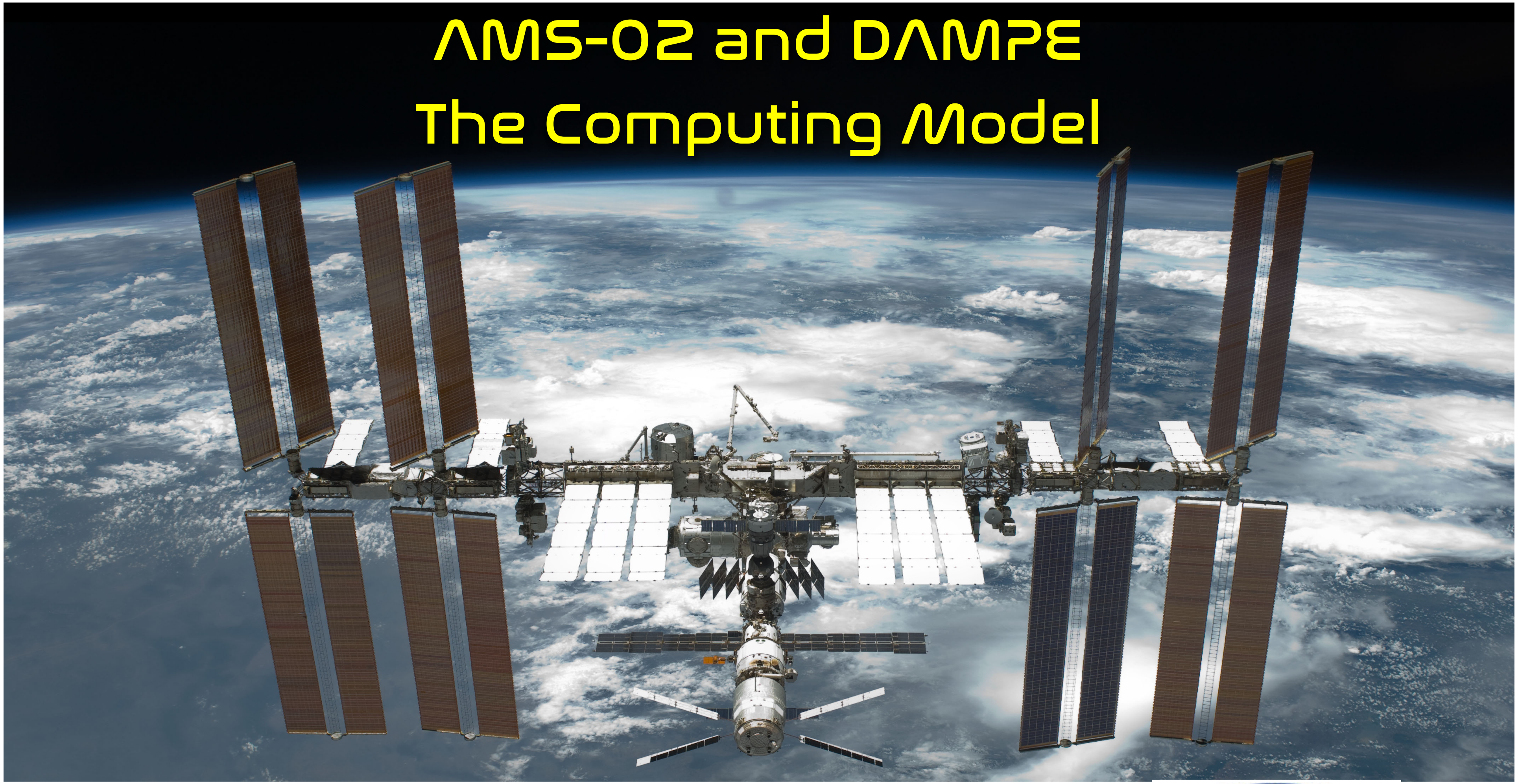
Workshop della CCR, L.N.G.S., 22-26 Maggio 2017

AMS-02 and DAMPE
The Computing Model



Matteo Duranti
Istituto Nazionale Fisica Nucleare, INFN Perugia

Outline

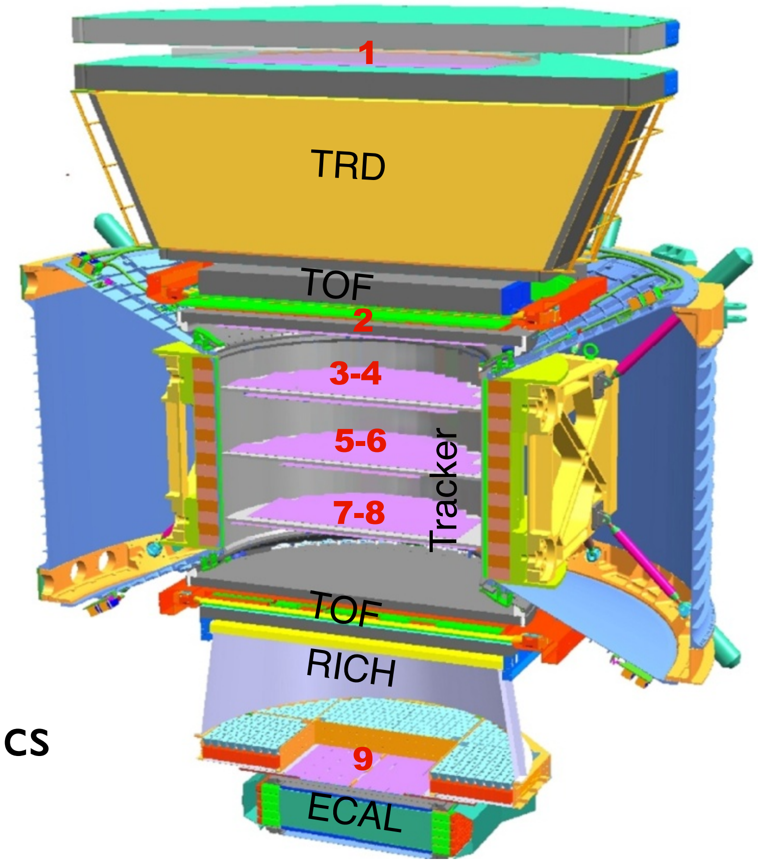
- The AMS experiment on the International Space Station:
  - the computing model
  - the INFN (i.e. CNAF) role in the computing network
- The DAMPE experiment on board a Chinese Satellite:
  - the computing model
  - the INFN (i.e. CNAF and ReCaS) role in the computing network
- The future: DODAS

*Material from: B. Shan, M. Tacconi, D. D'Urso, S. Zimmer and D. Spiga

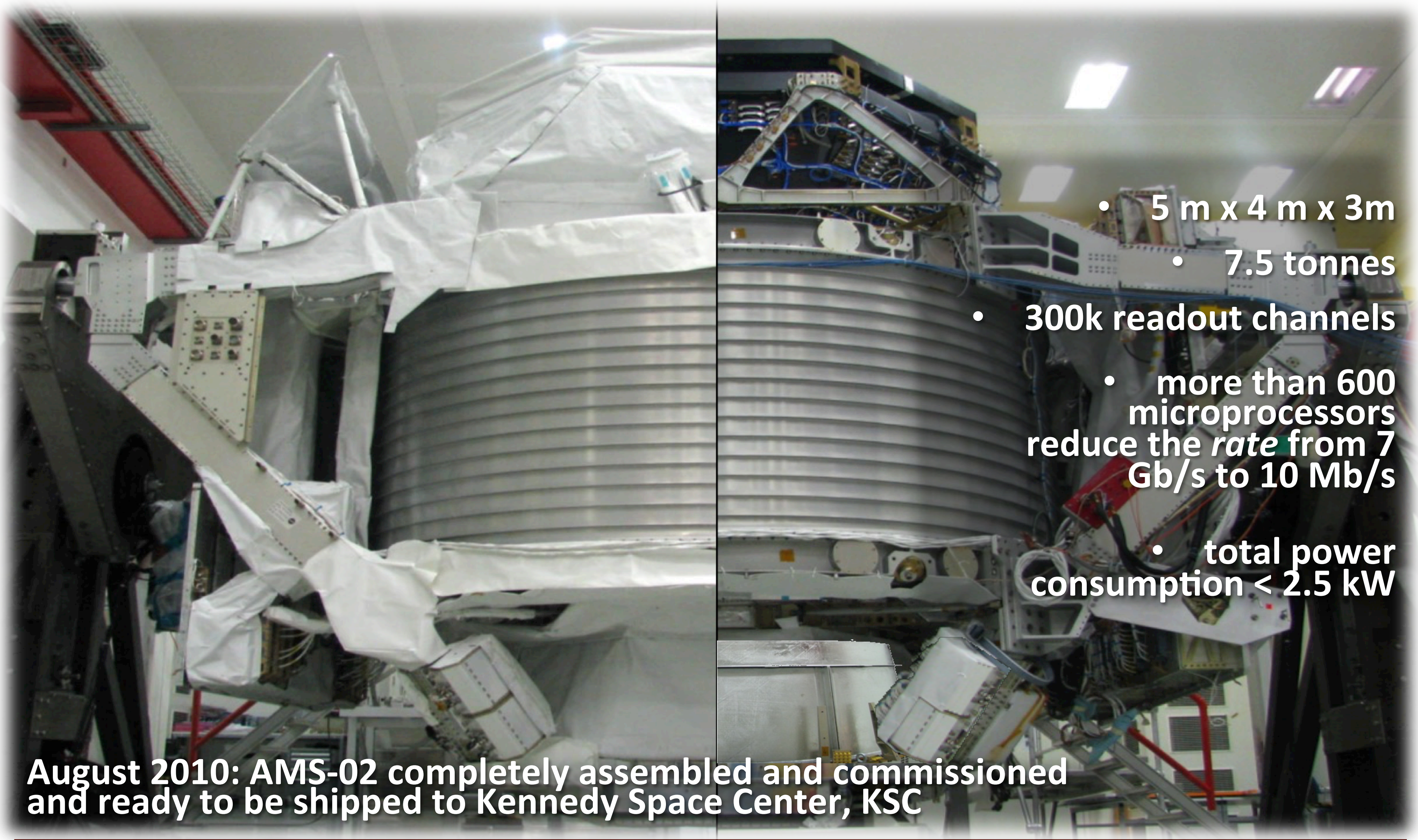
Alpha Magnetic Spectrometer experiment

The AMS experiment

- Fundamental physics and antimatter:
  - primordial origin (signal: anti-nuclei)
  - "exotic" sources (signal: positrons, anti-p, anti-D, γ)
- Origin and composition of CRs in the GeV-TeV range
  - sources and acceleration: primaries (p, He, C, ...)
  - propagation in the ISM: secondaries (B/C, ...)
- Study of the solar and geomagnetical physics
  - effect of the solar modulation
  - geomagnetic cutoff



The AMS-02 detector

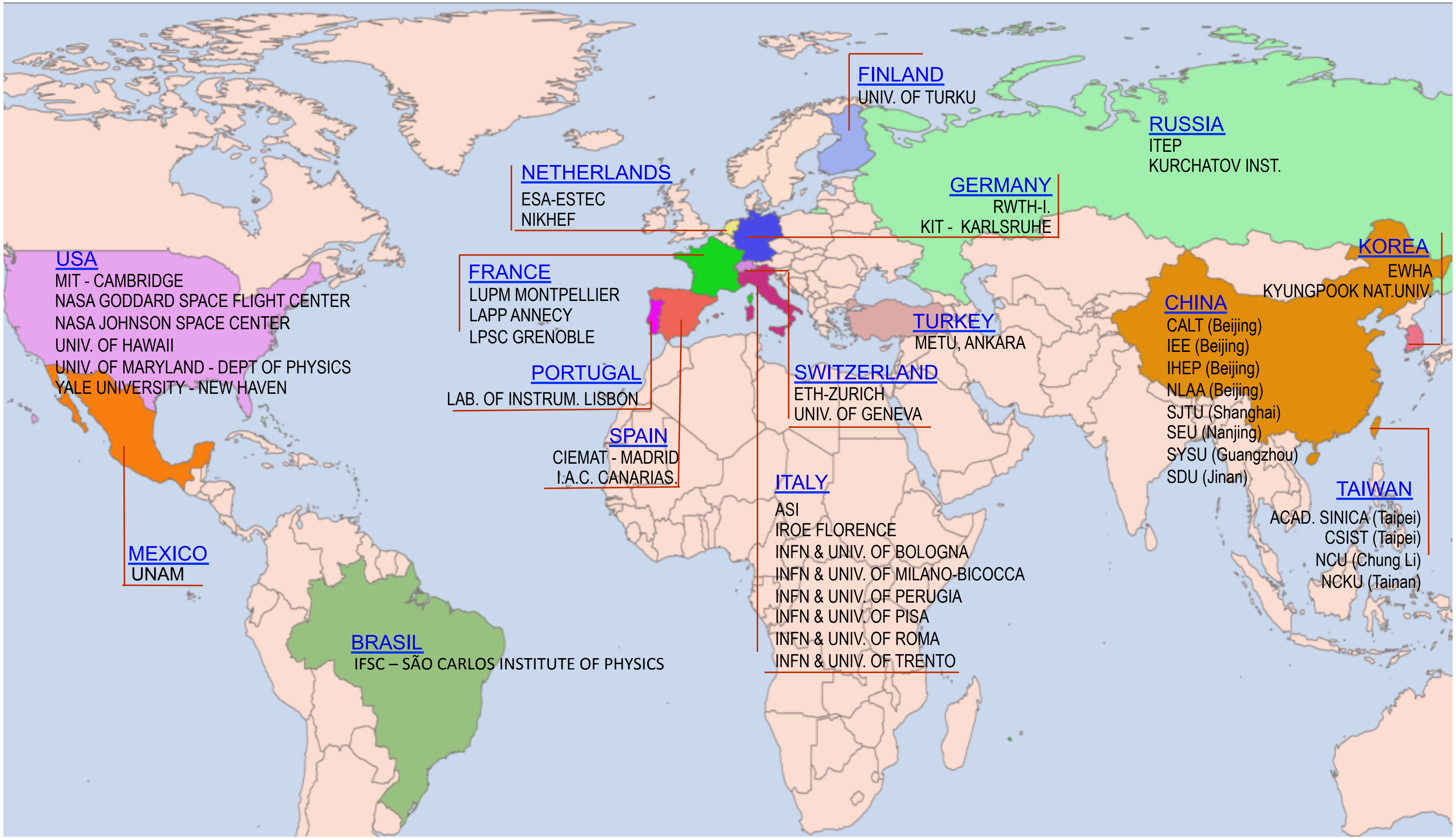


Sezione di Perugia
INFN Istituto Nazionale di Fisica Nucleare

The AMS experiment and the AMS-02 detector

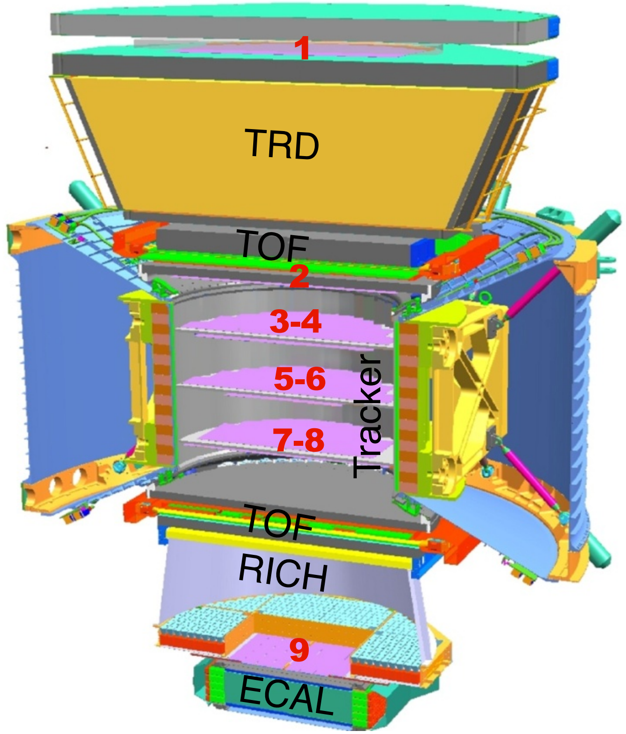
16th May 2011: AMS-02 launch

AMS collaboration: 16 countries

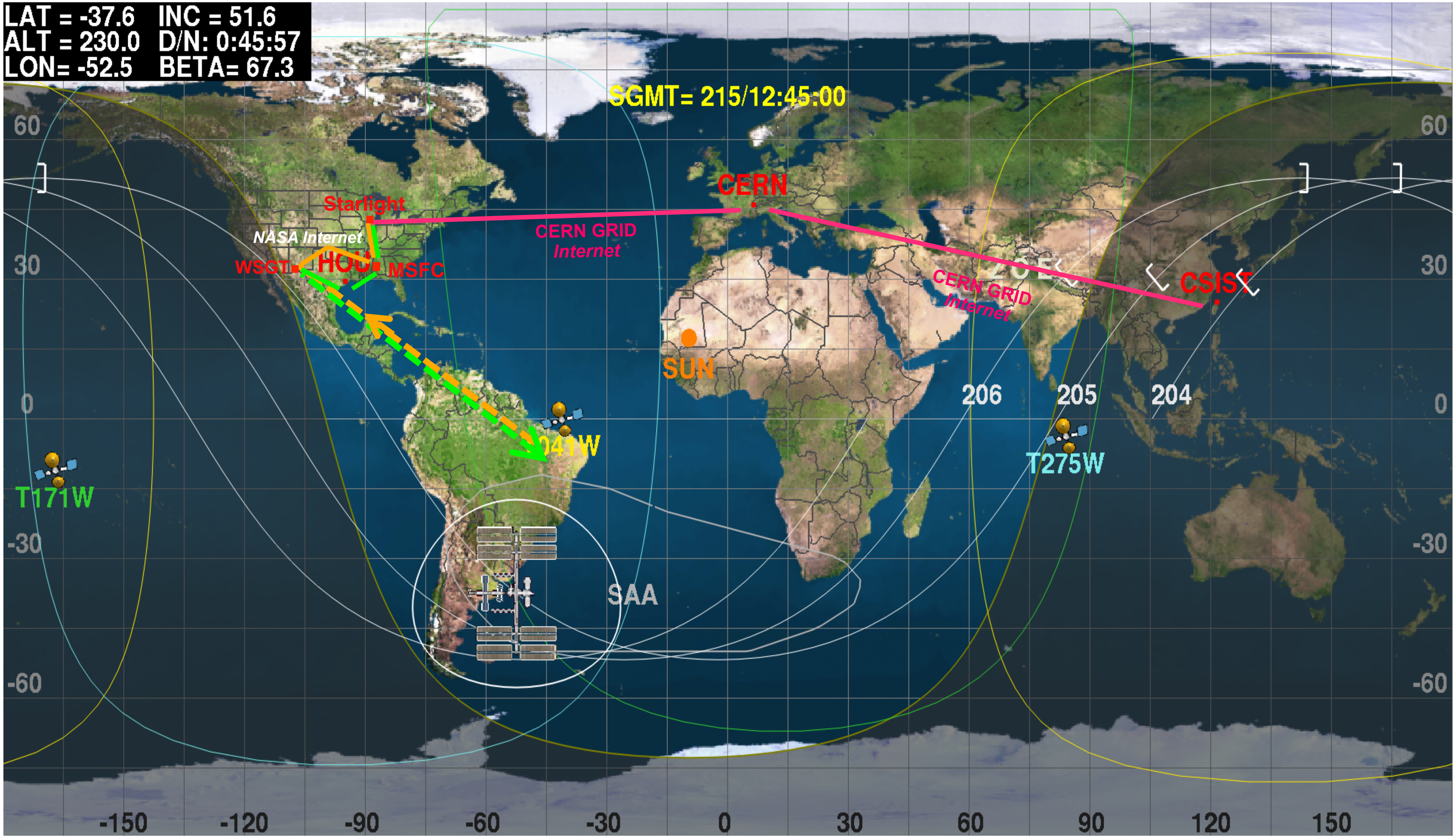


The AMS experiment and the AMS-02 detector

- installed on the International Space Station, ISS, on May 19, 2011
- operations 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 8th, 100 billion triggers acquired
- 35 TB/year of raw data



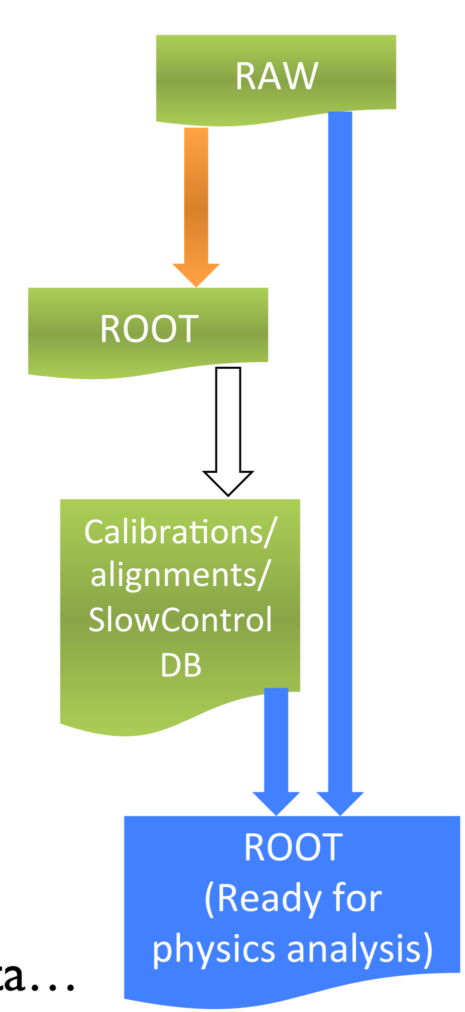
In orbit @ 400 km from ground



Payload Operation Control Center, POCC @ CERN

AMS ISS-data format

- First Production (a.k.a. "std")
  - 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 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, …

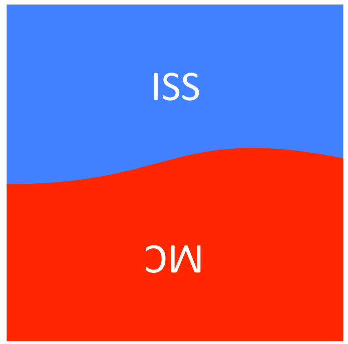
Reduced datasets

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

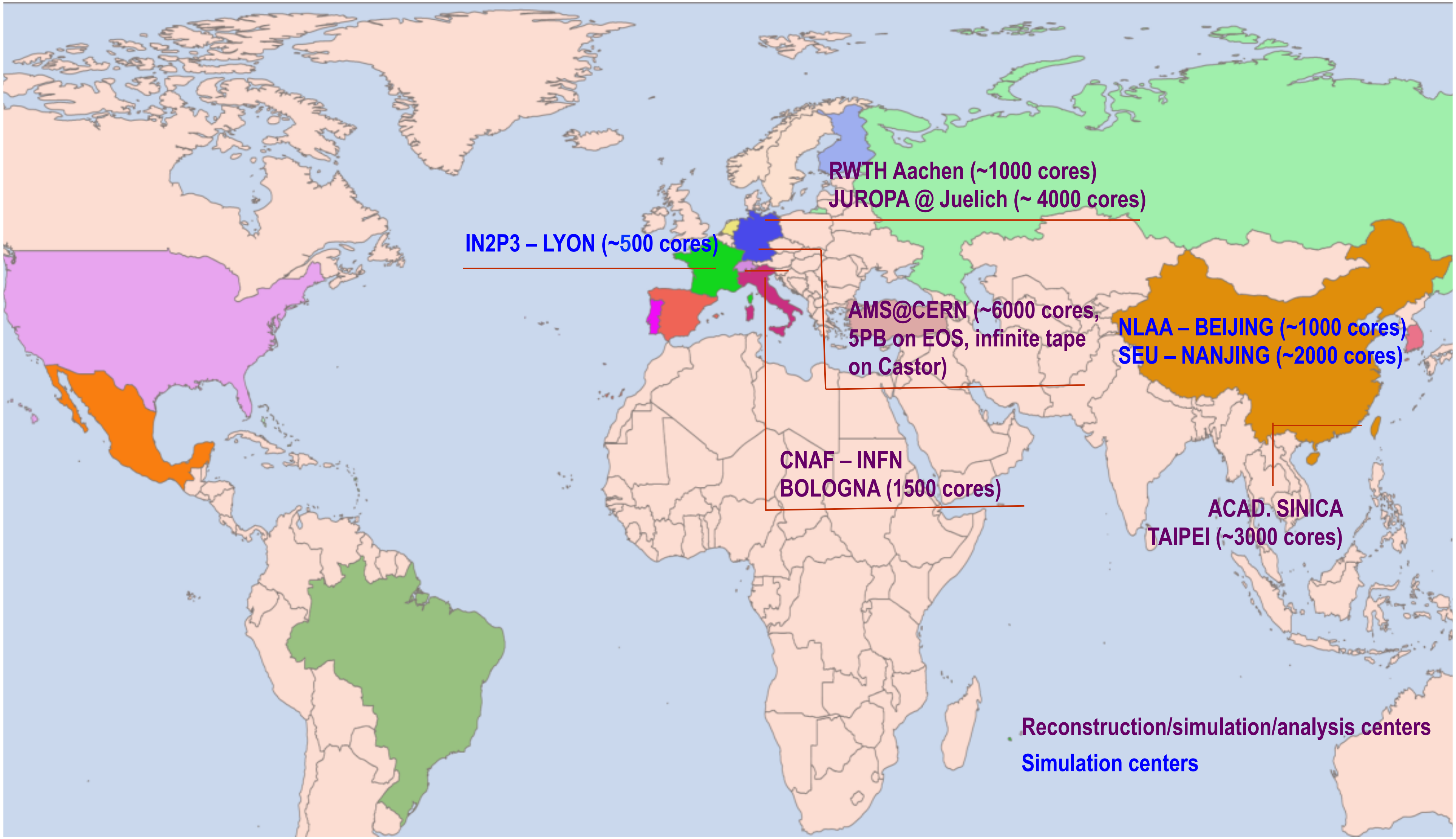
- reduced dataset or "stream": not all the triggers but only the events that most likely will contain the *signal* of the analysis under consideration)
  → each "study group" has its own production and its own data format (directly the complete one or easily permitting the access to it)
- "mini-DST": ROOT ntuples with a lightweight data format (i.e. ROOT ntuples) and with not 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



Major computing resources



AMS data handling and Computing network

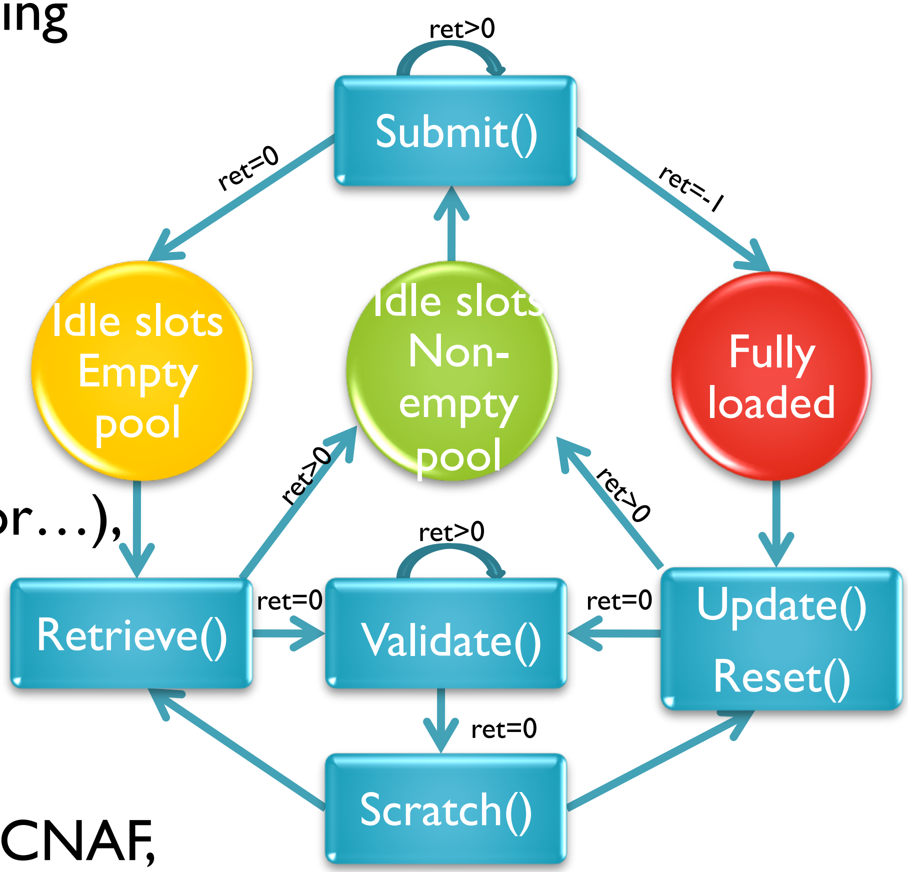
- the "std" production is done in the Scientific Operation Center, SOC, @CERN
  → 200 cores fully dedicated to *deframe*, *merge* & *deblock*, *reconstruct*, …
- the "one-minute ROOT file" production ("std" production prescaled and split in one-minute data files) is done in CERN OpenStack virtual machines
  → 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 an high speed connection
- MC production is done in the regional centers

AMS data handling and Computing network

- 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 an high speed connection
- MC production is done in the regional centers

Light-weight production platform

- Fully-automated production cycle
  - Job acquiring, submission, monitoring, validation, transferring, and (optional) scratching
- Easy to deploy
  - Based on Perl/Python/sqlite3
- Customizable
  - Batch system (LSF, PBS, HTCondor…), storage, transferring, etc…
- Running at:
  - LXBATCH, JUROPA and RWTH, CNAF, IN2P3, NLAA, SEU, AS, …



Sezione di Perugia
INFN Istituto Nazionale di Fisica Nucleare

INFN and CNAF role in the Computing Network

- CNAF joins the effort of the passX full reproduction
- CNAF joins the effort of the MC production
- RAW FRAMES and RAW are copied to tape@CNAF as the Master Copy of the Collaboration
  - Multi-threaded finite state automaton (written in Python) + state transition jobs (written in Perl)
  - It uses a database (Mysql/Oracle) for book-keeping
  - It relies on GRID's file transfer protocols.
  - Thanks to the direct *srm* to *srm* protocol, able to achieve 1.2Gbit/s throughput performance

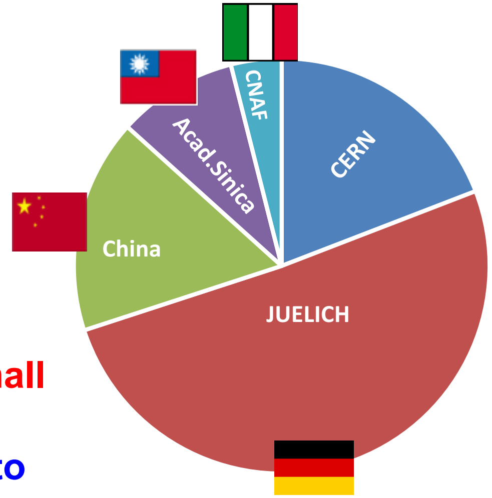
Helium MC campaign 2016

Antihelium and AMS

At a signal to background ratio of one in one billion,
detailed understanding of the instrument is required.

Detector verification is difficult.

1. The magnetic field cannot be changed.
2. The rate is ~1 per year.
3. Simulation studies:



Helium simulation to date:
2.2 million CPU-Days =
35 billion simulated helium events:
Monte Carlo study shows the background is small

How to ensure that the simulation is accurate to
one in one billion?

The few events have mass 2.8 GeV and charge -2 like $^{3}\overline{He}$.
Their existence has fundamental implication in physics.

It will take a few more years of detector verification and to collect more data to ascertain the origin of these events.

73

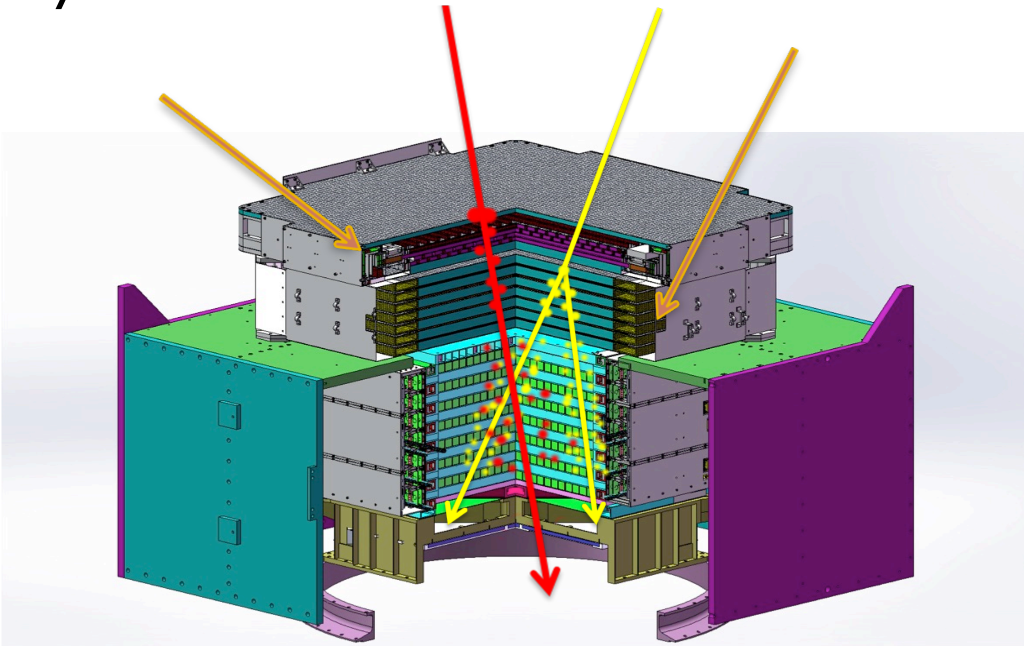
INFN and CNAF role in the Computing Network

- CNAF is also the main computing resource for the Italian Collaboration
  - ~ 12000 HS06
  - ~ 2 PB of storage on *gpfs* + 500 TB of storage on *tape*
  - queue for the production of the "Data Summary Tape" for the Italian analyses ("gold" users)
  - queues for the analysis (all users)
- Remote access of the data @ CNAF from the local farms in the various INFN structures
  - based on the use of the General Parallel File System (GPFS) and of the Tivoli Storage Manager (TSM) + a single, geographically-distributed namespace, characterized by automated data flow management between different locations has been defined (thanks to the Active File Management, AFM, of GPFS)
  - a "pre-selection" scheme permits the access to the full data format only transmitting the interesting events (or even just part of)

DArk Matter Particle Explorer experiment

The DAMPE experiment

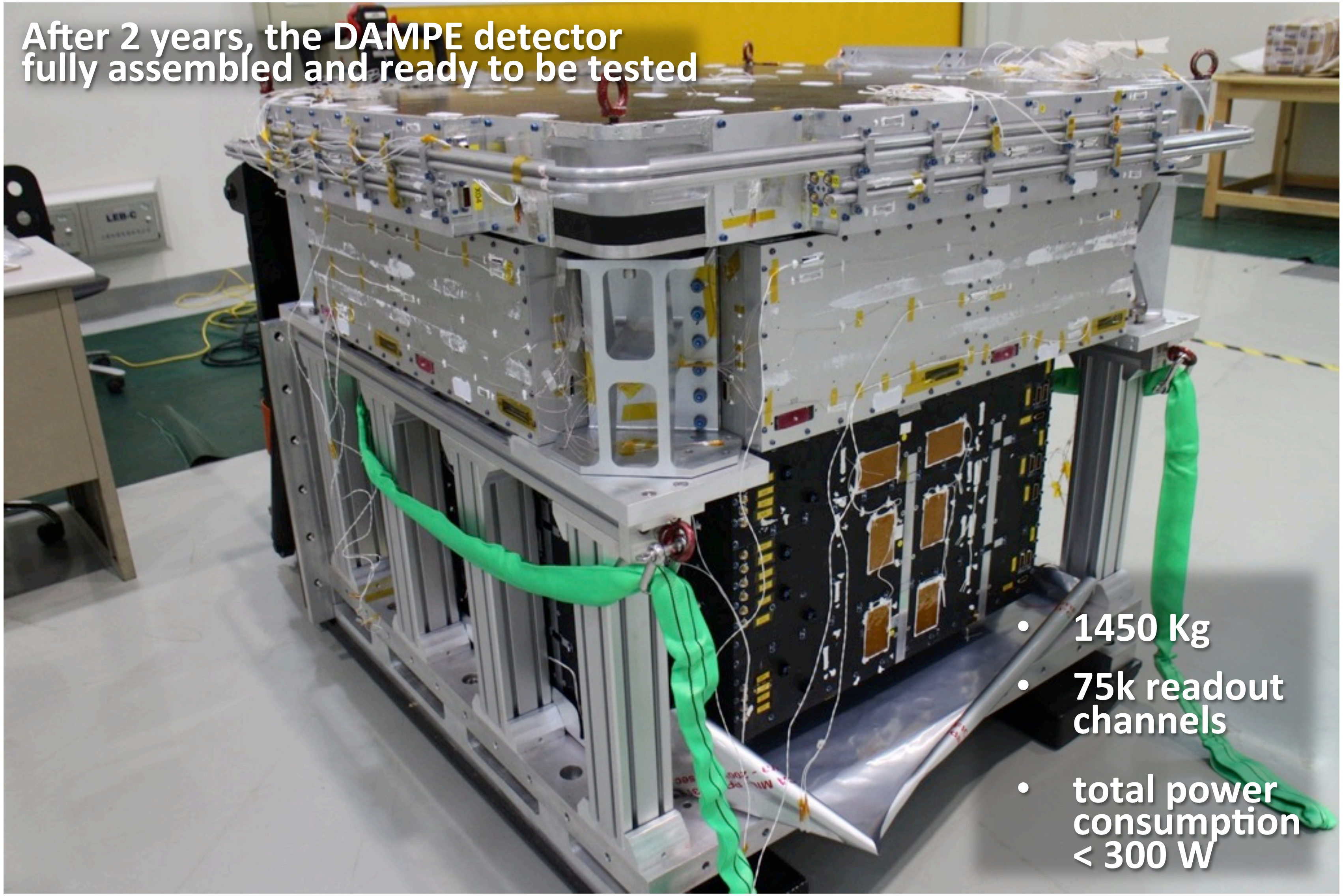
- Dark Matter indirect search (γ -rays and electrons in the GeV – 10 TeV energy range)



- Study of the composition and of the spectral features of CR's, in the GeV – 100 TeV range
- High energy photon astronomy

Sezione di Perugia
INFN Istituto Nazionale di Fisica Nucleare

The DAMPE detector

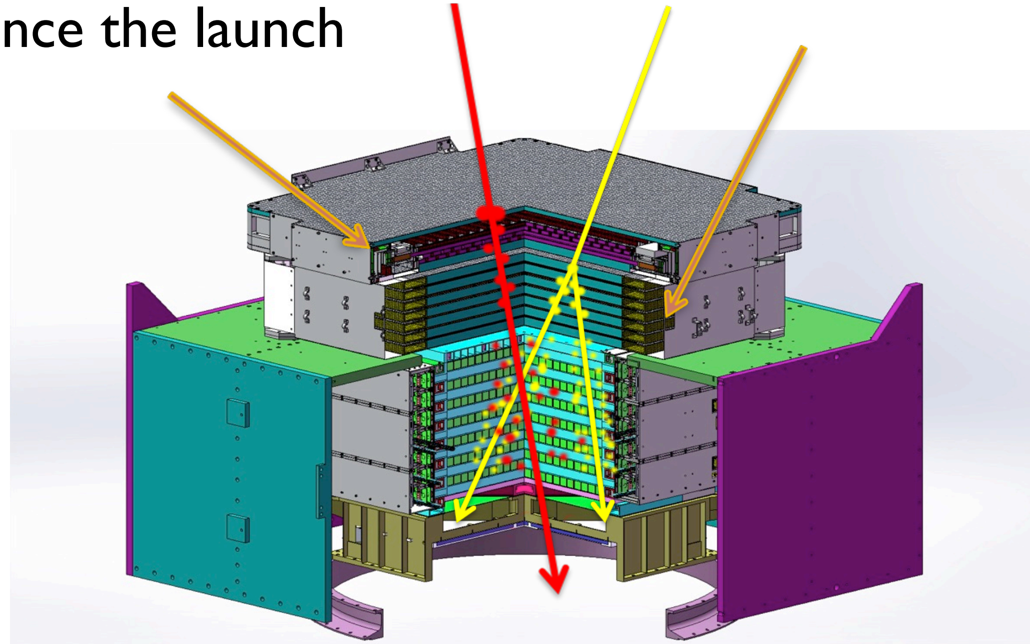


The DAMPE experiment

The DAMPE experiment

- operating in space, on board a Chinese satellite, since Dec 17, 2015
- operations 24h/day, 365d/year, since the launch
- 75k readout channels + temperature sensors
- acquisition rate up to 100Hz
- ~ 15 GB per day transmitted to ground:
  - ~ 15 GB/day raw data
  - ~ 15 GB/day raw data + Slow Control and orbit informations (ROOT format)
  - ~ 70 GB/day reconstructed data (ROOT format)

  → ~ 100 GB/day (35 TB/year) in total

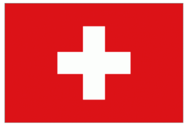


The DAMPE collaboration

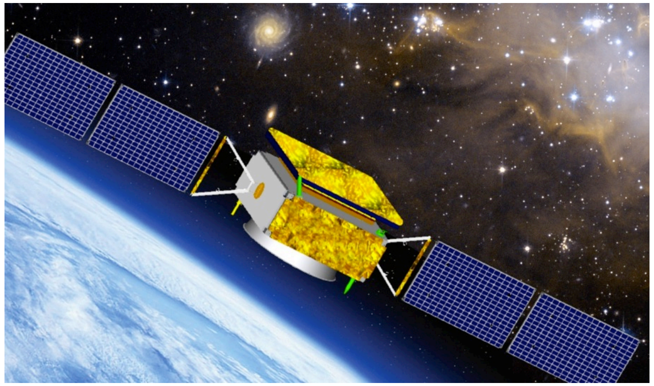
- **CHINA**
  - **Purple Mountain Observatory, CAS, Nanjing** — ***<u>Prof. Jin Chang</u>***
  - **Institute of High Energy Physics, CAS, Beijing**
  - **National Space Science Center, CAS, Beijing**
  - **University of Science and Technology of China, Hefei**
  - **Institute of Modern Physics, CAS, Lanzhou**

- **ITALY**
  - **INFN Perugia and University of Perugia**
  - **INFN Bari and University of Bari**
  - **INFN Lecce and University of Salento**

- **SWITZERLAND**
  - **University of Geneva**



DAMPE data path to Europe



Chinese Space Communication System

China National Space Administration (CNSA) center in Beijing

Purple Mountain Observatory (PMO) in Nanjing

CNAF (Italy)

Local Centers: Perugia, Geneva, Bari, Lecce

DAMPE data path

- Flight data handling reconstruction is done in the PMO cluster
  → 1400 cores that are designed to fully reprocess 3 years (expected mission duration) of DAMPE data within 1 month
- MC production is done in Europe (UniGe-DPNC cluster and, mainly, CNAF and ReCaS Bari)
  → 2016: 400 core-years used to produce all the datasets corresponding to ~ 1 year of flight data
- Data transfer China ←→ Europe is based on *gridftp* and limited to 100 Mb/s (the PMO connection to the China Education and Research Network, CERNET)
  → 6 cores @ PMO
  → during full reproductions: "by hand" (China to Europe and PMO to IHEP) protocol… ☺
- Data transfer Italy ←→ Geneva is based on *rsync*
  → 10 cores @ CNAF

MC computing model

MC production workflow manger:

- light-weight production platform
- web-frontend and command tools based on the flask-web toolkit
- influenced by the Fermi-LAT data processing pipeline and the DIRAC computing framework
- NoSQL database using MongoDB

MC simulation:

- MC based on Geant + custom simulations (digitization, …)
- run almost completely in Italy (CNAF and ReCaS Bari)

MC transfer:

- DAMPE server @ IHEP, Beijing and 'fast' transfer using the Orientplus link of the Geànt Consortium
- IHEP → PMO transfer done using the "by hand" protocol ☺

Sezione di Perugia INFN Istituto Nazionale di Fisica Nucleare

Data access and user analysis

- China and Europe essentially decoupled for connection limitations
- In Europe, MC and flight data are accessible via an XRootD federation (UniGe-DPNC, CNAF and ReCaS).
- The data analysis is done "locally": each institution is using its National resources
- Each study group is defining, producing and using its own "mini-DST" reduced dataset

INFN role in the Computing Network

- CNAF is the "mirror" of the flight data outside China
  - → 100 TB on gpfs (200TB for 2018)
  - → 0 on tape (100TB for 2018)
- CNAF and ReCaS are the main MC production sites
- CNAF is also the main computing resource for the Italian Collaboration
  - → 3k HS06 pledged… Obtained 13k HS06, mainly used for MC production (8k HS06 per 2018)
- ReCaS is also the XRootD redirector

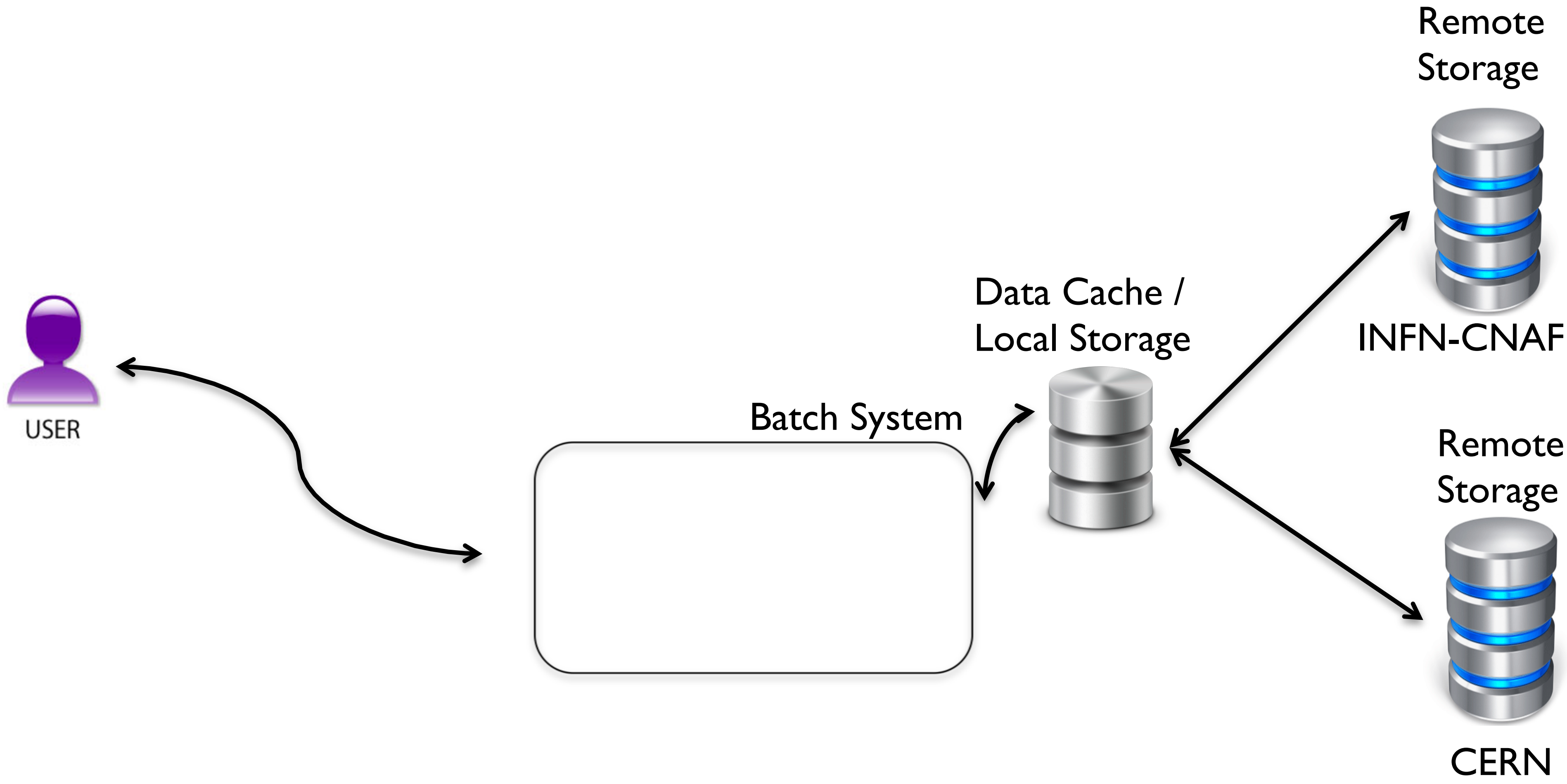
The future: DODAS

- Both the AMS and DAMPE production workflows need to be deployed in several and etherogeneous clusters
  → the workflow is, by design, lightweight and "simple" to allow to be adapted, by hand, to the various regional centers
  → deploying the workflow in "new" resources is not cost-less
- Both the AMS and DAMPE computing models are not fully compliant with the *cloud computing* paradigma
  →deploing the workflow in a modern computing infrastructure such as cloud IaaS is not trivial
- Medium/small size collaborations, such as AMS and DAMPE have not the man power to re-design and re-implement their sw

→ the answer can come from: DODAS

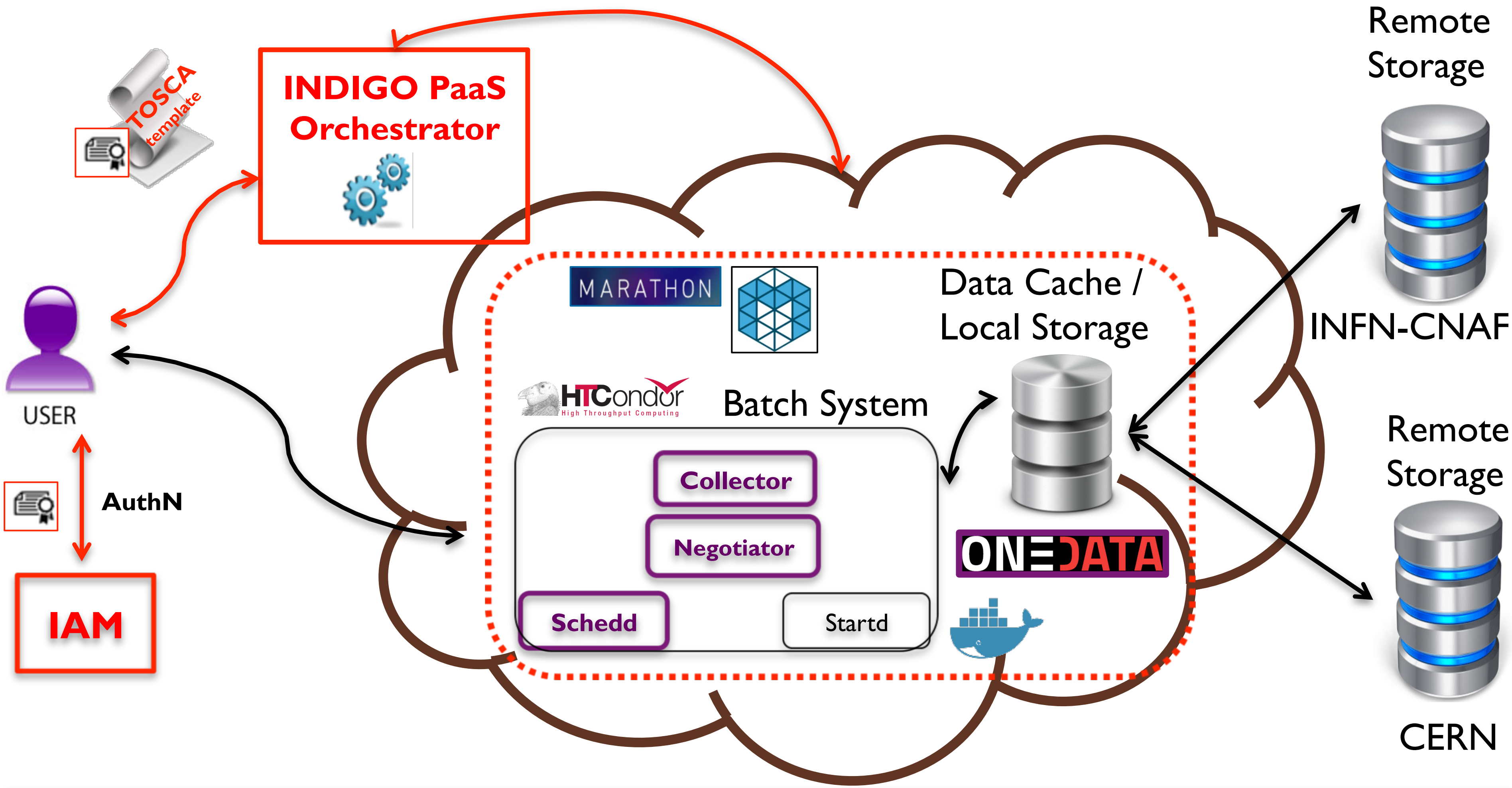
Dynamic OnDemand Analysis Service, DODAS

AMS requirements mostly match solutions provided by
CMS Analysis Cluster On-Demand



Dynamic OnDemand Analysis Service, DODAS

AMS requirements mostly match solutions provided by
CMS Analysis Cluster On-Demand



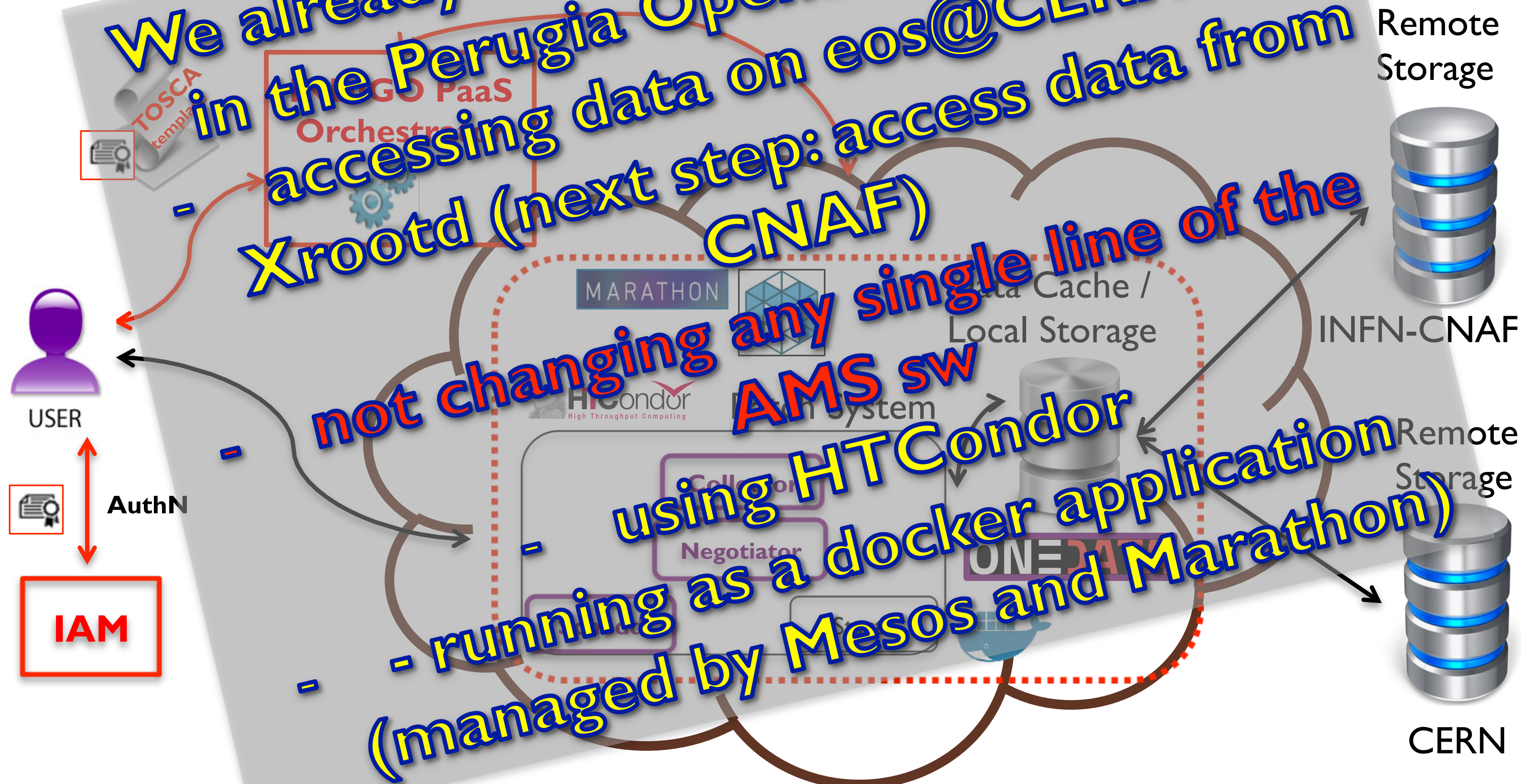
Dynamic OnDemand Analysis Service, DODAS

INDIGO - DataCloud

AMS requirements mostly match solutions provided by
CMS Analysis Cluster OnDemand

We already managed to make it work in the Perugia OpenStack cluster:

- accessing data on eos@CERN via Xrootd (next step: access data from CNAF)
- not changing any single line of the AMS sw
- using HTCondor
- - running as a docker application (managed by Mesos and Marathon)



Sezione di Perugia
INFN Istituto Nazionale di Fisica Nucleare

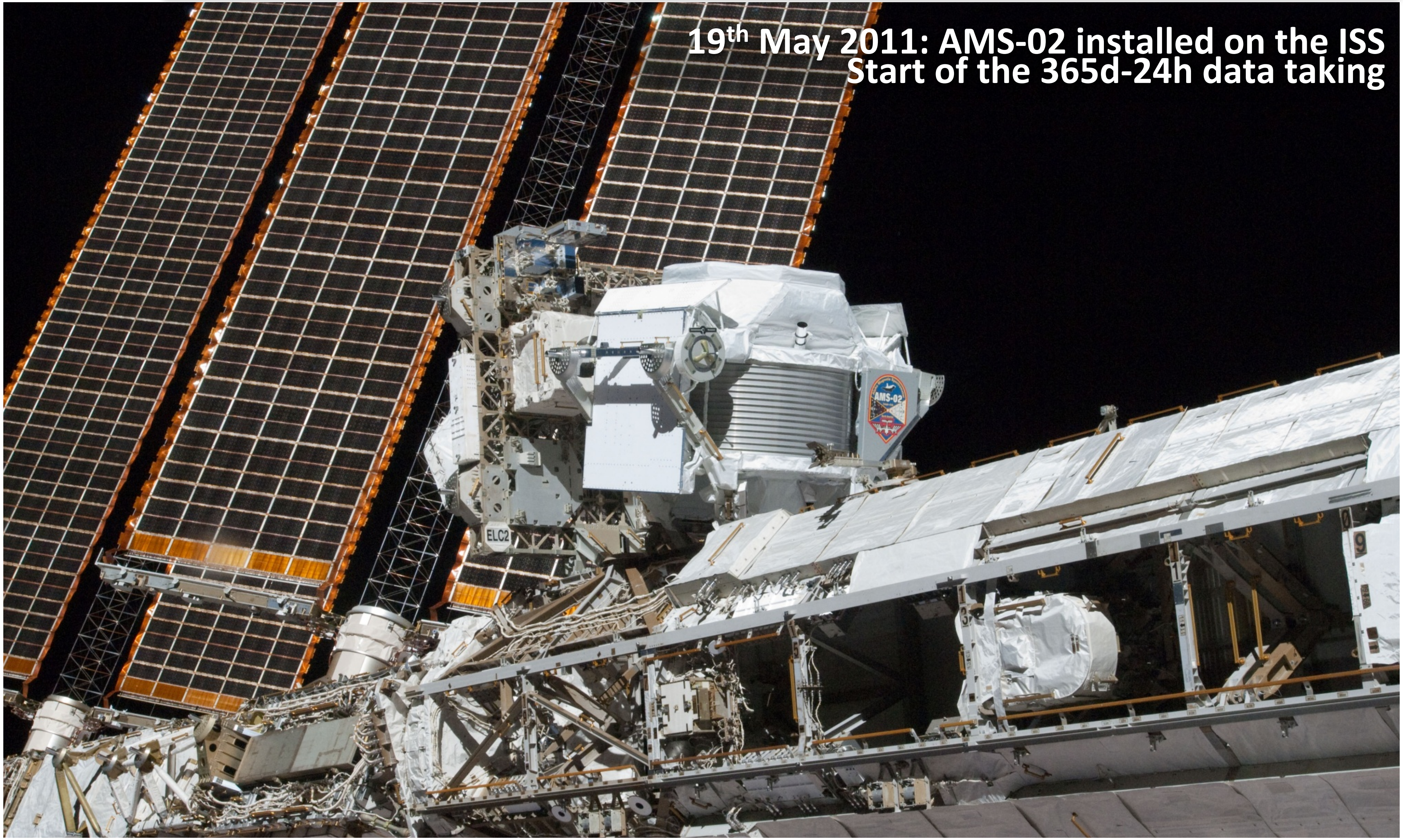
Conclusions

- Both AMS and DAMPE computing models are based on custom solutions developed to solve the unique needs of space-born experiments, with the man power of medium/small collaborations

- Both for AMS and DAMPE the computing resources are often etherogeneous and not tailored for their needs

- A solution to uniform the infrastructural level of the obtained resources and to give access to resources born for the cloud paradigm can come from DODAS

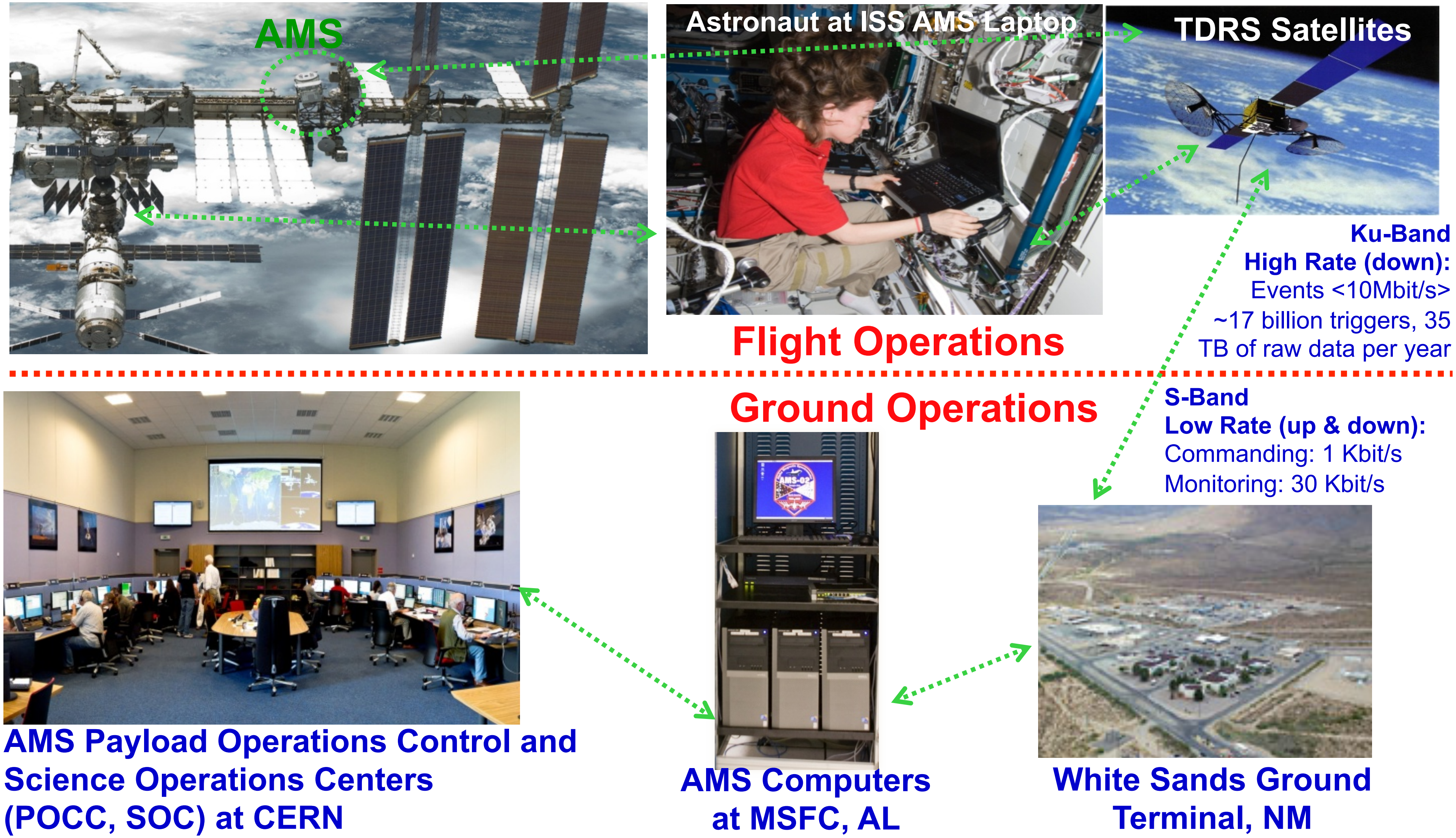
The AMS experiment and the AMS-02 detector



Payload Operation Control Center, POCC inside the BFCR (Blue Flight Control Room) at the MCC-H (Mission Control Center, Houston)

POCC @ CSIST (Taiwan)

AMS data path



AMS ISS-data format

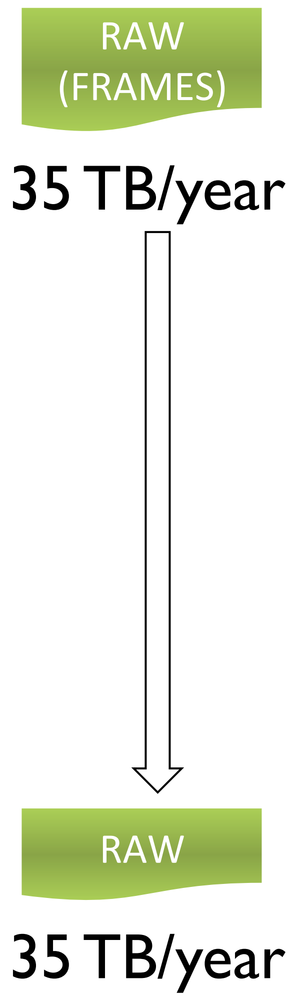
- RAW data from the NASA Marshal Space Flight Center, MSFC (Huntsville, AL) are packed in fixed-size *FRAMES*, uniquely identified by the triplet *(APID, Time, SeqNo)*.
- The data format and protocol are decided by Consultative Committee for Space Data System (CCSDS).

RAW (FRAMES)

35 TB/year

AMS ISS-data format

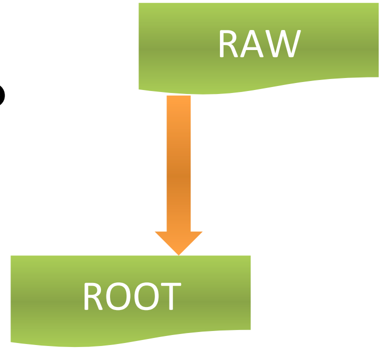
- RAW data from the NASA Marshal Space Flight Center, MSFC (Huntsville, AL) are packed in fixed-size *FRAMES*, uniquely identified by the triplet *(APID, Time, SeqNo)*.
- The data format and protocol are decided by Consultative Committee for Space Data System (CCSDS).
- The *FRAMES* contain, as payload, the real AMS *RAW* data, the AMS-BLOCKS
- Deframing/Merging
  - FRAMES are unpacked (*deframed*) to extract AMS-Blocks
  - AMS-Blocks are *merged* to build-up AMS Science Runs
  - Holes and transmission errors or corruptions are identified at merging time
    → playback from AMS Laptop on ISS



AMS ISS-data format

Reconstruction

- RAW data (i.e. sequences of AMSBlocks) are decoded to extract detector RAW signals
- Reconstruction applied: High level objects are created from the RAW signals
- ROOT files with the 'final' data format are created



Sezione di Perugia
INFN
Istituto Nazionale di Fisica Nucleare

MoU signed on April 2013