From AMS to HERD: status and exploration of CLOUD solutions for cosmic ray data analysis

V. Formato

Workshop CCR – 28/05/2021

INFN ROMA TOR VERGATA

Outline

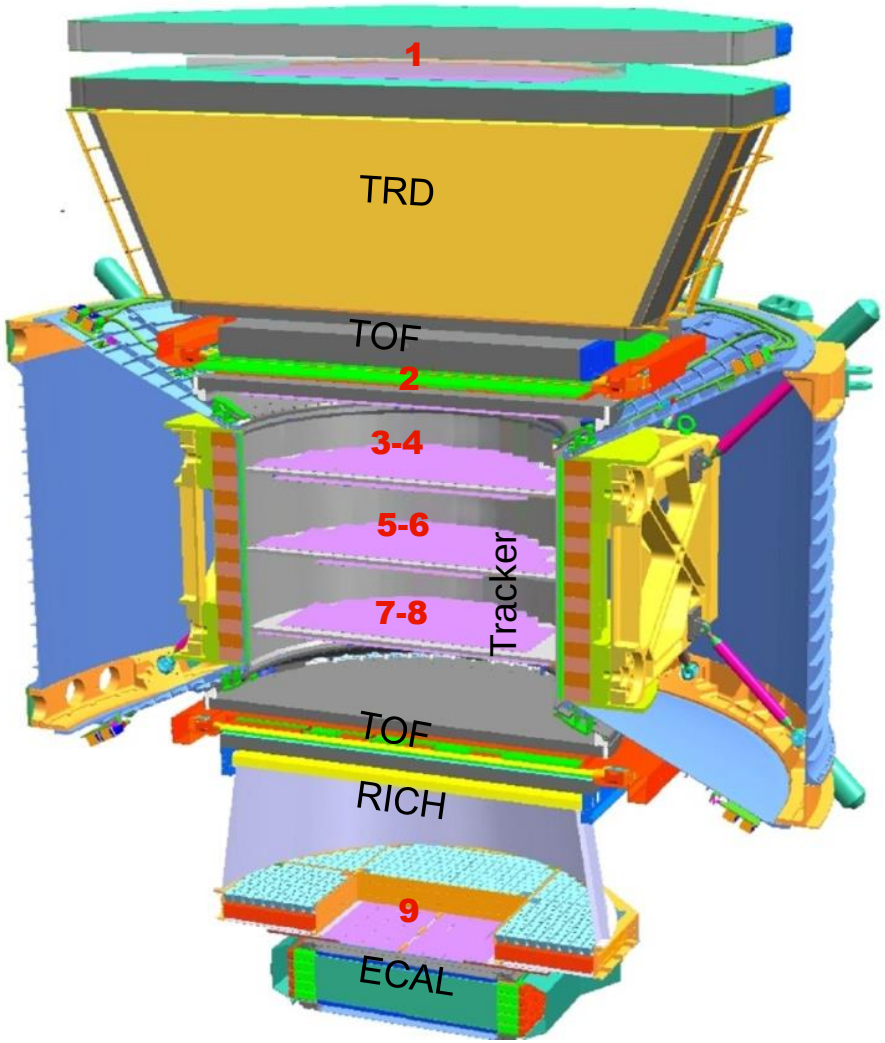
- The AMS (and HERD) experiment and its computing requirements
- The experience gained with federated cloud solutions
- The impact on / role of our main computing center (i.e. CNAF and RECAS)
- Exploration on data-access technologies and their cloud deployment

Who

- DODAS Perugia team:
  - D. Spiga
  - M. Tracolli
  - D. Ciangottini
  - M. Mariotti
- support from CNAF:
  - D. Cesini
  - L. Morganti
  - …
- AMS/HERD team:
  - V. Formato
  - M. Duranti
  - N. Mori
- support from ReCaS:
  - G. Donvito
  - M. Antonacci
  - S. Nicotri
  - …

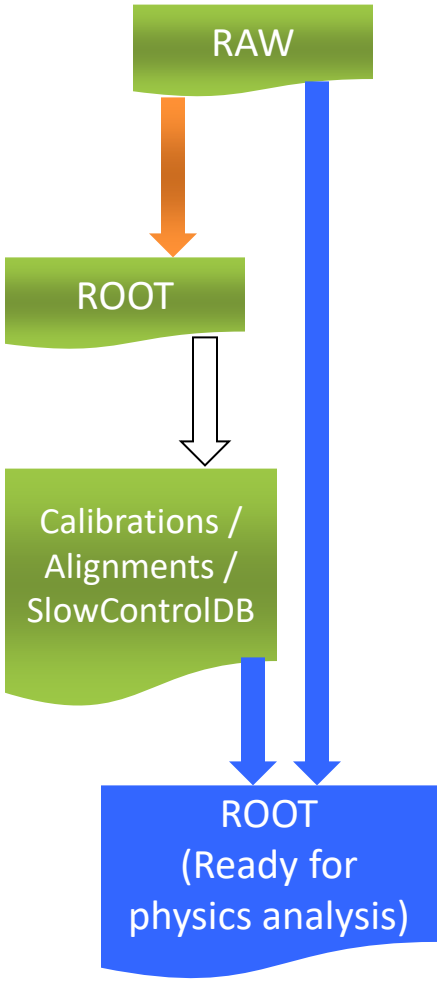
AMS-02

- Installed on the International Space Station, ISS, on May 19, 2011
- Operating 24h/day, 365d/year, since the installation
- 300k readout channels + 1500 temperature sensors
- Acquisition rate up to 2kHz
- More than 600 microprocessors to reduce the rate from 7 Gb/s to 10 Mb/s
- 4 Science Runs (DAQ start/stop + calibration) per orbit: 1 Science Run = ~ 23 minutes of data taking
- On May 2019, ~135 billion triggers acquired
- 35 TB/year of raw data



AMS-02 - Data

- First Production (a.k.a. "std", incremental)
  - Runs 365dx24h on freshly arrived data
  - Initial data validation and indexing
  - Usually available within 2 hours after flight data arriving
  - Used to produce calibrations for the second production as well as quick performance evaluation ("one-minute ROOT files", prescaled)
  - Used for non-critical on-line monitoring in the POCC
  - 100 cores (@ CERN) to keep up with the acquisition
- Second Production (a.k.a. "pass*N*")
  - Every 6 months, incremental
  - Full reconstruction in case of major software update
  - Uses all the available calibrations, alignments, ancillary data…
  - 100 core-years per year of data



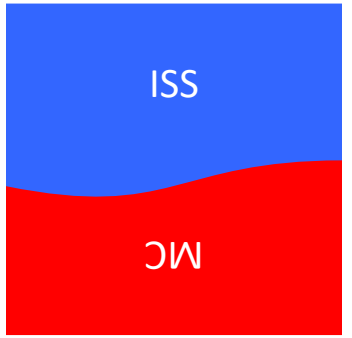
AMS-02 - MC

- In addition to ISS data, a full MC simulation of the detector with at least x10 statistics is needed:
  - To determine the Acceptance of the detector
  - To test the analysis flow
  - To test and train discriminating algorithms (for example MVA's)
  - To understand the irreducible background
  - The "beam" is unknown: in general all the CR species (at least according to their abundance), even if not directly under measurement, must be simulated (at all the energy, according to natural spectra [i.e. ~ power laws]) as possible source of background
  - MC based on Geant 4.10.1 (multi-thread, OPENMP) + custom simulations (digitization, capacitive coupling, …)
  - As the detector understanding improves, new updated MC is required. Statistics that must follow the data statistics: 2015: ~ 8000 CPU-years, in 2016: ~11000 CPU-years, …

AMS-02 – Reduced datasets

For both ISS-Data and MC it is necessary to produce:

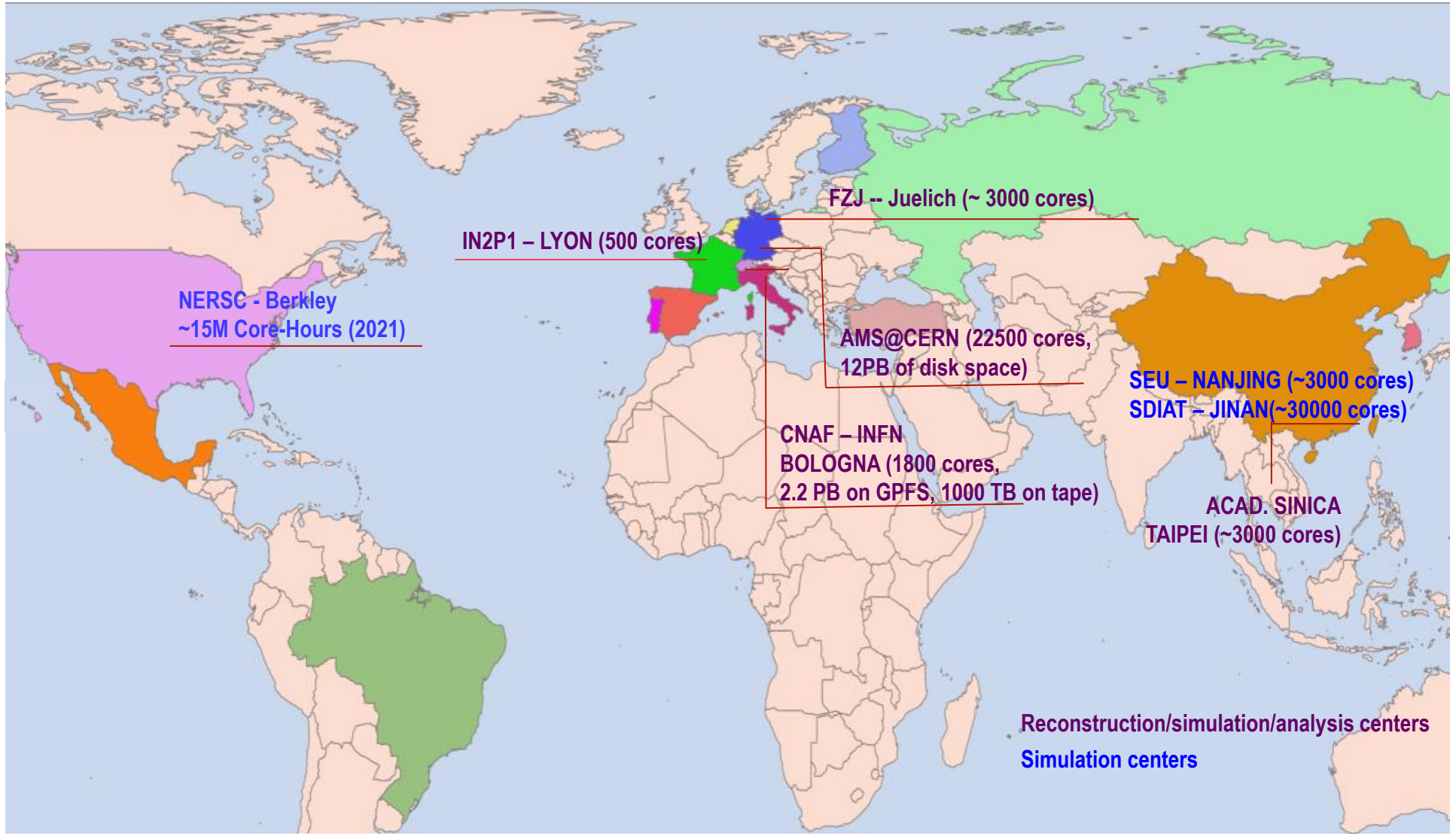
- Reduced dataset or "stream": not all the triggers but only the events that most likely will contain the *signal* of the analysis under consideration)
  → each "study group" has its own production and its own data format (directly the complete one or easily permitting the access to it)
- "mini-DST": ROOT ntuples with a lightweight data format (i.e. ROOT ntuples) and with a subset of all the variables
  - ✓ small size to allow the download also on local desktop/laptop and to permit the processing with a low I/O throughput
  - × must be updated and extended on monthly base



AMS-02 – computing overview

- The "std" production is done in the Scientific Operation Center, SOC, @CERN
  → 200 cores fully dedicated to *deframe*, *merge* & *deblock*, *reconstruct*, …
- The "one-minute ROOT file" production ("std" production prescaled and split in one-minute data files) is done in CERN OpenStack virtual machines
  → 6 single-core machines fully dedicated to this production and to the delivery of the files to the ASIA-POCC
- The "passX" incremental production is done @CERN, on *lxbatch*)
- The "passX" full reproduction is done in the regional centers with a high speed connection
- MC production is done in the regional centers
- mini-DST (i.e. "ntuples") and analysis are done in the regional centers

AMS-02 – computing overview



AMS-02 – additional needs

- "std" production has a well established pipe-line production and requires a limited amount of CPU resources;
- The "passX" incremental production has a well established pipe-line production and requires a limited amount of CPU resources;
- The full reproduction of the "passX" (i.e. the "passX+1") requires a big amount of resources, in a limited time, increasing with the mission time;
- The MC production must follow the "passX" statistics and sw and detector calibration updates;
- The "mini-DST" production and the analysis must follow the "passX" statistics and sw and detector calibration updates;
  - The "request" is intrinsically "peaky", due to the nature of the work
  - The "offer", in terms of computing centers, can be partially "peaky" (see next slide)

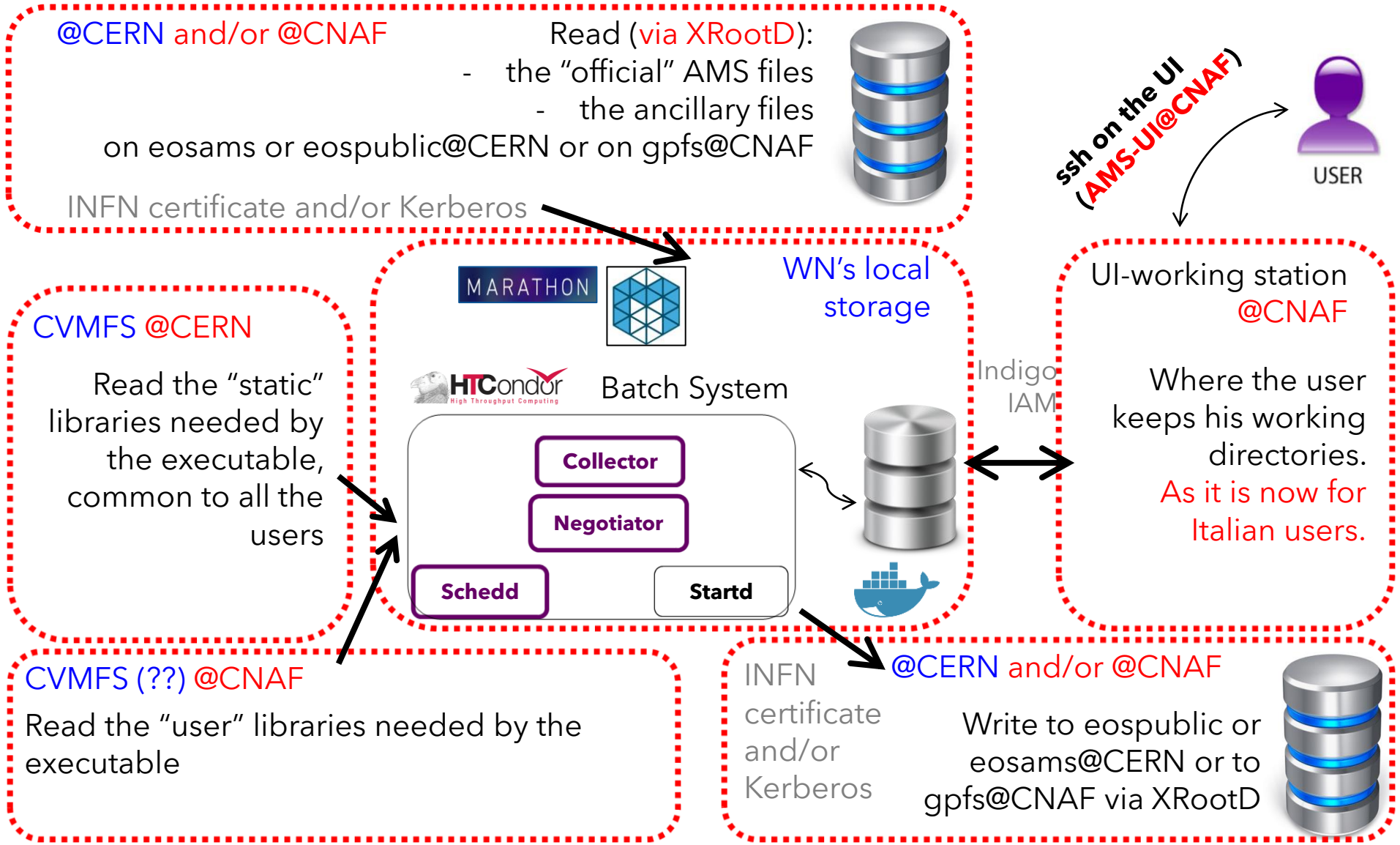
Typically available resources

- Stable, massive, resources:
  - CERN
  - CNAF
  - etc…
- Additional stable resources:
  - ASI (see next slides)
  - Small "farm di Sezione"
- Temporary "free" resources:
  - Chinese resources (see next slides)
  - Cloud resources obtained in the framework of grants, etc…

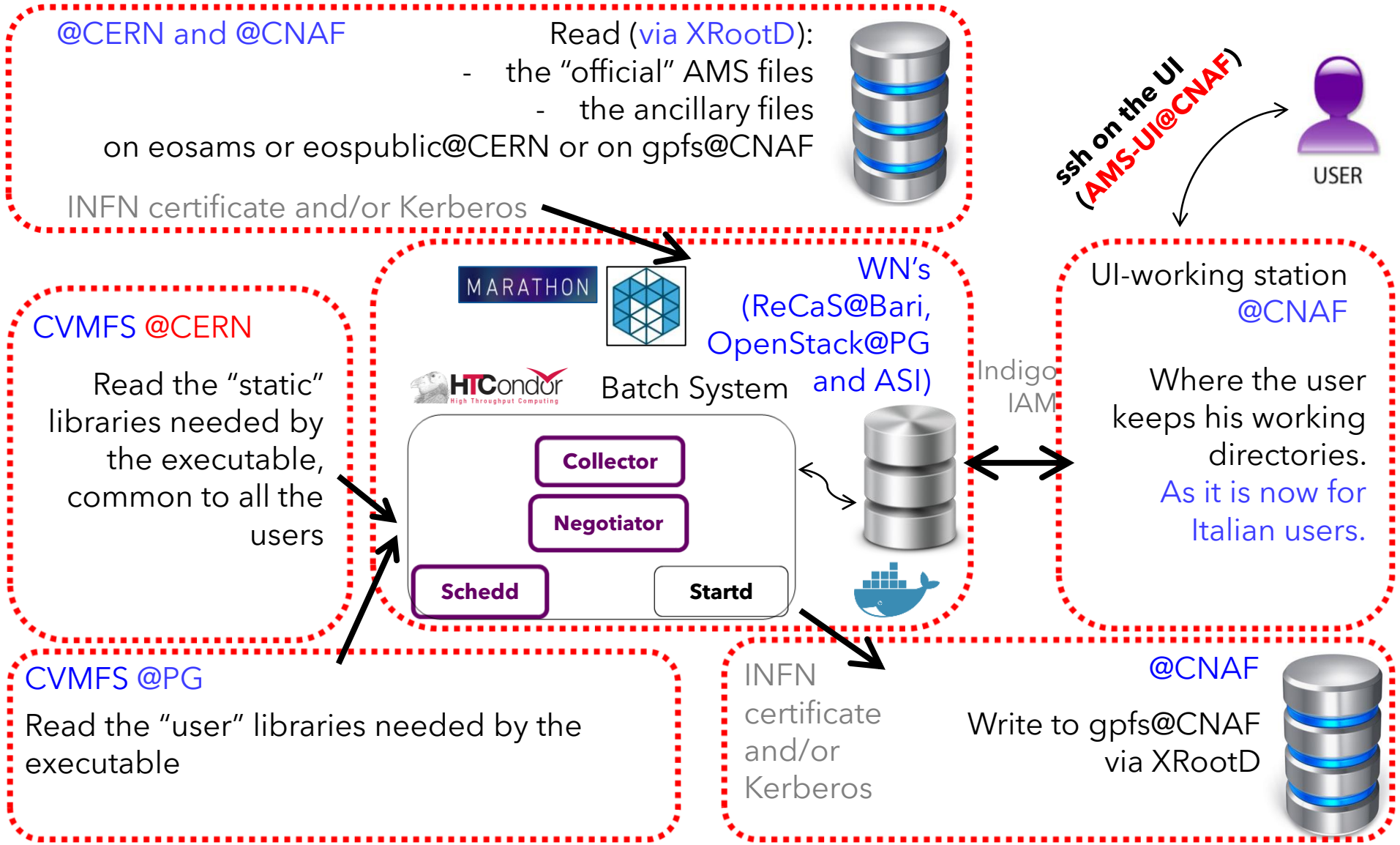
Typical analysis flow

- The job is running a “custom” executable, reading the “official” AMS ROOT files (few GB, @CERN on the ‘eosams’ space): "input files";
- The executable is linked against some libraries, common to all the users (for example the libraries of the AMS patched ROOT), that are needed in a “shared” place: “common static libraries”;
- The executable is linked against some libraries, specific for each user (for example the AMS-sw, that each user has in the required version and/or patched and other libraries from the same user sw framework) , that are needed in a “shared” place: “user libraries”;
- The job needs to read some text files (few KB, easy to transfer for every job) and “ancillary” ROOT files (few MB, @CNAF or @CERN on the user EOS space or ‘eosams/user’): “ancillary input files”;
- The job writes the “mini-DST” ntuples (few tens of MB, ~ 3TB for the total production) on the massive storage (i.e. CNAF storage): "output files";

Target “layout”



Current "layout": DODAS@AMS-02 ASI farm

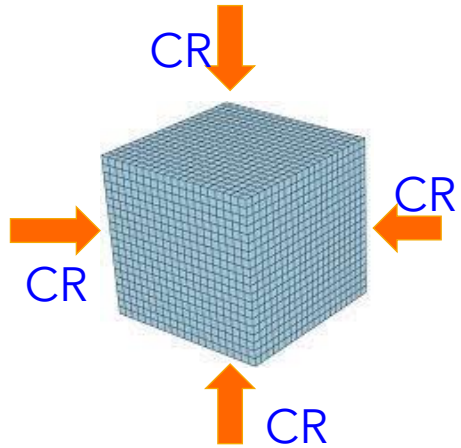


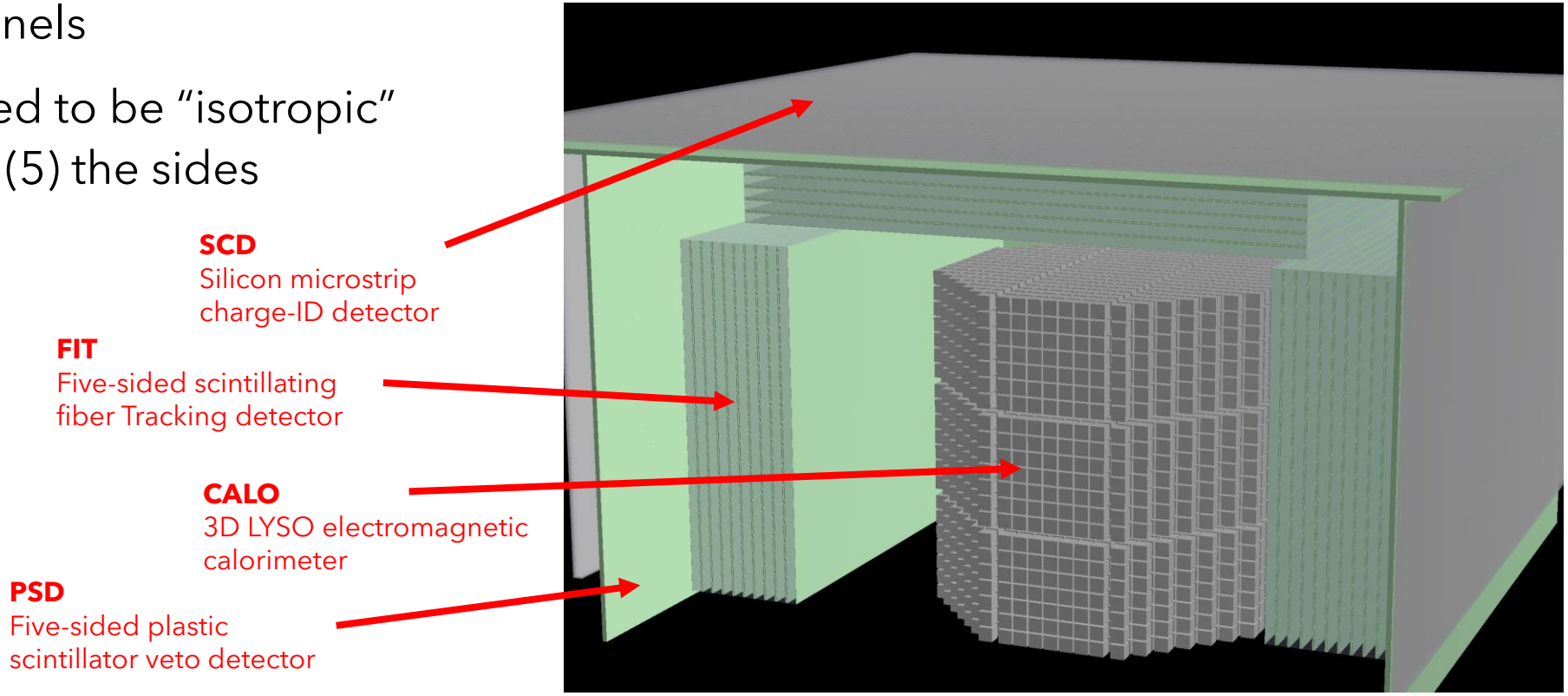
Desiderata

- Shared filesystem where to host the “libraries” (CVMFS)
  We are reaching a point where this is becoming more and more needed, we would like to investigate how to setup a CVMFS server @CNAF.
- INFN-Cloud integration to
  - Exploit external resources (e.g. ASI farm)
    Eventually we want this to be 100% transparent from the user point of view… They shouldn’t even worry about «where the jobs are going»
  - Provide quick access to computing resources for small side R&D projects

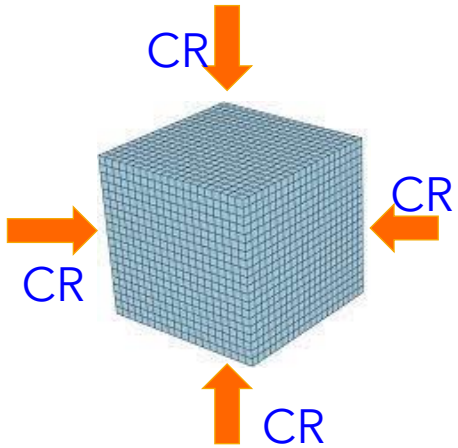
HERD

- Foreseen to operate in space, on board the Chinese Space Station starting from 2024
- Charged CR physics but also γ-ray physics
- ~ O(1M) read-out channels
- The detector is designed to be "isotropic" and accept CR from all (5) the sides

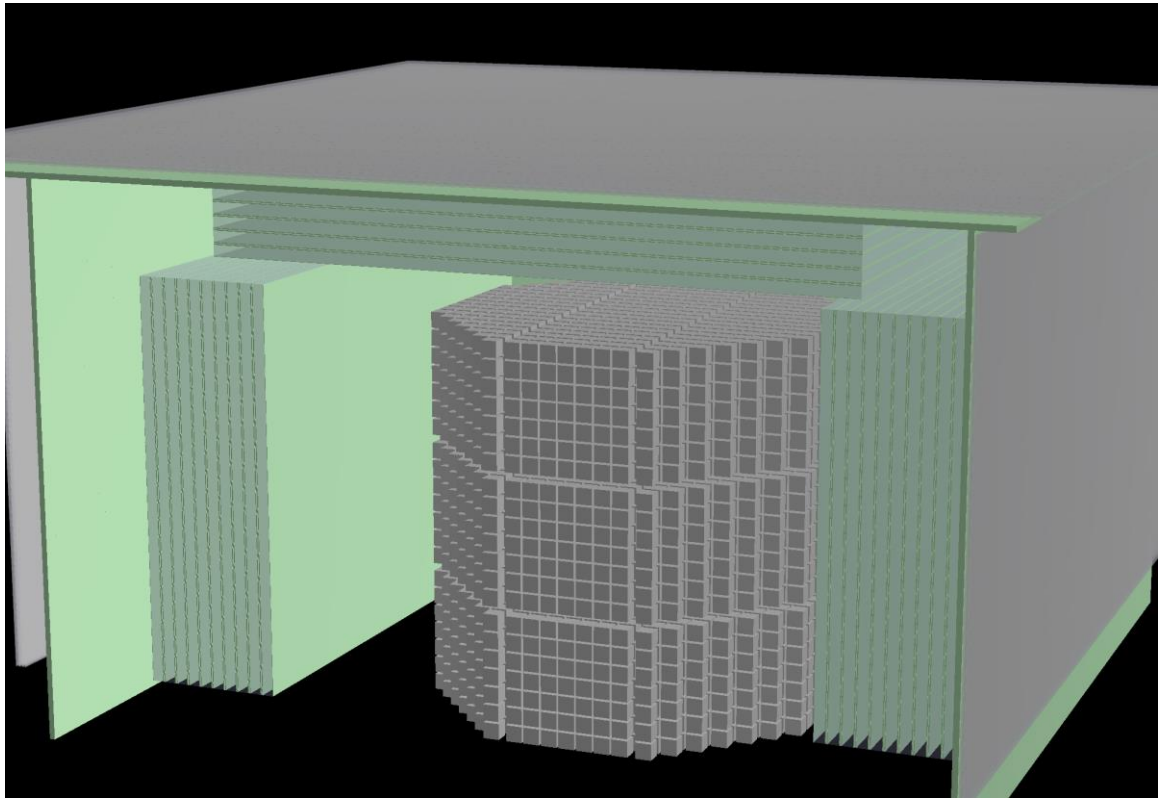




HERD



- Foreseen to operate in space, on board the Chinese Space Station starting from 2024
- Charged CR physics but also γ-ray physics
- ~ O(1M) read-out channels
- The detector is designed to be “isotropic” and accept CR from all (5) the sides

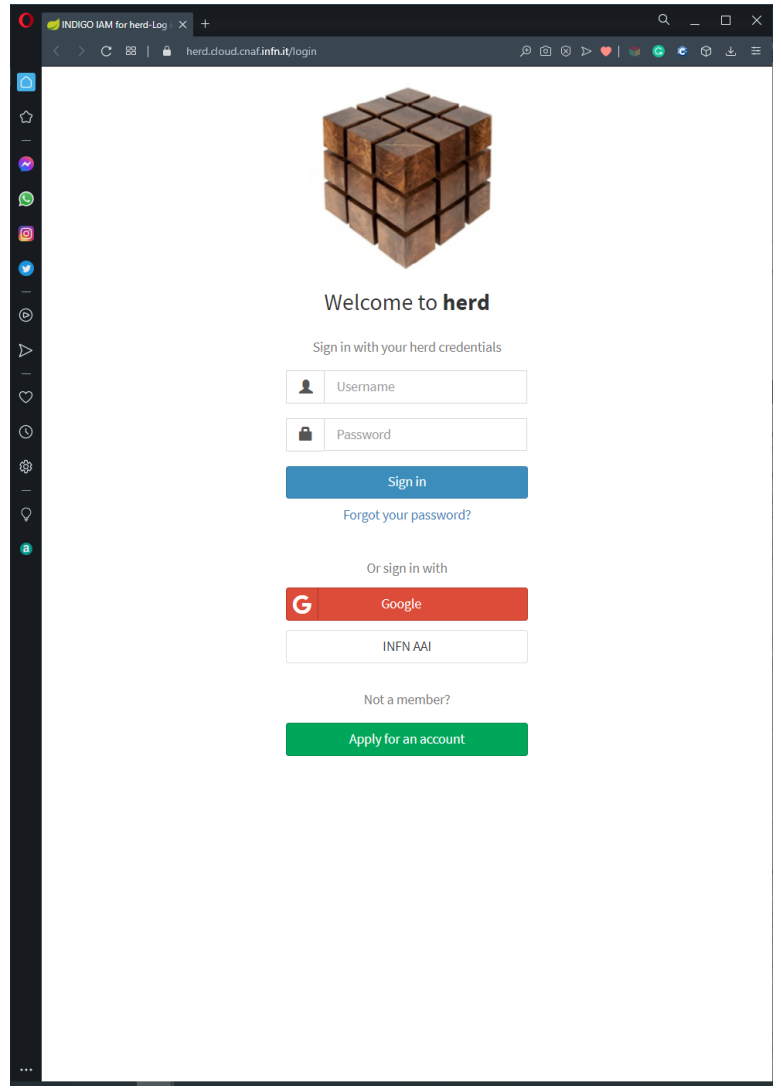


- We are currently designing the computing model of the experiment to exploit from the very beginning CLOUD resources and new solutions for data/workflow management.

HERD – scouting phase

In addition to the compute solutions explored for AMS, HERD can provide a use-case for:

- Cloud hosted services
  We are planning to deploy HERD-specific services (SW documentation, experiment calendar, …) on INFN-Cloud resources. (currently testing on catch-all accounts… we'll want to move to dedicated experiment resources soon)
- "Experiment" IAM instance
  We want to implement a single authentication point for all HERD users, who won't have to create multiple accounts to access multiple services available to the collaboration (especially for HERD users *outside* INFN)
- Data management tools
  We are currently ready to test solutions for data management/distribution provided by the Escape project (Rucio/XRootD/dCache) to be integrated with CNAF/ReCaS HERD storage and IAM as authentication service.



Conclusions

- We (i.e. "astro-particle in space") are a community eager of resources and poor in terms of man-power for computing: we're willing to test any solution to increase our pool of resources and to keep up with the software infrastructure developments, with a limited amount of effort;
- Given the nature of the partners we have for the various projects (ASI, Chinese collaborators, …) we can have small and/or temporary resources: merging them in a single batch system would be a big added value;
- As a new experiment in a design phase, HERD represents the perfect opportunity to migrate towards a deeper integration with INFN-Cloud provided services and work is already ongoing.