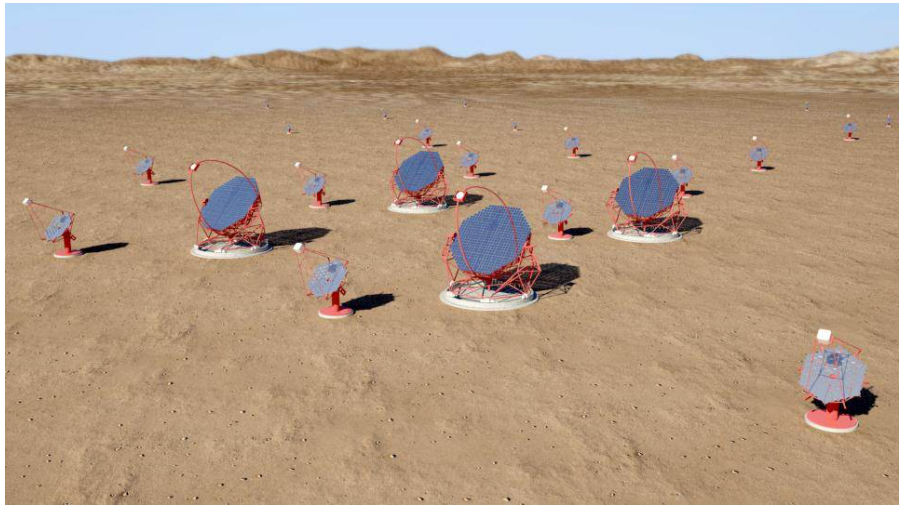
OBELICS Task 3.3
-
CTA resource requirements

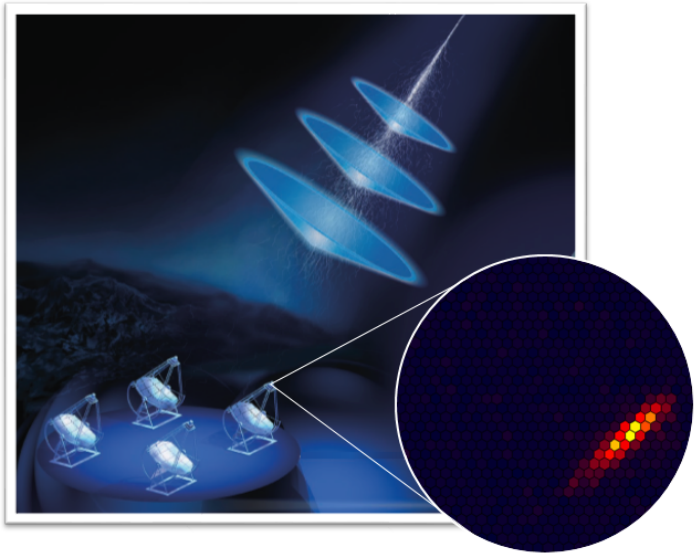
Thomas Vuillaume

CTA - Introduction

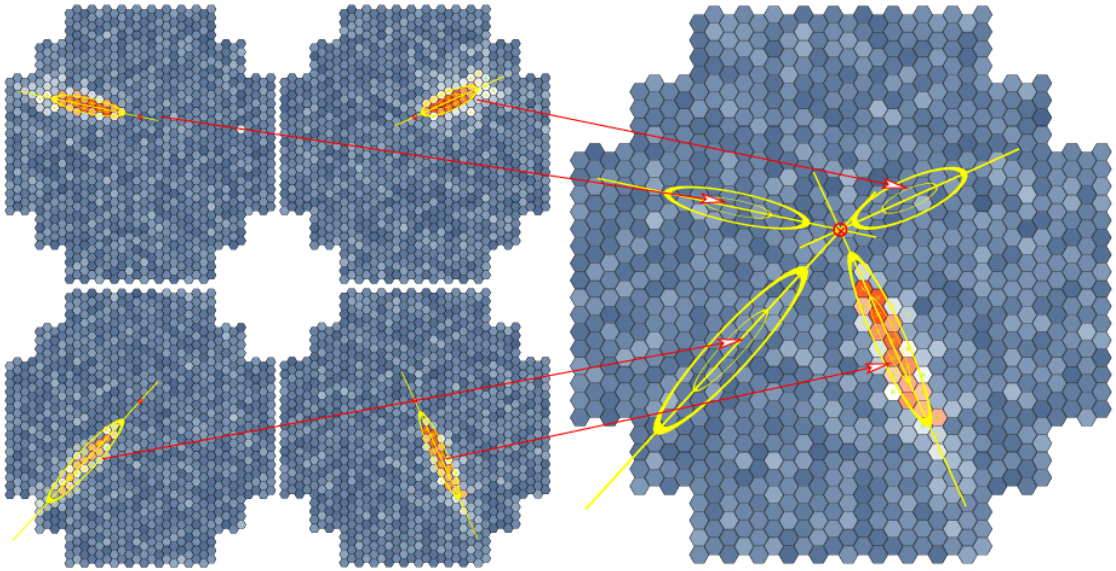
- Two sites : North (La Palma, Canary Island), South (Paranal, Chile)
- Array of telescopes (~ 100 on South)
  - 3 sizes
  - 7 types of cameras
- High-energy astronomy (GeV to TeV)
  - Supernova Remnants, Pulsars, X-ray binaries, AGNs, GRB, Galactic center…
- Construction starts in 2017 – production phase in 2019



CTA – Cherenkov showers



- Shower observed by the array = ellipse image in each camera
- Geometric and energetic reconstruction by sophisticated algorithms

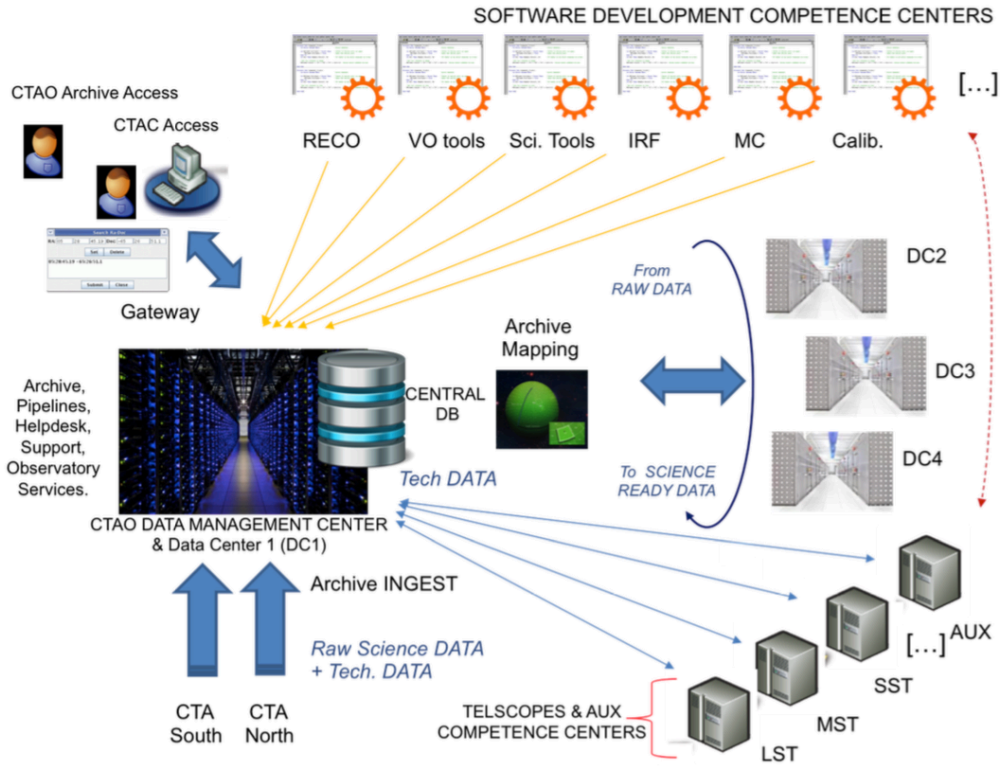


CTA – Data Model

- Events observed by array (= several telescope trigger at the same time)
- Pipeline steps <=> Data levels :
  - DL0 : raw data
  - DL1 : calibrated camera data
  - DL2 : reconstructed shower parameters (energy, direction, particle ID)
  - DL3 : Sets of selected events with instrumental response needed for science analysis
  - DL4 : High level binned data products (spectra, sky maps, lightcurves)
  - DL5 : Legacy observatory data (catalog, sky survey)

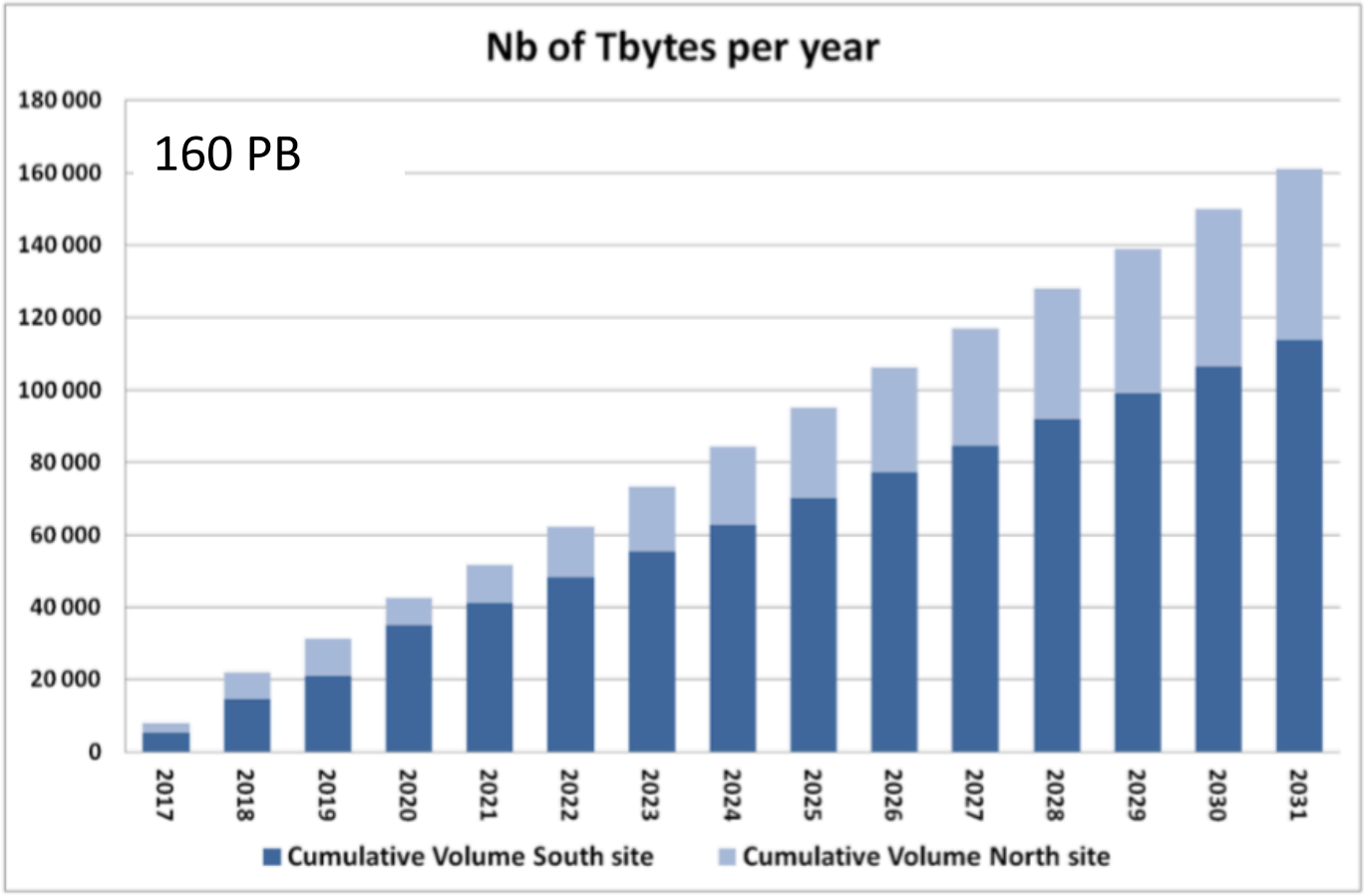
- Other data products:
  - Calibration
  - Technical data from telescopes (sensors temperature, tracking…)
  - Auxilliary data (e.g. Weather)
  - Monte-Carlo events data
  - Instrument response functions

CTA – Data Model



- 4 PB/yr DL0 => 25 PB/yr in total
- Transferred and/or generated by four off-site Data Centers
- Process data where they are stored
- Databases:
  - Proposal handling
  - Archives Management
  - Technical, engineering and monitoring DB

CTA – Storage Needs



CTA – Computation Needs

- one day of raw data acquisition must be processed in less than one day and annual MC simulations must be processed in less than one month.

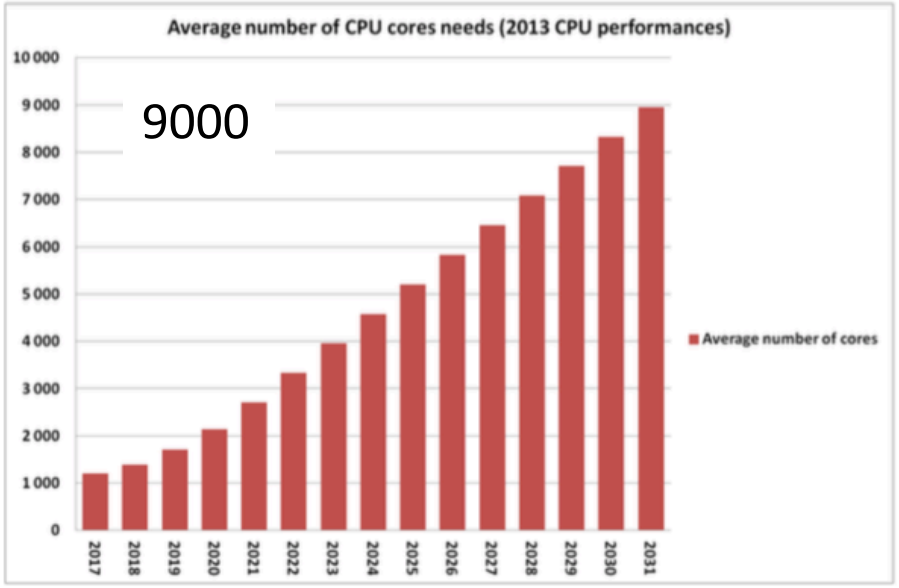


Figure 2.26 – Evolution of the average number of CPU core needs.

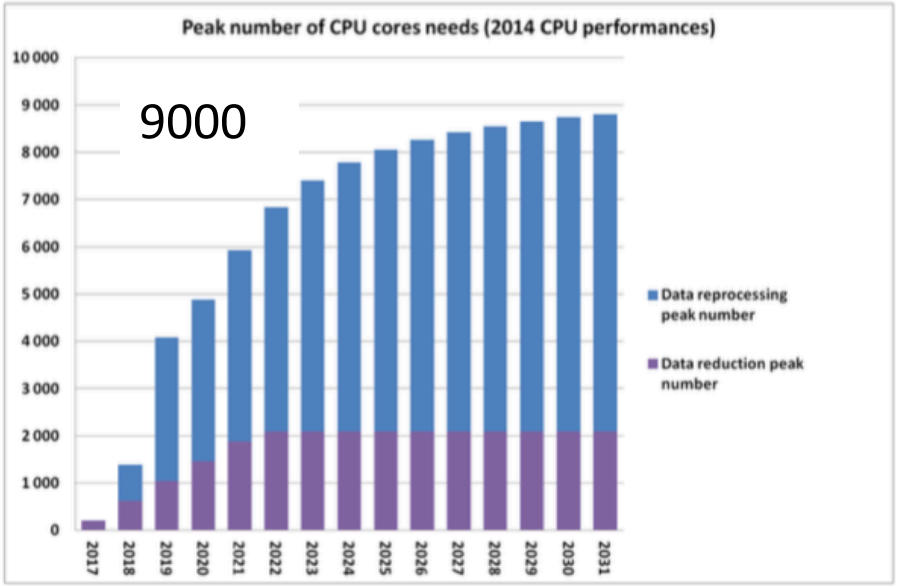


Figure 2.27 – Evolution of the peak number of CPU core needs.

*(2013 CPU performance)

CTA – Data Transfer Needs

- Data Management Requirements = max of 10 days to transfer daily raw data

- Network bandwidth = 1 Gb/s
- Data Reduction Factor = 10

} Requirements exceeded after 7 days of continuous observations

CTA – Database

- Distributed Data Centers
  - ⇒ Requires Complex and efficient Archive Management System
    - ⇒ To map and catalog data
    - ⇒ To optimize the data reduction and simulation pipelines
    - ⇒ To keep track of and synchronize changes applied to data during analysis chain – Data Provenance !

- Insights from ASTERICS other experiments are more that welcome here !