Facility for Antiproton and Ion Research (FAIR)

(under construction in Darmstadt, Germany)

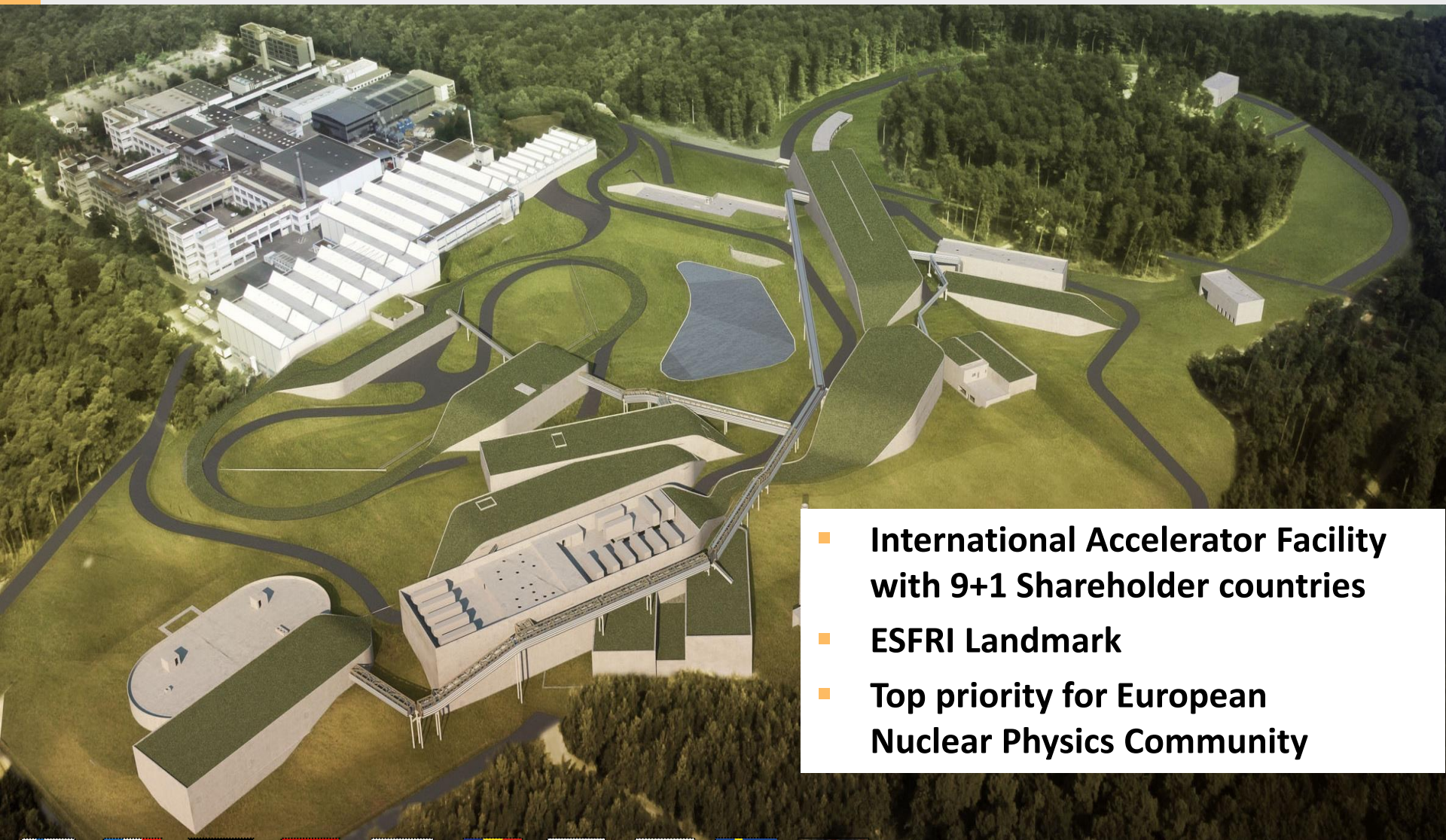
Status of FAIR

- challenges in data processing
in the context of EOSC and the *FAIR* principles

Kilian Schwarz
(with slides from J. Eschke and M. Al-Turany)
GSI GmbH

ESCAPE WP2/5 meeting, Amsterdam, 01 July 2019

FAIR: Facility for Antiproton and Ion Research – A World-Wide Unique Accelerator Facility



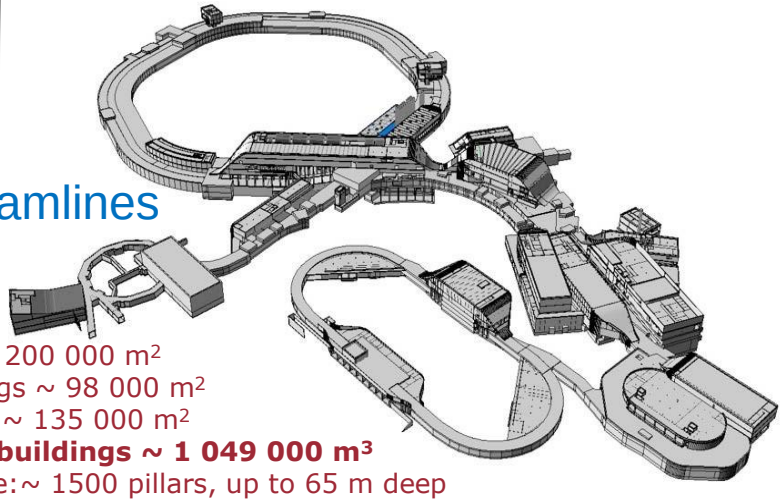
- International Accelerator Facility with 9+1 Shareholder countries
- ESFRI Landmark
- Top priority for European Nuclear Physics Community

Finland France Germany India Poland Romania Russia Slovenia Sweden UK

Status of FAIR Project: Civil Construction

Progress since official start on 4th of July 2017

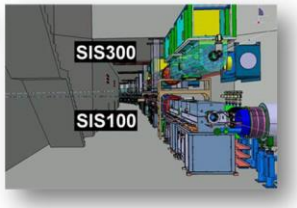
Inside the tunnel of the FAIR ring accelerator SIS100



3.2 km beamlines

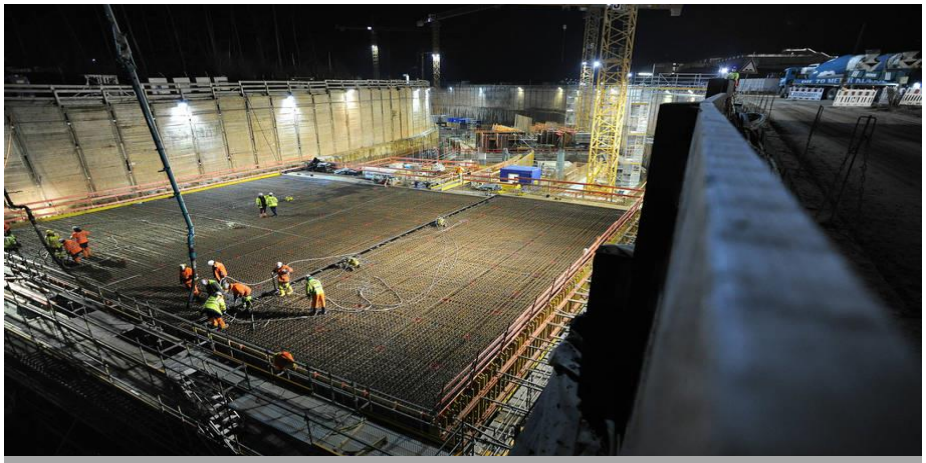
Total area > 200 000 m²
Area buildings ~ 98 000 m²
Usable area ~ 135 000 m²
Volume of buildings ~ 1 049 000 m³
Substructure: ~ 1500 pillars, up to 65 m deep

SIS100/300 tunnel



Status of FAIR Project: Civil Construction

Excavation SIS100 tunnel



First tunnel slab is being poured at night

Upgraded SIS18 completed ready for FAIR and FAIR phase 0

Excavation transfer building & CBM cave

construction timeline:

- civil construction completed in 2023
- installation of accelerators and experiments 2022 - 2024
- start of pilot beams in 2025

Facility for Antiproton & Ion Research

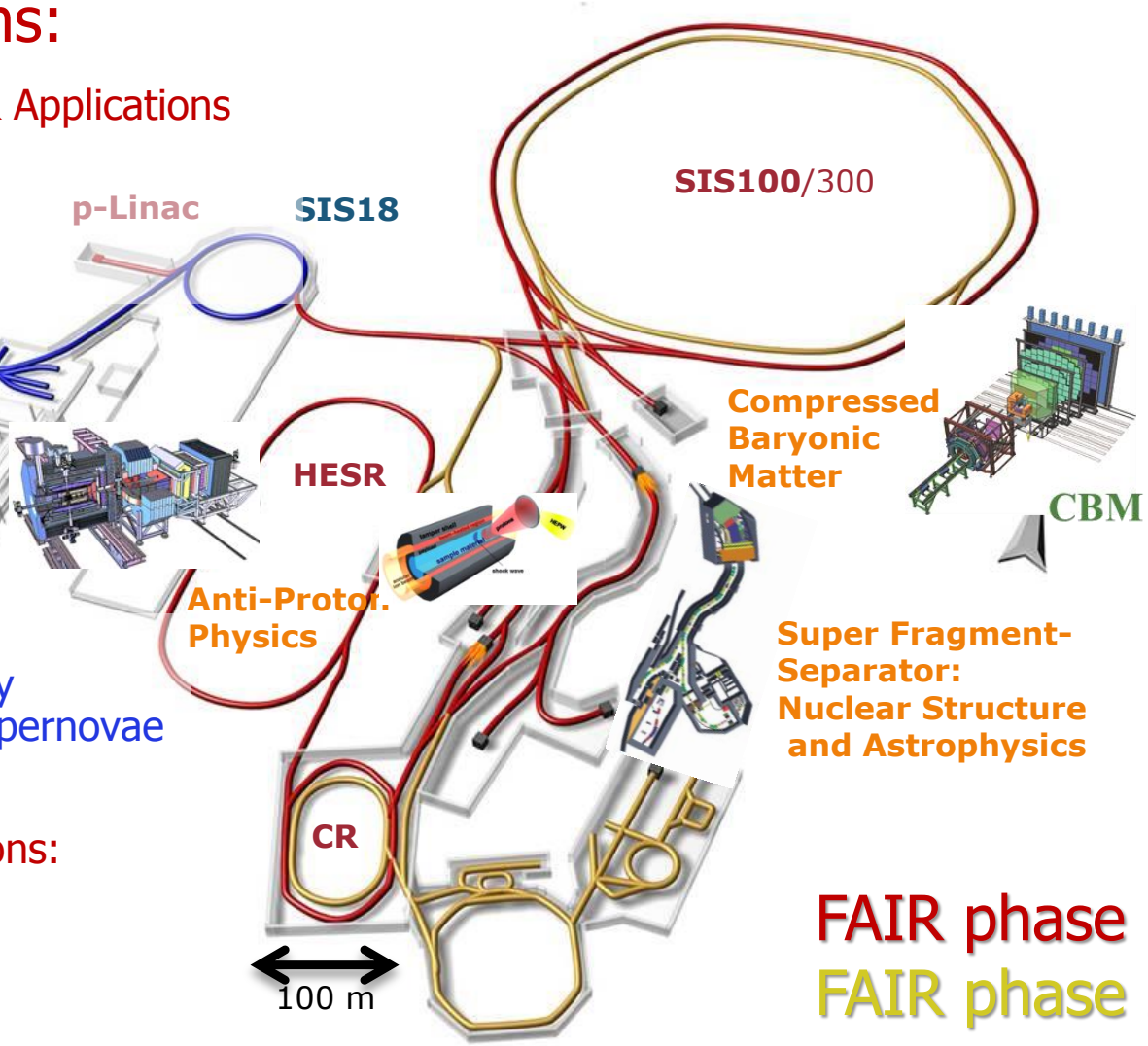
Experimental programs:

APPA: Atomic & Plasma Physics & Applications
- Highly charged atoms
- Plasma physics
- Radiobiology
- Material science

CBM: Nucleus-nucleus collisions
- Nuclear matter at neutron star core densities
- Phase transitions from hadrons to quarks

NUSTAR: Rare Isotope beams
- Nuclear structure far off stability
- Nucleosynthesis in stars and supernovae

PANDA: Antiproton-proton collisions:
- Charmed hadrons (XYZ)
- Gluonic matter and hybrids
- Hadron structure
- Double Lambda hypernuclei



FAIR phase 1
FAIR phase 2

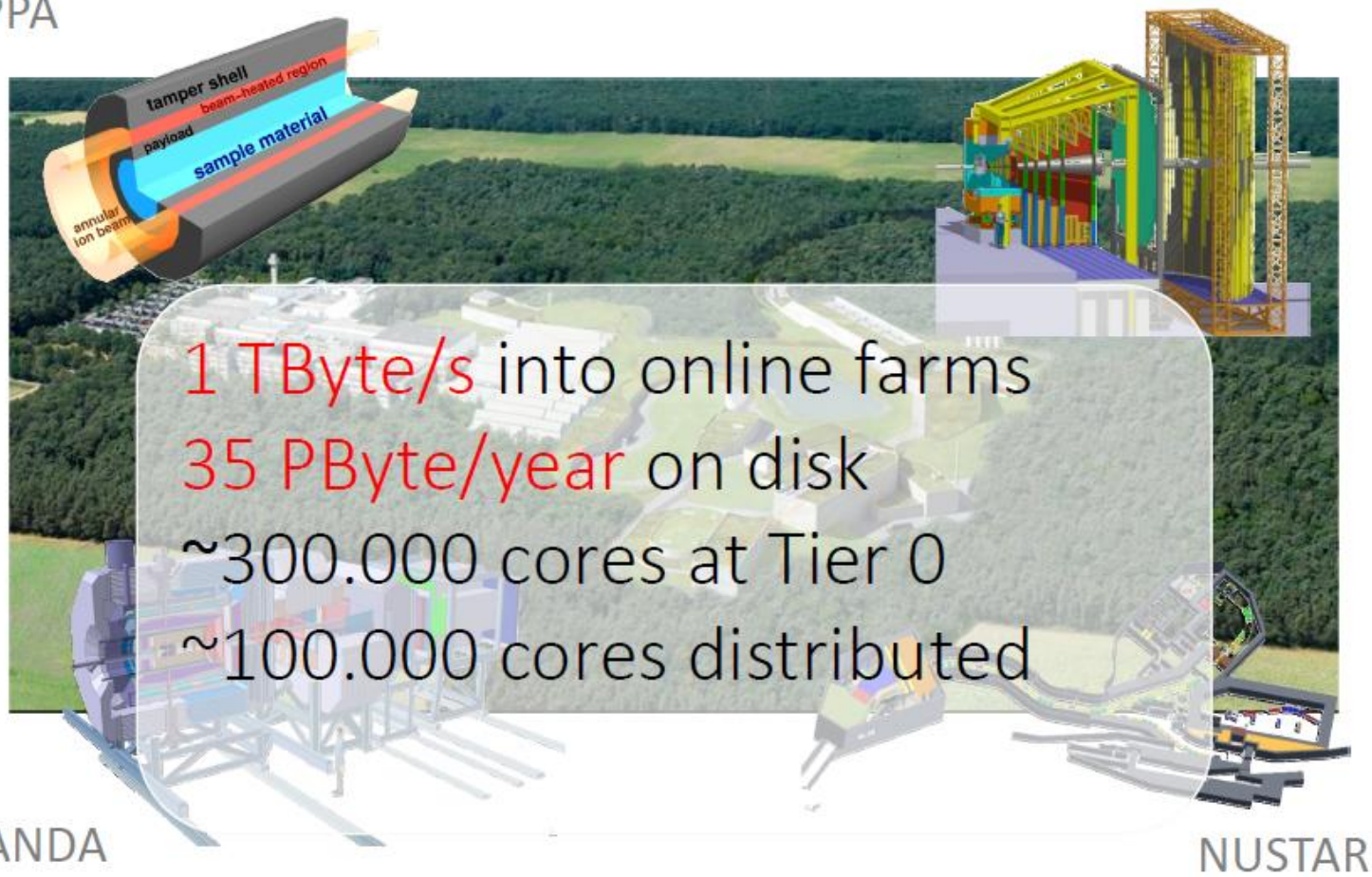
FAIR Collaborations

more than 2500 scientist from ~200 institutions in over 50 countries

Computing at FAIR

APPA

CBM



1 TByte/s into online farms

35 PByte/year on disk

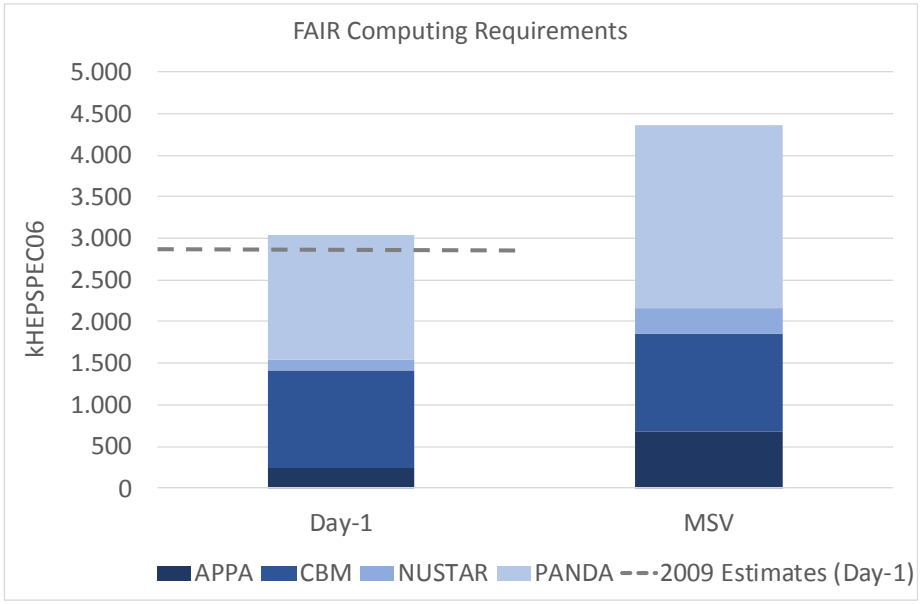
~300.000 cores at Tier 0

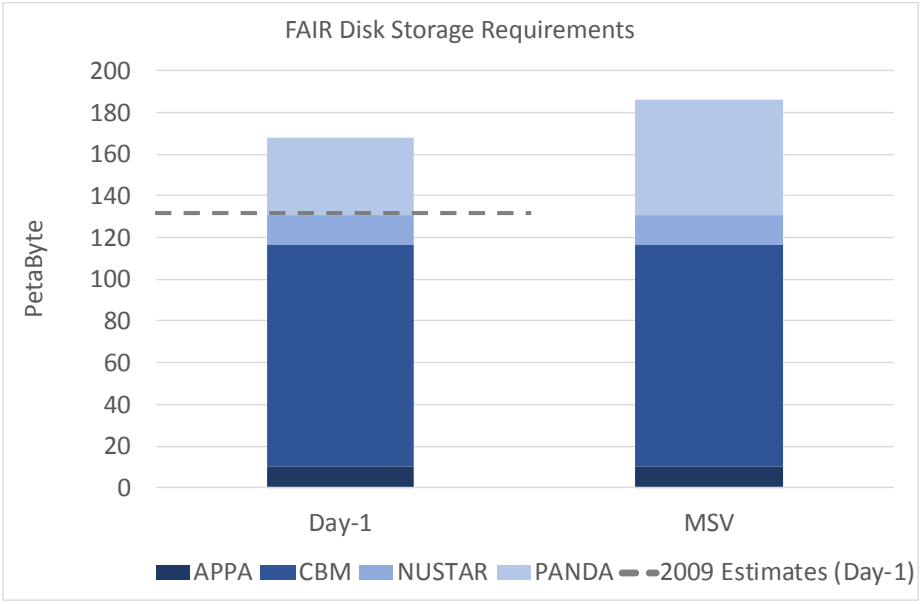
~100.000 cores distributed

PANDA

NUSTAR

Computing – step 1: Experiment requirements determined





Assumptions for resource requirements:
Day-1 and MSV detector setups, nominal accelerator performance, multi-year integrated values (data lifetime)

FAIR Data Center

A common data center for FAIR (Green IT Cube)

CBM FLES
+ 60'000 CPU cores
- To perform online a full event reconstruction on the 1 TB/s input data stream

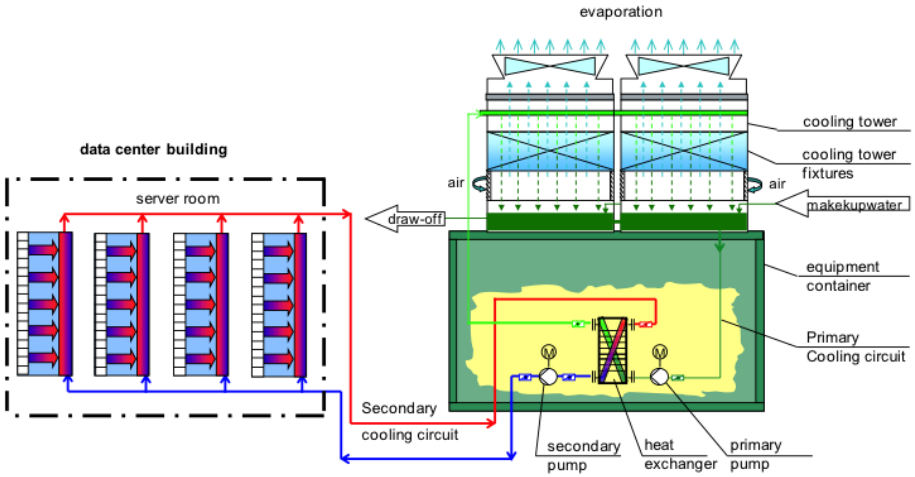
+ ? GPUs
- To speed up the reconstruction

Panda online

+ 66'000 CPU cores
- To perform online a full event reconstruction on the 300 GB/s input data stream

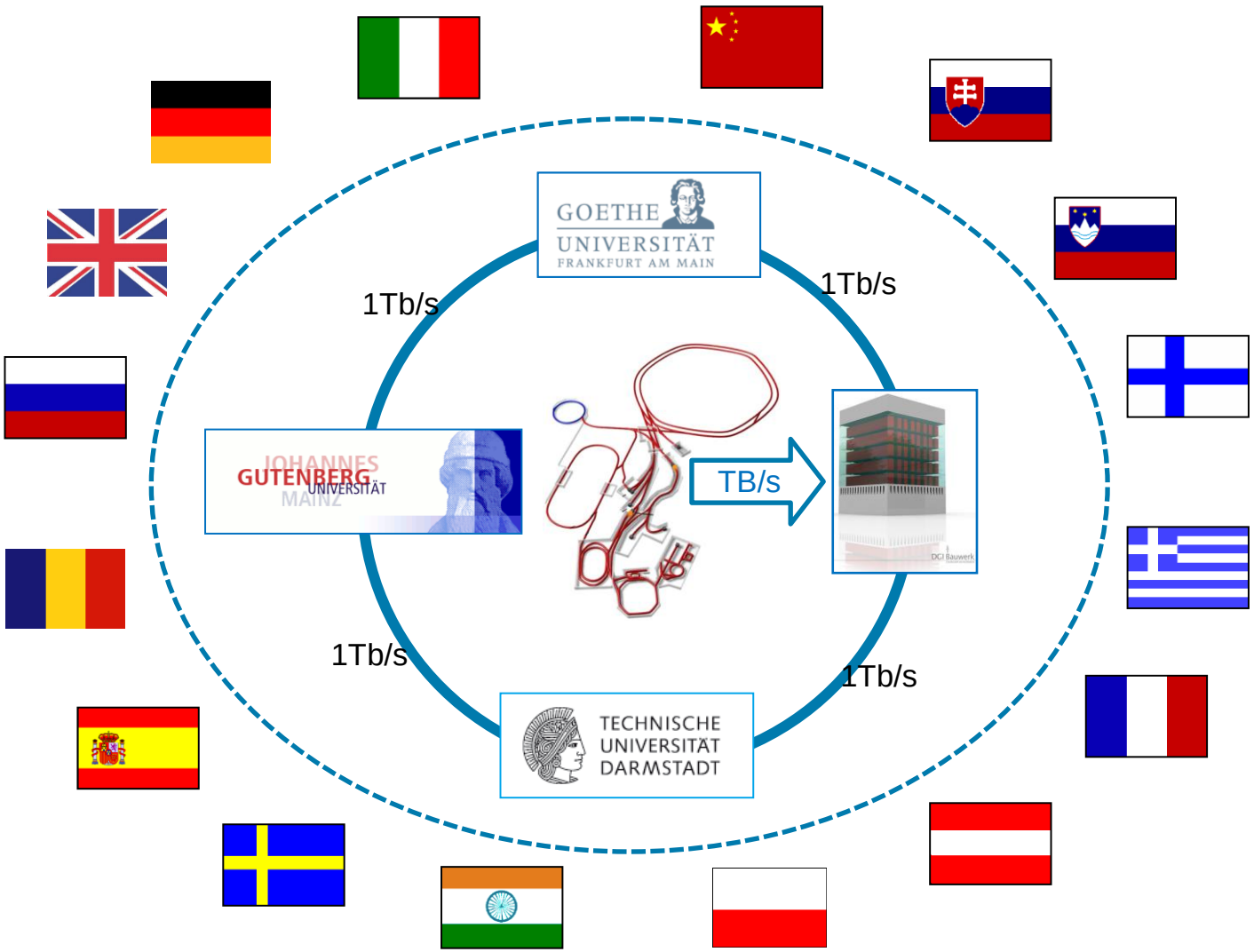
+ ? GPUs
- To speed up the reconstruction

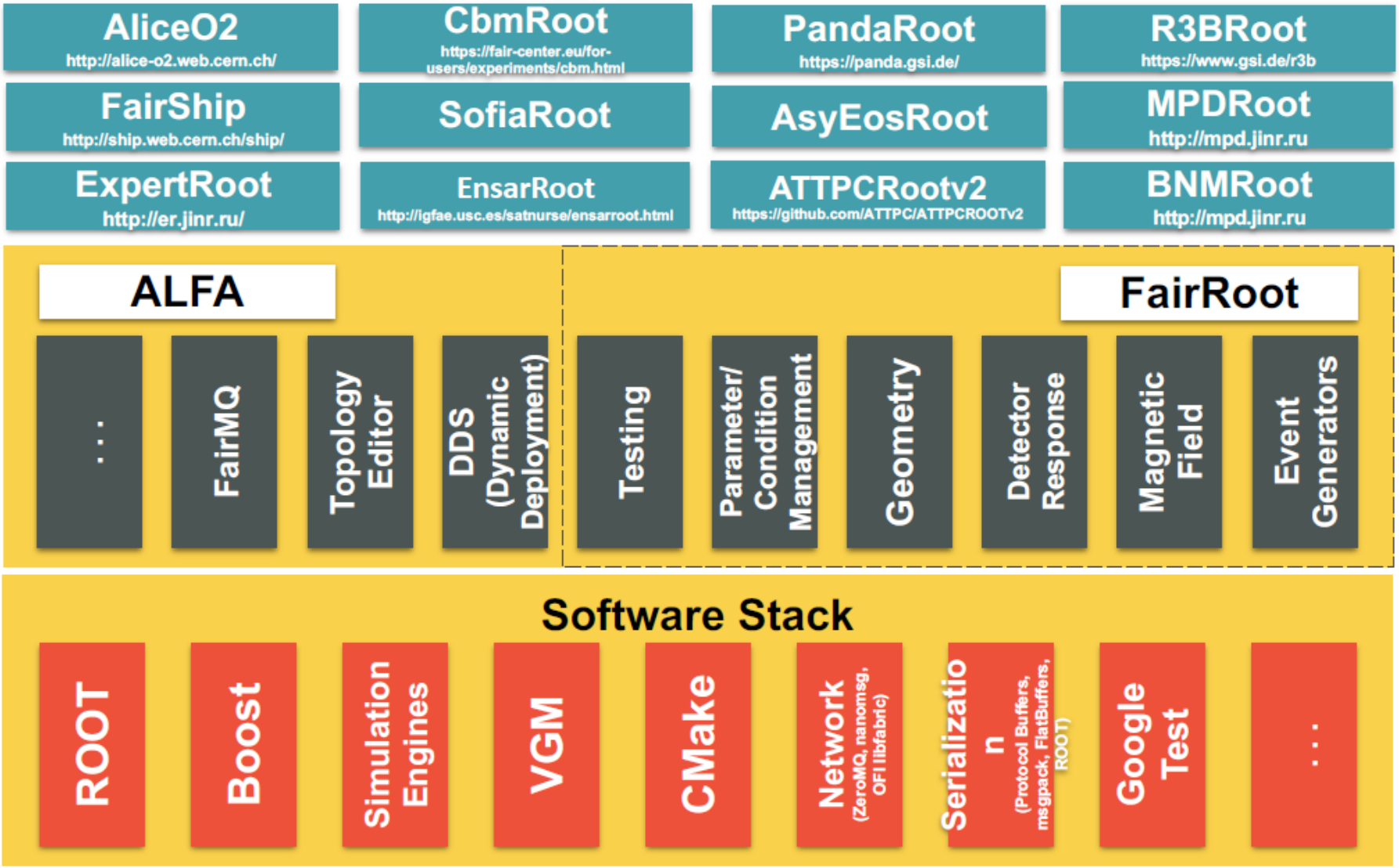
Dynamically allocated resources
for exclusive usage and limited time



6 floors, 4.645 sqm
room for 768 19" racks (2,2m)
4 MW cooling (baseline)
Max cooling power 12 MW
Fully redundant (N+1)
PUE <1.07

FAIR Computing: T0/T1 MAN (Metropolitan Area Network) & Grid/Cloud




AliceO2
http://alice-o2.web.cern.ch/
CbmRoot
https://fair-center.eu/for-users/experiments/cbm.html
PandaRoot
https://panda.gsi.de/
R3BRoot
https://www.gsi.de/r3b
FairShip
http://ship.web.cern.ch/ship/
SofiaRoot
AsyEosRoot
MPDRoot
http://mpd.jinr.ru
ExpertRoot
http://er.jinr.ru/
EnsarRoot
http://igfae.usc.es/satnurse/ensarroot.html
ATTPCRootv2
https://github.com/ATTPC/ATTPCROOTv2
BNMRoot
http://mpd.jinr.ru
ALFA
FairRoot
. . .
FairMQ
Topology Editor
DDS (Dynamic Deployment)
Testing
Parameter/ Condition Management
Geometry
Detector Response
Magnetic Field
Event Generators
Software Stack
ROOT
Boost
Simulation Engines
VGM
CMake
Network (ZeroMQ, nanomsg, OFI libfabric)
Serialization (Protocol Buffers, msgpack, FlatBuffers, ROOT)
Google Test
. . .

FAIR status/requirements computing:

- data management
  - 2 large experiments with similar requirements as LHC and several smaller non HEP like experiments targeting many different areas of research
  - a common data management infrastructure has to be created which fulfills requirements of all experiments.
- software development
  - FairRoot is already being used by all FAIR experiments and additionally by some non FAIR experiments.
  - software needs to support continuous data read out and complex online processing for event selection at high data rates.
  - online and offline processing needs to become faster and more efficient, also by using new architectures and algorithms.

FAIR status/requirements computing:

- data/software access in the context of EOSC
  - In order to be able to publish at least parts of the data FAIR is in the process of developing corresponding MoUs.
  - The FAIR analysis software (FairRoot) should be made accessible via the software and service repository developed in the context of ESCAPE.
- FAIR paradigm
  - the FAIR paradigm is planned to be introduced (at least to a large extend) for a consistent data management system which is being developed based to a large extend on common systems and available technologies. Also a meta data system under consideration of the DOI/data cite requirements is under development.

FAIR status/requirements computing/ESCAPE WP2/questionaire:

- Data production
  - 1 TB/s into online farm, 10 GB/s on disk
  - file size: ALICE: TF-10 GB, CTF-2 GB (FAIR not yet decided)
    - simulated event size (CBM): 250 kB
  - primary data compressed by factor 7 (zip)
  - primary data will be archived
- Data model
  - at least 2 replicas: 1 at T0, 1 at T1
  - data formats: RAW, (ESD), AOD
  - data lifecycles still to be defined
  - processing campaigns are planned, how often still tbd
  - raw data will be reprocessed periodically once per year
  - searchable meta data will most likely be needed

FAIR status/requirements computing/ESCAPE WP2/questionaire:

- Data access and processing
  - raw data processing in quasi-online mode for event selection and data reduction. Combination of FPGAs (1st stage) and CPU/GPU (2nd stage). 1/100th-1/1000th stored on disk. Data are reprocessed offline to generate AOD needed for final analysis.
  - protocols: xroot and root, Apache arrow
  - data cache yes, read ahead cache no
  - QoS only disk (production/non production) and tape
  - file popularity management service needed
  - data are read directly from storage
  - files are read fully or random parts
  - impact of remote data access evaluated
  - no data processing from tape
  - most likely intelligent data lake will orchestrate data movement. Several WMS under investigation
  - CLIs, APIs, Web Interfaces needed
  - HPC facilities used, no MPI jobs but MQ
  - interested in event-driven data processing
  - raw data have replicas, derived formats can be reproduced
  - temporary unavailable data access leads to loss of time

FAIR status/requirements computing/ESCAPE WP2/questionaire:

- Data access control
  - open access after a period of 1 year
  - no anonymous data access
  - all members of collaboration should have access
    - reading no groups with specific privileges, only selected users can write to critical data directories
  - most likely no access right information via database
  - Token based authentification planned (Sci-Tokens)
- AAI contact person
  - not yet clear. For the time being take me
- > 1000 active end-users
- user access via web browser (smaller experiments), via terminal and native application
- currently user authentication via ssh keys and X509, EduGain or similar planned
  - authorisation via Linux groups
- user registration via LDAP & DB and experiment hosted Web services
- data access via http and xrootd protocol. Privileges as user have on local Linux file system

FAIR interest in ESCAPE

- ESCAPE takes place right before the official start of FAIR.
- within ESCAPE essential IT ingredients are being developed, especially infrastructures for distributed data management and computing, which are needed by FAIR.
- FAIR hopes to profit from taking part in ESCAPE by getting important support and ideas for setting up their own infrastructure for distributed computing.