MAGIC
Major Atmospheric
Gamma Imaging
Cerenkov Telescope

MAGIC-TDAS 00-14
000809/H.Kornmayer

A version-zero program for digesting FADC and control data

H. Kornmayer, R.K.Bock, Th.Bretz, W.Wittek
<harald@mppmu.mpg.de>

August 9, 2000

Abstract

Data from all different detector parts and control subsystems of MAGIC will be collected by a process running in delayed real time or at regular intervals. The task of this process is to collect and relate data from the data acquisition (mostly FADC data) and from the control system, and put all data into a single ROOT file for permanent storage, possibly applying some compaction and calibration algorithms.

This document sums up the actually available information about the data and their properties, and proposes a mode of running for early data taking in 2001. The test runs with a very partial camera in late 2000 are also considered.

Comments and criticisms are welcome!

Contents

1 Introduction

During the last month there were some discussions about strategies how the raw data from the MAGIC telescope should be stored. It had been decided in the collboration to use ROOT (http://root.cern.ch) as the tool for permanent storage and data presentation (also the gui interface for analysis can be realized with ROOT). We discuss in this note how a merger step will merge and store the data from the different MAGIC data sources, and which preprocessing steps might be run with it, initially. We call this first step of analysis the merging and preprocessing program, MERPP, being a part of MARS (= MAGIC analysis and reconstruction system).

During the software meeting in the Eng (see http://sarastro.ifae.es/internal/computing/meeting-3-2000/minutes/node6.html), two main streams of data were defined:

- the high-rate event data from the Data Acquisition System (DAQ); these consist nearly exclusively of FADC data.
- the slow data from the Central Control System (CCS).

The high-rate stream will be stored on an intermediate (binary) format (pure raw data). Initially, this information will be recorded during the night on a DAQ-dedicated set of disks. The output from the CCS may be in ROOT format already, but we also have to consider the possibility that some or even all control data are presented in their own intermediate format (ASCII). Again, this information will be on a CCS-owned disk. The merger and preprocessing program will synchronize this intermediate information, and merge the streams into a single file (ROOT raw data) containing all data taken by the MAGIC system during one run. Some obvious corrections (preprocessing) may be applied at this time; the content of output files and the extent of preprocessing being left to future decisions. As there are no data copies made during data taking, the MERPP will run during day time, and will copy onto tape any information which is meant to be kept permanently.

We next have a closer look at the two main input streams to MERPP.

2 The event/FADC stream

This stream contains all data coming from the readout of the FADC system[1]. In this stream, events will be of different types:

- the events or, more precisely, cosmic ray events:
  these events are triggered by the event trigger. They contain the physics that we are interested in, and will be analysed. The rate of these events depends on the source and on the settings of trigger, discriminator thresholds, etc. The maximum rate we design the system for is $\mathcal{R} =$ 1000 Hz, unless some serious data compaction algorithm is invoked in real time.
- the pedestal events:
  the proposed initial stategy for getting the pixel-by-pixel offsets (pedestals) are special pedestal events, triggered at a given constant rate by an artifical (clocked) trigger. These events contain information about the PM offsets and about the light of the night sky background (NSB) including the light from stars. A constant trigger rate of about 10 Hz is proposed, although it should depend on the rate of apparent image rotation (see TDAS note 00-11). At the beginning

[1] The persons responsible for the DAQ are Juan Cortina, cortina@mppmu.mpg.de and Toni Coarasa, coarasa@mppmu.mpg.de.

of a run, a certain number of pedestal events will be recorded to have a statistically meaningful starting sample; subsequent pedestal events will be merged into the mean/rms information with a given weight (e.g. at 10 Hz and an individual pedestal weight of 0.01, pedestals will be dominated (63%) by measurements during the last 10 seconds)

- the absolute light calibration (ALCAL) events:
  there will be a dedicated system on the telescope dish that allows an absolute calibration from FADC digitizings to the number of Cherenkov photons for each channel[2]. This system will send a given number of light pulses with a fixed brightness to all pixels. The normal trigger of the MAGIC telescope will cause the data acquisition to read out these events, from which one obtains the conversion factors from ADC values to number of photons. These events will be taken only once per run or with a maximum rate of $\mathcal{R} = 1/10$ min ≈ 0.002 Hz

The three different kinds of events, appearing at different rates in the event stream, will have the same format; they will occur on the pure raw data stream in chronological order. There will be no structure in the file (i.e. no way to read in only ALCAL events); these events are distinguishable by a flag.

3 The controls system stream

This stream is defined by the Central Control System [3] which collects data from all control subsystems of MAGIC [2]. In general, the frequency of events in the controls system stream is much lower than that of the event data. Control data may have been preprocessed and brought into a suitable form (in ROOT format) by the CCS; this, however, is not presently defined: some or all of the preprocessing of control data may be left to the standard merging/preprocessing step. There are a several different MAGIC control subsystems producing data at different frequencies (e.g. low rate of weather information, much higher rate for oscillation monitoring). Different subsystem can be grouped by data frequency [1], resulting in the following classification:

- run stream
- slow stream
- medium stream
- fast stream

As all data should appear in the final output structure, they will be briefly described in the following.

3.1 The run stream

This data stream contains data from all subsystems creating only on record per run. So one very obvious member of this stream will be the run header and footer (terminator). The RunHeader contains information about the run (run number, project ID, starting time, shift members, run type, coordinates of the source, ... etc.). The RunFooter contains some global information about the run (stopping time, number and rate of triggers, etc....).

[2] The responsible person is Thomas Schweizer, tschweiz@ifae.es.

[3] The responsibility for the development of Central Control is presently vacant

3.2 The slow stream

There are some subsystems that will create data with a frequency of ≈ 0.1 Hz. In [1] the following possible event (data) types are listed:

- HV settings:
  assuming a 2 byte word for each PMT channel and 600 PMT channels one gets a size for a HVevent of 9600 bits. A frequency of ≈ 0.1 Hz leads to a data transfer rate of 960bits/sec.
- weather information
- LIDAR information
- starfield information
- general status

3.3 The medium stream

There are some subsystems that will create data with a frequency of ≈ 1 Hz. In [1] the following possible event (data) types are listed:

- trigger settings
- scalers:
  this will include single pixel rates, trigger rates (without the influence of dead time), and possibly other scalers desired for monitoring.

3.4 The fast stream

There are some subsystems that will create data with a frequency of ≈ 10 Hz. In [1] the following possible event (data) types are listed:

- DC currents:
  assuming a 2 byte word for each PMT channel and 600 PMT channels one get a size for DCevent of 9600 bits. A frequency of ≈ 10 Hz leads to a data transfer rate of 96kbits/sec.
- Telescope position:
  there will be different systems monitoring the tracking. One expects to record the values of the shaft encoders and the sky coordinates (right ascension and declination), with their errors, along with a time stamp (note: careful synchronization of time is crucial, we assume here that it is provided by a general GPS time distribution).
- Camera oscillation:
  realistically, we should assume that the time analysis of raw oscillation data has to be made in the preprocessing step; the raw data consist of the displacements dx, dy, dalpha, together with

a time stamp (see note on synchronization in previous paragraph), taken at regular intervals (up to frequencies of 10 Hz, depending on weather conditions). More optmistically, we could assume that oscillation data are time-analysed in real time, inside the control system. The shorter output of this analysis will then be in parameterized form (e.g. a few terms of Fourier analysis), allowing interpolation of corrections using the time tag of each event.

4 Summary on input rates

The different streams and types of events (data) can be found in table 1.

stream	Event Type	rate	size of event	transfer rate	data / hour
FADC stream	CR events	< 1000 Hz	20 kB	20 MB/s	72 GB
	Ped events	≈ 10 Hz	20 kB	200 kB/s	720 MB
	ALCAL events	≈ 0.001 Hz	20 kB	20 B/s	72 kB
fast stream	DC currents	$= 10$ Hz	1.2 kB	12 kB/s	43 MB
	telescope position	$= 10$ Hz	??	??	??
	camera oscillation	$= 10$ Hz	??	??	??
medium stream	Trigger settings	$= 1$ Hz	??	??	??
slow stream	HV settings	$= 0.1$ Hz	1.2 kB	0.12 kB/s	430 kB
	weather	$= 0.1$ Hz	??	??	??

Table 1: **structure of all data streams** that should be presented in the final raw data file of MAGIC

5 Requirements for the merging and preprocessing program

The requirements for the raw data file created by the Merger Program are listed here.

1. **only one single raw data file**
   After the process of the Merger Program the different event types (CR events, Ped events, DC currents events, Camera oscillation events, etc,...) should be merged in one single file. This file contains ROOT trees with a different event type each.

2. **fast access to the different event types**
   The access to different types should be very easy and - more important - very fast. So the access to data of the slow stream (e.g. weather data) must be direct, accessing only the selected information without reading through a large volume of (unselected) data. ROOT files allow this selective access.

3. **quick plots of distribution of measured values**
   For fast checks of data quality it is important to create the distributions of some quantities (i.e. humidity during night, etc...), beyond the monitoring that will go on during data taking both in the DAQ system and the CCS. Such plots can be done easily using the standard tools of ROOT, once the data exist in ROOT format.

4. **logging information included**
   Logging information from the different control subsystems must exist as one of the ROOT trees in the output of the merger program step.

5. **Monte Carlo and raw data must follow the same structure**
   The stucture of the raw data file and the structure of the Monte Carlo data file must be the same. MC events, however, will have additional information (e.g. a class containing the generation details), and will not have, in general, all the additional control information like weather, LIDAR, etc. A flag (and the file name) will indicate that a file is from simulation, and that some calibrations may already have been applied.

6. ...

6 The structure of the file after the Merger Program

The binary format of FADC data is described in [3]. The ROOT format has yet to be supplied.

7 Initial policy of data storage (first light in 2001)

The pure raw data from both DAQ and CCS will be kept permanently, but will not typically be re-read after passing through the Merger Program. Initially, the standard merger step will not suppress any events (of whatever type), will maintain all pixels and all time slices of the FADC, and will not apply any corrections to the data; it will, however, apply loss-less compaction to FADC data (e.g. suppressing low-gain channels if not defined). The output of the merger step is kept permanently (ROOT raw data).

A separate preprocessor step will subsequently correct and substantially reduce FADC data by time analysis, pedestal subtraction, oscillation and position corrections, conversion to photons and, possibly, thresholding (zero suppression). The common execution of merger and preprocessor is seen as the standard first step. Once confidence in the programs has been established, ROOT raw data may no longer have to be kept, but these decisions are for a later date. The possibility to combine programs in different order will be foreseen from the beginning.

8 First implementation (test runs in Oct/Nov 2000)

A first merging and minimal preprocessing program will be implemented using the proposed MARS framework. It will assume that pedestals are obtained from specially triggered pedestal events (not a suggestion to maintain this forever!). This first implementation could have the following specification (to be revised at the workshop in Goettingen?):

- read in raw data from the DAQ system, and build a ROOT tree with camera-triggered events and pedestal events separated (no absolute light calibration will be available, and no data stream from the control system is envisaged). These raw ROOT data are preserved.

- analyze pedestal events forming the mean and rms error for each pixel, with an update calculated upon arrival of each new pedestal event. The weight factor to be used in updating could be set at .01 for a single event, should be dependent on time interval. The analysis will allow to define outliers, develop a graphical output of histograms for ADC values in the same pixels, for means/rms errors of different pixels, etc.

- idem for pixels in 'cosmics' events (this will serve to see if pixels without a hit in triggered events could be used for pedestals).
- exercise the graphical interface for setting up program versions, if it is available by October.

References

[1] Petry D., Minutes of MAGIC Control Software Meeting, ENG, 15/16 March 2000

[2] Petry D. et al., The Control System of the MAGIC Telescope, TDAS 00-07, 11 March 2000

[3] J. Cortina and J.A. Coarasa, Description of the DAQ raw data format, TDAS 00-09 DRAFT 2, 8 August 2000

[4] Oestereich B., "Objektorientierte Softwareentwicklung", Oldenburg Verlag, 1998