Data management challenges of the Cherenkov Telescope Array observatory

Giovanni Lamanna

LAPP - Laboratoire d'Annecy-le-Vieux de Physique des Particules,

Université de Savoie, CNRS/IN2P3, Annecy-le-Vieux, France

ASPERA-Workshop,
Hannover 3-4 May 2012

Outline

- Introduction about CTA
- Data, IT and Computing challenges:
  - data acquisition
  - data transmission
  - data pipelines
  - data access
- Conclusions

Introduction

The content of this talk reflects the status of the activities in the context of the "Data Management" organization for CTA.

I wish to acknowledge all members of the CTA Consortium Work Packages who have (directly and indirectly) contributed to it.

In a few cases I will provide
some PERSONAL views and ideas
in line with the aim and the nature of this workshop

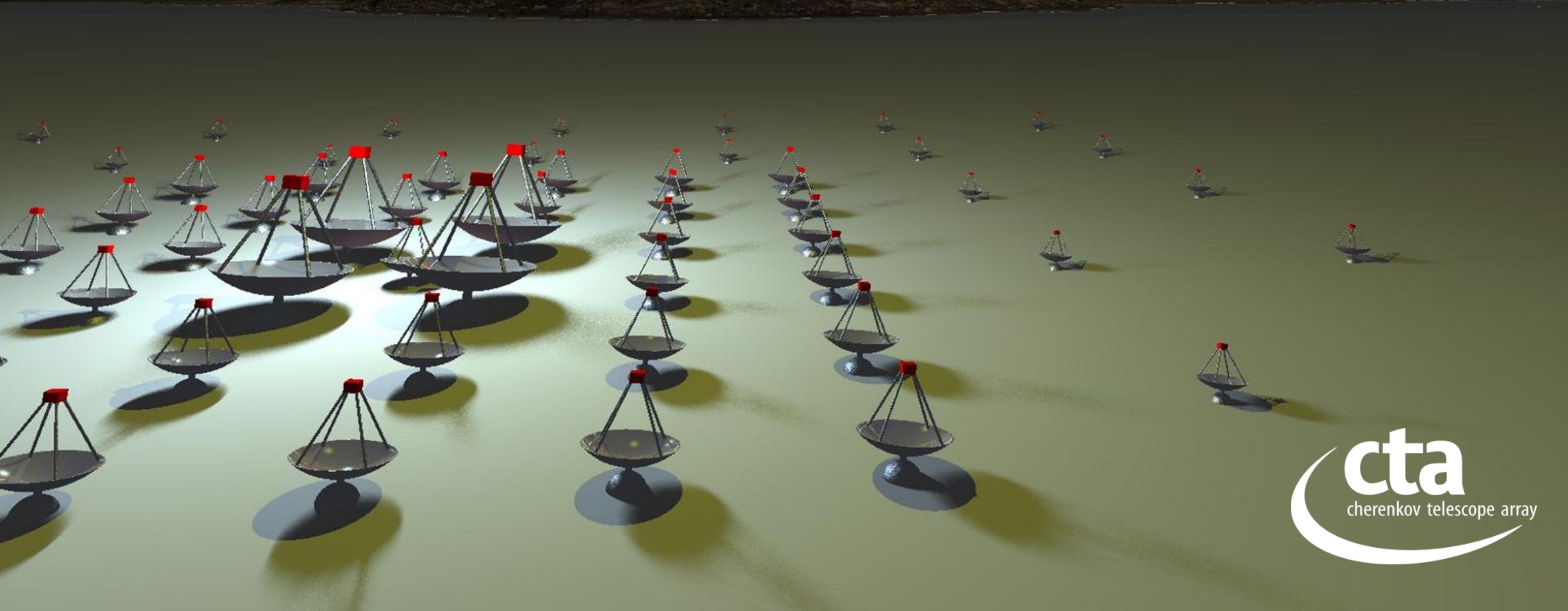
CTA

The first world-wide
ground-based
Very High Energy
γ-ray Observatory

The next decade
Astroparticle infrastructure

An international consortium of
25 countries and
> 800 scientists

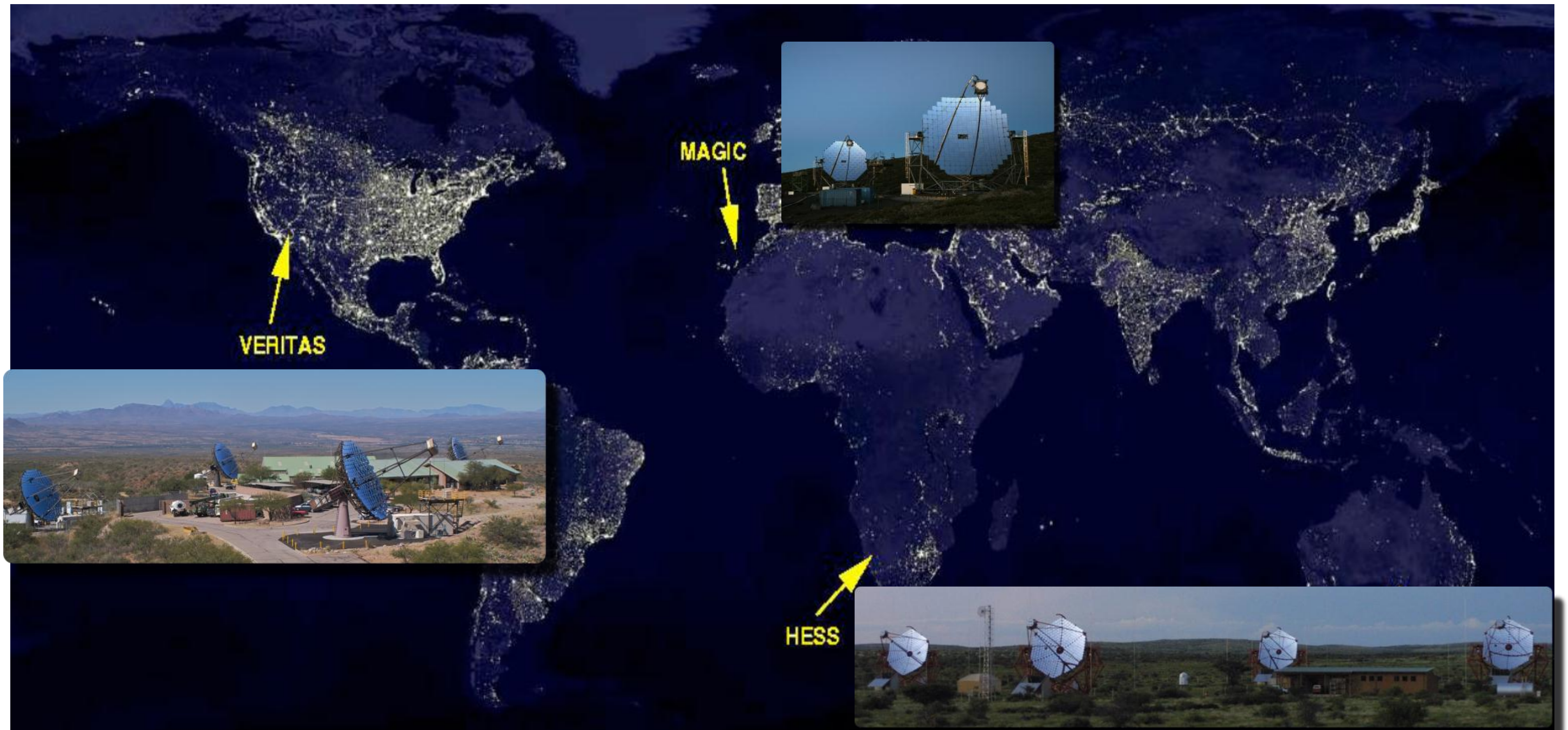
In the ESFRI roadmap since
2008 and
an ASPERA priority



cta
cherenkov telescope array

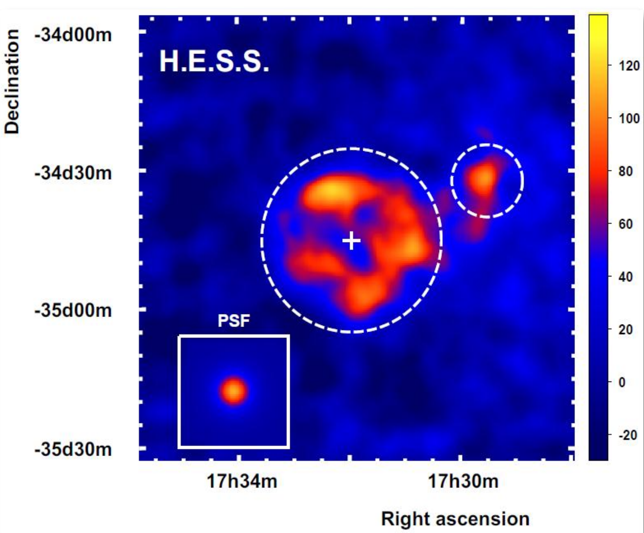
Imaging Air Cherenkov Telescopes (IACT)

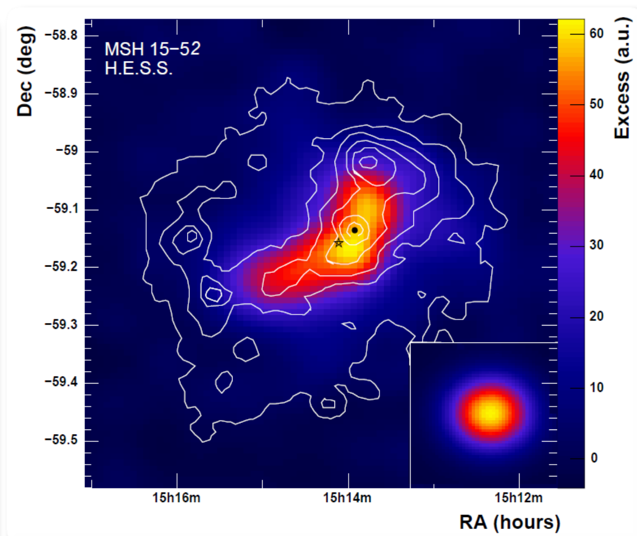
- CTA represents the beginning of VHE gamma-ray Astronomy at TeV (10^{12} eV) energy scale (billion times energy of X-rays)
- Currently 3 (IACT) experiments with 2 – 5 telescopes

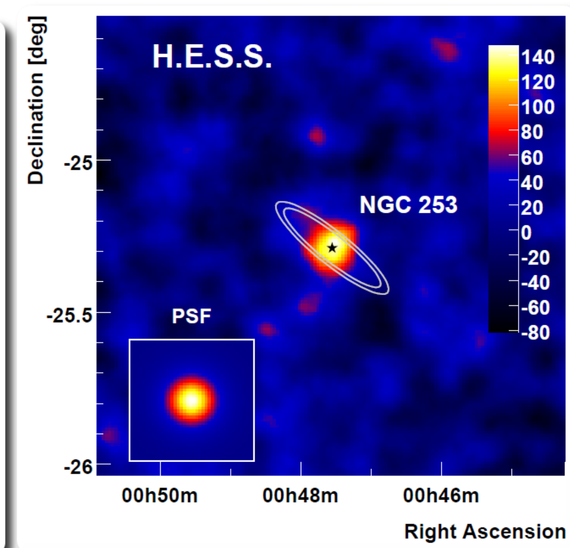


VHE gamma-ray Astronomy

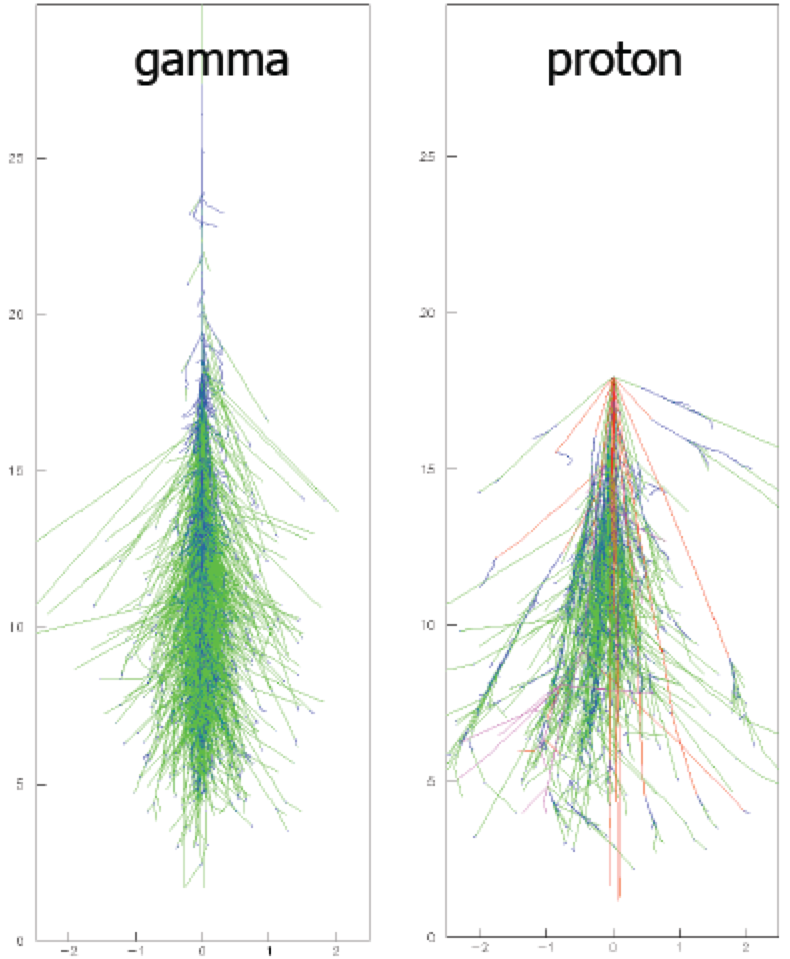
- 1st discovery in 1989
- Current instruments operational since 2003 (H.E.S.S.)
- Discovery of hundred of sources in the last 10 years
- Observation of non-thermal universe
  - Supernova remnants, Pulsar Wind Nebulae, galaxies, Active Galactic Nuclei
  - “Astroparticle”: Dark matter, Cosmology, fundamental physics







IACT detection technique



- CR (protons dominate over gammas) creates air shower of tens of thousands of secondary particles

- Cherenkov cone emission light from shower particles

  1º at 10 km height;

  100 m radius light pool;

  ~ 5ns light 'flash'.

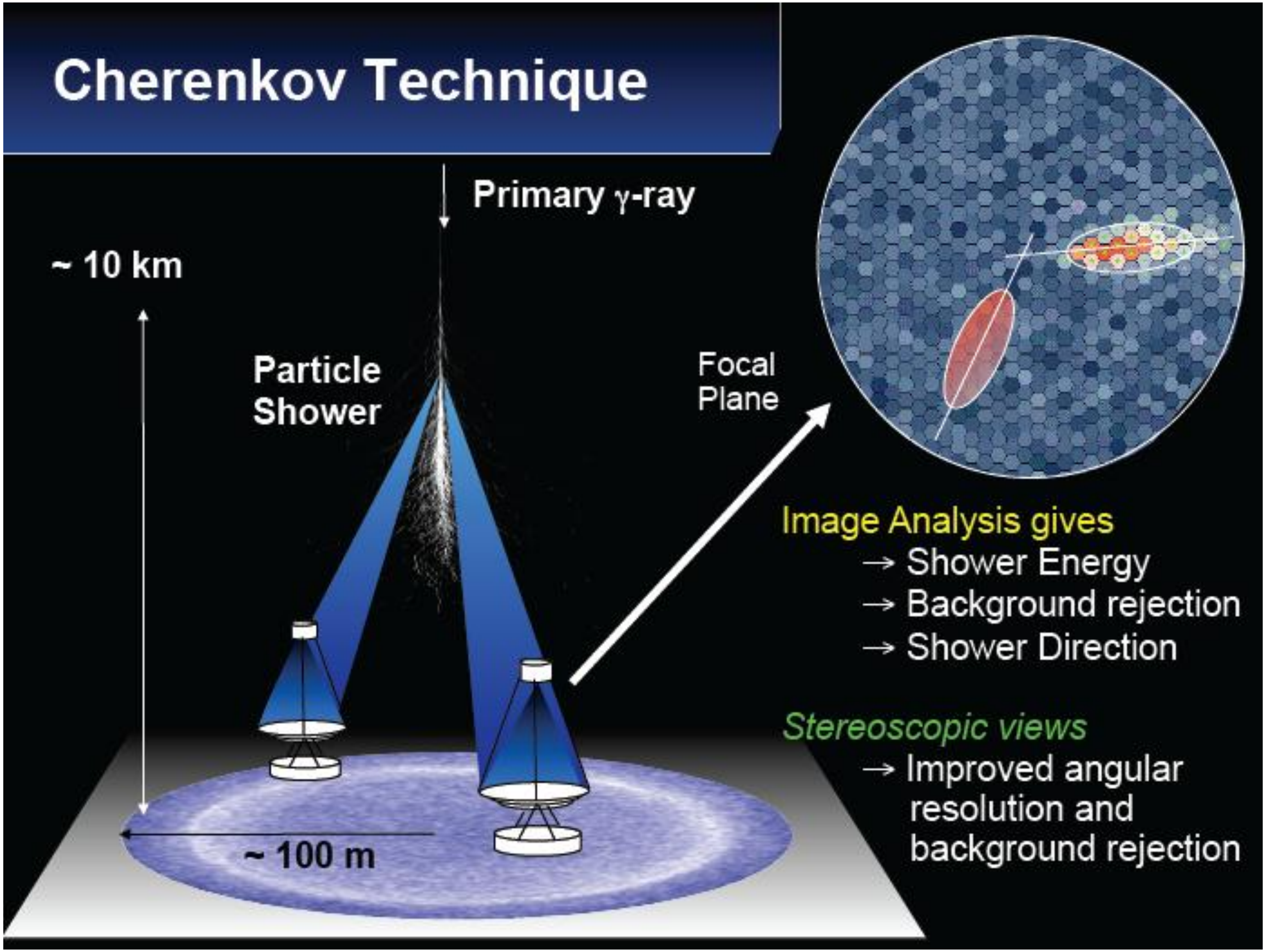
- Detection with multiple optical telescopes

- Total area
  → sensitivity

- Mirror surface
  → sensitivity, energy threshold

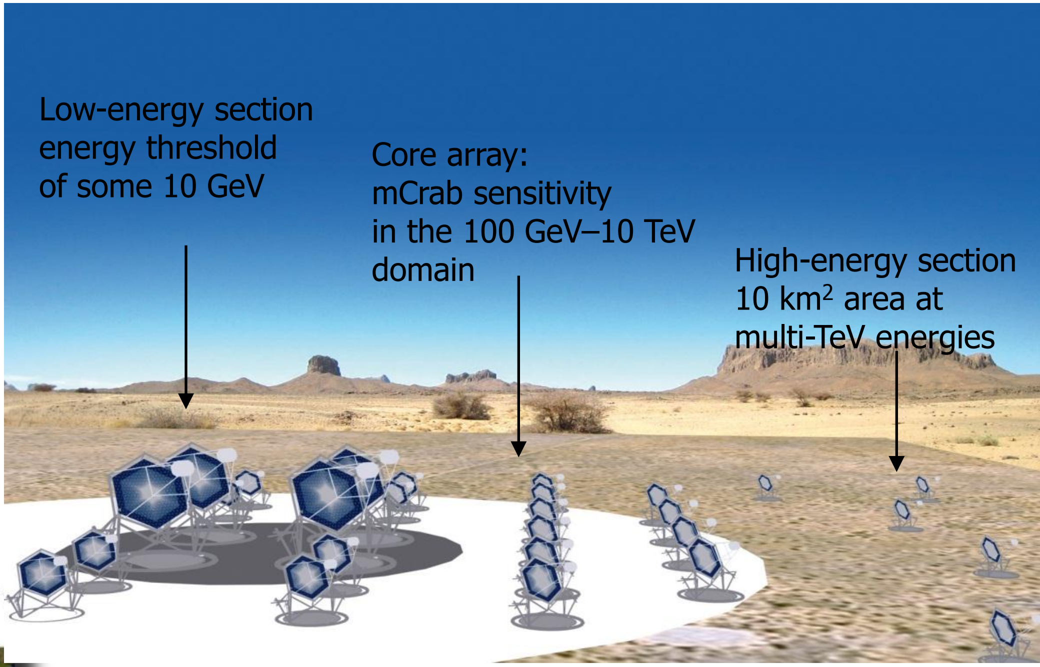
- Number of telescopes
  → angular resolution and background rejection

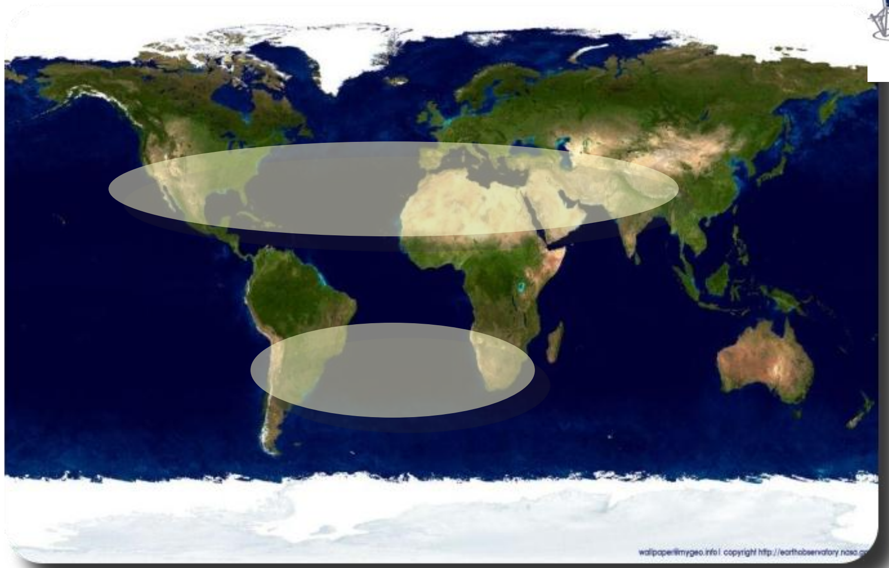
IACT detection technique



CTA- Cherenkov Telescope Array

- Cherenkov Telescope Array
  - Aim: 10x sensitivity and precision of current instruments
  - 2 sites: south and north
  - Operate as an observatory



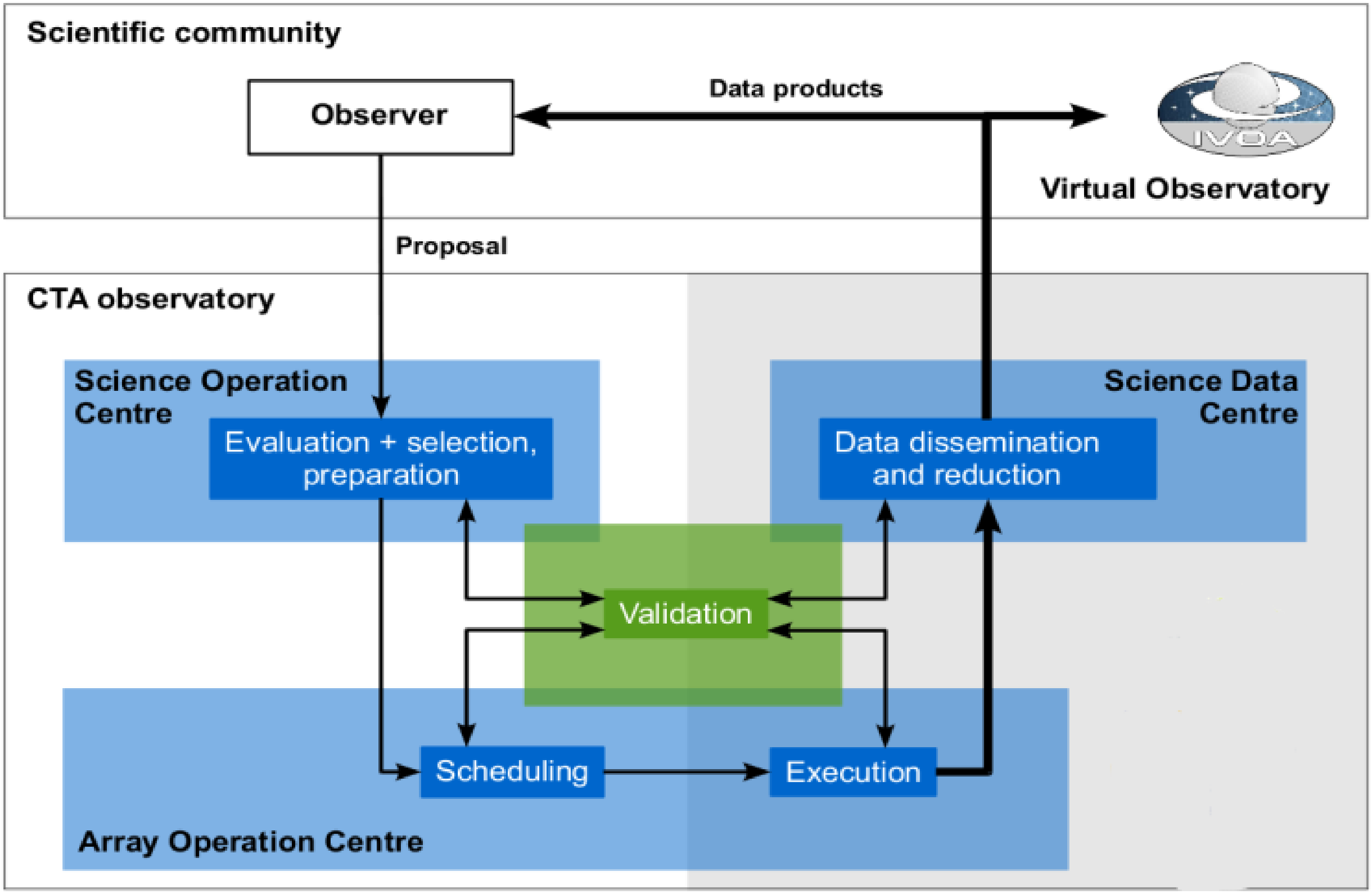


An array of > 60 telescopes of 3 classes for a wide energy range of sensitivity

(~ 20 GeV – hundreds of TeVs)

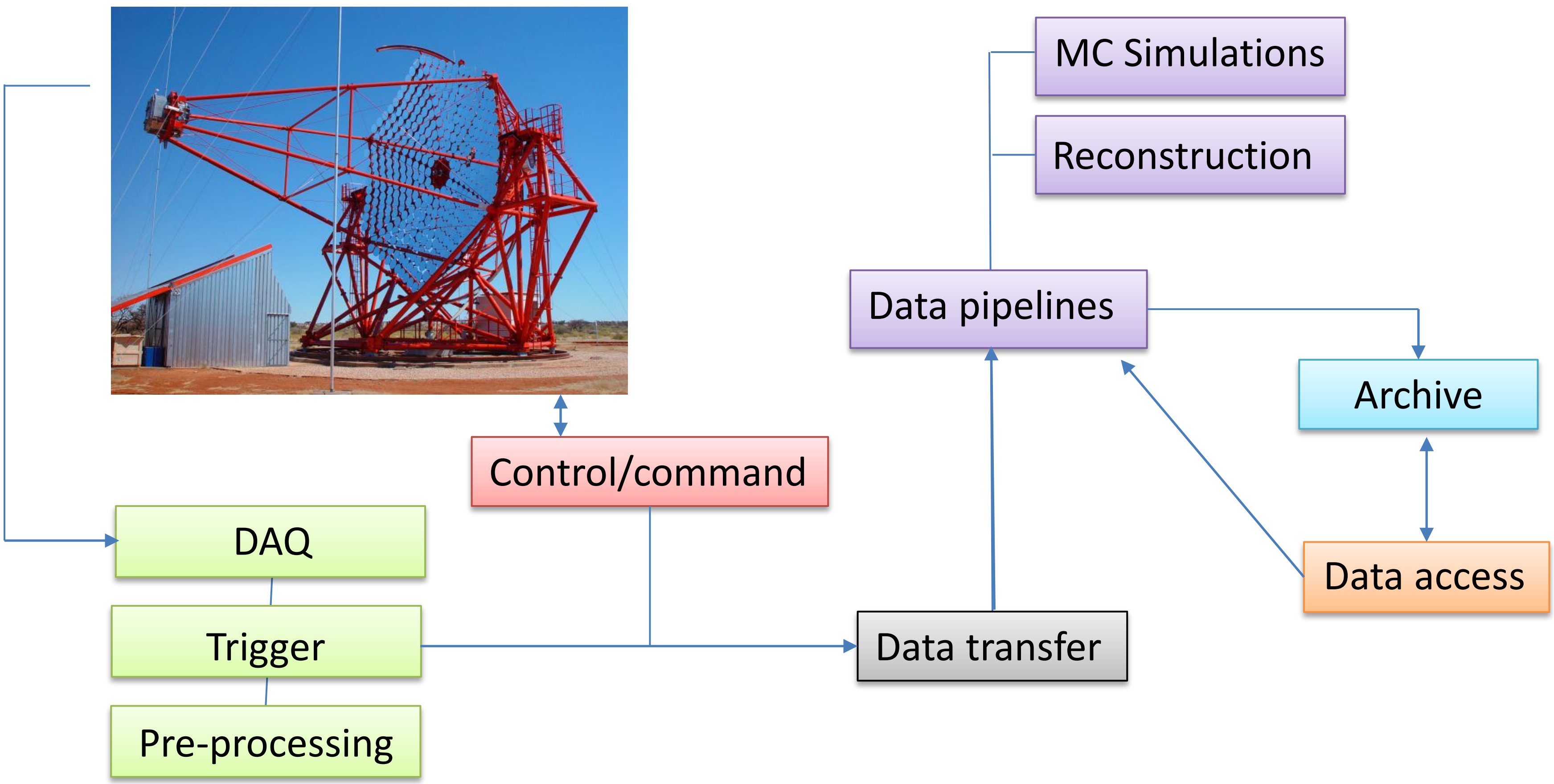
CTA in FP7- Preparatory Phase (2010-2013)

CTA- Cherenkov Telescope Array



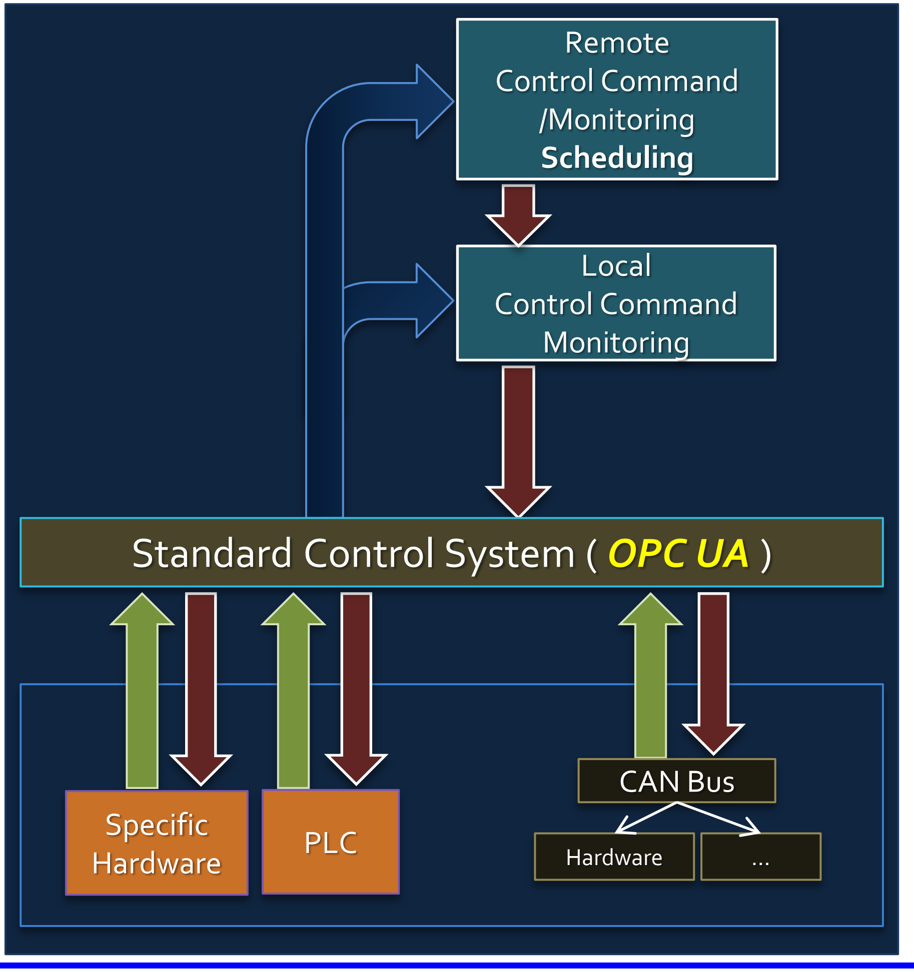
ICT challenges

The “data management” requirements set up a series of possible challenging questions and ICT applications in CTA.



DAQ & Control-Command

A modern and “commercial” approach of telescope control and command is <u>under evaluation</u>: **OPC-UA** (OPen Connectivity-Unified Architecture).



- Slow Control/ Monitoring
  - **On site and remotely (Web-services)**
  - **“Client-server” approach to control a specific device**
  - **Knowing the current status of :**
    - **The array**
    - **A telescope**
    - **A(ny) device**
- Access information in a <u>standard way</u> and independently on the context (standalone mode or through DAQ components).

(see www.opcfoundation.org)

Data acquisition

- Cameras pixels see mostly Night Sky Background (90 to 250 MHz)

- A camera-level trigger (e.g. minimum *pe* threshold + minimum pixels multiplicity) must match the CR triggers ~ 0.5 to 5 kHz while reducing NSB to ~100Hz

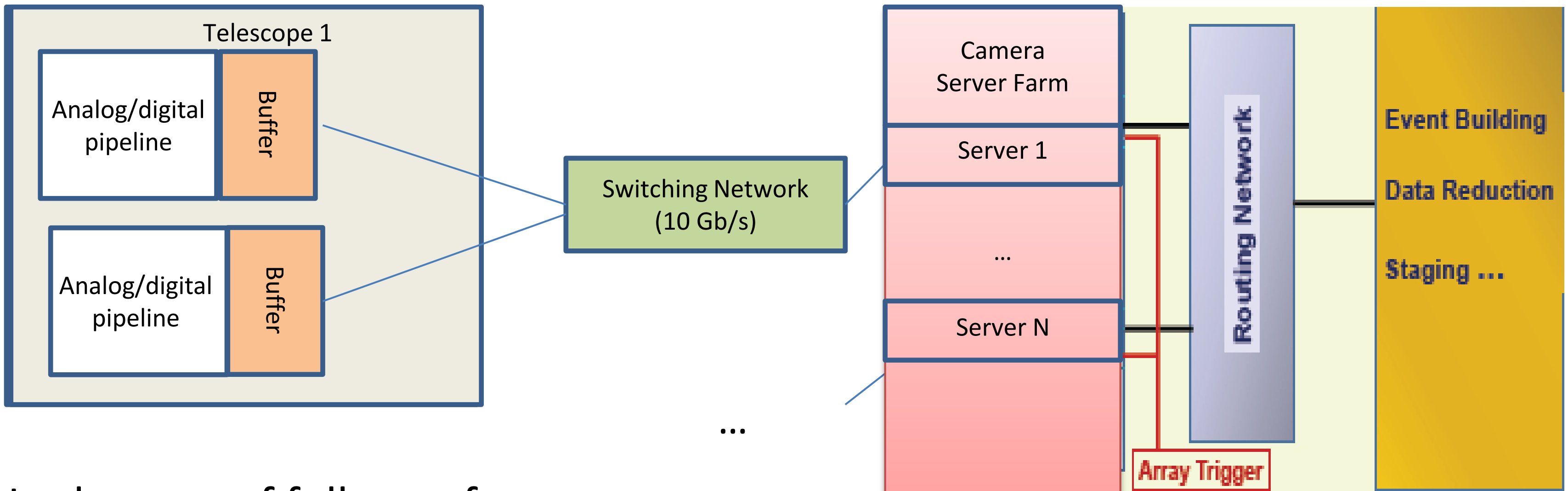
- Array trigger: a stereoscopic coincidence among telescopes in a several 10 ns long time window.

- Complete sampling of waveform event signal would be optimal (scientifically) *vs* integration (default CTA mode): data rate varies from 0.4 to >5 GB/s (@ trigger rates of ~ 15 kHz and ~2000 camera pixels) .

Data acquisition

Severe requirements :

- Multiple Gigabit Ethernet lines that are available today can be used to transport the data from the individual cameras to a central place for event building. 0.5 GB/s should be managed using commercially available switching networks.



- In the case of full waveform 5 GB/s are too large for event-building over a standard switching network. Here a potential computing challenge ?

Data transfer (from camera)

The definition of the raw data format is also important to optimize data transfer management (e.g. by reducing transfer time from camera to on-site computing):

- acting on waveform (non-standard approach) number of slices ;

- compressing (e.g. gzip) on camera embedded CPUs

- comparing CPU and disk management between data transfer and compression + transfer + decompression
  (preliminary estimation of new technology against compression / decompression
  e.g.: Xeon E5650 @ 2.4Ghz, 3ware-9750 6gbps SAS, LVM2 stripping)

- an idea is to investigate timing performance of GPU technology for possible low latency applications (Real Time Analysis), performing some (more advanced ?) pixel clustering recognition and then dealing with smaller data size.

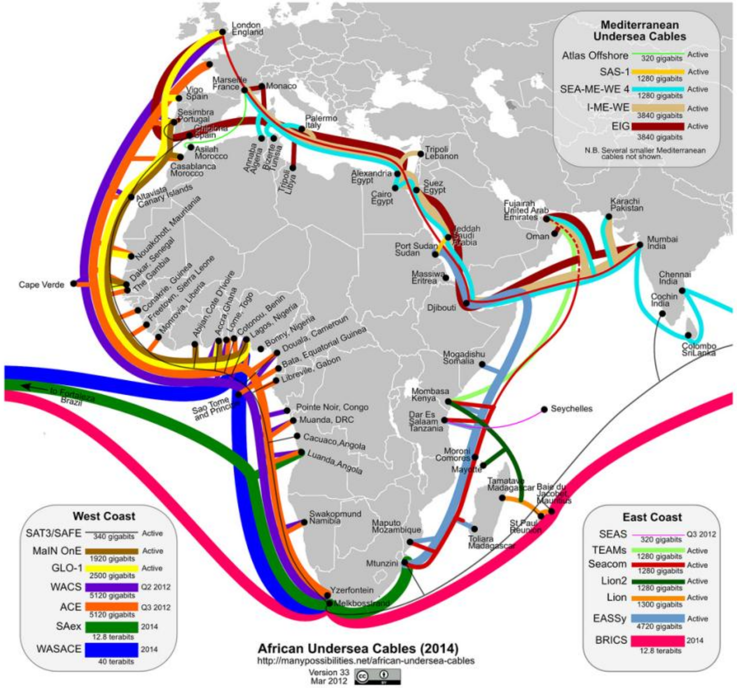
(PERSONAL view)

Data transmission (pipeline)

(PERSONAL view)

Web-service server/client paradigm and the W3C technology deserve high-efficient data transmission:

- A data stream of 5 GB/s (dominated by science raw data) can be rapidly processed with a dedicated high-bandwidth and reliable network.
- A "challenging" approach is the remote monitoring, control and observation scheduling… (just a little fraction of data stream)



- Remind: potential CTA hosting sites are Namibia canaries Islands, South America, …
  The new and powerful intercontinental oceanic networks open new perspectives for the computing and data management models of CTA.
  - ACE/WACS (5.1 Tb/s) Namibia-Canaries-Europe is under tests.
  - WASACE (40 Tb/s in 2014) SA-Brazil-US-Europe

Data (analysis) pipelines

Computing for raw data calibration and reco. (*in some cases extrapolated from H.E.S.S.*)
Some numbers for a possible configuration:
CTA-South : 69 Telescopes, 13 kHz trigger rate, 300 kB per triggered event
CTA-North: 21 Telescopes, 7 kHz trigger rate, 200 kB per triggered event

MC simulations Full waveform sampling (partial w.f.)	CTA SOUTH	CTA NORTH	CTA
Main data stream	4.0 GB/s	1.3 GB/s	5.3 GB/s (0.5 GB/s)
Data volume per month (15% duty cycle)	1.5 PB	0.5 PB	2 PB (0.18 PB)
Data volume per year	18 PB	6 PB	24 PB (2.2 PB)
Data volume in 10 years	180 PB	60 PB	240 PB (22 PB)
For 1 hour observation: CPU time for calibration and reconstruction (HS06) (extrapolated from HESS)	12 Ms + 18 Ms (reduced to a few with hillas Reco.)	5 Ms + 7 Ms (reduced to a few with hillas Reco.)	43 Ms = ~ 1400 hours (of a single CORE) (50% uncertainty)
CPU time for calibration and reconstruction (HS06) per year	40 Gs	17 Gs	56 Gs = ~ 200 years (of a single CORE)

(PRELIMINARY)

Data (analysis) pipelines

(PERSONAL view)

1) Calibration:
   - the full set of procedures are applied per each camera at pixel level
2) Reconstruction:
   - The most sophisticated statistical approaches are based on Likelihood, BDT, multi-PDF, and detailed comparison data/MC.
   - The event is built based on stereoscopy and telescopes combinatory (depending on observations).

Some ways to be explored reducing computing time:

a) Calibration optimization through parallelization. Parallelization of reco. processing at camera level before stereo convergence ?
(*requiring new data structure conception, avoiding tails effect in multithread…)*

b) Software to be conceived/designed to be able to scale to "many-core" computing.
(Lack of standards, code generation not well understood)

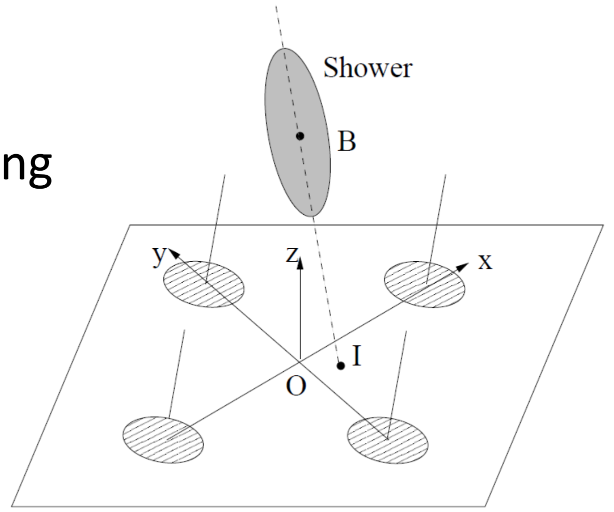
c) GPUs parallel computing faster and at a tenth of the cost of conventional systems.
(translating algorithms into Graphics processing ? need of professional coders !)

Data (analysis) pipelines

(PERSONAL view)

Further "personal (non-GPU expert) ideas" about possible GPU applications based on current IACT science data analysis:

a) From likelihood (H.E.S.S.) 3D semi-analytical modeling approach
(Secondary particles in the shower, light yield, detector response are modeled)



- To a faster graphical modeling (reducing time for fit convergence, e.g. MINUIT)

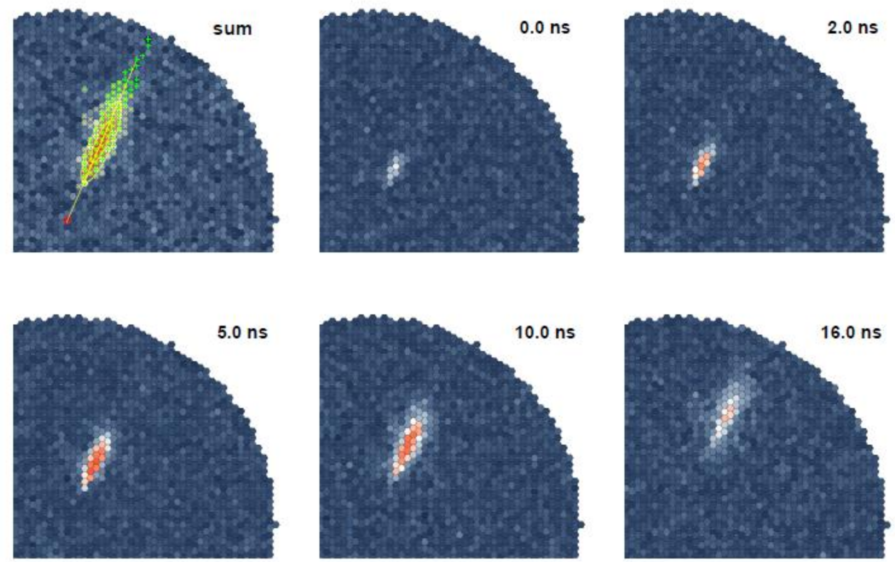
or

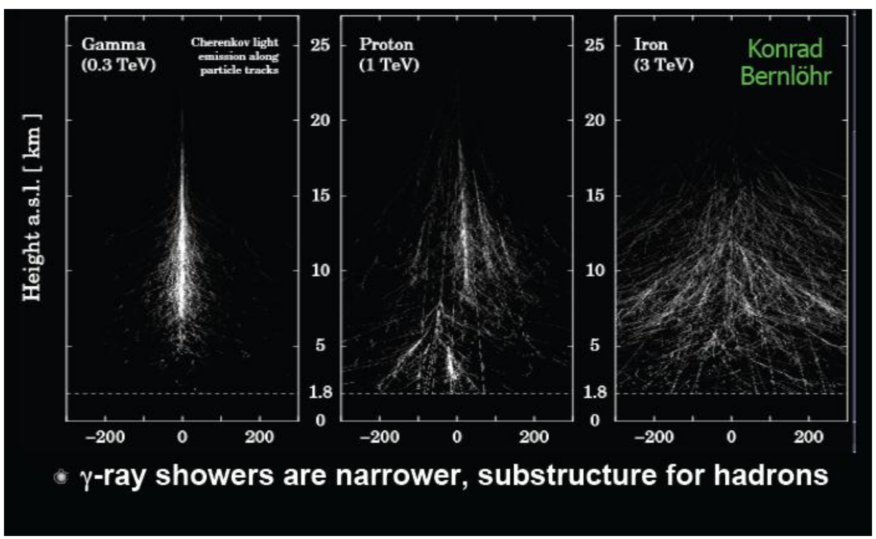
- To **GPU Multi Camera Stereo Matching Algorithm**
3D video coding technology and GPU based image processing.
(*Does a GPU-based 3D stereo reconstruction exist ?)*
It would bring the rapidity of processing and a higher sensitivity (higher background rejection)

Data (analysis) pipelines

(PERSONAL view)

b) The graphical matrix (with GPUs) could be filled in with:
"cubic values" to register and treat "shower flash":
pixel id, light intensity (t), time (t).
(the shower evolutionary path brings information on the nature, energy and the direction of the primary: gamma against protons)

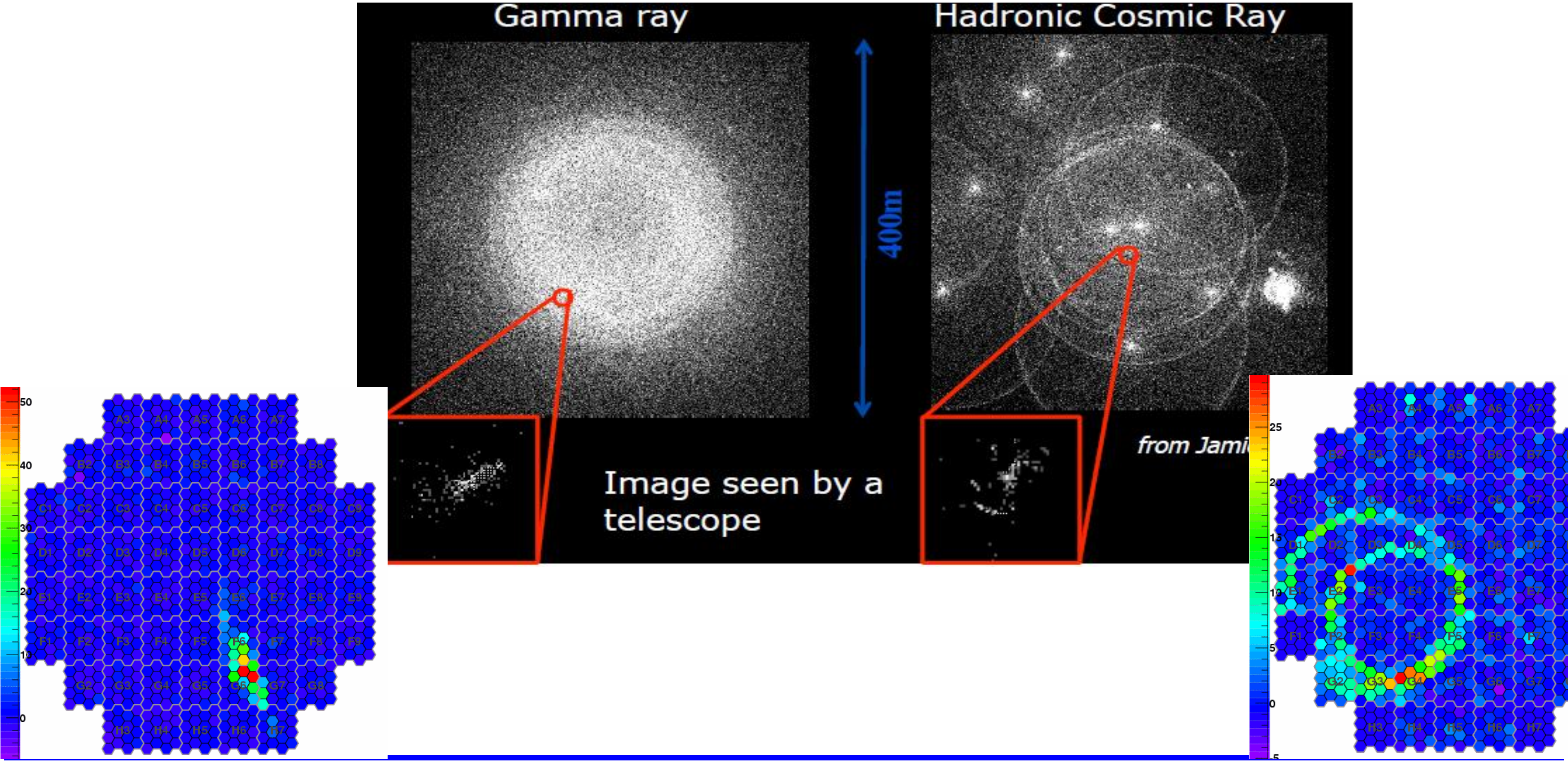




Data (analysis) pipelines

(PERSONAL view)

c) The stereo trigger with graphical matrix (with GPUs) over a large sub-array of telescopes could be combined investigating the (total) Cherenkov light pool.
In particular rings from muons can help to reduce the proton background and the trigger rate to gamma data rate.



Monte Carlo Simulations

Two software packages:

1) CORSIKA for the air showers;
2) SIM_TELARRAY for the array, the telescope and the detector response.

Massive Monte Carlo simulations are an essential tool:

- to optimize the best configuration and performance of CTA and for data analysis
- to discriminate between gamma and hadron showers
- to calibrate the scientific data

Huge numbers of background showers have to be simulated.

Processing, access and analysis of Monte Carlo data demand massive computing, storage resources as well as an efficient worldwide access of the scientific community.

Monte Carlo Simulations and GRID/DCI

The EGI DCI/GRID approach has been applied:

- ❖14 sites in 6 countries: 2000 CPUs used by CTA simultaneously
- ❖ > 500 TB of storage distributed over sites

- The Distributed Computing Infrastructure is optimal for the CTA computing needs.
- The simulation software can evolve to match parallelization in phase with the evolution of "worker nodes" of future DCI-"CLOUD".
- Multi-threading of SIM_TELARRAY has not been efficient with actual EGI-Grid: "tails-effect" and limited allocated RAM (4 GB) per job.

(PERSONAL view)

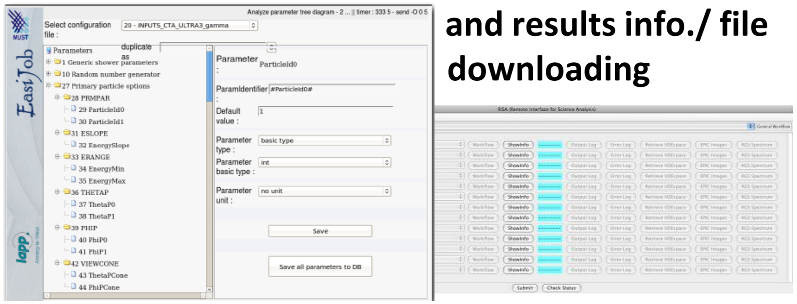
- Adapting the MC software to modern computing (many-core) cannot be a CTA priority.
- Better to consider (high-statistics) simulated data (produced in the GRID) as a tool to validate the potential "computing challenges" for the data pipelines (previously mentioned).

Data access and Science Analysis System

SAS under investigation based on a client/server DCI-Grid computing model as conceived for MC simulations in CTA Computing Grid

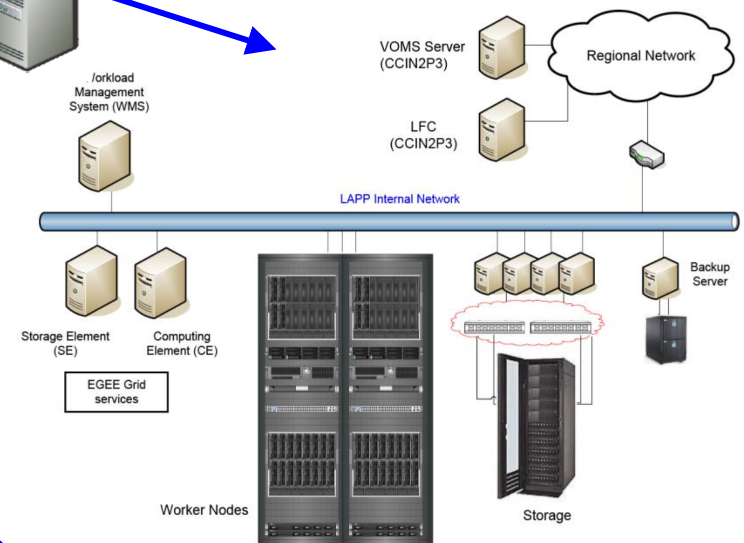
User Web Client and/or VM for data searching

Client application for analysis job submission and results info./ file downloading



Web Server and Grid jobs broker

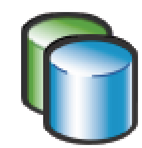
Grid jobs management



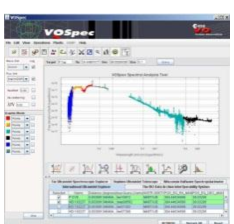
High-level data request

FITS and VObs data access

Interface layer with archive DB



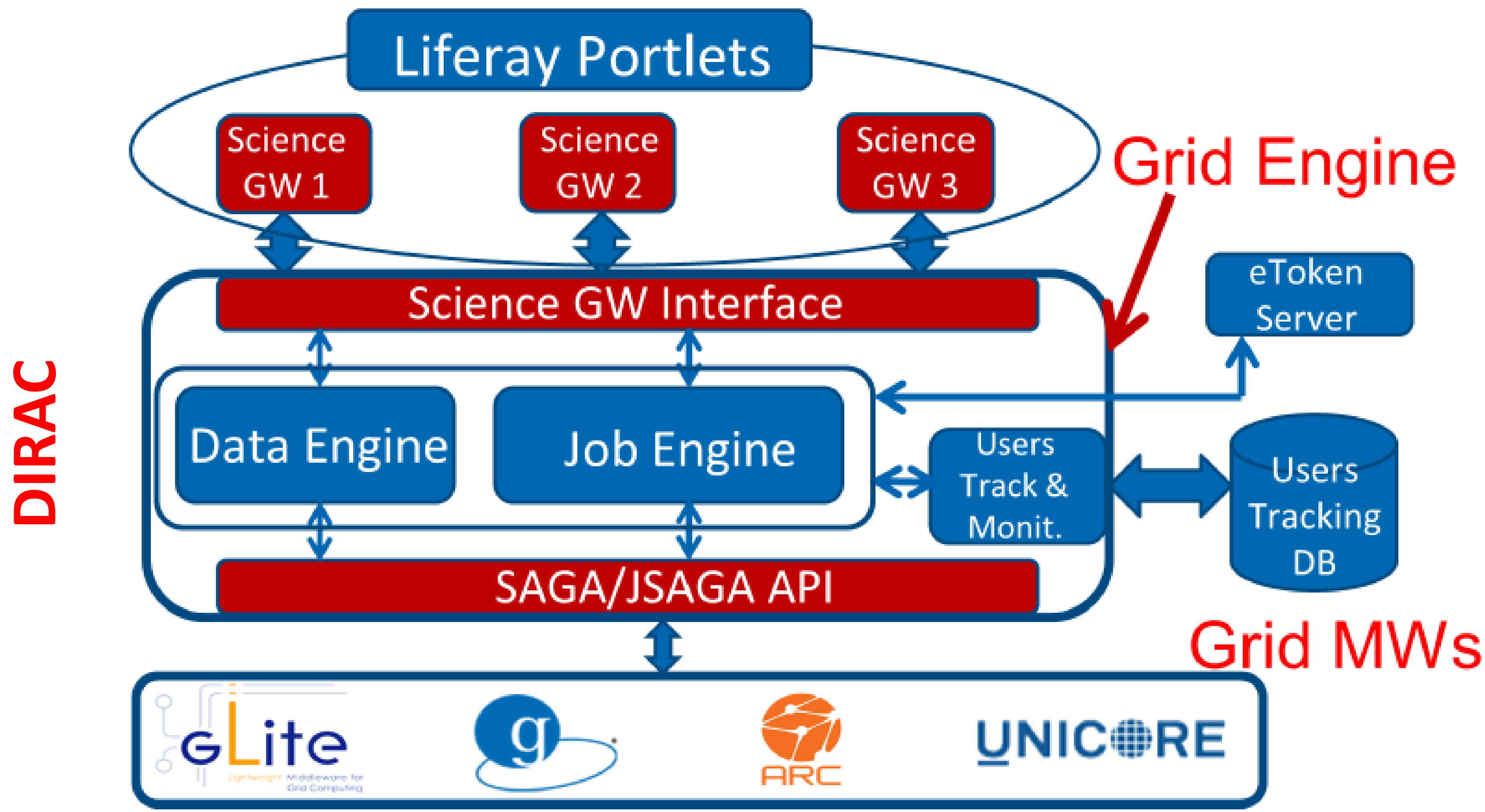
Science Archive, DB and MetaDB systems, notification and results uploading



FITS and VO compliant results VObs tools (VOSpec, Aladin)

(Preliminary)

A layout which we are going to test



Conclusions

ABOUT CTA:

CTA data stream (0.4 up to 8 GB/s), archive (some PB per year) and processing (20-to-50 Gs (HS06) per one year of observations) deserve software, data model and computing optimization and represent important ICT challenges.

IN GENERAL: *(PERSONAL view)*

For the (ASPERA) scientist:

Challenging development around GPUs and towards "many-core" computing are interesting (and promising) for "scientific" perspectives.

For the organizers of this workshop:

These challenges deserve a transverse cooperation which could start and/or be reinforced now and here.

For the "International Collaborations" as CTA:

In some cases (as for the computing) adapting objectives to current technologies are more safe and have less impact on the schedule.

…. However new solutions could also reduce costs. Computing model should be considered well in time in a project

Thank you !