Data Management

Giovanni Lamanna

LAPP - Laboratoire d'Annecy-le-Vieux de Physique des Particules,
Université de Savoie, CNRS/IN2P3, Annecy-le-Vieux, France

CTA Workshop,
Paris, June 30, 2014
CTA Workshop - Paris, 30 June, 2014

Design overview

INSTRUMENT CONTROL AND DATA ACQUISITION

Scheduling/External Trigger
- Central Control
- Array Trigger/Clock Distribution
- Realtime Analysis Pipeline

DAQ
- Camera
- Data Buffer
- Event Builder

Telescopes

DATA MANAGEMENT OFF-SITES

Archive
- Data access

Transfer Unit

Processing

Data

Commands
Trigger
Data
Clock
Data management project

Covering full breakdown components of the data processing (hardware, software and middleware):

- All levels “data” model and format
- Pipelines: reconstruction, simulation and analysis
- Archives: engineering and science
- Observer access: tools and support
- ICT-infrastructure: backbone elements and e-infrastructures
Requirements

Some key requirements:

**Availability:**
>98% of the Real Time Analysis during observations  
>95% of full CTA data-processing capability  
>98% of the CTA Data Archive  
Data products to Guest Observers <2 months the data taking

**Larger phase-space:**
- Instrument response information, providing CTA response to the required precision for the full phase-space of allowed CTA observing modes, directions and conditions

**Higher performance:**
- Reprocess 1 year of CTA data within 1 month.  
- The communications network to the sites must provide sufficient capacity to transfer event-level reconstruction parameters (but not the raw data) within 5 hours, for one night data.  
- Transfer of all raw data from the site to the off-site Data Centre in < 10 days.
Design concept

Distant antennas and a worldwide community: reliable high-bandwidth intercontinental connection

Data rate: ~6 GB/s, 250 TB/day, 30 PB/year: challenges for archive, pipelines processing, and long term sustainability/preservation

Open access to Observatory data and cooperative worldwide community: a global “Gateway” to upload users into a “Scientific Analysis System” integrating together Data Centre, Archive, Software and e-infrastructures

User Web Client

Client analysis job submission and/or Archive access

Publishing results into the Science Archive

FITS and VObs data

High-level products
### DATA MODEL

<table>
<thead>
<tr>
<th>Data Level</th>
<th>Short Name</th>
<th>Description</th>
<th>Data reduction factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0</td>
<td>RAW</td>
<td>Data from DAQ written to disk.</td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>CALIBRATED</td>
<td>Physical quantities measured in the camera: photons, arrival times etc. Preliminary image shape parameters could be also included within.</td>
<td>1-0.2</td>
</tr>
<tr>
<td>Level 2</td>
<td>RECONSTRUCTED</td>
<td>Reconstructed shower parameters such as energy, direction, and particle ID. Several sub-levels increasingly sophisticated are envisaged.</td>
<td>10^{-1}</td>
</tr>
<tr>
<td>Level 3</td>
<td>REDUCED</td>
<td>Sets of selected (e.g. gamma-candidate) events.</td>
<td>10^{-2}</td>
</tr>
<tr>
<td>Level 4</td>
<td>SCIENCE</td>
<td>High Level binned data products like spectra, skymaps, or lightcurves.</td>
<td>10^{-3}</td>
</tr>
<tr>
<td>Level 5</td>
<td>OBSERVATORY</td>
<td>Legacy observatory data, such as CTA survey sky maps or the CTA source catalog.</td>
<td>10^{-5} - 10^{-3}</td>
</tr>
</tbody>
</table>

**Table 1:** Data levels foreseen in CTA.

Extended use of FITS format is considered
CTA Raw Data (L0) is the bulk data coming from the cameras

- File format in block structure: sequential records of incoming data

  - Files -> FITS compliant
  - Binary blocks

PROTOTYPES and TESTS under evaluation

<table>
<thead>
<tr>
<th>Packet Header</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>TypDHF</td>
</tr>
<tr>
<td>SF</td>
<td>Source Sequence Counter</td>
</tr>
<tr>
<td>Packet Length - 1 (32 bits)</td>
<td></td>
</tr>
<tr>
<td>CRC flag</td>
<td>Packet Type</td>
</tr>
</tbody>
</table>

- FITS header provides textual meta-data
- Fields can be updated if required
- FITS Datasum provide full file verification

- Individual block handling and integrity checking
- Payload compressed with optimal algorithm

“packet” (from AGILE) data format for:
- Different types of cameras.
- Zero suppressed or not.
- Also data generated in the different stages of the pipeline
No standard yet in VO to archive high energy astronomical data

Ex.: IVOA “Spectral Data Model” requires physical units and not instrument counts: it does not describe completely X-ray and gamma-ray spectra

→ calibration and a transfer model are needed

Proposal submitted to Euro-VO and discussed at the IVOA interoperability meeting: “High Energy Spectral Data Model Extension” – IVOA Note

Working with IVOA...

→ Discuss/Validate suggested Data Model with IVOA experts

→ Define characterization parameters (axis and properties) for each product: Event Lists + IRFs, Images, Spectra and LightCurves

→ Define an IVOA Data Model for CTA source Catalogue
It is composed of 5 main sub-components: Calibration, MC, Reconstruction, Analysis, Real Time Analysis. Based on current IACT experience (HESS, MAGIC, VERITAS)
MC SIMULATION PIPELINE

MC critical for detector response and calibration
MC MAIN USAGE SCENARIO

MC pipeline

Calibration, observing conditions

AUX0/ENG0/CAL0 archive

DL0 archive (observations)

IRF archive

reconstruction pipeline

reconstruction pipeline

how many simulation sets are needed to meet performance requirements depend on

> detector (e.g. mirror quality),
> site (e.g. atmospheric variability)
> number of observing modes (e.g. pointed vs divergent)
> ...

IRFs are functions of

> elevation
> azimuth
> NSB
> telescope combinations
> time (yearly, monthly, weekly, ..?)
> ....
CTACG: Distributed Grid e-infrastructure for CTA MC simulation production:
- About 20 sites from 7 countries
- More than 1 PB produced and accessed
- About 3.4 M executed jobs
Examples of new developments and prototypes

Testing “Big Data” new technologies, e.g. software frameworks and file systems. Exploring the parallelization of pipelines

Data Parallel Systems such as Hadoop.
- clusters are built with commodity hardware
- each node takes the roles of both computation and storage
- good data locality reduces cross-switch network traffic - one of the bottlenecks in data-intensive computing.

First real-time analysis prototype
- Concept: modern framework

Calibration/reco. pipelines (accelerating hardware)

[Diagram of RTA test bench]

Kayla Devkit To develop CUDA applications for GPU-accelerated systems with an ARM processor.

NVIDIA Quadro K2000 graphics card
ARCHIVES: the reference model

Defining:
- Standards.
- Functionalities.
- Information flow.
- Implementation roles.
Current proposed baseline:

**Main Centralized Storage System**: with one (or more) archiving resources, a database of meta-data and a back-end service for user access plus minimal (one to a few more resources) *distribution for replicas*:

- they will operate **hybrid distributed AS** by a common (central) resource manager (thus each resource is not completely independent from each other).

- a back-up archive (independent)

- expandable in **scheduled distributed AS** where different entities work like a relay team.
Any single entity takes care of partial data storage, handling and processing and provides higher level data products to a centralized/common database system.
User Archive ACCESS

SOFTWARE DEV. CENTERS

RECO
VO tools
Sci. Tools
IRF
MC
Calib.

CTA South
CTA North

ACTL CENTER

CENTRAL DB

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

Gateway

Archive INGEST

Raw DATA

Tech DATA

Archive, Pipelines, Helpdesk, Support, Observatory Services.

CTA-DATA CENTER (Tier 0)

Sci. Tools IRF MC Calib.

Gateway

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

Gateway

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

Gateway

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

Gateway

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)

From RAW DATA

To SCIENCE READY DATA

DCI (a few Tier 1s)
Archive context diagram
Network infrastructure

The network is a major issue for intercontinental but also pan-European data transmission.

Consulting from DANTE/GEANT International research network O(Gbps) vs CTA site candidates:
- Feasibility
- Operation costs
- Local installation costs (independent evaluation)
- Reliability

GÉANT is the pan-European research and education network that interconnects Europe’s National Research and Education Networks (NRENs).

DANTE (Delivery of Advanced Network Technology to Europe) plans, builds and operates advanced networks for research and education. It is owned by European NRENs (national research and education networks), and works in partnership with them and in cooperation with the European Commission. DANTE provides the data communications infrastructure essential to the development of the global research community.
CTA south array:
Type E, 69 telescopes: 4 LST, 23 MST and 32 SST
Cosmic-ray trigger rate: ~ 32kHz
Mean trigger tel. mult.: 1.7(LST), 2.3(MST), 1.9(SST)
Array data rate: 4.15 GBytes/s

CTA north array:
Type NA, 19 telescopes: 4 LST and 15 MST
Cosmic-ray trigger rate: ~ 18kHz
Mean trigger tel. mult.: 1.9(LST), 2.0(MST)
Array data rate: 1.72 GBytes/s

Data reduction options must be studied
• Suppression of pixels with low signals.
• Compression.
• Software trigger to reject hadrons.

27.7 PB/Year

With a goal of at least a factor 10 reduction to be studied during commissioning:
Raw data rate: 0.415 GBytes/s (S) & 0.172 GB/s (N)
Raw data volume: 2.77 PB/year
Network bandwidth (100% duty time): 562 Mb/s & 233 Mb/s
Cumulated data volumes - Factor 10 compression

Nb of Tbytes per year - Close to minimum

Commissioning
Based on HESS CPU annual resources needs and CTA simulation current needs
Science Tools

The main functionalities of the Science Tools will cover the generation of sky images, the extraction of source spectra, the determination of light-curves, and the adjustment of parametric source models. The Science Tools will be distributed together with the data and instrument response functions to the CTA Guest Observer, and will be provided freely for download to CTA Archive Users.
Conclusions

Data Management

- A major project: to evolve from current IACT experiment systems to a more reliable Observatory-based approach.

- A few interesting and frontier challenges.

- Sustainability and reliability are mandatory.

- Interesting opportunity for cross-fertilization both in e-Science and Astronomy.

- Final goal is the community-based “Scientific Analysis System” and the implementation of the CTA-Observatory Data Center.
BACK-UP SLIDES
ICT-Infrastructures (new developments during last year)

Gateway prototypes

- Liferay technology
- Federated Single Sign-On (SSO) system (Shibboleth)
- Accessible to all CTA members (CTA LDAP integration)
- CTA MC analysis tools/applications integrated
- CTA data browsing (based on EGI Catalogue)
- CTA Grid resources monitoring (based on NAGIOS)
- DIRAC Grid engine and metadata interface included

Cooperation with Geant and Dante providing information on connectivity availability and pricing for potential sites to be used for the construction of the CTA array.
### Raw data volumes assumptions (Step 1)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values</th>
<th>Unit</th>
<th>From</th>
<th>Monte-Carlo production version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger Rates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTA-S protons</td>
<td>22800</td>
<td>Hz</td>
<td>Konrad</td>
<td>Prod2</td>
</tr>
<tr>
<td>CTA-N protons</td>
<td>13000</td>
<td>Hz</td>
<td>Konrad</td>
<td>Prod2</td>
</tr>
<tr>
<td>All nuclei/protons</td>
<td>1,4</td>
<td></td>
<td>Konrad</td>
<td>Prod2</td>
</tr>
</tbody>
</table>

Percentage of useful pixels / all pixels: 3%
## Data rates and annual data volume (Step 1)

### South site:

<table>
<thead>
<tr>
<th>CTA-South MC Prod 2</th>
<th>Rate (Hz)</th>
<th>Hours/Year</th>
<th>All/Protons</th>
<th>LST/evt</th>
<th>MST/evt</th>
<th>SST/evt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22800</td>
<td>1314</td>
<td>1,4</td>
<td>0,71</td>
<td>2,22</td>
<td>0,54</td>
</tr>
</tbody>
</table>

### Array E trig. rate

<table>
<thead>
<tr>
<th>Tel. type</th>
<th>Array E</th>
<th>NTel/Trigger</th>
<th>Trigger Rate/Tel</th>
<th>Current Array</th>
<th>Window (ns)</th>
<th>Samples/ns</th>
<th>Bytes/samples</th>
<th>HDR</th>
<th>Bytes/pixel</th>
<th>Nb Pixels</th>
<th>wavef. pixels</th>
<th>Data rate (GBytes/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>31920 LST</td>
<td>4</td>
<td>0,71</td>
<td>5 666</td>
<td>4</td>
<td>20</td>
<td>1,0</td>
<td>4</td>
<td>5</td>
<td>85</td>
<td>2 000</td>
<td>60</td>
<td>1,08E-01</td>
</tr>
<tr>
<td>31920 MST-NECT</td>
<td>12</td>
<td>1,11</td>
<td>2 953</td>
<td>12</td>
<td>60</td>
<td>1,0</td>
<td>4</td>
<td>5</td>
<td>245</td>
<td>1 800</td>
<td>54</td>
<td>4,37E-01</td>
</tr>
<tr>
<td>31920 MST-FL</td>
<td>12</td>
<td>1,11</td>
<td>2 953</td>
<td>12</td>
<td>60</td>
<td>0,25</td>
<td>2</td>
<td>5</td>
<td>35</td>
<td>1 800</td>
<td>54</td>
<td>6,24E-02</td>
</tr>
<tr>
<td>31920 SST-2M-G</td>
<td>24</td>
<td>0,18</td>
<td>239</td>
<td>24</td>
<td>100</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>205</td>
<td>2 048</td>
<td>61</td>
<td>6,74E-02</td>
</tr>
<tr>
<td>31920 SST-2M-AS</td>
<td>24</td>
<td>0,18</td>
<td>239</td>
<td>24</td>
<td>100</td>
<td>0,01</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>2 400</td>
<td>72</td>
<td>5,78E-03</td>
</tr>
<tr>
<td>31920 SST-1M-DG</td>
<td>24</td>
<td>0,18</td>
<td>239</td>
<td>24</td>
<td>100</td>
<td>0,25</td>
<td>2</td>
<td>5</td>
<td>55</td>
<td>1 300</td>
<td>39</td>
<td>1,15E-02</td>
</tr>
</tbody>
</table>

**Vol/Year (1314h)**: 3,12E+00 PB

### Integrated

<table>
<thead>
<tr>
<th>Tel. type</th>
<th>Array E</th>
<th>NTel/Trigger</th>
<th>Trigger Rate/Tel</th>
<th>Current Array</th>
<th>Window (ns)</th>
<th>Samples/ns</th>
<th>HDR</th>
<th>Bytes/pixel</th>
<th>Nb Pixels</th>
<th>wavef. pixels</th>
<th>Data rate (GBytes/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>31920 LST</td>
<td>4</td>
<td>0,71</td>
<td>5 666</td>
<td>4</td>
<td>15</td>
<td>2 000</td>
<td>1940</td>
<td>6,14E-01</td>
<td>4,14E+01</td>
<td>6,14E-01</td>
<td>6,14E-01</td>
</tr>
<tr>
<td>31920 MST</td>
<td>24</td>
<td>2,22</td>
<td>2 953</td>
<td>24</td>
<td>15</td>
<td>1 800</td>
<td>1746</td>
<td>1,73E+00</td>
<td>1,73E+00</td>
<td>1,73E+00</td>
<td>1,73E+00</td>
</tr>
<tr>
<td>31920 SST</td>
<td>72</td>
<td>0,54</td>
<td>239</td>
<td>72</td>
<td>15</td>
<td>1 800</td>
<td>1746</td>
<td>4,20E-01</td>
<td>4,20E-01</td>
<td>4,20E-01</td>
<td>4,20E-01</td>
</tr>
</tbody>
</table>

**Vol/Year (1314h)**: 1,25E+01 PB

### Total Vol South/year

<table>
<thead>
<tr>
<th></th>
<th>15,58 PB</th>
</tr>
</thead>
</table>

### Plus 20% CAL + TECH

<table>
<thead>
<tr>
<th></th>
<th>18,70 PB</th>
</tr>
</thead>
</table>

### Bandwidth (100% duty time)

|                      | 5,62 Gbits/s |

**Total**: 18.7 PB/year, Data rate: 4.15GBytes/s
# Data rates and annual data volume (Step 1)

## North site:

<table>
<thead>
<tr>
<th>CTA-North MC Prod 2</th>
<th>Rate (Hz)</th>
<th>Hours/Year</th>
<th>All/ protons</th>
<th>LST/evt</th>
<th>MST/evt</th>
<th>SST/evt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13000</td>
<td>1314</td>
<td>1.4</td>
<td>1.28</td>
<td>1.65</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Array E trig. rate</th>
<th>Tel. type</th>
<th>Array E</th>
<th>NTel/Trigger</th>
<th>Trigger Rate/Tel</th>
<th>Current Array</th>
<th>Window (ns)</th>
<th>Samples/ ns</th>
<th>Bytes/ samples</th>
<th>HDR</th>
<th>Bytes/ pixel</th>
<th>Nb Pixels</th>
<th>wavef. pixels</th>
<th>Data rate (GBytes/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples</td>
<td></td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18200 LST</td>
<td></td>
<td>4</td>
<td>1.28</td>
<td>5824</td>
<td>4</td>
<td>20</td>
<td>1.0</td>
<td>4</td>
<td>5</td>
<td>85</td>
<td>2000</td>
<td>60</td>
<td>1.11E-01</td>
</tr>
<tr>
<td>18200 MST</td>
<td></td>
<td>24</td>
<td>1.65</td>
<td>1251</td>
<td>15</td>
<td>60</td>
<td>1.0</td>
<td>4</td>
<td>5</td>
<td>245</td>
<td>1800</td>
<td>54</td>
<td>2.31E-01</td>
</tr>
<tr>
<td>18200 SST</td>
<td></td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>1.0</td>
<td>1</td>
<td>5</td>
<td>105</td>
<td>1300</td>
<td>39</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>Vol/Year (1314h)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated</td>
<td></td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18200 LST</td>
<td></td>
<td>4</td>
<td>1.28</td>
<td>5824</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>2000</td>
<td>1940</td>
<td>6.31E-01</td>
</tr>
<tr>
<td>18200 MST</td>
<td></td>
<td>24</td>
<td>1.65</td>
<td>1251</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>1800</td>
<td>1746</td>
<td>4.58E-01</td>
</tr>
<tr>
<td>18200 SST</td>
<td></td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>1800</td>
<td>1746</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>Vol/Year (1314h)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Total:
- **7.75 PB/year**
- **Data rate: 1.72 GB/s**

**27.7 PB/Year (2.6 PB presented in PTDR) Base 10**
Assumptions:

- Observation time per year: 1314 hours
- Simulation: Equivalent to the yearly cumulated data volume (Min: 1PB, Max: 20PB)
- Data reduction: 25% of fully integrated scenario prior to any compression. (Calibrated data are not archived)
- One full reprocessing of all archived data but only two versions kept in Archive (Last & previous ones).

Reminder: Replying to 1st review Rec.
# Computing Model (Step 3)

**Assumptions:**

<table>
<thead>
<tr>
<th>Data occupying storage – Planned Model (Data)</th>
<th>Event Size</th>
<th>Data Access</th>
<th>Disk replicas of each version</th>
<th>Number of versions</th>
<th>Number of Tape Copy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAW (DL0)</td>
<td>27270 to 39420 bytes depending on camera</td>
<td>Write once, low read rate</td>
<td>10% kept on disk (cache)</td>
<td>1</td>
<td>1 + 1 (Backup)</td>
</tr>
<tr>
<td>CALIBRATED (DL1)</td>
<td>100% of RAW</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RECONSTRUCTED (DL2)</td>
<td>10% of integrated RAW</td>
<td>New version per year, low read rate</td>
<td>100% kept on disk</td>
<td>2</td>
<td>1 + 1 (Backup)</td>
</tr>
<tr>
<td>REDUCED (DL3)</td>
<td>1% of integrated RAW</td>
<td>High read rate</td>
<td>1 (100%)</td>
<td>2</td>
<td>1 + 1 (Backup)</td>
</tr>
<tr>
<td>SCIENCE (DL4)</td>
<td>0.1% of integrated RAW</td>
<td>High read rate</td>
<td>1 (100%)</td>
<td>1</td>
<td>1 + 1 (Backup)</td>
</tr>
<tr>
<td>OBSERVATORY (DL5)</td>
<td>0.1% of integrated RAW</td>
<td>High read rate</td>
<td>1 (100%)</td>
<td>1</td>
<td>1 + 1 (Backup)</td>
</tr>
<tr>
<td>MONTE-CARLO DATA</td>
<td>100% of all Observation data (Min 500 TB, Max 20 PB)</td>
<td>Read/Write</td>
<td>100% during commissioning phase (3 years), 1 PB afterwards</td>
<td>1</td>
<td>1 + 1 (Backup)</td>
</tr>
</tbody>
</table>

Reminder: Replying to 1st review Rec.
Based on HESS CPU annual resources needs and CTA simulation current needs