The DArk Matter Particle Explorer (DAMPE) - and the High Energy Radiation Detection Detection facility (HERD)

Xin Wu
DPNC, University of Geneva, Switzerland

Science Data Center Meeting
29 September 2016, ISDC, Geneva, Switzerland
The DAMPE detector

- Plastic Scintillator Detector (PSD)
- Silicon-Tungsten Tracker (STK)
- BGO Calorimeter (BGO)
- Neutron Detector (NUD)

- Thick imaging calorimeter (BGO of 32 $X_0$)
- Precise tracking with Si strip detectors (STK)
- Tungsten photon converters in tracker (STK)
- Charge measurements with PSD and STK
- Extra hadron rejection with NUD

First time 3 in 1!

High energy electron, $\gamma$-ray, and cosmic ray telescope
Scientific objectives of DAMPE

- High energy particle detection in space
  - Measure the high energy cosmic electron and gamma spectra and search for Dark Matter signatures
  - Study of cosmic ray spectrum and composition
  - High energy gamma ray astronomy

Detection of 1 GeV - 10 TeV e/γ, 100 GeV - 100 TeV cosmic rays with excellent energy resolution, direction reconstruction (γ) and charge measurement
Detection of 1 GeV - 10 TeV γ, 100 GeV - 100 TeV cosmic rays with excellent energy resolution, direction reconstruction (γ) and charge measurement.

DAMPE 3 years, projection
Protons and nuclei: DAMPE 3 years

- Proton
- Helium
- B/C
- p to Fe
The DAMPE detector in real

Weight : 1450/1850 kg (payload/satellite)
Power: 300/500 W (payload/satellite)
Readout channels: 75,916 (STK 73,728)
Size: 1.2m x 1.2 m x 1.0 m
Data rate: 16 GB/day
The DAMPE Collaboration

- China
  - Purple Mountain Observatory, CAS, Nanjing
  - University of Science and Technology of China, Hefei
  - Institute of High Energy Physics, CAS, Beijing
  - Institute of Modern Physics, CAS, Lanzhou
  - National Space Science Center, CAS, Beijing

- Switzerland
  - DPNC, University of Geneva, Switzerland

- Italy
  - INFN Perugia and University of Perugia
  - INFN Bari and University of Bari
  - INFN Lecce and University of Salento

First launch of the 5 approved scientific satellites of the Chinese Academy of Sciences
STK: DPNC responsibility

DPNC is the leading house of the silicon-tungsten tracker project, in collaboration with INFN Perugia and IHEP, Beijing.
Successfully launched on Dec. 17 2015!

- Altitude: 500 km
- Inclination: 97.4065°
- Period: 95 minutes
- Orbit: sun-synchronous
- Pointing precision 0.005°

- Dec. 20: all detectors powered on, except the HV for PMTs
- Dec. 24: HV on!
- Dec. 30: stable trigger condition
- Smooth operation since!
• 15 orbits/day
• ~50 Hz average trigger rate
  – Main high energy trigger and prescaled low energy and MIP triggers
• ~5 M events recorded per day, trigger rate very stable
  – Data down-link: 16 GB/day
  – ~100 GB/day data produced on ground

Mostly stable data-taking since Dec. 30, 2 weeks after launch!
Lots of high energy particles collected

- ~465 M events collected for 3 months!
  - ~9.4% in SAA, mainly <1 GeV
  - Bulk of data collected by high E trigger, efficient for > 10 GeV
  - Also collect large amount of low E (for $\gamma$) and MIP (for calibration) data

~200k TeV (raw) events/year!

Reaching few 100 TeV for high energy CR physics: connect to ground-based extended air-shower experiments
Charge measurement with PSD and STK

- Proton
- PSD high gain readout
- PSD low gain readout
- Helium
- STK charge measurement

PRELIMINARY
• ~24 M events in fitted signal e/γ (2σ) in full energy range
  – ~2.2 M events in e/γ signal region above 10 GeV

• ~15k events in fitted signal (2σ) with E>100 GeV
  – Signal stable with track match cut, s/b improved
Gamma rays observed!

DAMPE 165 days
E > 1 GeV
Counts / (0.5°)^2 pixel
σ_θ ≈ 0.2° @ 3 GeV

FERMI 5 years
E > 1 GeV

Vela
Geminga
Crab
DAMPE Data Path

- **L0**: telemetry packets from the satellite received by 3 ground stations twice per day (morning and afternoon, ~16 GB/day)
  - Automatically transferred to Ground Support System (GSS) in Beijing
- **L1**: payload data extracted from L0, then merged, duplicated fragments removed and CRC checked
  - 1 day of data can be processed at GSS within 1 hour
  - New L1 data synchronized to Science Application System (SAS) in Nanjing
- **L2/L3**: unpacked and processed payload data in root format
  - Performed at SAS in Nanjing (PMO)
- **L1/L2/L3** data automatically transferred from PMO to Europe (CNAF in Italy)
  - Synchronized to local cluster at DPNC
**DAMPE Data Processing Pipeline**

- **Scientific data**
  - 1 file per data transfer
  - *0013*.fits

- **Housekeeping data**
  - 13 files per data transfer
  - e.g. *011A*.fits

- **Merged files**
  - *0013*.frd

- **Data**
  - Pedestal calibration
    - *PED*.frd
  - Pedestal update
    - *DLD*.frd
  - Gain calibration
    - *DAC*.frd

- **Calibration data**
  - Scientific data
    - *OBS*.frd

- **Data conversion**
  - Raw data -> ADC values
  - Housekeeping data
    - *root
  - Calibration Data Base
    - *cal

- **Reconstruction**
  - Reconstructed data
  - Housekeeping data
  - Data Base

- **Data description**
  - Housekeeping Data Base

---

**Reconstruction done at DAMPE SDC at PMO (>1400 cores, can reprocess 3 years of data in ~1 month)**
Collaborative Software Development

- Base in a SVN code repository DAMPE software is stored in the SVN servers in China and Europe
  - `svnsync` is invoked in PMO server after each commit request
  - cron job is running in Geneva checking if servers are synchronized

Master: SVN server PMO (China)

http://119.78.211.2:81/SVN DAMPE/rep1/

User in China

http://dpnc.unige.ch/SV NDAMPE/DAMPE1/

User in Europe
Data transfer status

• PMO -> CNAF: gridftp
  – 6 gridftp processes running all the time on a dedicated server at PMO
  – Mean transfer rate ~ 6 MB/s, daily transfer: ~ 100 GB, ~110 TB transferred
    • Main L2 science data stream (2A) ~60 GB/day
  – Leased connection between PMO and the China Education and Research Network (CERNET)
    • Limited to 100 Mb/s, OK for daily transfer ~100 GB/day
    • For reprocessed data (~11 TB for 6 months of data) more efficient to transfer by hard drives!

• CNAF -> Geneva: rsync
  – 10 lsf jobs running all the time at CNAF
  – Dedicated data transfer server at DPNC
  – 70-80 MB/s, so basically no latency
  – New data transfer scheme based on XRootD is under development
DAMPE DPNC local cluster

- Share with the ATLAS Tier 3 cluster at DPNC (also IceCube, AMS, T2K)
  - Total: ~750 (+240 soon) cores, ~360 (soon ~600) TB
  - DAMPE: ~60 TB, >200TB by the end of the year
- DAMPE data stored in XRootD and NSF disk servers
  - Main science and MC data sets in XRootD servers for intensive access
    - Need a grid certificate
    - Smaller and temporary datasets in NSF disk
- Local data activities
  - STK monitoring, calibration, alignment, track reconstruction
  - Proton and heavy ion fluxes, electron spectrum, gamma ray analysis
- MC simulation is centrally handled from Geneva
  - a custom distributed workflow submission system which interacts with various computing resources across Europe (CNAF, INFN-Bari, DPNC)
- Two DELL servers running instances for ELOG, Doxygen, workflow submission, data catalog, database server, ...
Data transfer integrity assurance

- A job running at PMO evaluates md5 checksums and pushes the results to the DPNC data server.
  - ~30 seconds per file, ~30 files per day’s data
- On the DPNC rsync server, “agents”
  - Evaluates md5 sums of local files;
  - Marks as “good” files having matched md5 sum with PMO
  - Checks if "good" files in XRootD server with the same checksum
    - If file doesn't exist (or has different checksum), it is uploaded (re-uploaded) to XRootD
- Check independently each file on XRootD has the same md5 sum as in PMO
- A loop is done through the files in PMO, to ensure that each of them was copied to Geneva.
- In addition, every file on XROOTD is open with ROOT by a "crawler", and basic trees are checked, to see if file is not corrupted.
Next generation high energy particle detector on board the Chinese Space Station
  - Similar to DAMPE, but with larger acceptance
  - 5-side sensitive!
    • Advance the frontiers of DM search, cosmic-ray physics and γ-ray astronomy
  - Payload ~2000 kg, launch ~2025, lifetime ~10 years

International proto-collaboration formed: China, Switzerland, Italy, Sweden
  • X. Wu is the Tracker Working Group co-leader
  • R. Walter (ISDC) is the Gamma-ray Astronomy Working Group co-leader

Bigger, better, longer!
Conclusions

• DAMPE is a powerful space telescope for high energy electron, γ-ray, and cosmic rays, in operation!
  – Detector has been well calibrated before launch
• On-orbit commissioning has completed
  – A short commissioning period thanks to the robust detector technologies and the high quality of construction
• Smooth continues and a large amount of good quality data is being collected
• Rather robust data processing/reprocessing and data transfer schemes have been implemented
• Science data infrastructure at DPNC is in place
  – Data production and analysis are in full swing!
• Similar infrastructure is needed for HERD
DAMPE Offline Software Framework

• DAMPE offline software framework is inspired by the GAUDI philosophy ([http://proj-gaudi.web.cern.ch/proj-gaudi/](http://proj-gaudi.web.cern.ch/proj-gaudi/))
  – The core part is implemented in C++
  – Management part is done in Python

• The central component of framework is the algorithm base class
  – All algorithms are inherited from the base class, including algorithms for the conversion of binary data, simulation, and reconstructions

• Several core services:
  – algorithm manager, input-output service (based on ROOT), geometry manager, ...

• Geometry manager allows for loading DAMPE read-out geometry from the XML database without using Geant4 libraries
  – Reducing significantly the overhead of initializing the reconstruction and analysis jobs
Some Software Framework Details

- Simulation is implemented as a dedicated algorithm, based on Geant4
- All the supporting structures of the detector are implemented in the geometry, directly from the CAD drawings, using a custom CAD-to-GDML convertor
- All DAMPE offline jobs are configured via job option files written in Python
  - Boost-python libraries are exploited for accessing the C++ code from Python
- Algorithms can be stacked together inside a job option, forming a sequence
- IO-service acts as a data buffer for communication between algorithms, and for reading (writing) the input (output) data to the ROOT files
  - IO-service is agnostic of whether data comes from the input file or as an output of preceding algorithm in a sequence
- Configuration files are used to store the configuration parameters
  - Allows for explicit versioning (tagging) of data-production campaigns
**Algorithm Base Class C++**

### Algorithm Sequence
- `Core.Add(Alg1)`
- `Core.Add(Alg2)`
- ...
- `Core.Add(AlgN)`

### Ancillary Services
- `Core.Add(GeometryMgr)`
- `Core.Add(ConfigMgr)`

### Code Execution
- `Core.Initialize()`
- `Core.RunEventLoop()`
- `Core.Finalize()`

---

**Input-Output Service (ROOT)**

- **Input data (ROOT)**
- **Output data (ROOT)**

---

**Block Diagram of DAMPE Software**

**Job Description Files (Python)**

```python
import DampeCore as Core

# Algorithm sequence
Core.Add(Alg1)
Core.Add(Alg2)
...
Core.Add(AlgN)

# Ancillary services
Core.Add(GeometryMgr)
Core.Add(ConfigMgr)

# Code execution
Core.Initialize()
Core.RunEventLoop()
Core.Finalize()
```

---

**Configuration Manager C++**

- **Simulation.cfg**
- **Reconstruction.cfg**
- etc.

---

**Geometry Manager C++**

- **Detector Geometry (GDML)**
- **CAD drawings (.STEP)**
<table>
<thead>
<tr>
<th>#</th>
<th>Info</th>
<th>Folder</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single transfer data</td>
<td>1AM</td>
<td>Test binary data</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1E</td>
<td>Raw data (ADC hits, HC data). Root format</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2Q</td>
<td>Reconstructed data for fast monitoring (clusters, tracks, etc.)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1BS</td>
<td>Binary data</td>
</tr>
<tr>
<td>5</td>
<td>Multi transfer merged data</td>
<td>1A</td>
<td>Binary data</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>1C</td>
<td>Pedestal, MIPs and Observation Science Packets</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>1B</td>
<td><strong>Binary data</strong></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>1F</td>
<td><strong>Raw data (ADC hits, HC data), ROOT format</strong></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>2A</td>
<td><strong>Reconstructed data (calibrated energies, hits, tracks, etc.)</strong></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1D</td>
<td>Pedestal, MIPs and Observation Science Packets</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>1Q</td>
<td>BGO calibration data and plots</td>
</tr>
<tr>
<td>12</td>
<td>Calibration</td>
<td>CAL</td>
<td>Calibration constants</td>
</tr>
<tr>
<td>13</td>
<td>Configuration</td>
<td>Config</td>
<td>Job options used for data processing</td>
</tr>
<tr>
<td>14</td>
<td>Position information</td>
<td>EE</td>
<td>Satellite position (longitude, altitude, X, Y, Z), speed (Vx, Vy, Vz)</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>EF</td>
<td>Test position data</td>
</tr>
<tr>
<td>16</td>
<td>Orbit simulation information</td>
<td>OA</td>
<td>Orbit information 16h-00h of the previous day</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>OI</td>
<td>Orbit information 00h-08h of the current day</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>OL</td>
<td>Orbit information 00h-08h of the previous day</td>
</tr>
<tr>
<td>19</td>
<td>----</td>
<td>Productlist</td>
<td>Empty</td>
</tr>
<tr>
<td>20</td>
<td>----</td>
<td>Sub</td>
<td>Empty</td>
</tr>
<tr>
<td>21</td>
<td>----</td>
<td>TE</td>
<td>Satellite test data</td>
</tr>
<tr>
<td>22</td>
<td>----</td>
<td>TM</td>
<td>Documentation</td>
</tr>
</tbody>
</table>