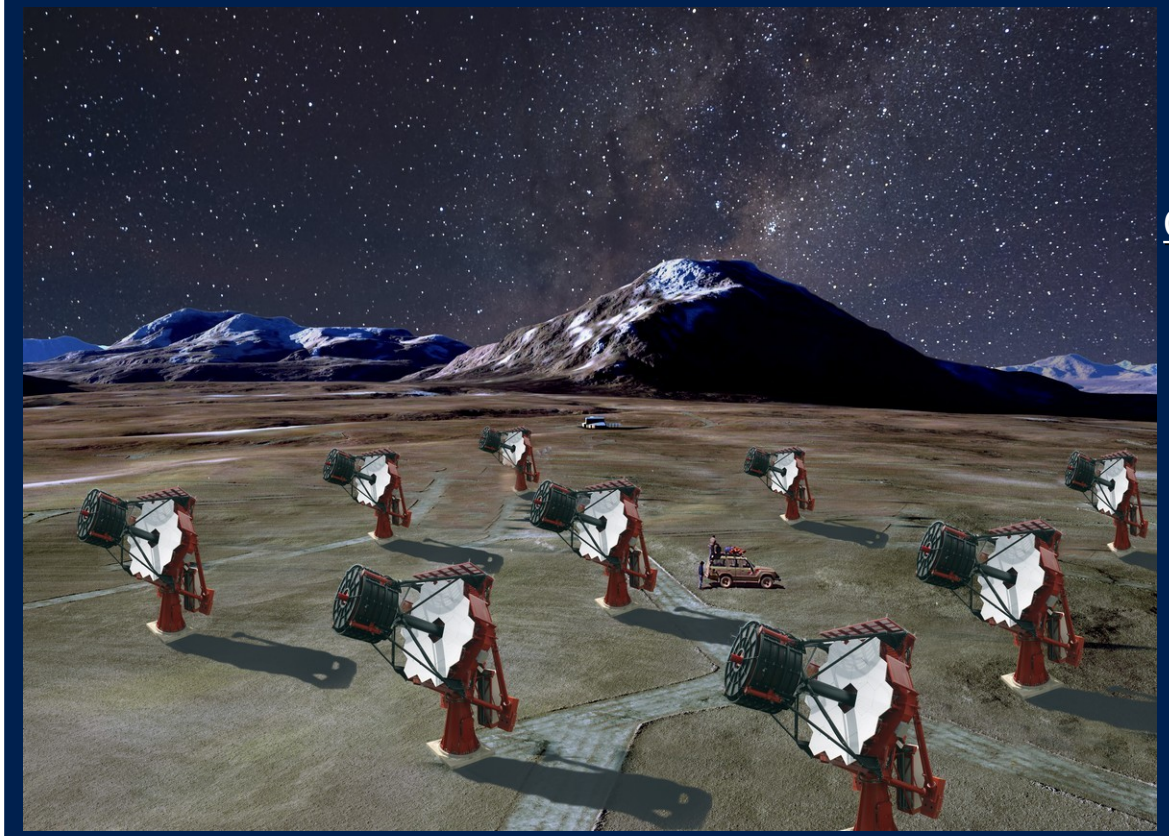
INAF – ICT for ASTRI, ASTRI-MiniArray & the Cherenkov Telescope Array



S. Gallozzi, M. Mastropietro, C. Bigongiari

On behalf of the ASTRI & CTA Team:

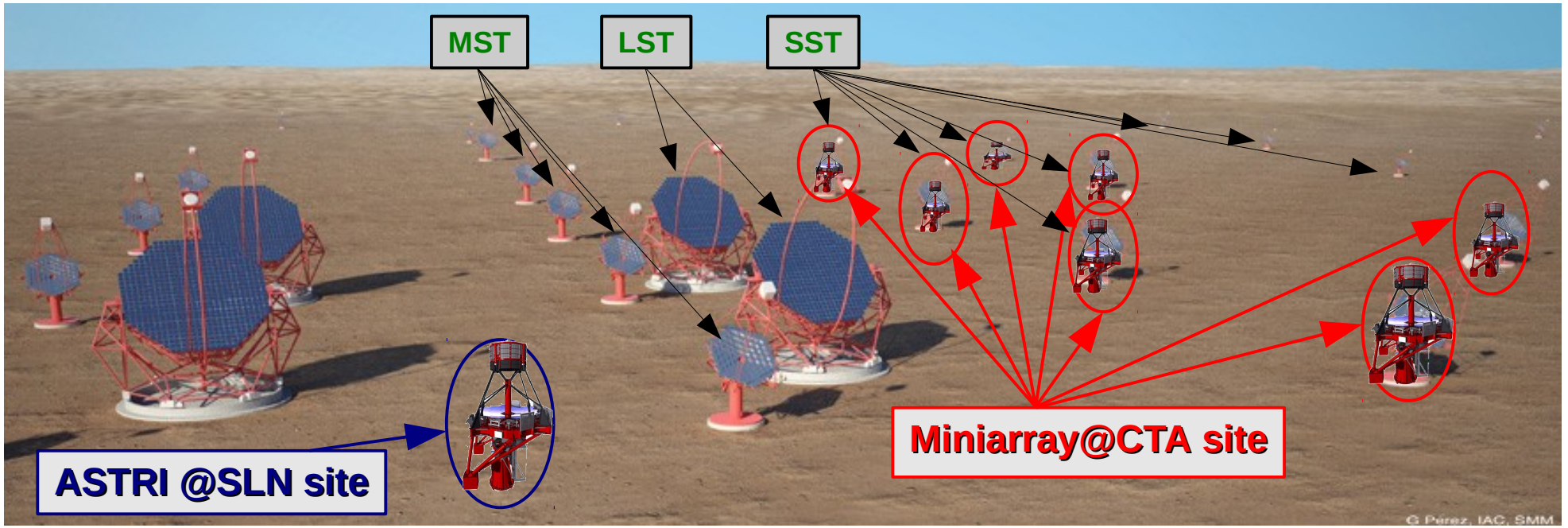
L.A. Antonelli, D. Bastieri, A. Carosi, A. Di Paola, I. Donnarumma, S. Lombardi, F. Lucarelli, P. Munar, A.Stamerra, M. Perri, V. Testa

- **INAF-O.A.Roma**
- **A.S.D.C.**
- **Università di Padova**
- **INAF-O.A.Torino**
- **INAF-I.A.P.S. Roma**

Index

- **ASTRI SST-2M prototype, ASTRI mini array & CTA project**
- **Archive & Data Management**
- **Data Reduction and Computations**
- **Monte Carlo Simulations**

ASTRI & CTA



The **Cherenkov Telescope Array (CTA)** will explore our Universe in depth in Very High Energy (VHE, E > 10 GeV to above 100 TeV) gamma-rays

~20 telescopes for the North-site (Canarie)
~100 telescopes for the South-site (Chile)

An expected RAW data rate of **0.4-5.3 GB/s** for both sites.

CTA will contain **~120 telescopes** of three different sizes (**S**mall / **M**edium / **L**arge):

SST(∅~4m) E>10 TeV FoV~10deg
MST(∅~12m) E~[0.1;1] TeV FoV~6-8deg
LST(∅~24m) E<0.1 TeV FoV~4-5deg

Archive

Without data compression and assuming 165 operational nights/yr:

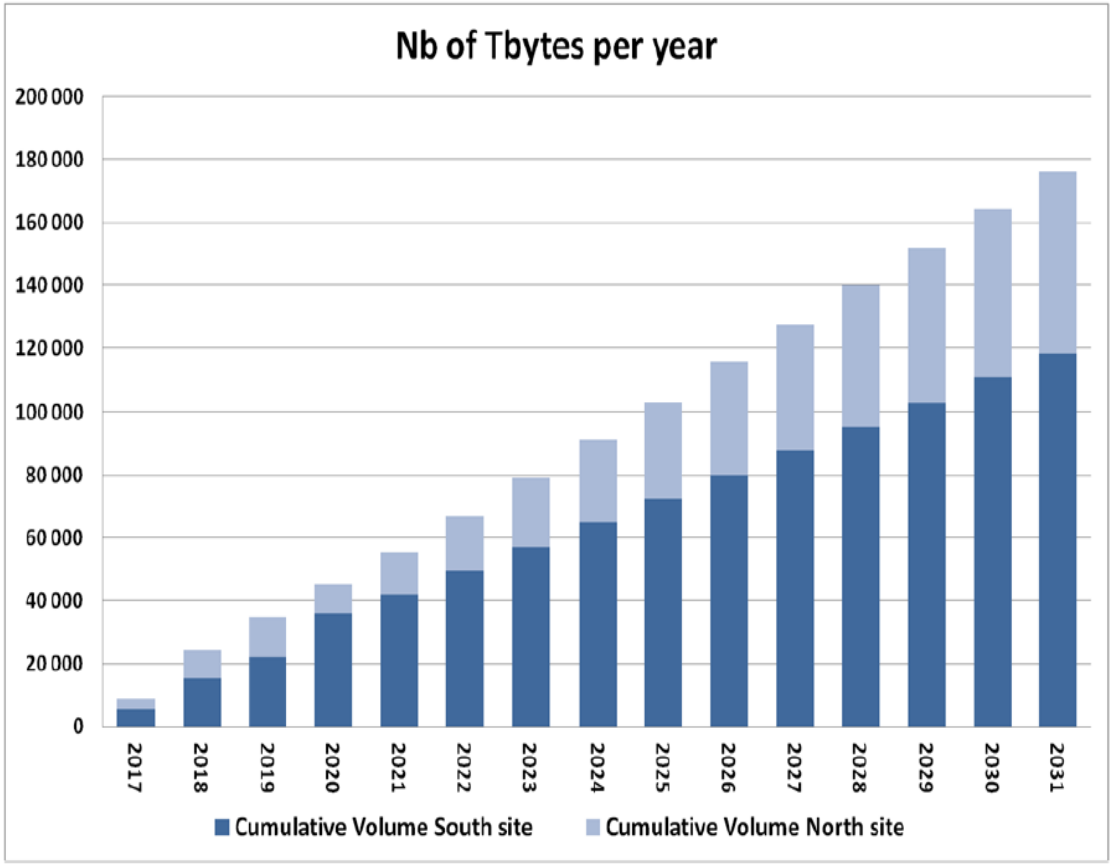
ASTRI/Prot. → ~0.8 TB/night
→ **~0.3 PB/year**
Mini-Array → ~3 TB/night
➔ **~6.1 TB**/night A.R.
→ **~1.0 PB/year** A.R.

CTA → ~8.5 GB/s
→ **~40 TB/night**
➔ **~4 PB/year**
➔ **~20 PB/year** A.R.

(A.R. = After Reduction → input+processed data including calibs, intermediate reduction and MC simulation data)

this is the **OPTIMISTIC SCENARIO**

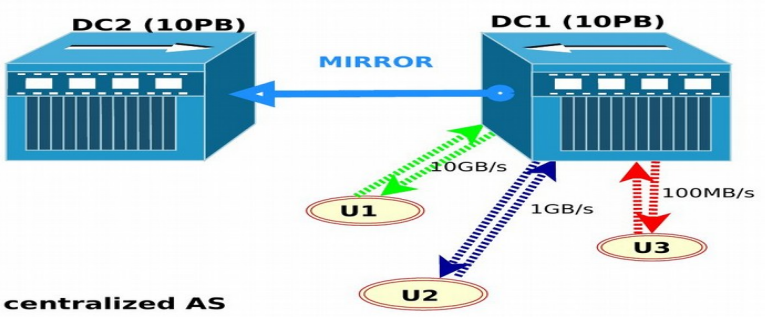
The pessimistic one can take ~>100PB/year !



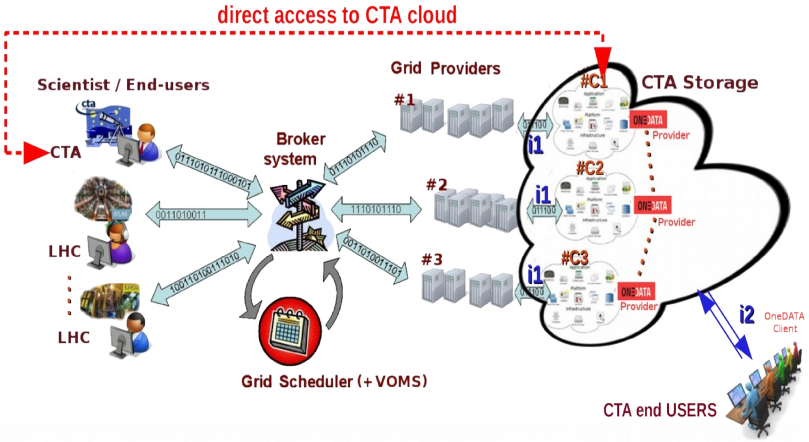
The CTA Archive system must store, manage, preserve and provide easy access (IVOA Access) to a such huge amount of data for a long time.

Archive

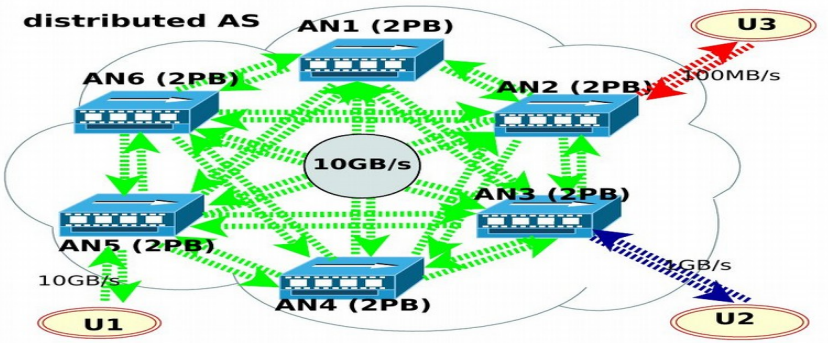
-For **ASTRI prototype** → **Single Mirrored Archive System** for the off-site Archive is enough solution for the management of a the whole data-flow chain. The AS has to be mirrored.



-For **ASTRI-Miniarray project** → **Distributed Archive System** is recommended in order to test all relevant technologies for the whole CTA Archive system. A demo data-model has been Identified and will be produced within the **E.U. INDIGO-DataCloud collaboration**.

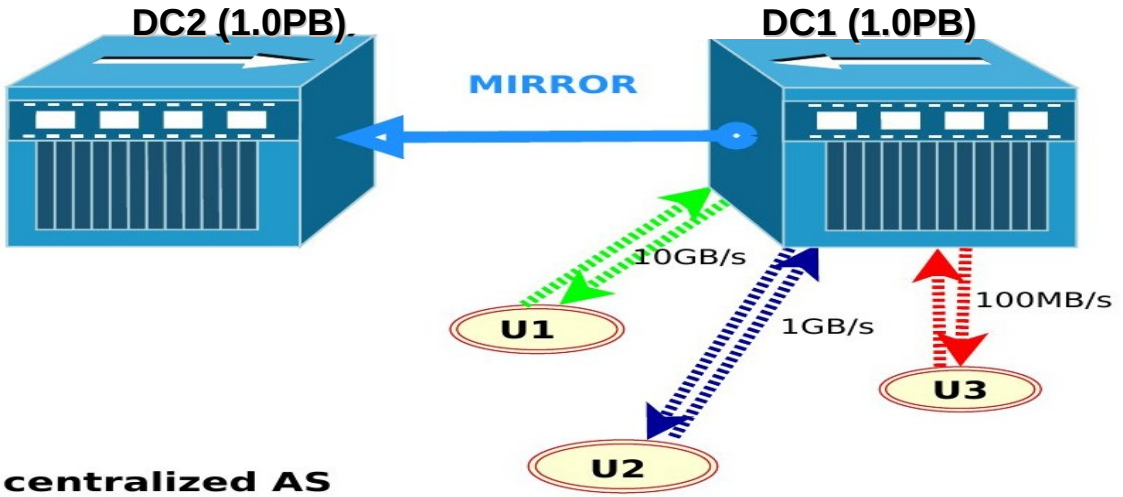


-For **CTA Observatory** → **Distributed Archive System** for the off-site Archive is the only solution for the management of such a large amount of data and can be deployed by the federation of different storage entities.



Archive

-For **ASTRI prototype** →
Single Mirrored Archive System
for the off-site Archive is enough solution for the management of the whole data-flow chain.
The AS has to be mirrored.



Hosted at INAF-Oss.Roma on a dedicated RACK with Storage Server, Computing Nodes and a service frontier server.
Mirrored at ASI-Science Data Center in Rome (n.b. to be installed).

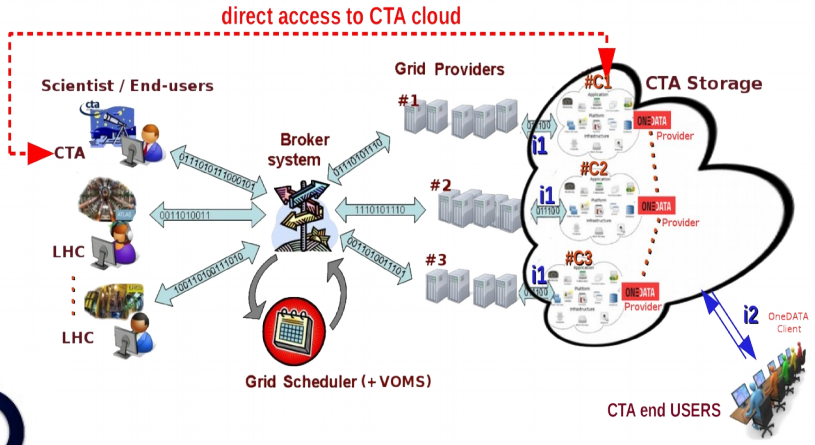
IN-HOUSE solution!

→ NO particular hardware need

Better to enhance network connectivity of INAF-Oss-Roma and ASDC (at least **2-4GB/sec**).

Archive

-For **ASTRI-Miniarray project** → **Distributed Archive System** is recommended in order to test all relevant technologies for the whole CTA Archive system. A demo data-model has been Identified and will be produced within the **E.U. INDIGO-DataCloud collaboration**.



INDIGO - DataCloud

Distributed in three different sites:

For testing technologies (INDIGO-DEMO):

1) INAF-Oss.Roma **2)** ASDC@Rome **3)** INAF-Oss.Catania

For Production (CTA Testbed pre-production) → **INFN Memorandum of Understanding**:

1) INFN-LNF@Frascati **2)** CNAF??? **3)** other....???

NO Mirror needed **INFN resources: (storage: 0,5 PB + computing: 2650 HepSpec06)**

Computation (pipelines and simulation) performed on GRID using CTA VO

CTA Dedicated Storage Nodes on GRID providers

CTA Storage Cloud federation using One-Data-Providers within the INDIGO-DataCloud prj.

CTA External users access using OneData-Clients and APIs.

~2PB of storage servers to dedicate for each GRID node to the CTA Project

~10GB ethernet connection guaranteed by the LHC-One GRID node connections.

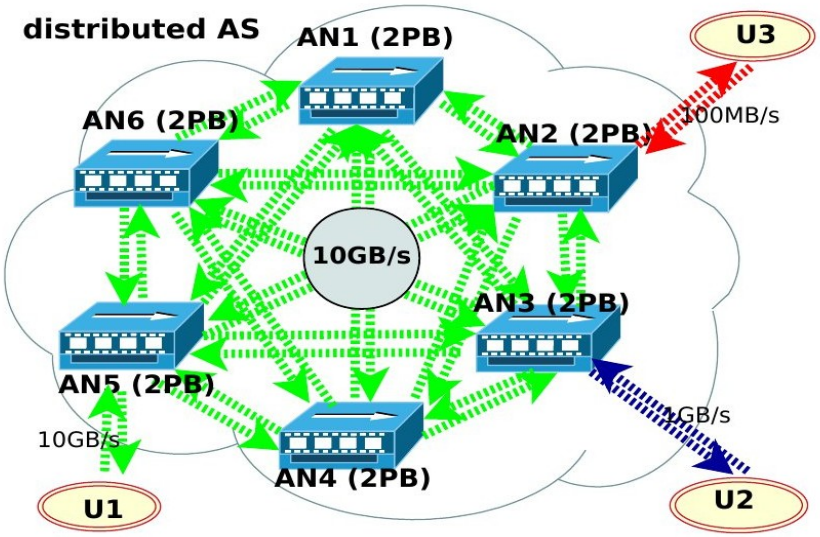
Archive

-For **CTA Observatory** → **Distributed Archive System** for the off-site Archive.

Distributed in different EUROPEAN sites:

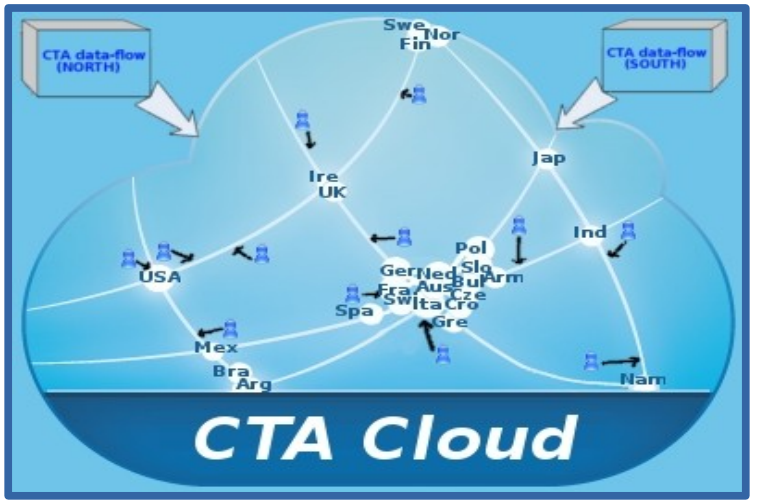
CTA Production (expand the mini-array solution):

1) CNAF@Bologna
2) INFN-LNF@Frascati
3) INFN-Catania
4) LAPP/CNRs (Lione/Annecy)???
5) CERN???
6) DESY???
……



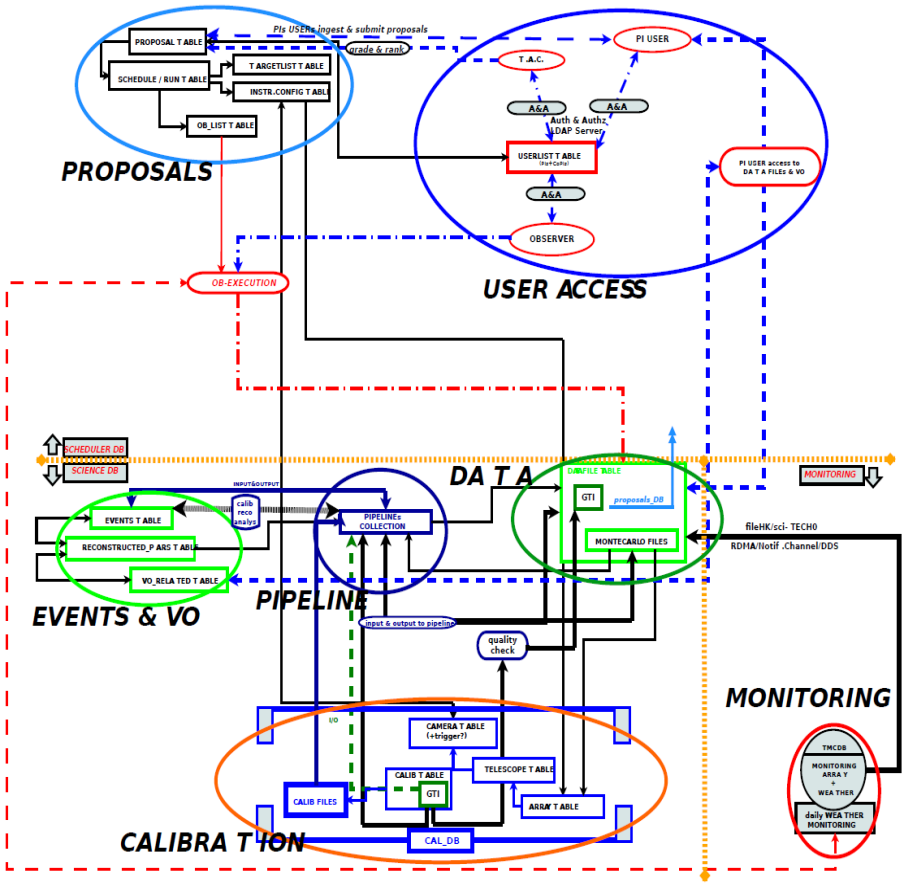
NO Mirror needed

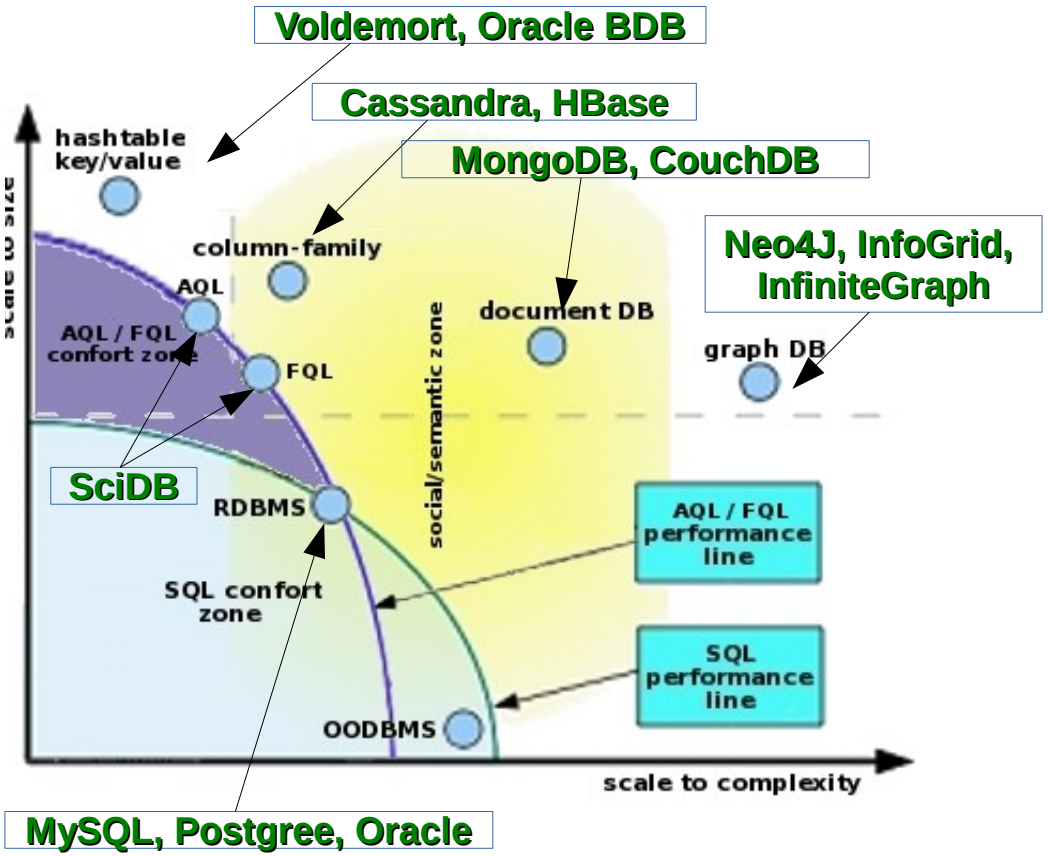
Computation (pipes&simuls) run on GRID using CTA VO
Almost 90% of CTA hardware from the CTA community
Storage acquired year by year upgrading each node
Start with **~2PB of storage servers** to dedicate for each GRID node to the CTA Project
~10GB ethernet connection guaranteed by the LHC-One GRID node connections.



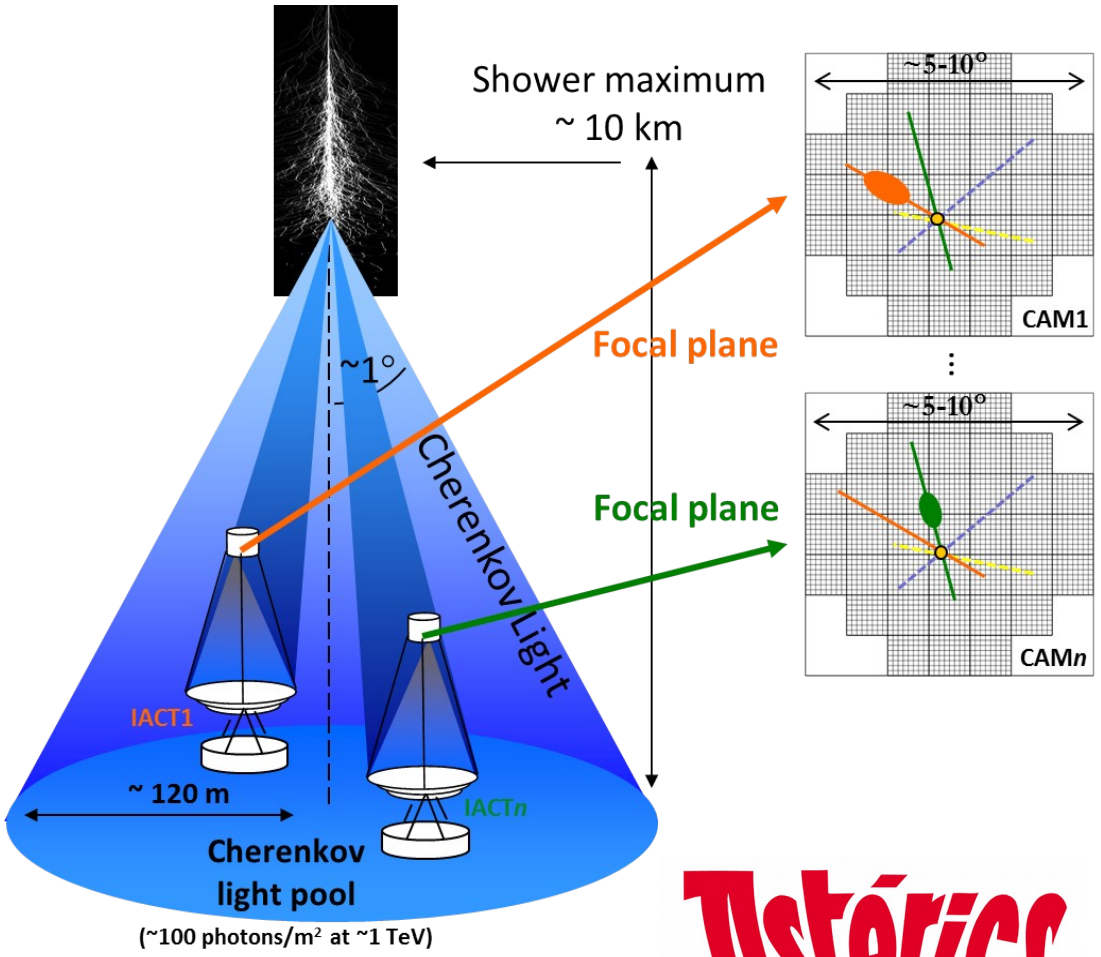
Databases

Testing prototype-databases using ASTRI data-model





Pipelines



Calibration (telescope-wise)
Apply data calibration algorithms and perform data quality checks

Reconstruction (telescope-wise)
Image cleaning and parameterization + single-telescope event reconstruction
(array-wise)
Stereo parameters and reconstruction (event direction, energy, classification)

Analysis (event-wise)
Production of gamma event list,
IRFs and good time intervals (GTI)

Science
Production of detection plots, spectra, sky maps, light curves

Astronomy ESFRI & Research Infrastructure Cluster

ASTRI Pipeline

- The ASTRI Scientific Software is highly parallelizable.

  On some algorithms it may exploit:

  - shared-memory multiprocessors using OpenMP
  - NVIDIA GPUs using CUDA programming model
- Our current computing node is a:
  - dual **Intel® Xeon® E5-2650** CPU 16 physical cores at 2 GHz,
  - 128 GB of RAM,
  - 1 NVIDIA Tesla K20c.
- Our server is able to process **one night of mini-array data taking in 4.5 h**.
  - time is acceptable,
  - facility slightly outdated, (arch. *Sandy Bridge*, no FMA support) but more than enough for our goals.
- For its nature, the A-SciSoft pipeline is highly IO bound:
  - data should be near the computing node.

CTA Pipelines

- Computing needs in construction phase. All values are in (Billion HS06 · sec)

Year	2017	2018	2019	2020	2021
Data pipeline needs	19.7	59.1	98.5	138	177
Simulation needs	360	360	360	360	360
Re-processing needs	0	19.7	78.8	177	315
Cumulative needs	**380**	**439**	**537**	**675**	**853**

- Computing needs in operation (from t0 to t0+10 years).

Year	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
Data pipeline needs	197	197	197	197	197	197	197	197	197	197
Simulation needs	360	360	360	360	360	360	360	360	360	360
Re-processing needs	493	690	887	1084	1281	1478	1675	1872	2070	2267
Cumulative needs	**1050**	**1247**	**1444**	**1641**	**1838**	**2035**	**2232**	**2430**	**2627**	**2824**

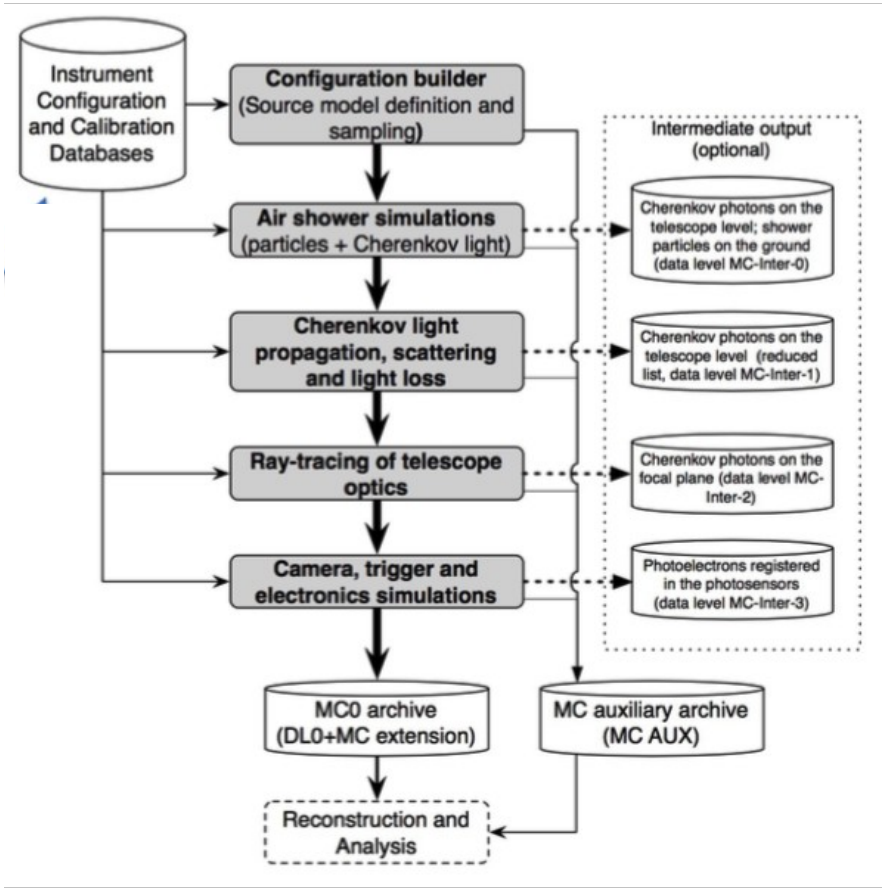
Uge amount of computing power needed for re-processing

Numbers from CTA TDR DATA 1.5 20160226

Simulations

Intensive **Monte Carlo simulations** are an essential tool for Imaging Air Cherenkov Telescopes during both the design/pre-construction phase and the data-taking phase

- Optimize many design and technical details
  - Array site
  - Telescope layout
  - Pixel signal extraction
  - Trigger logic
  - …….
- Optimize signal extraction strategy
  - Observing strategy
  - Zenith angle
  - Atmospheric conditions
  - …….
- Estimate the Instrument Response Functions



Since the background rejection is very high huge numbers of background events need to be simulated ≈ 10^9 – 10^{10}

Distributed Computing

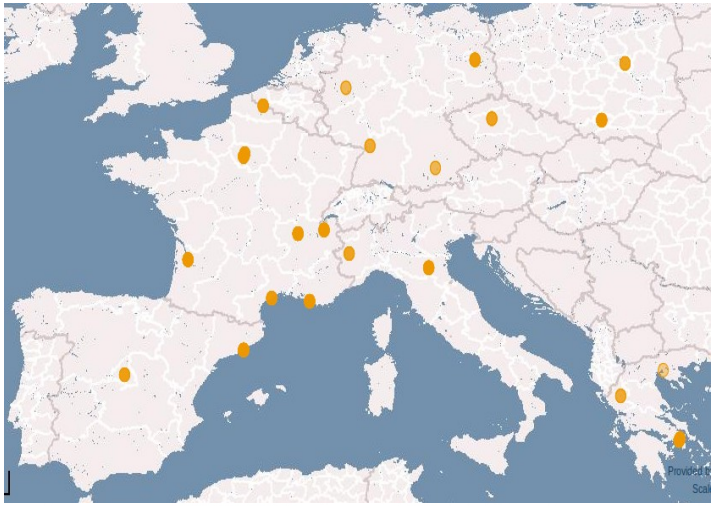
- Use of an existing and reliable e-Infrastructure: EGI grid
- Use of DIRAC for Workload and Data Management

CTA Virtual Organization:

- Active since 2008
- 19 EGI sites in 7 countries and 1 ARC site
- About 100 members

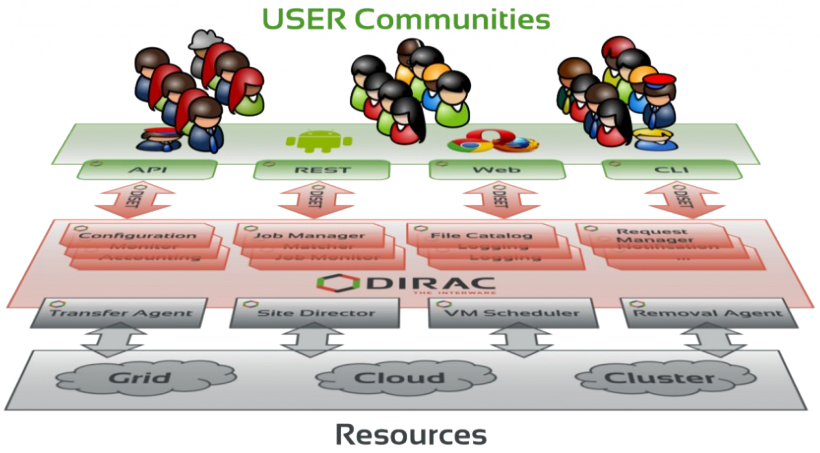
Resources:

- Dedicated storage:
  - Disk: 1.3 PB in 6 sites:
    CC-IN2P3, CNAF, CYFRONET, DESY, GRIF, LAPP
  - Tape: 400 TB in 3 sites
- CPU: 8000 cores available on average



Computing Model:

- CTA simulation exploits CPU resources of all GRID sites supporting the CTA VO
- Output Data are stored at 6 main Storage Element
- MC data analysis takes place at the 6 sites where MC data are stored



The story so far

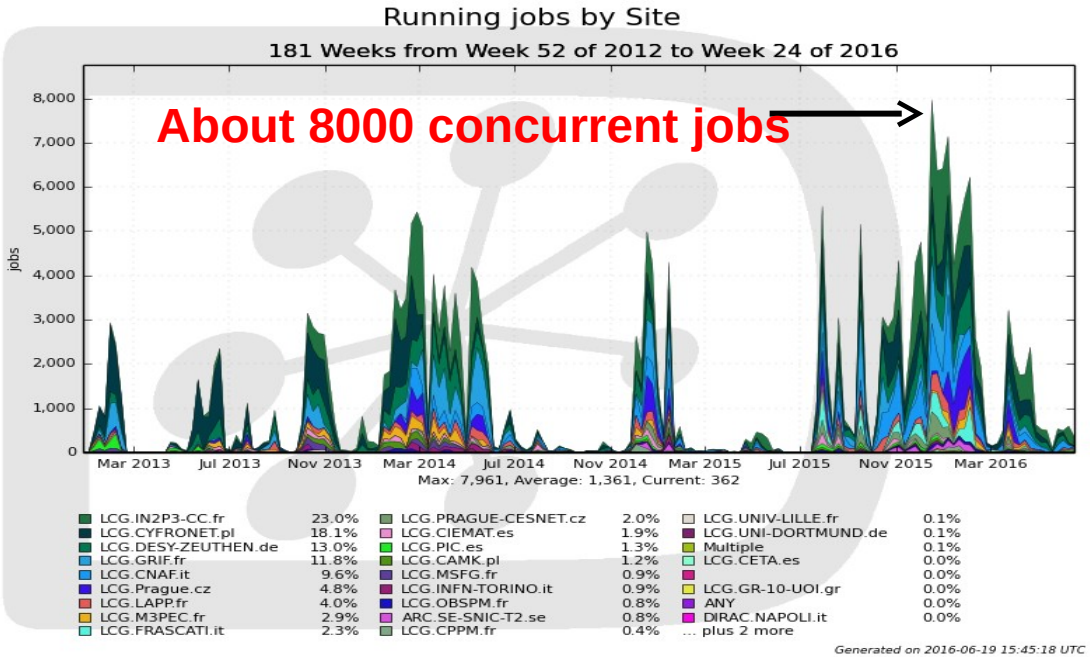
CTA is in the “Pre-Production” phase.

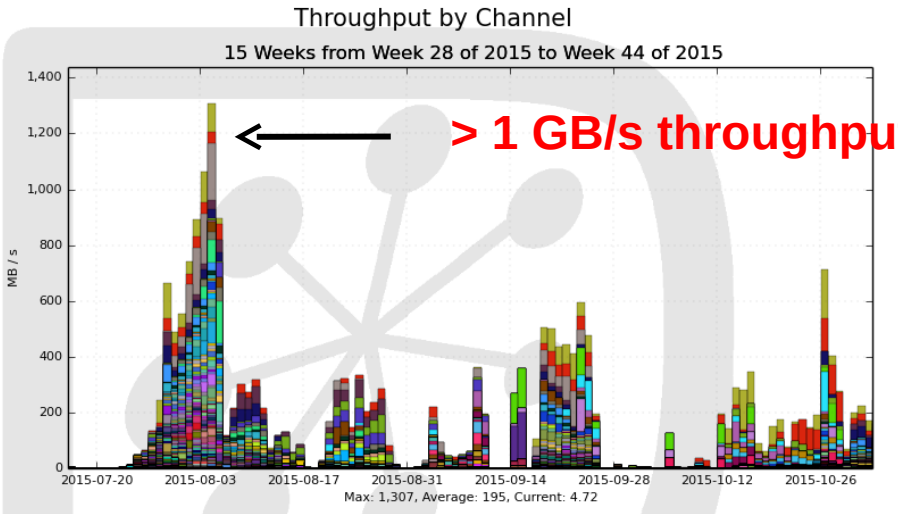
Massive MC simulations and Analysis campaigns

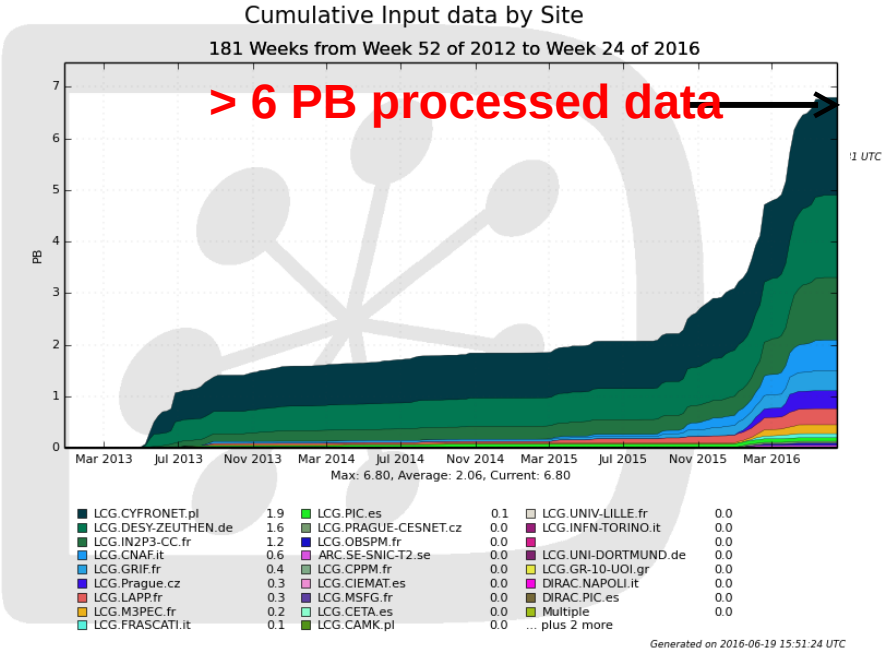
- ‘Prod1’(2012)
  - Preliminary studies
- ‘Prod2’ (2013-2014)
  - Characterize all site candidates to host CTA telescopes
  - 4.6 billion events generated for all 8 site candidates
    - 2 different zenith angles and 3 different telescope configurations
- ‘Prod3’ (2015) in progress:
  - For the 2 selected sites: study different telescope layouts
    - 800 telescope positions
    - 7 telescope types,
  - Run 3 different Analysis chains on the simulated data
    - Each one processing about 500 TB and 1 M of files for 36 different configurations

Computing is already a challenge for CTA !

Resource Usage







Average number of concurrent running jobs around 1400

Average throughput around 200 MB/s

- ~ 2 PB data produced
- ~ 7 PB data processed
- > 20 M files registered in the Catalog

Data taking phase

Two approaches under discussion:

- Classical Approach: generate MC sets for each point in a large phase space
  - Fixed zenith, azimuth, source offsets, night-sky background, array layout, atmospheric profiles, etc.
  - Any change in the instrument (e.g. trigger settings, HV changes) requires a new MC set
  - Phase space is huge if one aims to meet systematic uncertainty requirements

- Run-wise Approach: produce MC sets for certain observation periods/directions
  - Simulate (sub-)array of telescopes that are tracking a sky position consider e.g. broken pixels, calibration, night-sky background, atmospheric model for this run
  - By definition much closer to reality than full phase-space approach;

Many science goals of CTA will be limited by systematics:

A number of MC events of the same order of magnitude of real data is needed for any point in the phase space or any period of observations !

Needed resources

Year	2017	2018	2019	2020	2021
Cumulative raw data volume(PB)	4	10	12	15	19
Cumulative event data volume(PB)	4.5	12	17	25	35
Monte-Carlo data volume(PB)	5	12	17	20	20
Technical data volume(PB)	0.01	0.04	0.09	0.16	0.26
Cumulative Data (PB)	**9**	**24**	**35**	**45**	**55**

The amount of needed MC data has been estimated in 5 PB at the beginning of the data taking phase increasing

Extrapolating from present resource usage for MC data simulation we can estimate that about 300 M HS06 hours of CPU at the beginning increasing later to about 1200 M HS06

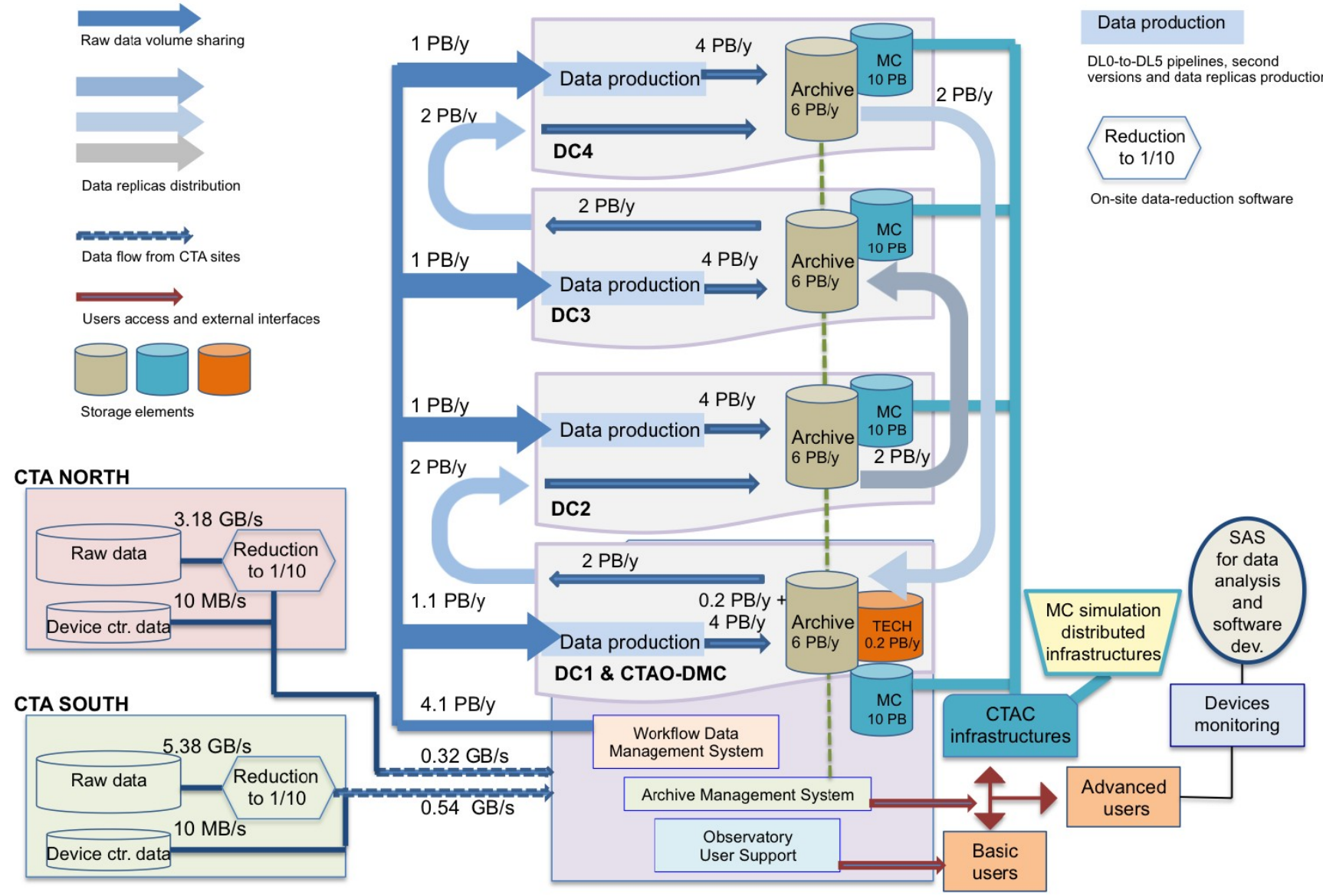
	2013	2014	2015	2016
CPU used (M HS06 hours)	49	110	93	59
Disk (TB)	874	1004	1426	2281

The expected ASTRI Mini-Array amount of real data is about a factor 20 less that that of the entire CTA, so we can estimate that about 250 TB o MC data will be needed yearly which will require about 15 M HS06 hours.

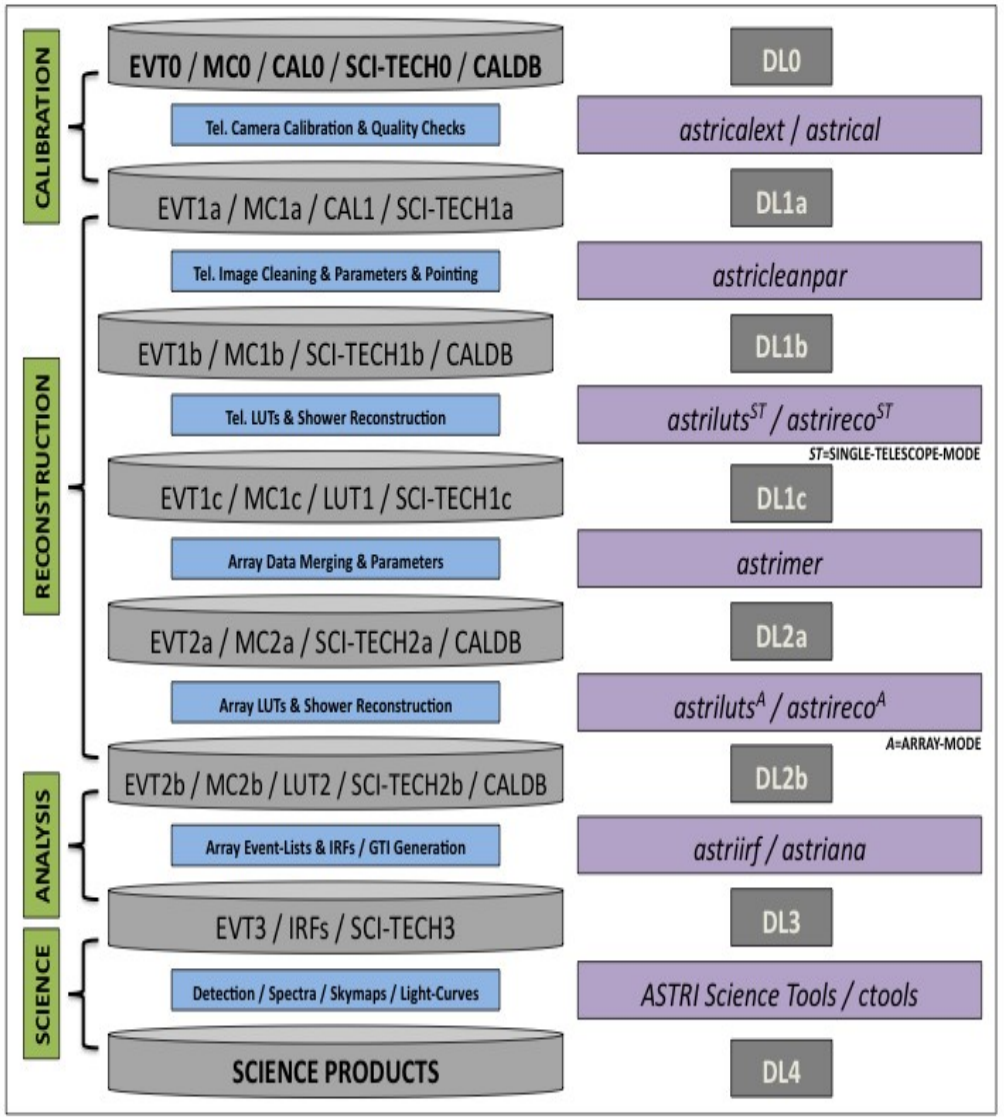
Questions?

Backup slides

CTA Data Flow



ASTRI Pipelines



The ASTRI Scientific Software:

- can handle both real and simulated Monte Carlo data
- is written in C++/Python with some CUDA accelerated algorithms
- manages FITS data from DL0 to DL4 (/ libraries) (for mini-array, DL0 in RAW format)
- can run on x86 & NVIDIA GPUs (use of low-power ARM CPUs is envisaged)
- is developed in independent software modules linked by pipelines written in Python
- follows the general CTA design and data model scheme defined in CTA Data Management
- makes use of ad hoc and adopted CTA Science Tools