (LK II)
- Solution for Community Specific Demands



ST2: The Digital Scientific Method

- Method Research in Data Analysis & Simulation
- Machine Learning, Visual Analytics, Scientific Workflows
- Heterogeneous HPC, HTC, I/O
- New technologies, e.g. Quantum Computing



ST3: The Digital Experiment and Machine

- Start-to-End Simulations (Machine/Physics/Detector)
- Fast Feedback („Human in the Loop")
- Determination of Data Quality, Handling of Metadata
- Control Systems

Scope

> Involves all HH-Departments

> Truly interdisciplinary effort

Interdisciplinary Data Analysis Facility – IDAF

Evolution of the Tier-2 Centre into Matter And Technology

Changes with the Upcoming PoF-IV Funding Period

> The Tier-2 centre becomes the *LK2* **IDAF** within *DMA* and *Matter And Technology*

> Consolidation and sharing of the existing computing infrastructure with all pros/cons

1. Access to the NAF for PETRA-III/FLASH users
2. Extended access to HPC resources for ATLAS/Belle II/CMS beyond backfilling with Grid Jobs
3. Retain essentially the Grid cluster and the HPC buy-in model

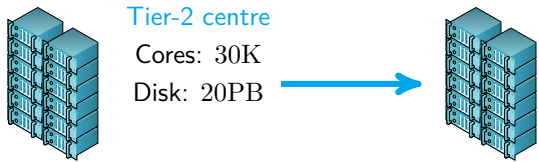
> Gain additional funding for personnel/hardware – did not initially materialise

Interdisciplinary Data Analysis Facility – IDAF

Evolution of the Tier-2 Centre into Matter And Technology

Changes with the Upcoming PoF-IV Funding Period

- The Tier-2 centre becomes the *LK2* **IDAF** within *DMA* and *Matter And Technology*
- Consolidation and sharing of the existing computing infrastructure with all pros/cons
  1. Access to the NAF for PETRA-III/FLASH users
  2. Extended access to HPC resources for ATLAS/Belle II/CMS beyond backfilling with Grid Jobs
  3. Retain essentially the Grid cluster and the HPC buy-in model
- Gain additional funding for personnel/hardware – did not initially materialise



National, European and Wordwide Projects

Personel and Hardware Resources for Inhouse R&D

Participation in National Projects

- National research data initiative (NFDI) *e.g.* Pan-Pahn (HEP) and ASTRO (Astroparticle)
- The Helmholtz Data Federation (HDF)
- Helmholtz Incubator "Information & Data Science"
  - Helmholtz Infrastructure for Federated ICT Services (HIFIS)
  - Helmholtz Imaging Platform (HIP)
  - Helmholtz Information & Data Science Academy (HIDA)
  - Helmholtz Artificial Intelligence Coordination Unit (HAICU)
  - Helmholtz Metadata & Knowledge Platform
- PoF-IV Preparation
  - Data Management and Analysis
  - Cross Community Activities (CCA) with Jülich on Artificial Intelligence and Quantum Computing

National, European and Wordwide Projects

Personel and Hardware Resources for Inhouse R&D

Participation in Worldwide Projects

- Close collaboration between Helmholtz Centres with Canadian Laboratory TRIUMF on Machine Learning, Big Data and Quantum Computing
  → 16-17 September 2019 Workshop on QC, ML & infrastructure DESY
- World-Wide-LHC-Computing Grid
  - ESCAPE – Datalake and new AAI methods

Participation in European Projects

- European Open Science Cloud
  - PaNOSC – Close collaboration with EuXFEL
- Archiver as follow-up to HNSciCloud – data on commercial clouds
- eXtreme DataCloud – storage quality of service
- EXPANDS

Summary

PRC 88 DESY-IT

- DESY continues to deliver significant resources for the HEP and Astro-Particle Physics
  - Grid – part of a global effort
  - NAF – part of the local effort
- DESY continues to investigate new methods of data access and analysis
  - Jupyter-Hub
  - Adopt modern workflows and tools
- Continue to represent our user communities in international projects shaping the computing environment
  - Challenges of upcoming new experiments
  - Use new technologies to improve user experience
  - Major source for personnel for in-house R&D projects
- DESY adapts long-term plans with regards to computing
  - Integration of topic DMA into Matter & Technology
  - Migration of NAF to IDAF