MAGIC Major Atmospheric Gamma Imaging Cerenkov Telescope

MAGIC-TDAS 01-07
D. Petry

Proposal for the MAGIC Telescope Datacenter

D. Petry

Draft1, 21 February 2001

Abstract

This proposal describes a possible design of the organizational structure and the hardware setup for bringing the MAGIC Telescope raw data from La Palma to the central archive and processing center ("datacenter"), performing the basic reduction and making the data accessible to the users.

The main features of the proposal are the use of tape libraries both at La Palma and at the datacenter, centralized calibration and FADC data reduction, and access to the data via a password-protected webserver at the datacenter. Using the webserver, the users can upload preselection scripts and either download the resulting reduced datasets via the Internet or request data on tape. The final analysis is done at the home-institutes of the individual users.

The suggested tape medium is LTO, Sony AIT-2/3 or Exabyte M-2.

Contents

1 Overview

Figure 1 shows schematically the proposed setup. The data is acquired on La Palma and written on two tapes. One tape stays at La Palma. The other one is sent to the *datacenter* where it is archived. The datacenter reads the data on a large hard disk and performs the calibration and FADC reduction. Users are able to access the datacenter via a webserver. They can submit *preselection scripts* which are executed on the calibrated data and result in *reduced datasets* which are smaller in volume than the original raw-data by about a factor 70. These datasets are then downloaded via the Internet or sent on tapes by mail. The final fine analysis is performed by the users in their home institutes.

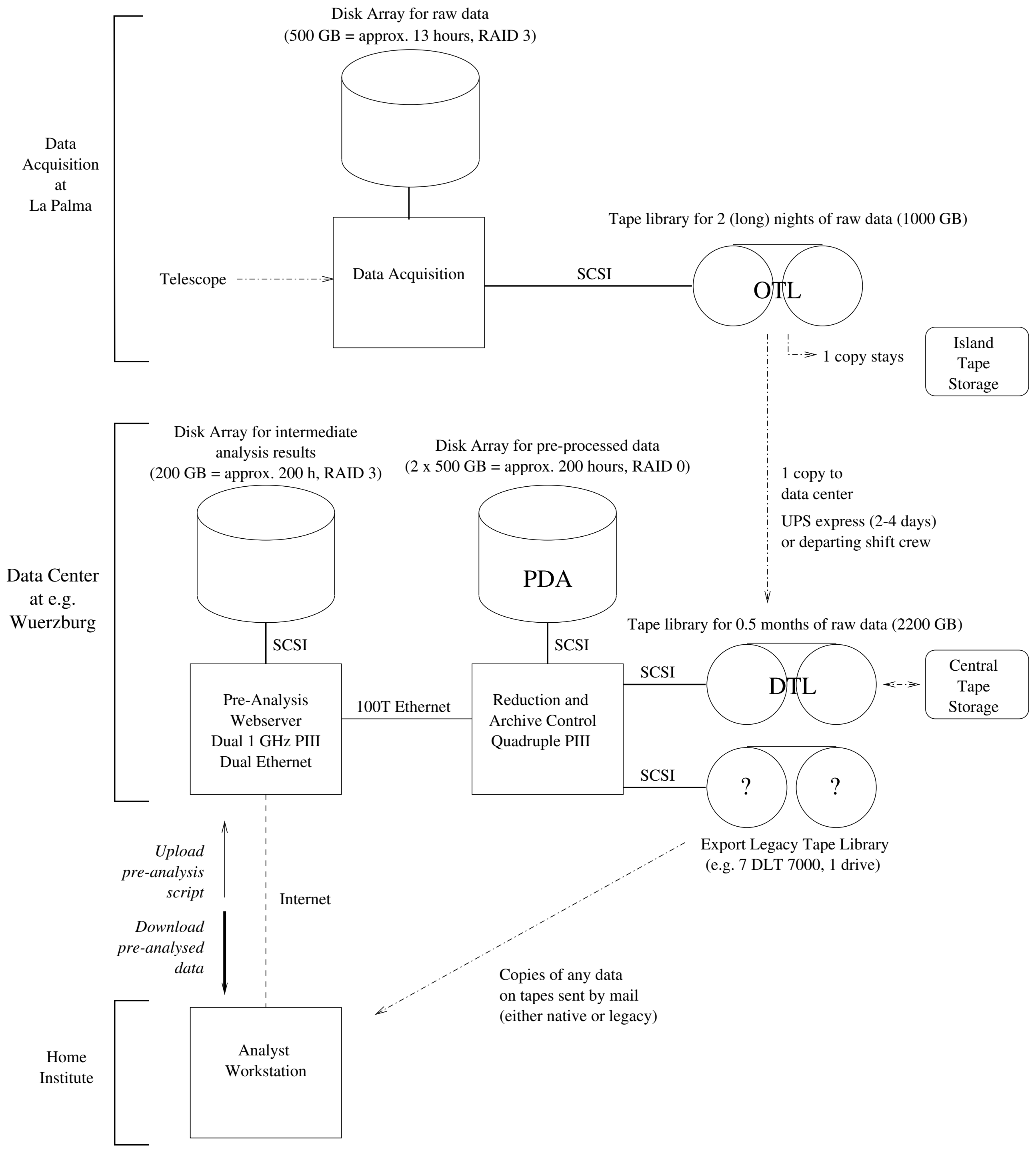

Figure 1: Overview of the setup proposed in this text.

2 Tape media

The choice of media is difficult and more information has to be gathered before a final decision can be made. The three candidates seem to be Sony AIT-2/3, HP/Seagate/IBM LTO (Ultrium) and Exabyte M-2.

All three media can - according to their nominal properties - do the job, but the question is whether they are reliable enough to provide smooth data taking and handling.

If one of the media provides higher sustained transfer rates, it still has to be demonstrated that these can really be achieved in connection with our computer hardware. Otherwise, purely from the speed point of view, one should decide for the fastest, i.e. LTO, although also Exabyte M-2 would be sufficient. For both technologies there seems to be Linux support, especially Exabyte M-2.

From the reliability point of view (and with the arrival of the third generation also from the capacity point of view), Sony AIT-2/3 is probably the optimum. Linux seems to be no problem here either since SuSE itself offer AIT systems.

From the capacity and the media cost point of view, LTO seems to win, but it is the youngest technology. Prices for AIT-3 are not yet available.

All features of the MAGICT data system are independent of the tape media choice. This choice is therefore not made in this proposal.

3 Data recording and archiving at La Palma

The data is acquired at the predicted rate of 10 MB/s and written to a RAID disk. Already in parallel with the acquisition, an asynchronous process writes the data to tapes which are kept in a tape library.

The Observatory Tape Library (OTL) is dimensioned to hold the maximum data volume expected from two maximum duration nights (2×12 h). Since one cannot expect major compression on the binary MAGICT data, this will be 2×432 GB, i.e. the OTL should hold tapes for 1000 GB of uncompressed data.

The data disk from which the tapes are written should be able to hold 500 GB to avoid data loss in the case of tape problems.

The OTL has two tape drives for redundancy and for fast change to the next tape.

After the data is taken and written on tape, a second copy of it is written on a second set of tapes.

The second set of tapes is for shipping to the datacenter. The first set is to be kept in a dedicated dust-free storage place at the observatory.

The tapes are marked with unique barcodes and Memory In Cartridge (MIC) labels. A database of these labels is kept at La Palma.

4 Tape media transport to the datacenter

The tapes are sent as often as possible but at least once per week to the datacenter. The shipping is done either with members of the collaboration or with UPS. The latter is the preferred way since it is probably the fastest. The additional cost is not very high since in most cases the collaboration member will not travel directly to the datacenter and therefore will have to send the tapes from his/her home institute anyway.

5 Data archiving at the datacenter

Upon receipt of new data tapes, the "keepers" of the datacenter register the new tapes with the datacenter tape label database. Then they insert the tapes into the Datacenter Tape Library (DTL) and start the preprocessing such that the preprocessed data is available to the collaboration as soon as possible.

The DTL is similar to the OTL but holds enough tapes to accommodate the data from half an observing period (120 h / 2 = 60 h) which is 60 h $\times$ 36 GB/h = 2160 GB. This is the size of a typical medium-size tape library.

When the DTL is filled with tape cartridges, and new tapes arrive from La Palma, the oldest cartridges are removed and put into dedicated storage places for long-term storage. This means that the raw-data remains in the DTL for about 14 days which should be enough for all necessary processing.

The DTL has 3 tape drives for redundancy and versatility. This will also be useful for copy jobs.

In case calibration parameters change and data has to be re-processed, the necessary tapes are returned to the DTL. If this should be necessary too often, the DTL has to be extended.

6 Data pre-processing at the datacenter

The DTL is directly (Ultra wide SCSI 2 = LVD) connected to a high-performance computer (the Preprocessor, a 4-processor machine) with two large RAID disk arrays (500 GB each).

The processing consists mainly of FADC analysis and calibration. A group of members of the collaboration should take the responsibility to ensure that the correct calibration methods and parameters are used.

The pre-processed data are kept on the 2 $\times$ 500 GB disk array (the Preprocessing Disk Array, PDA) until new incoming data makes it necessary to delete them. They are not archived. If we expect a factor 7 data reduction during preprocessing, this means that about 200 h of data can be stored on the PDA. The individual dataset can remain on the disk for between 1.5 to 3 months depending on the season (length of nights). During this time, users can access the data via the pre-selection webserver or ask for tape copies.

After the pre-processed data is removed from the PDA, the data will be less easy to access and a special requests will have to be made. This encourages swift data analysis but still leaves enough time for analysts to react.

7 The pre-selection webserver at the datacenter

Using standard web-server technology and programming techniques like cgi-scripts, a webserver is set up at the datacenter on a high-performance computer (2-processor machine). This computer is connected to the Internet and, via a separate Ethernet connection, to the Preprocessor and its PDA.

Users with a password can log in on the webserver and submit pre-selection scripts. This starts jobs on the web-server which access the preprocessed data on the PDA and select suitable subsets applying at least a data reduction of 10.

This pre-selection is done using simple gamma-hadron separation methods or other cuts on arbitrary parameters which can be defined by the user. The resulting pre-selected dataset is stored on the RAID disk array of the webserver (200 GB) and can then be downloaded via ftp to the home-institute of the user who can then do the final analysis without further interaction with the datacenter. The pre-selected dataset remains on the webserver until it is released by the user or until the system administrator finds the disk too full. In this way, failed downloads can be repeated.

In case the pre-selected dataset is too big for download or the connection to the user too slow, a tape copy can be requested and will be sent within a week.

With the dataset, the user also obtains all statistical information of the data he processed. If the pre-selection is made adequately, this statistical information (which includes histograms etc.) may in some cases be enough and no dataset has to be downloaded at all.

The webserver is scalable: In case the processing speeds are too low, the webserver can easily be duplicated by installing a second computer with identical setup. Different users can then submit to different servers and speed up their request.

8 Cost calculation

8.1 Hardware costs - La Palma

The necessary hardware consists of a tape library for 1000 GB of uncompressed data, a 500 GB RAID array and tape cartridges.

Media choice	Sony AIT-2	Ultrium	Exabyte M-2
Library type	Qualstar TLS42222-2L	Qualstar TLS8211-2L1	Exabyte 215M
Number of slots	20+4	11	15
Number of drives	2	2	2
Barcode reader	yes	yes	yes
Cost excl. VAT	34430.- DM	38610.- DM	35860.-

Media choice	Sony AIT-2	Ultrium	Exabyte M-2
Tape cartridges for one year (37800 GB)	2 sets à 756 tapes	2 sets à 378 tapes	2 sets à 630 tapes
Cost excl. VAT	270648.- DM	202608.- DM	288414.- DM

The cost of a 500 GB RAID array is approx. 12800.- DM excl. VAT. In addition, storage cupboards for the longterm storage of the tapes are necessary.

8.2 Hardware costs - Datacenter

The equipment needed at the datacenter consists mainly of a medium-size tape library, a 4-processor PC, a 2-processor PC, two 500 GB RAID arrays and a 200 GB RAID array. Furthermore some network equipment like an Ethernet Switch and storage cupboards for the tapes will be necessary but their cost is not given here.

Media choice	Sony AIT-2	Ultrium	Exabyte M-2
Library type	Qualstar TLS4440-2L	Qualstar TLS8433-2L1	Exabyte X80
Number of slots	42	33	40
Number of drives	3	3	3
Barcode reader	yes	yes	yes
Cost excl. VAT	60020.- DM	68500.- DM	85879.-

The necessary RAID arrays cost 12800.- DM (500 GB) and 6200.- DM (200GB) respectively.

The cost of the two necessary computers can only be estimated at the moment since the requested offer was not yet available at the time of writing. An upper limit is probably 20000.- DM.

The cost for an additional legacy tape drive will be e.g. 8000.- DM for a DLT7000 drive.

8.3 Other costs

The cost for sending data tapes from La Palma to the data center by UPS will be approximately 52 $\times$ 120.- DM = 6240.- DM per year.

8.4 Cost summary

For the first year of operation, the following costs will have to be envisaged in case the Ultrium media choice which is the cheapest over all:

Item	Minimum Cost
La Palma OTL	38610.- DM
Tapes	202608.- DM
Datacenter	120300.- DM
Tape Shipping	6240.- DM

Not included are storage cupboards at La Palma and at the datacenter (estimation: 2000.- DM), network equipment for the datacenter (estimation: 4000.- DM) and VAT.

Half the cost is caused by the tape cartridges. Loss-less online data-reduction will therefore be the main way to reduce cost.