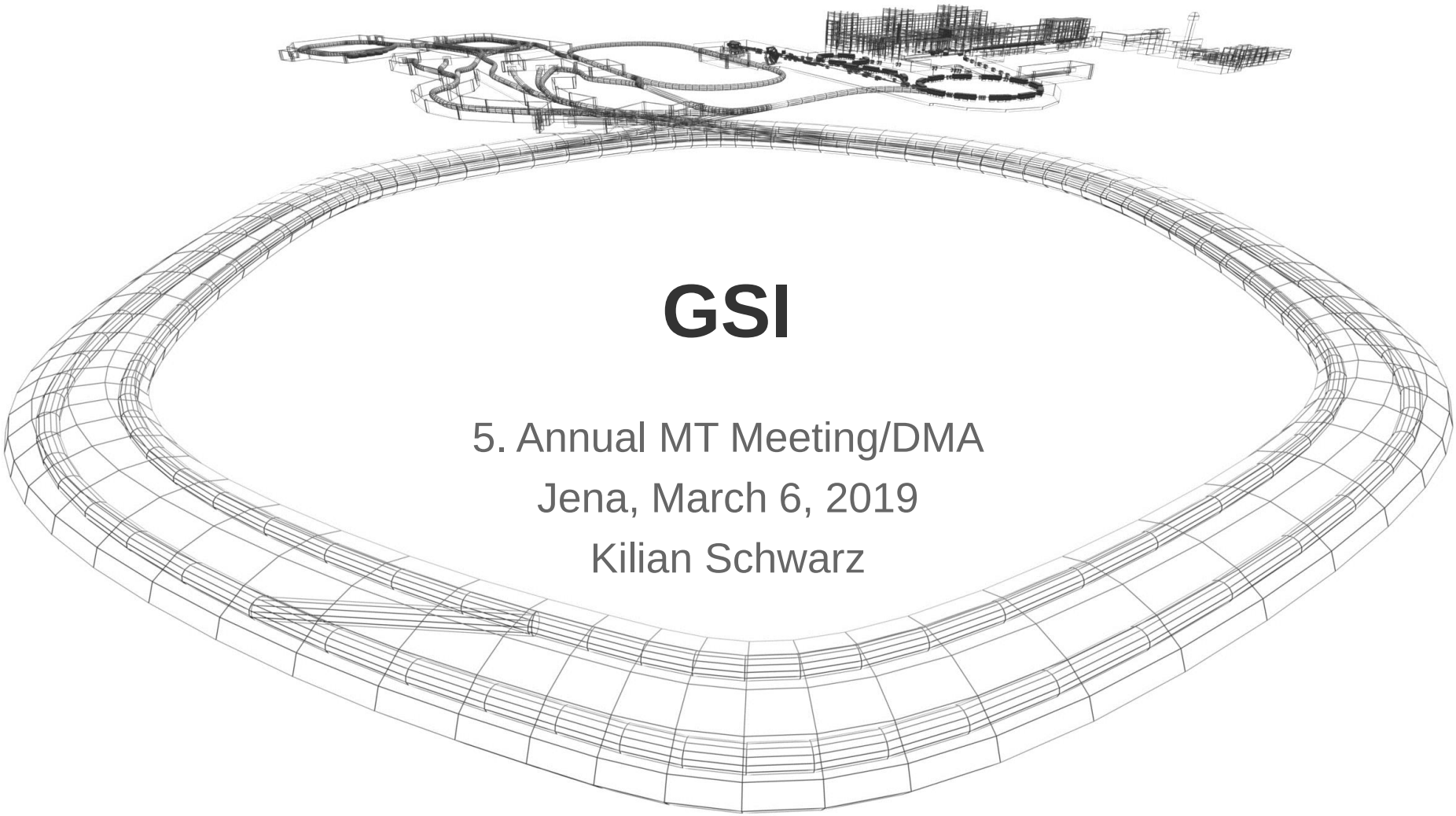


GSI

5. Annual MT Meeting/DMA

Jena, March 6, 2019

Kilian Schwarz

Agenda

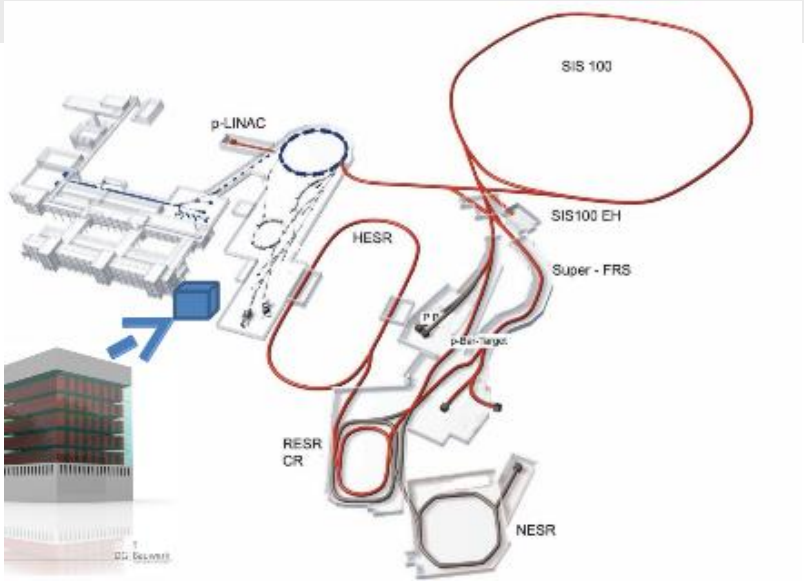
- GSI/FAIR
- activities in ST1
- activities in ST2

GSI: Helmholtz Centre for Heavy Ion Research
FAIR: Facility for Antiproton and Ion Research

GSI computing 2019

ALICE T2/NAF
HADES
LQCD (#1 in Nov‘ 14 Green 500)
~30000 cores
~ 30 PB lustre
~ 9 PB archive capacity

Green IT Cube
Computing
Centre



FAIR computing at nominal operating conditions

CBM
PANDA
NuSTAR
APPA
LQCD
5 M HS06
170 PB disk
40 PB tape/y

View of construction site

- **Open source and community software**
- **budget commodity hardware**
- **support of different communities**
- **manpower scarce**

FAIR construction pictures

Tunnel for SIS100

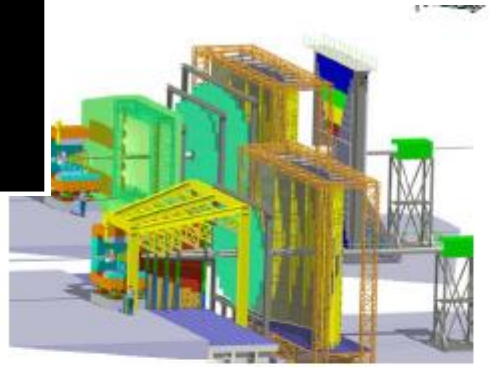
in the tunnel of SIS100

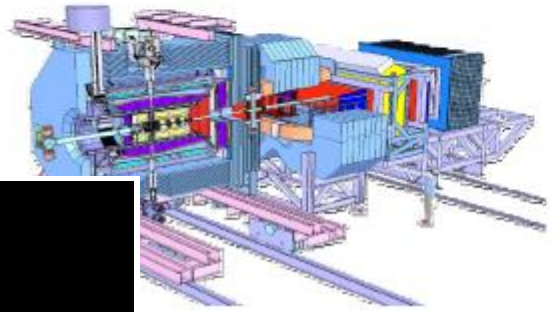
Cover of the first tunnel segment of SIS100

FAIR Computing

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



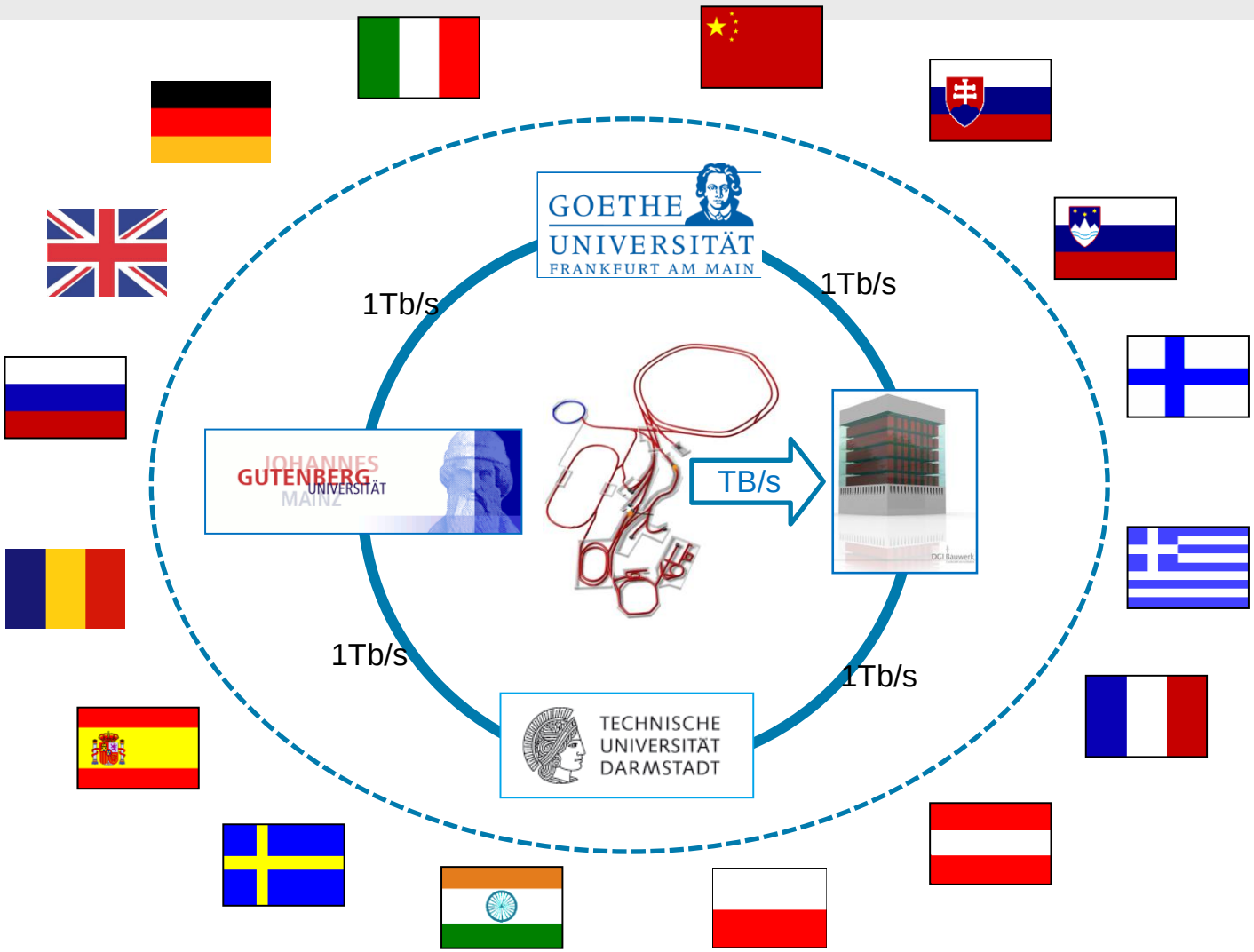


Panda online

+ 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

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



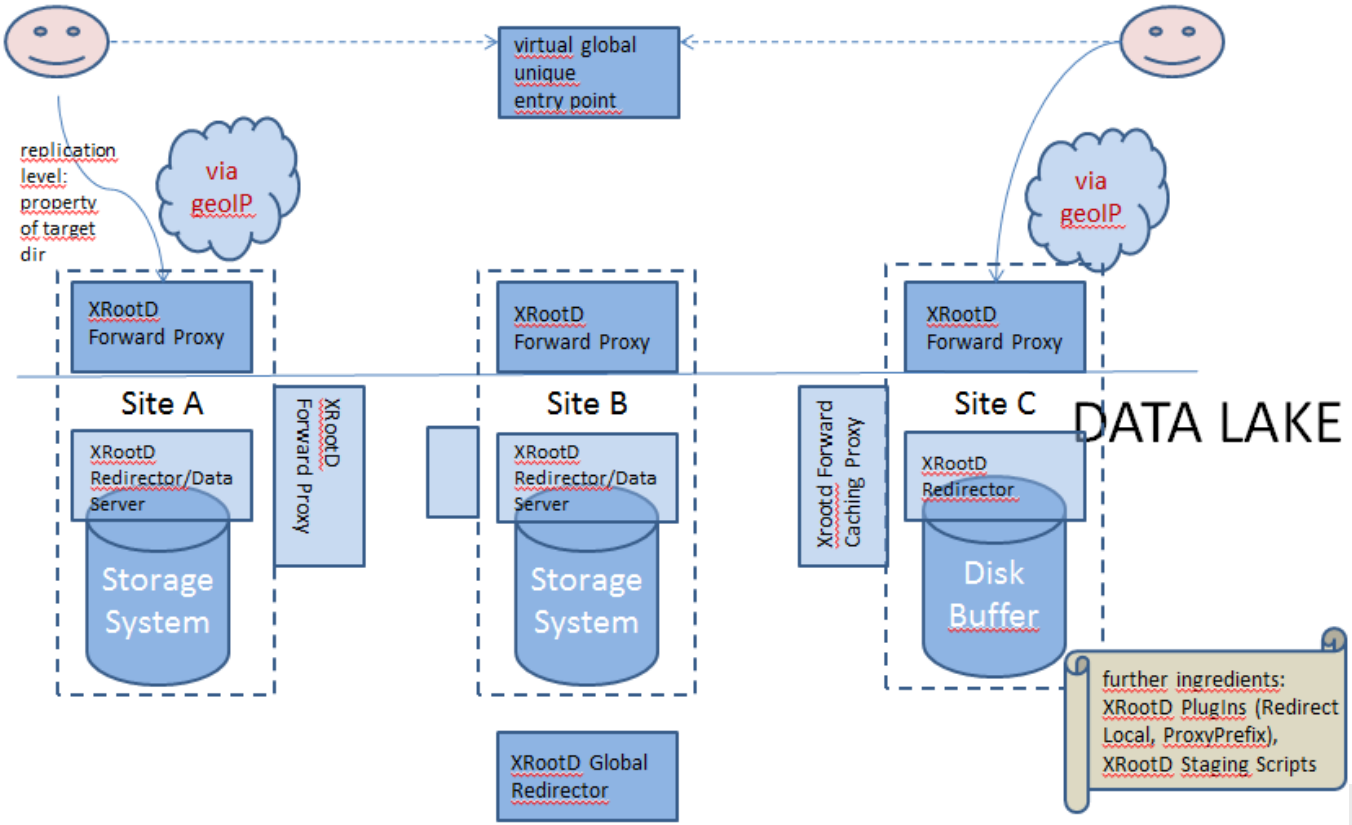
Agenda

- GSI/FAIR
- activities in ST1
- activities in ST2

ST1/Large Scale Scientific Data Lifecycle Management

- Federated storage infrastructures: data lakes
- in context with EU project ESCAPE

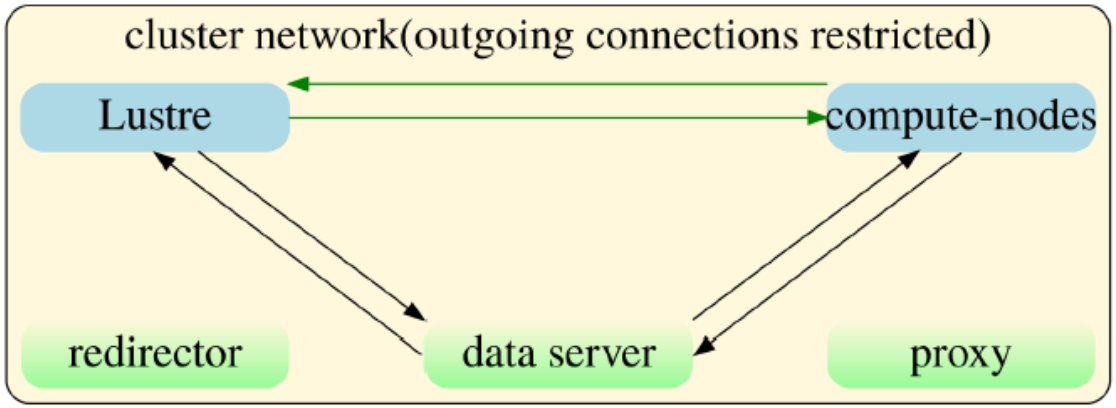
XRootD based
Data Lake design



ST1/Large Scale Scientific Data Lifecycle Management

- efficient high bandwidth data access
- tested with ALICE Analysis Facility Prototype @ GSI

XRootD Redir Plug-in :
Reading via XRootD data servers doubles the network traffic inside the infiniband network. This is, especially with a limited number of XRootD servers, a bottleneck in CPU & bandwidth to our setup.
➔ Clients should open a file directly from Lustre if at GSI by circumventing XRootD data servers

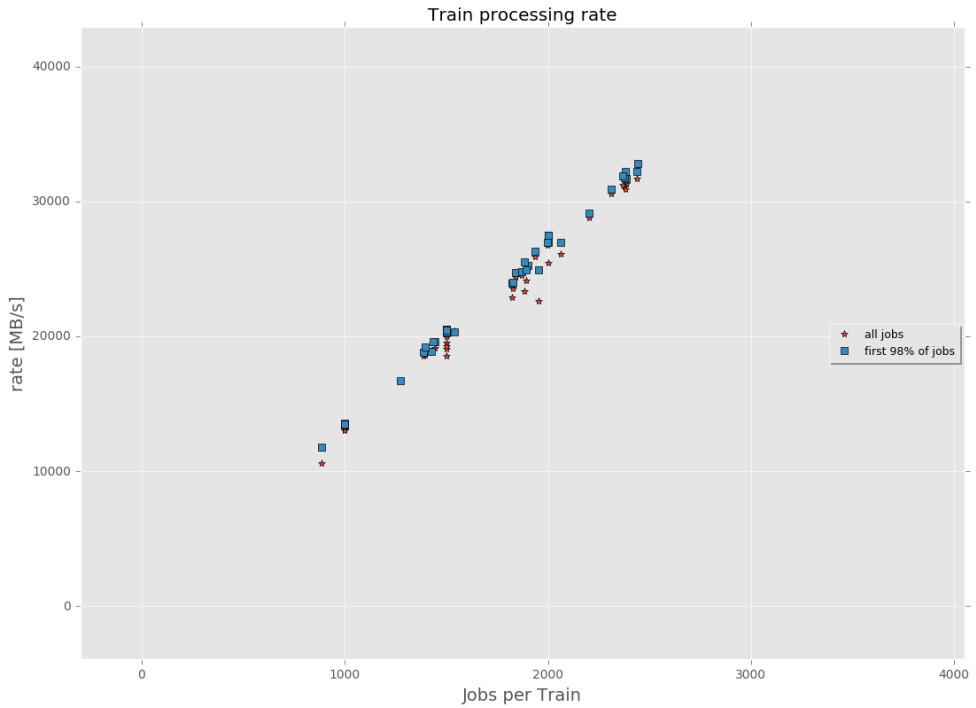


ST1/Large Scale Scientific Data Lifecycle Management

- efficient high bandwidth data access

Server (Redirector) Plugin (cms.ofslib).
v4 Client API (XrdCl) needs to be used
Needed Client code in XRootD base starting with version 4.8
(see commit 76108af & ef28e28 on xrootd/xrootd Github)
pre v4.8.0 clients are redirected to XRootD data servers

Redirection with ROOT client works as long as TNetXNGFile(new XrdCl) is used and ROOT is compiled against XRootD > 4.8.0



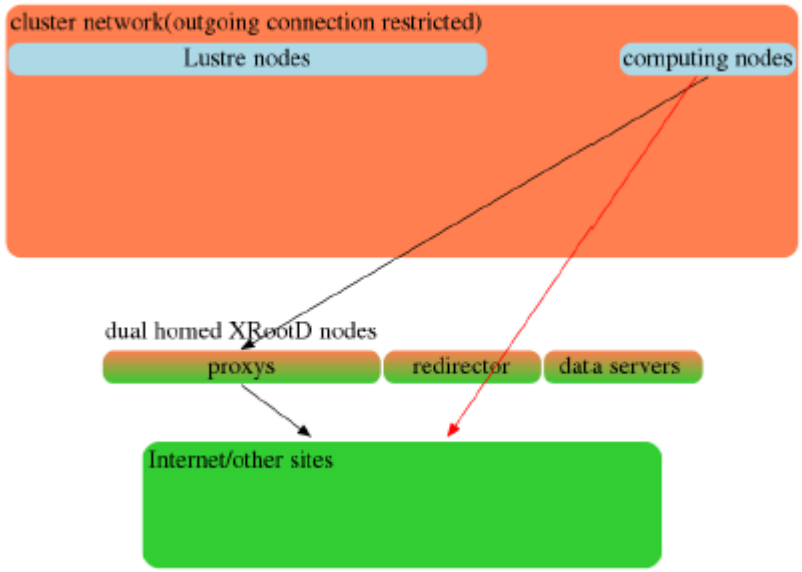
Scaling Test:
reading from Lustre (30 OSS).
Maximum rate: 32 GB/s due to limitation to 2500 concurrent jobs
Target rate (10 PB/24 h = 115 GB/s) should be achievable by scaling number of jobs and OSS accordingly.

ST1/Large Scale Scientific Data Lifecycle Management

- efficient high bandwidth data transport
- in context with BMBF Pilotmaßnahme ErUM-Data

a prototype of an XRootD based dynamic data cache for heterogenous resources is being developed.
A job which needs data from an external data server requests these data via an XRootD Forward proxy. During this process the data are being cached on a local data cache.
In case further jobs would require the same data set XRootD would recognise that the data are already locally available and would redirect these jobs via an XRootD Plug In to the local file system.

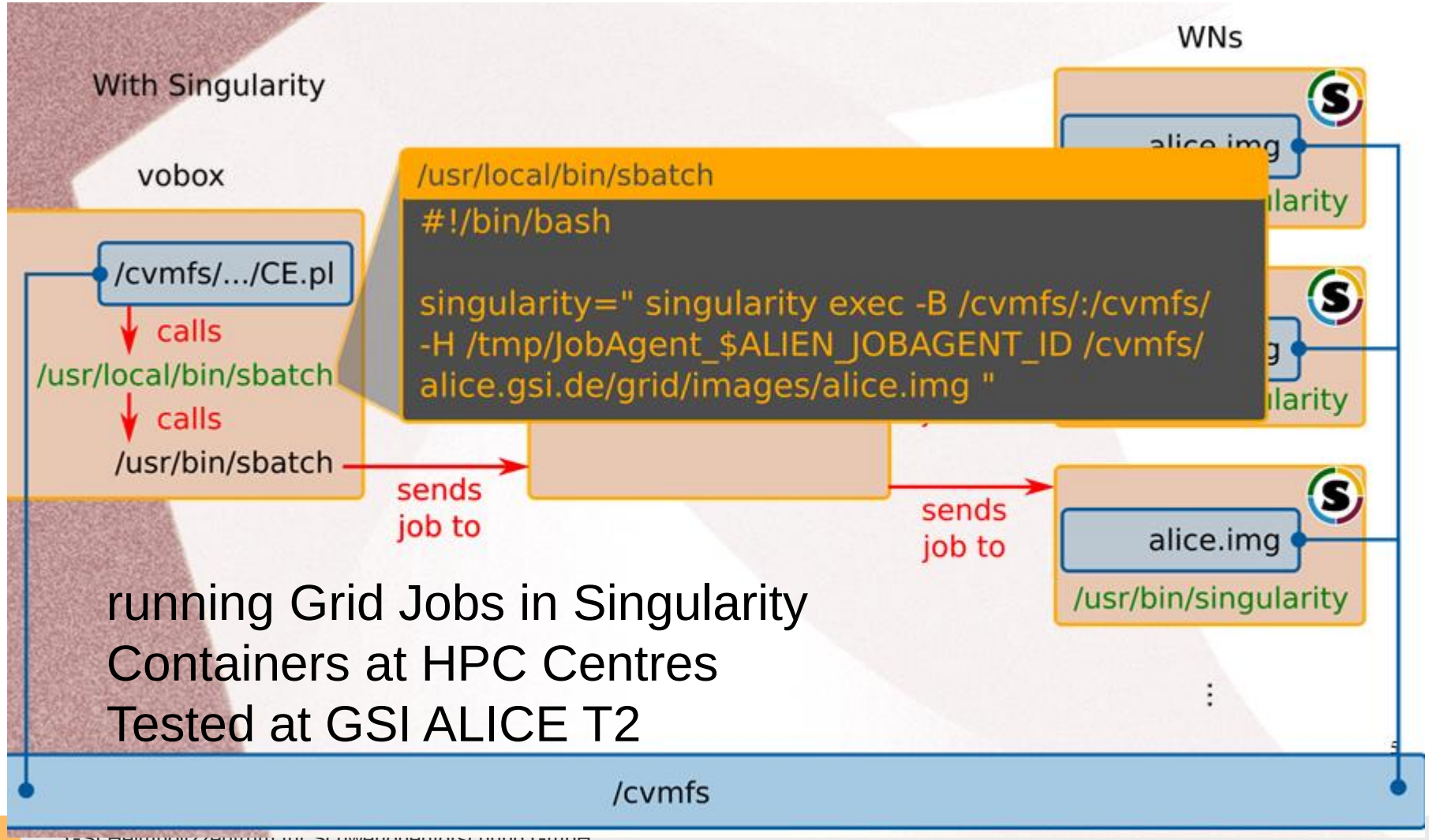
used by ATLAS & CMS at KIT & NEMO Cluster in Freiburg



ST1/Scalable & Distributed Services for Science

- distributed computing workflows

in context with BMBF Pilotmaßnahme ErUM-Data



running Grid Jobs in Singularity
Containers at HPC Centres
Tested at GSI ALICE T2

ST1/Scalable & Distributed Services for Science

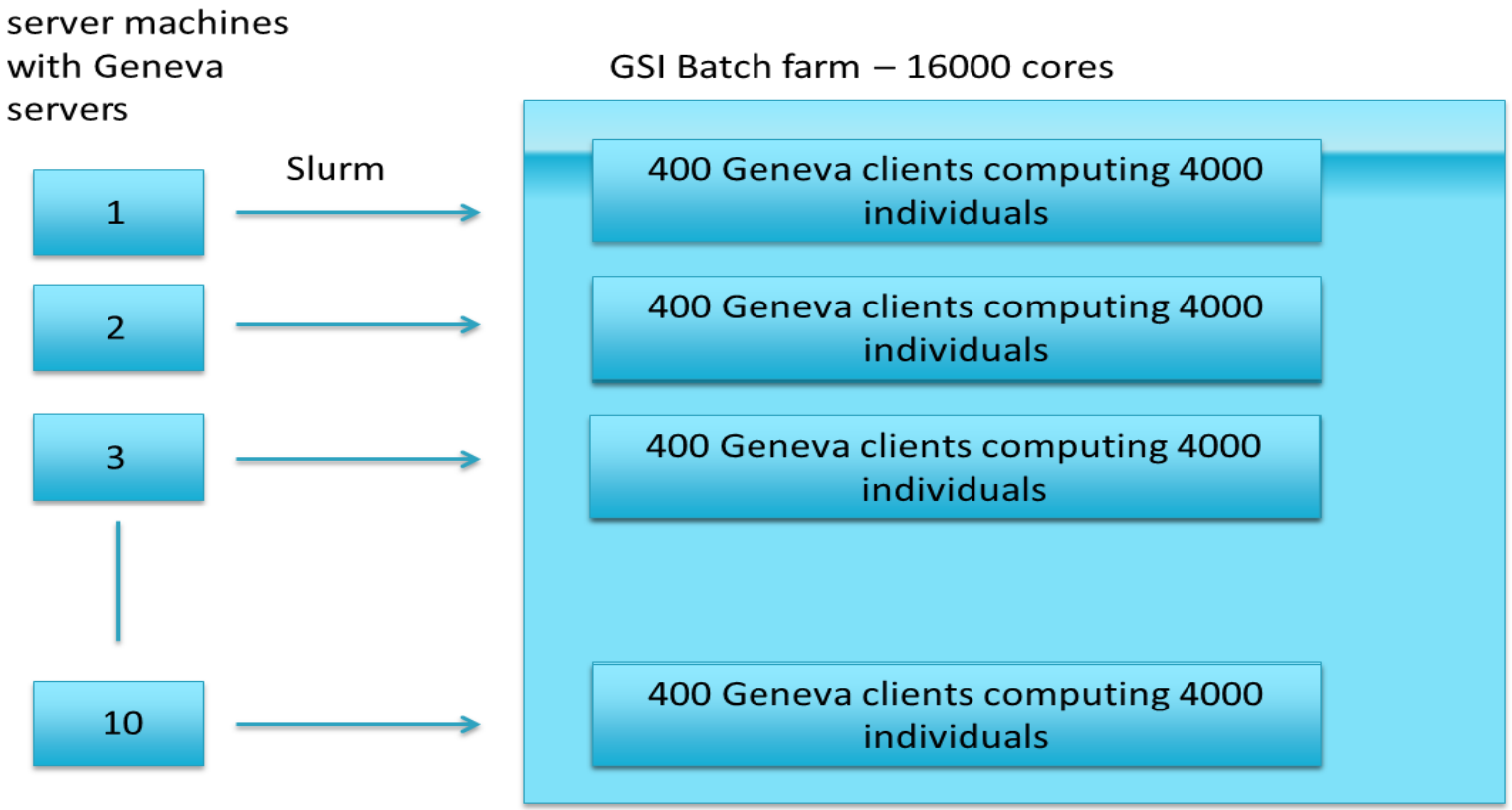
- distributed optimisation
  - Hadronic reaction amplitudes for FAIR

- a framework for predicting and analysing final-state interactions for the FAIR experiments is beeing developed
- this requires massive parallel computing, up to 50 and more coupled-channels needed
- reaction amplitudes are derived from effective Lagrangians where coupled-channel unitarity and the implications of micro-causality (dispersion relations) are implemented (isobar models are not good enough)
- parameter space is reduced significantly by using constraints from chiral and heavy-quark symmetry but also large-Nc QCD
- a subset of the parameters can be derived from the quark-mass dependence of existing QCD lattice data and/or fits to existing data
- conventional fitting routines like Minuit are not suitable for such problems – gradients are expensive and not stable
- in order to avoid local minima and to be able to find the best possible solution an Evoluationary Algorithm with reasonably high population is under investigation

ST1/Scalable & Distributed Services for Science

- distributed optimisation

Geneva Cluster @ GSI
example case: 10 minutes compute time for one solution, 10 x 400 clients, 10 x 4000 population, 1000 iterations, one week of compute time in total



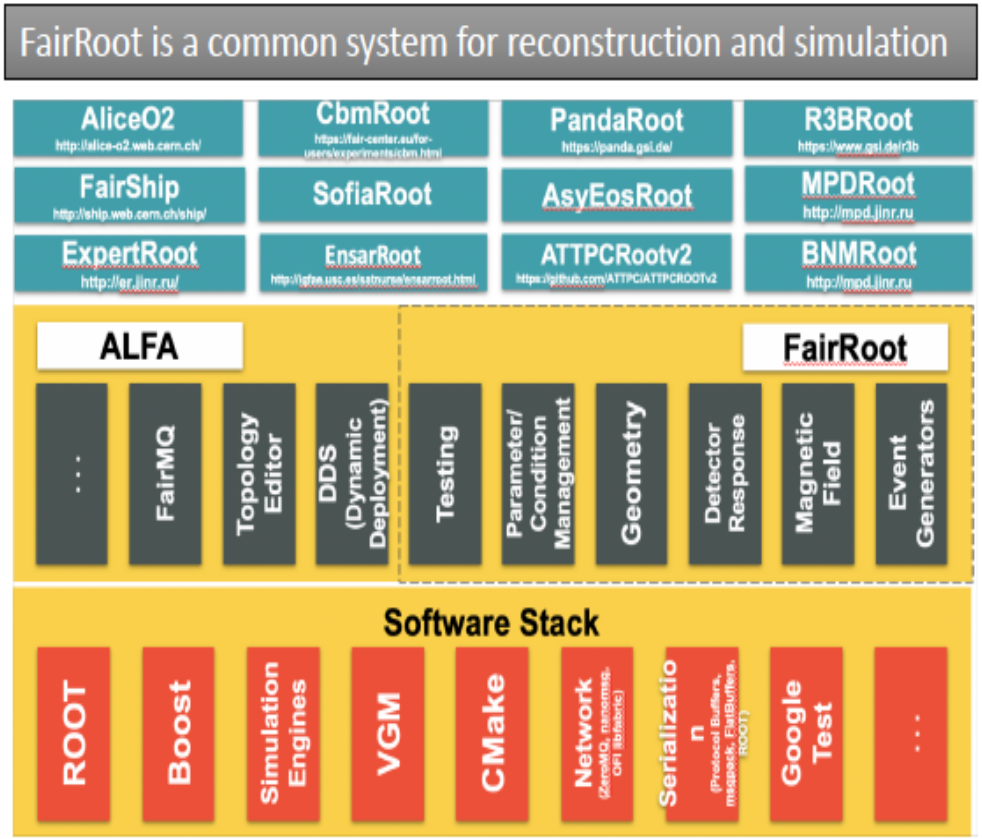
Agenda

- GSI/FAIR
- activities in ST1
- activities in ST2
  - further details see presentation of M. Al-Turany today

The GSI-IT collaborate with different experiments to enhance the synergy

ALFA framework that delivers:

- A data-flow based model (Message Queues based multi-processing)
- Transport layer (FairMQ, based on: ZeroMQ, nanomsg, shared memory and OFI)
- Dynamic Deployment System: deploy processing gragh on a laptop, few PCs or a cluster



ALICE-FAIR project aiming to massive data volume reduction by (partial) online reconstruction and compression

M. Al-Turany

Simulation

Parallel high-performance simulation framework

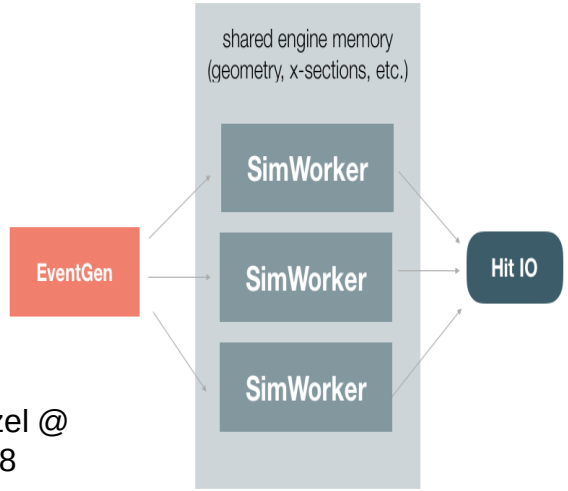
Development of a scalable and asynchronous parallel simulation system based on independent actors and FairMQ messaging

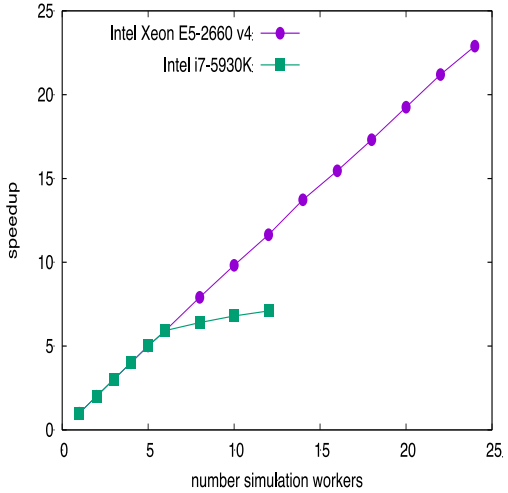
Supports parallelization of simulation for any VMC engine

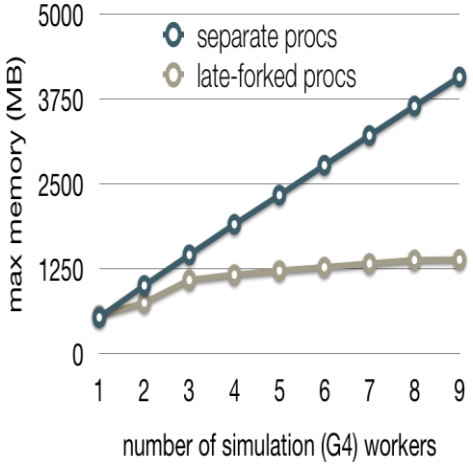
Supports sub-event parallelism

- Make simulation jobs more fine-granular for improved scheduling and resource utilization

- Demonstrated strong scaling speedup (24 core server) for workers collaborating on few large Pb-Pb event
- Small memory footprint due to particular ”late-forking” technique (demonstrated with Geant4)
- In result, reduce wall-time to treat a Pb-Pb events from O(h) to few minutes and consequently gain access to opportunistic resources







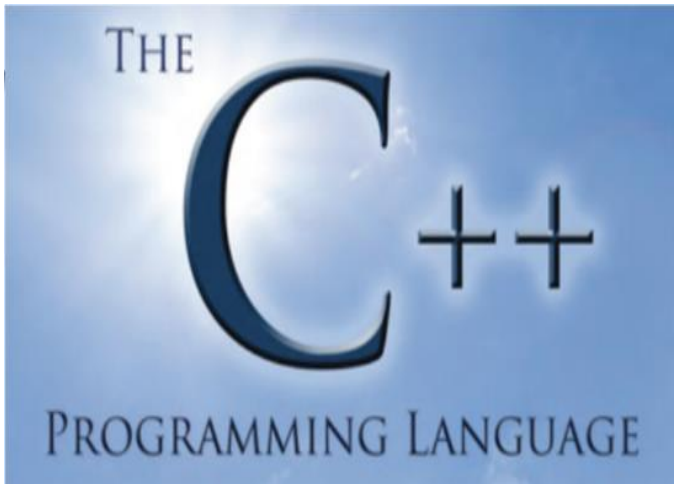
S.Wenzel @ CHEP18

Vc: Portable and intuitive data parallelism

- Vc implements SIMD types, making SIMD programming easy and portable
- Vc is used and taught internationally

Vc to improve C++

- A proposal based on Vc is making progress in the C++ committee (Matthias Kretz)
- Explicit SIMD programming is a stated goal for a future C++ standard



Summary and conclusion

- GSI participates in ST1 to
  - Large Scale Scientific Data Lifecycle Management
  - Scalable & distributed Services for science
- GSI participates in ST2 to
  - next generation computing for simulation & analysis
  - complex data analysis & fusion
  - knowledge extraction & data reduction
  - hight throughput tansport
- many of these activities are already enbedded in national or international projects or collaborations