ALICE Computing Model in Run3

Predrag Buncic
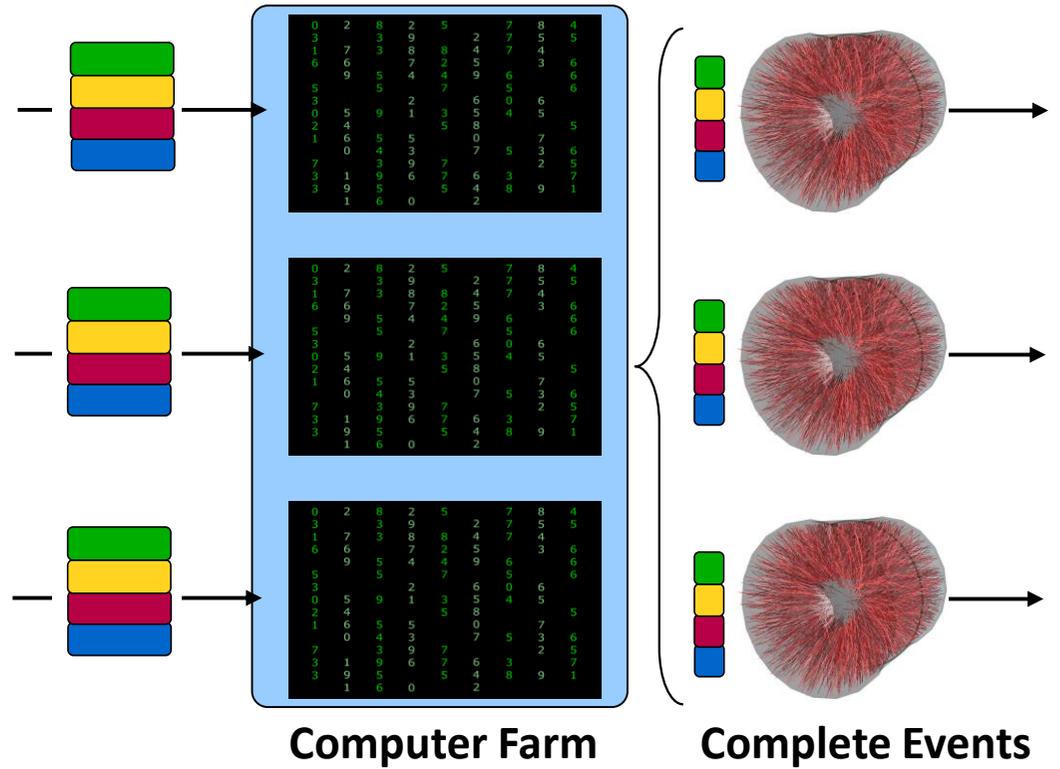
CERN

MULTIPLEXER

Computer Farm

Complete Events

Memory Matrix

Continuous detector reading: replace events with time windows (20 ms, 1000 events).

https://cds.cern.ch/record/2011297

6 TB/s

2.4 MW computing facility
100 k CPU cores, 5000
GPUs, 500 FPGAs
60 PB local disk buffer at P2

90 GB/s

x100
Event rate

x10
Storage

Software performance
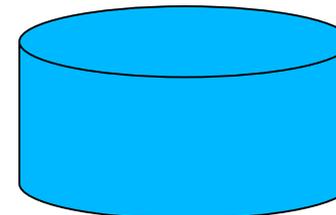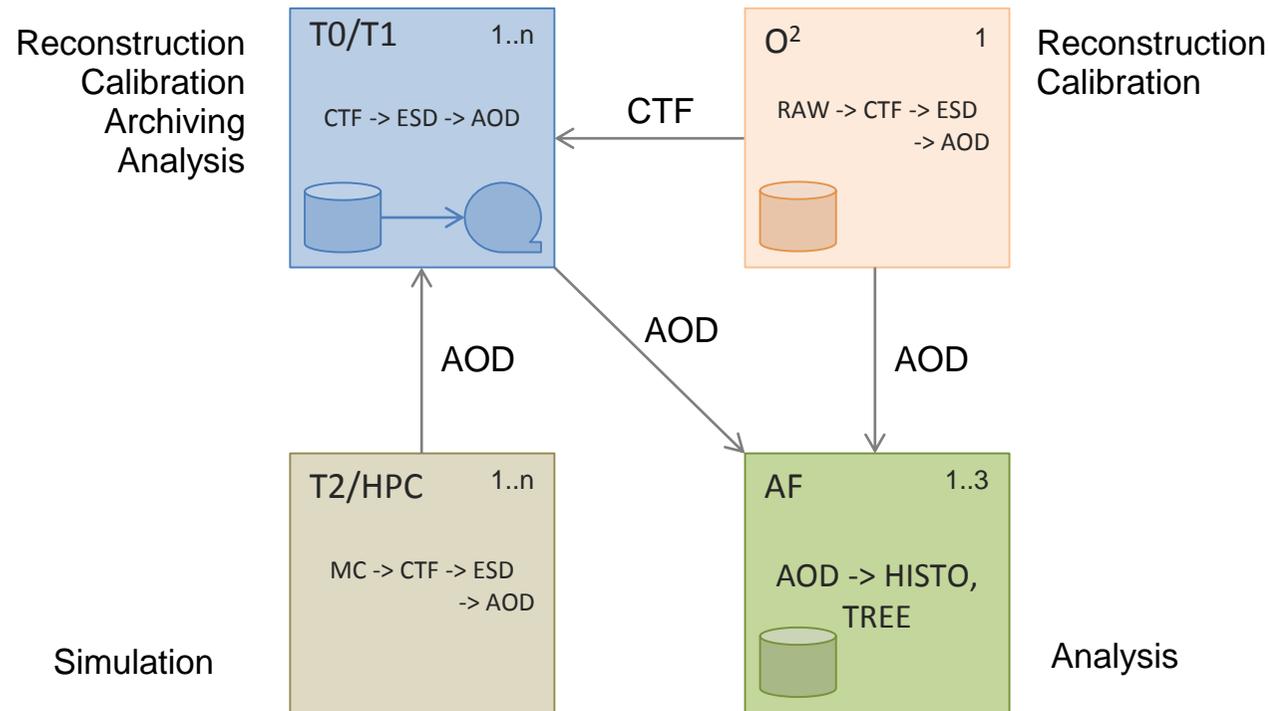Use of hardware accelerators
Parameterized MC

Minimize remote I/O
Specialize sites for given workflow
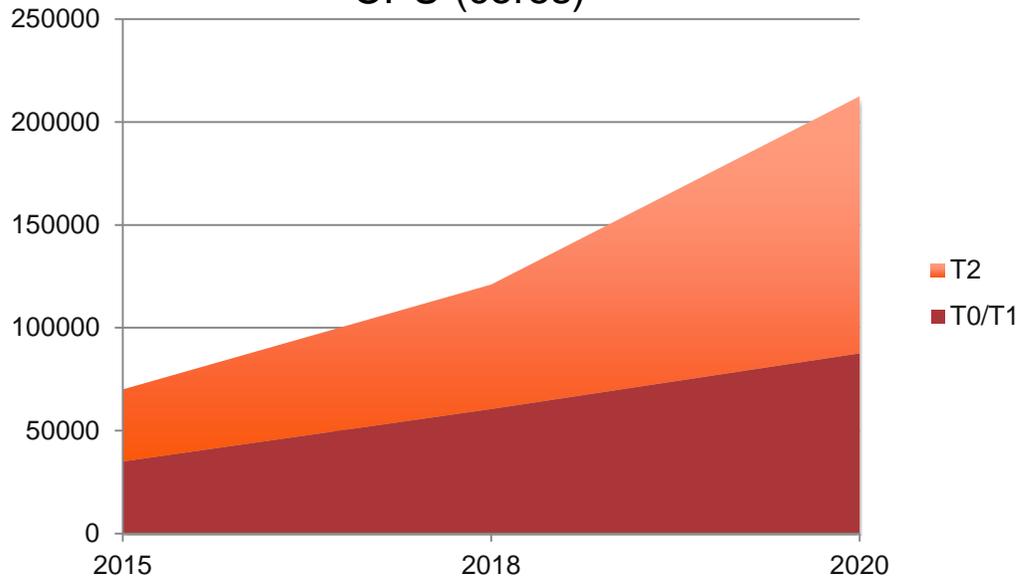
Grid Tiers will be mostly specialized for given role

- O2 facility (2/3 of reconstruction and calibration), T1s (1/3 of reconstruction and calibration, archiving to tape), T2s (simulation)
- All AODs will be collected on the specialized Analysis Facilities (AF) capable of processing ~5 PB of data within ½ day timescale
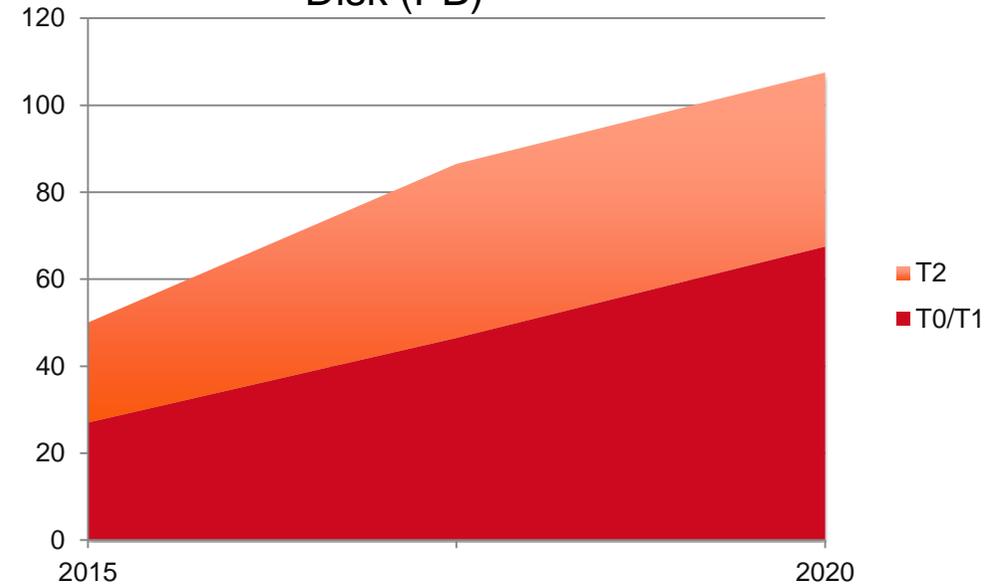
The goal is to minimize data movement and optimize processing efficiency
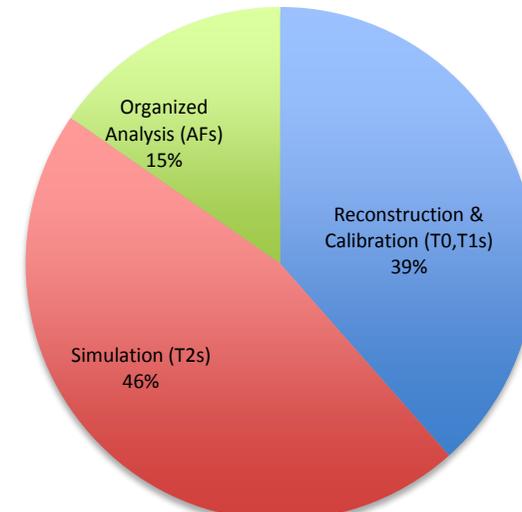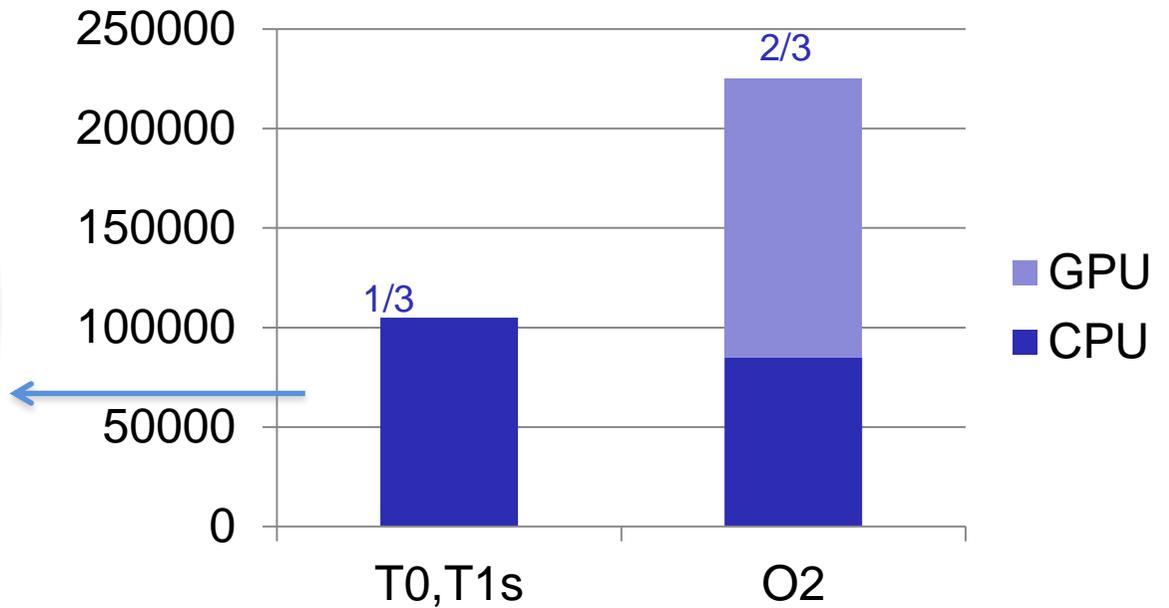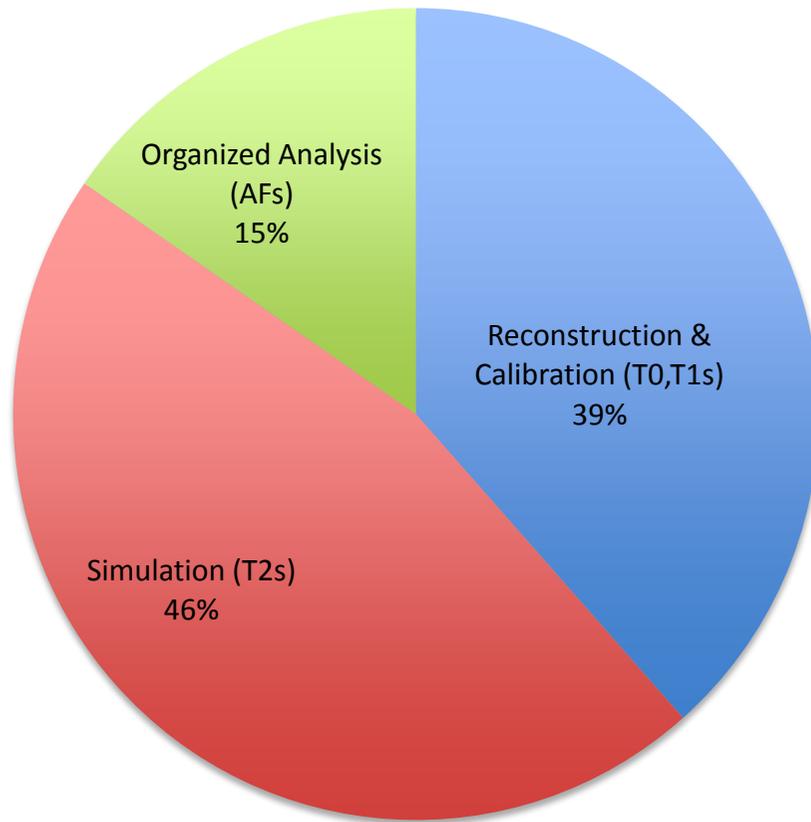
CPU (cores)

Disk (PB)

- Expecting Grid resources (CPU, storage) to continue growing at 20% rate per year

- Since T2s will be used almost exclusively for simulation jobs (no input) and resulting AODs will be exported to T1s/AFs, we expect to significantly lower needs for storage on T2s

- Simulation will remain the main consumer of CPU cycles

Pie chart:
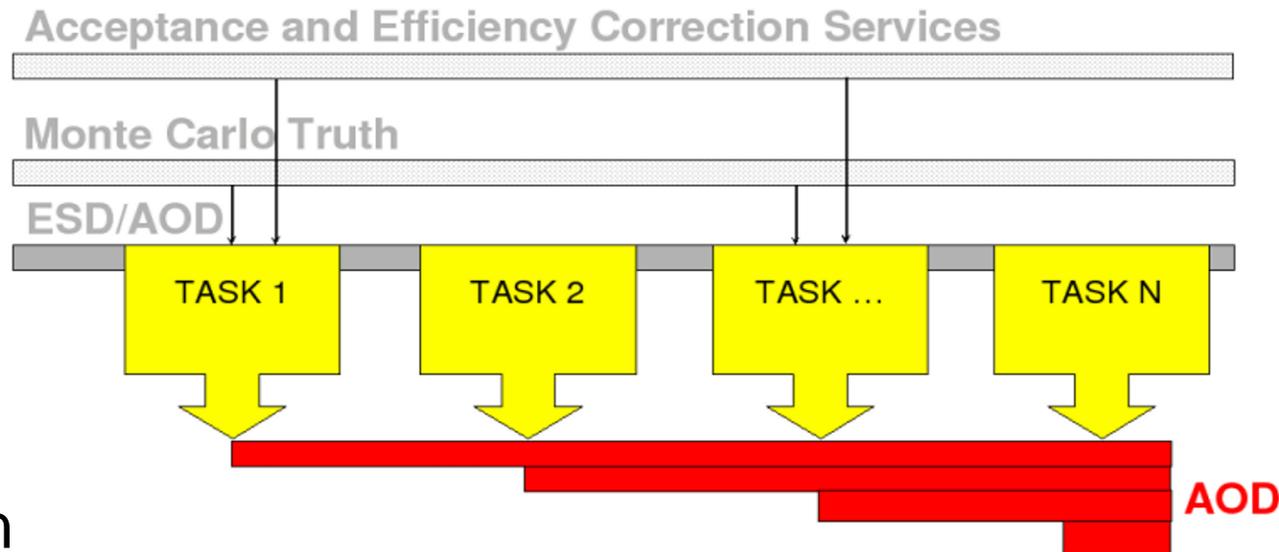- Organized Analysis (AFs) 15%
- Reconstruction & Calibration (T0,T1s) 39%
- Simulation (T2s) 46%

Bar chart (GPU / CPU):
- T0,T1s — 1/3
- O2 — 2/3

- Relative importance of O2 vs Grid may change with time if Grid resources continue to grow as expected
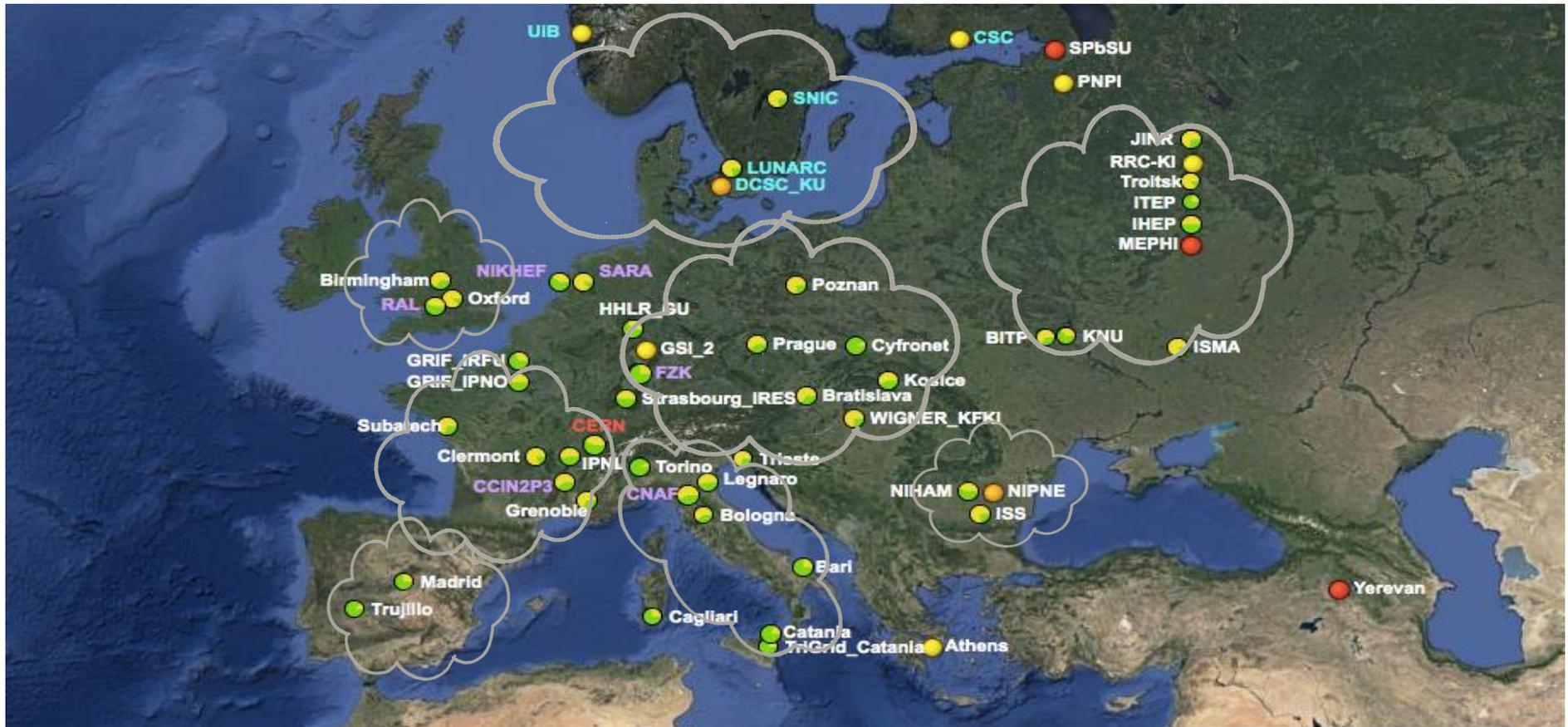
- ## Motivation

  - Analysis is the least efficient of all workloads that we run on the Grid

  - I/O bound in spite of attempts to make it more efficient by using the analysis trains
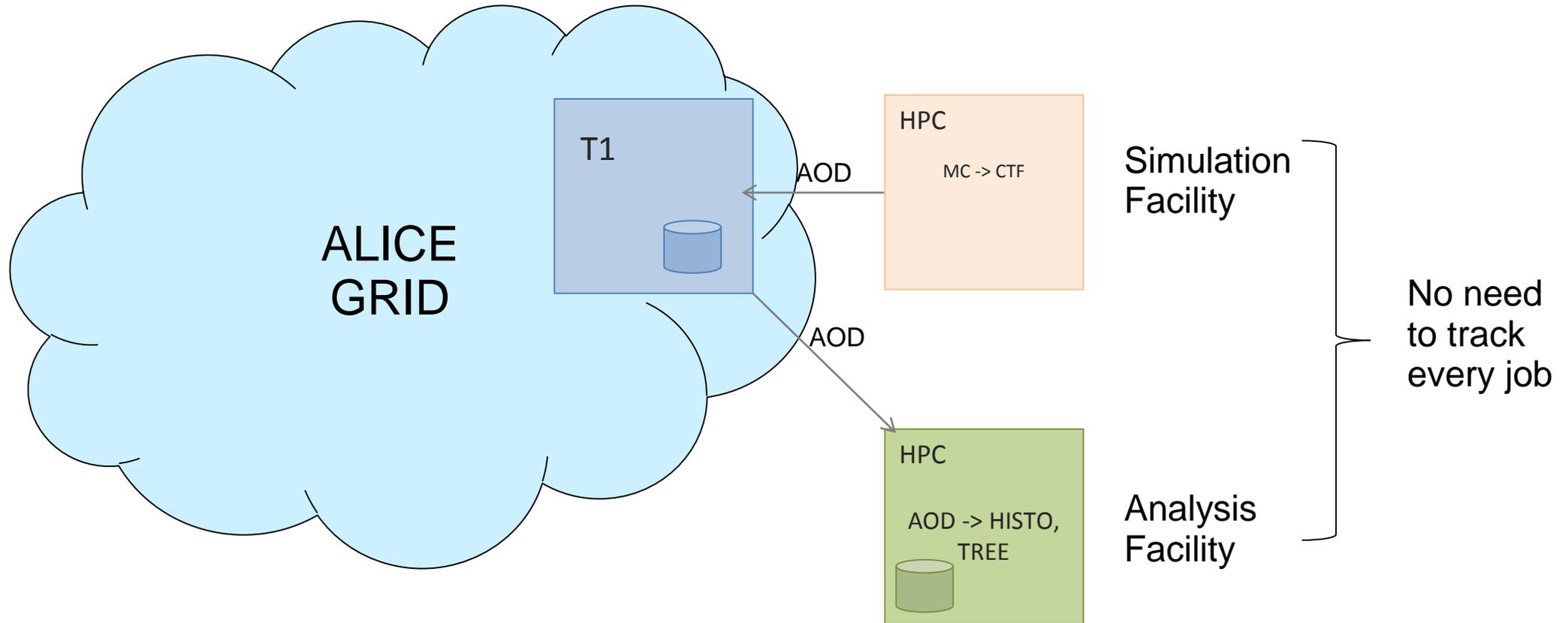
- ## Requirements

  - 20-30'000 cores and 5-10 PB of disk on very performant file system

  - In order to run the organized analysis on local data like we do today on the Grid efficiently using available resources we need to adapt and develop our analysis software

- Federating storage
  - Each cloud/region provides reliable data management and sufficient processing capability
  - Whatever happened in a region, stays in that region

- Localizing execution of complex workflows
  - For this to work, batch systems would have to be able to execute the complex workflows (many of them can but we do not use them like that in the grid context)

Predrag Buncic   The ALICE Online-Offline Project