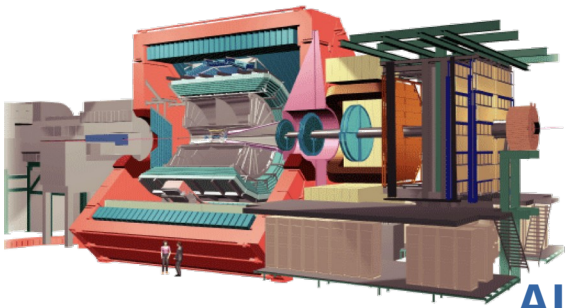
Innovative Workflows in Astro and Particle Physics
March 8 – 12, 2021

# Tools for distributed data processing at FAIR and in ALICE experiment

D. Kresan on behalf of SDE Group (GSI, Darmstadt)

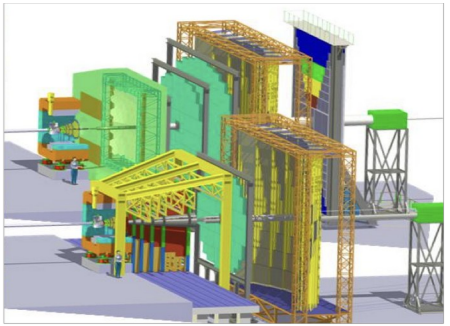


ALICE

3.4 TB/s

+ 463 FPGAs
+ 100’000 CPU cores
+ 3000 GPUs

90 GB/s

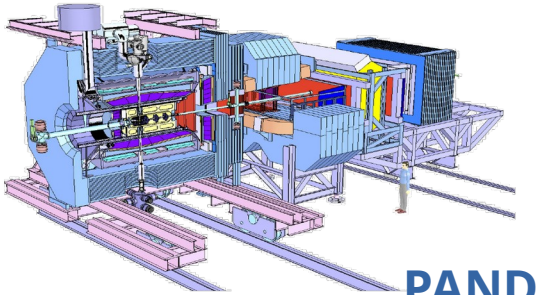


CBM

1.0 TB/s

+ ??? FPGAs
+ 60’000 CPU cores
+ ??? GPUs

1 GB/s



PANDA

0.2 TB/s

+ ??? FPGAs
+ 66’000 CPU cores
+ ??? GPUs

0.2 GB/s

2021: 50 PB/year

2025: 22 PB/year

2025: 24 PB/year

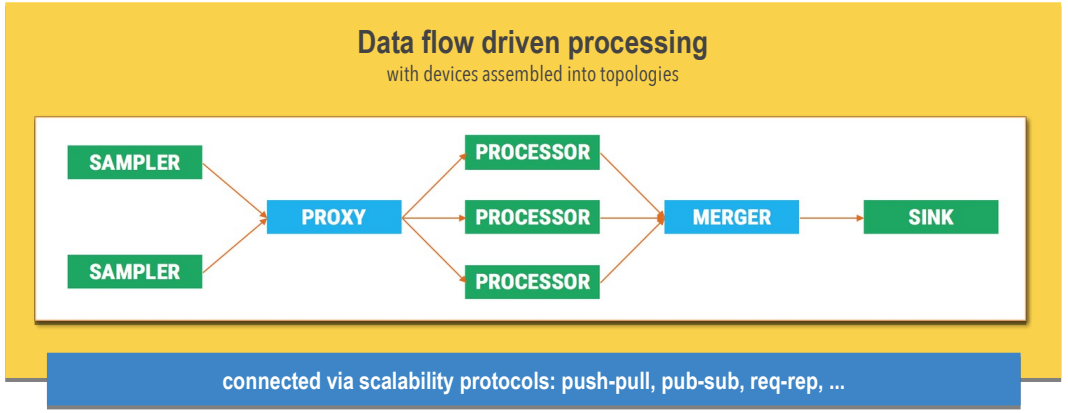
# Software for distributed data processing

- FairRoot: simulation, reconstruction and data analysis
- FairMQ: framework for high throughput distributed data analysis
- Dynamic Deployment System (DDS): management and controlling of distributed workflows

| | |
|---|---|
| 10.5281/zenodo.3896282 | https://fairroot.gsi.de |
| 10.5281/zenodo.1689985 | https://github.com/FairRootGroup/FairMQ |
| 10.5281/zenodo.3854975 | http://dds.gsi.de |

# Use case

- Provides building blocks
  - for experiment-specific algorithms
  - distributed data analysis on small to very large data scale
  - in heterogeneous hardware
  - multiple languages environment
- DDS is used on computing cluster to start, initialize and manage large amount of user processes
- Each process is a separated reconstruction/analysis step
  - communicate over network or shared memory via message passing




A. Rybalchenko

## ALICE EPN workflow

D. Rohr, G. Eulisse