Data handling and control for the European Solar Telescope

Ilaria Ermolli[a], Felix Bettonvil[b], Gianna Cauzzi[c], Lluis Cavaller[d], Manuel Collados[e], Paolo Di Marcantonio[f], Frederic Paletou[g], Paolo Romano[h], Jean Aboudarham[i], Roberto Cirami[f], Rosario Cosentino[h], Fabrizio Giorgi[a], Martine Lafon[g], Didier Laforgue[e], Kevin Reardon[c], Guus Sliepen[b]

[a]INAF Osservatorio Astronomico di Roma, via Frascati 33, Monte Porzio Catone, Italy;
[b]Astronomical Institute, Utrecht University, Princetonplein 5, 3584CC Utrecht, The Netherlands;
[c]INAF Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy;
[d]Grantecan, S.A., via Lactea, 38205 La Laguna, Tenerife, Spain;
[e]Instituto de Astrofisica de Canarias, via Lactea, 38205 La Laguna, Tenerife, Spain;
[f]INAF Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, 34131 Trieste, Italy;
[g]Laboratoire d'Astrophysique de Toulouse-Tarbes, Université de Toulouse, CNRS, 14 av. E. Belin, 31400 Toulouse, France;
[h]INAF Osservatorio Astrofisico di Catania, via S. Sofia 78, 95123 Catania, Italy;
[i]Observatoire de Paris, 5 Place Jules Janssen, 92190 Meudon, France

ABSTRACT

We introduce the concepts for the control and data handling systems of the European Solar Telescope (EST), the main functional and technical requirements for the definition of these systems, and the outcomes from the trade-off analysis to date. Concerning the telescope control, EST will have performance requirements similar to those of current medium-sized night-time telescopes. On the other hand, the science goals of EST require the simultaneous operation of three instruments and of a large number of detectors. This leads to a projected data flux that will be technologically challenging and exceeds that of most other astronomical projects. We give an overview of the reference design of the control and data handling systems for the EST to date, focusing on the more critical and innovative aspects resulting from the overall design of the telescope.

Keywords: Data handling, control, large telescope

1. INTRODUCTION

The EST (European Solar Telescope) is a 4-meter class telescope optimized for studies of the magnetic field and dynamics from the deep photosphere to upper chromosphere of the Sun (Collados et al. 2010 [1]). The EST is designated with the highest priority among the ground-based, medium term (2016-2020), medium size new projects in the ASTRONET Roadmap for European astronomy.

The conceptual design study of the EST, aimed to demonstrate the scientific, technical and financial feasibility of this project, has been financed by the European Commission in the seventh framework programme. This study, that covers all key aspects of EST design, from optical configuration to building and enclosure, is presently close to completion. It has led to the concepts of the various sub-systems of the EST presented on various papers of this issue (Collados et al. 2010 [2]).

We focus here on the control and data handling systems of EST. We describe the main functional and technical requirements for the definition of these systems, and the outcomes from the trade-off analysis between the technical alternatives identified to date. We also give an an overview of the current design of these systems.

Further author information: (Send correspondence to I. Ermolli)
Ilaria Ermolli: E-mail: ermolli@oaroma.inaf.it, Telephone: 39 06 94286470

Software and Cyberinfrastructure for Astronomy, edited by Nicole M. Radziwill, Alan Bridger,
Proc. of SPIE Vol. 7740, 77400G · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.856938

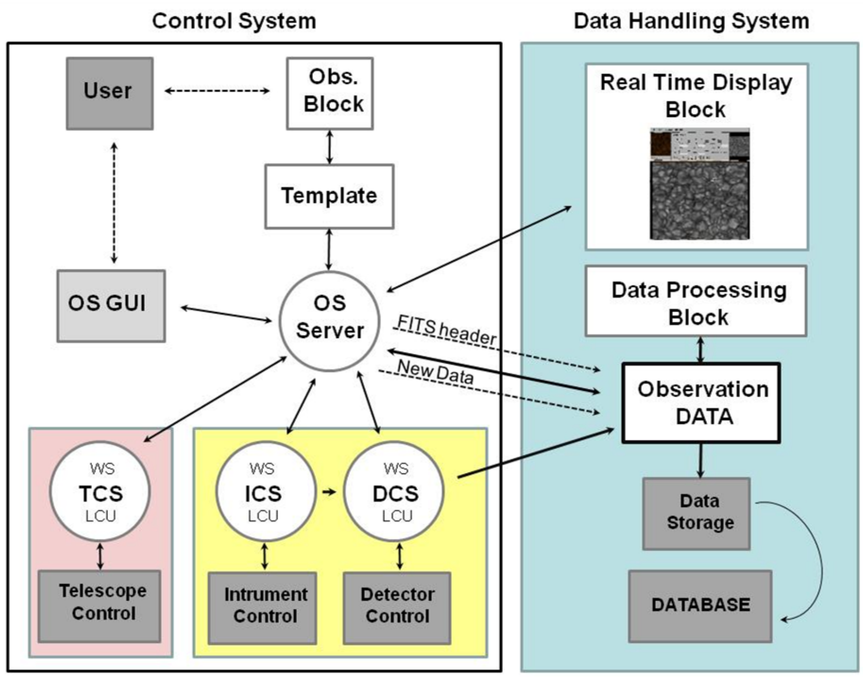

Figure 1. Observatory system model. Details are given in Sect. 2.2.

2. CONTROL

2.1 Requirements

The system in charge of EST control (hereafter referred to as EST Control System, CS) is required to provide for various modes of operation (day-time solar observation, day-time non-solar observation, night-time observation, engineering, maintenance) by various users (engineers, staff astronomers and technical operators, visiting astronomers), for pointing and tracking a position in various coordinate systems (none, sideral, heliocentric, heliographic), for operating, monitoring, and controlling various sub-systems. Within the various observing modes, the system shall allow users to choose among different observing methods. In addition, the system is required to provide for various modes of operation of single instruments, for simultaneous and synchronized use of the various instruments, and for coordinated observing campaigns of EST with other telescopes.

The analysis of functional requirements shows that CS shall have performance requirements for control similar to those of current medium-sized night-time telescopes. The technical alternatives considered to date indicate that the software architecture of CS could be based on: 1) systems specifically developed for this project, 2) complete open source systems developed for other projects, e.g. for data handling and control of other astronomical facilities such as ALMA, ATST, GTC, that could provide both architecture and tools to realize a system fulfilling the specifications and requirements of the EST, 3) complete commercial solutions, 4) hybrid solutions, with an architecture based on both open source existing systems and commercial solutions, the latter utilized for the low-level control of elements in the various sub-systems. These alternatives are still under study. On the other hand, the general strategy and main architecture of CS have been outlined already, though they may suffer modifications with project development.

2.2 Reference design

The EST CS shall consist of a set of physically distributed sub-systems, which have to work in a coordinated manner, and a common software architecture adopted across the whole observatory. Figure 1 shows the current reference design of EST CS. The various sub-systems shall include central common services (CCS), telescope control (TCS), instrumentation control (ICS), detector control (DCS), observation operation (OS), and maintenance operation (MS). These sub-system shall operate through standard commands and common software and

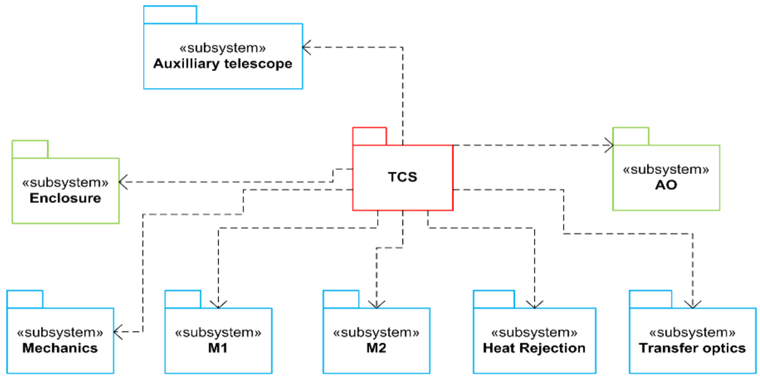

Figure 2. Telescope Control sub-system model. Details are given in Sect. 2.2.

have identical interfaces to other observatory systems and sub-systems, e.g. the data handling system (hereafter referred to as Data Handling Control System, DHCS) described in the following.

The physical architecture of CS shall consist of a set of interconnected computers, electronic equipment, sensors and actuators (Local control units, LCU); all these components shall be responsible for the direct control of the various sub-systems. The various workstations (WS) shall be connected via one or more LANs, that shall also provide access to a group of centralized services. Fiber optic connection shall be used in order to secure high rate of data transfer. At the moment it seems feasible to utilize current technology for most of the components of CS sub-systems, except for some elements of DHCS.

The sub-system of Central Common Services (CCS) shall consist of libraries and files for the development, usage, and debugging of observatory operations. Some of these files shall have graphical user interfaces. The CCS shall include utilities for the development of applications, communication software and drivers. These shall form a family of software modules interfacing the low-level control and hardware boards of various elements in the sub-systems of CS directly. These modules shall consist e.g. of utilities common to all drivers for log-on and graphical interface debug, analog I/O drivers, digital I/O drivers, time board drivers, servo amplifier, encoders, engineering user interfaces. Drivers shall not be accessed directly by applications, but via a higher level software acting as a service layer between the application level and the low-level device driver. This should provide a standardized, hardware independent interface between applications and devices, motors. The software of this intermediate module should allow for an application to operate a specific motor without knowledge of details about drivers or boards. The CCS shall provide also for basic services (e.g. logging, event, handling of error, message, and time, alarm management, access control, command handling, common access interface), time synchronization, database monitoring, command interpreter, debugging tools (engineering interfaces), real time access to a local data base. Special modules shall be developed for handling applications dealing only with specific observatory components, e.g. the telescope sub-systems described below.

The Telescope Control sub-system (TCS) shall control all devices which belong to the main telescope structure, including auxiliary telescope and observatory enclosure. It shall consist of various modules or sub-systems. Figure 2 shows some of the modules constituting the TCS. These include telescope structure, primary mirror, secondary mirror, AO, heat stop, transfer optics, polarization optics, auxiliary telescope, enclosure. The design of hardware and software components of these modules depends on their specific elements and mechanisms, that are presently under study.

The Instrument Control sub-system (ICS) shall control all devices which belong to the various instruments, except the detectors, which shall be controlled by a dedicated detector control sub-system. This sub-system shall consist of various modules for the operation of the three instruments planned at the Coudé focus and for their coordinated use. Apart from having a standard structure, the ICS shall also have a number of functions that are identical across the different instrument sub-systems, e.g. to set up, to run, and to monitor the status of each instrument.

Both TCS and ICS shall provide the CS with general purpose utilities and library functions employed by various telescope and instrument sub-systems. They shall look also at system services (e.g. monitoring of disk space and of telescope status, transfer of metadata, real time rapid data display) and allow for data transfer and protocol converter between various observatory sub-systems. The TCS and ICS shall maintain all parameters required for operation in a local database, the central repository, that should be is updated when the status of hardware components is read or changed during operations. These parameters, that shall include configuration, setup, and status parameters, could be automatically copied, cyclically or on change of values, also in a second repository of the DHCS, so that they could be directly accessible by the DHCS for data production and status information display. Control and data information shall be transferred over one or more LAN connecting the various WS and LCUs of TCS and ICS. The parameters handled by the TCS and ICS depend on the specific elements and mechanisms of both the telescope and instruments at a given time and could vary with time.

The Detector Control sub-system (DCS) shall carry out all the tasks required to control instrument detectors, to manage also real-time image processing, if needed, and to transfer the data from detectors to the central row data store of the DHCS. This sub-system shall consist of various modules. They shall include LCUs responsible for the interface to the low-level control and camera hardware and WSs utilized for the interface to the Observation sub-system.

The Observation sub-system (OS) shall be responsible for the execution of observations for a given observing mode and for the final archiving of the observations with standard format files. It shall provide for setup and coordination of the various CS sub-systems such as instrument, detector, telescope and also interfaces to other observatory components (e.g. the data storage sub-system). It shall also contribute to creation of the final FITS header of the science data file. The OS shall not access hardware functions of the instruments and telescope, but it shall have the knowledge of how to coordinate the various sub-systems of CS to perform observations and operations for the various observing and operation modes. This sub-system shall consist of various modules responsible for the execution of single exposures, for the storing of the results of exposures in the data store with FITS files, and for the definition and running of sequences of exposures, that shall be varied out for acquisition, calibration and science purposes. It shall represent the highest layer of the DHCS. The result of an observation shall consist of the complete set of data from the read-out operations and of its full description, with all the metadata and logging information. An observation may also require two or more different instrument setups, overlapping observations (e.g. by managing the next instrument setup in parallel to the readout of detectors or by managing synchronized observations by two or more instruments, each one running with its one setup). The archiving operation shall not affect the observing cycle, this requires a separate module of DHCS shall manage the archiving operations.

The Maintenance sub-system (MS) shall carry out instrument configuration, check-out and troubleshooting. It also shall provide technical templates, e.g to verify the instrument calibration, which involve operations of other sub-systems of CS.

Finally, the Observer Support Sub-system (OSS) shall consist of tools to support observers during the preparation of observing runs and the creation of observation templates. These shall be sequencer scripts that then shall allow to setup and execute one or more observations. The end-user who operates an instrument through these templates should only be exposed to a limited set of sub-system variables, instead of all the parameters of setup and configuration files.

3. DATA HANDLING

3.1 Requirements

The Data Handling Control System of EST (DHCS) shall provide for the recording of all the data and metadata acquired by the telescope and instruments, for the access and display of the data needed to facilitate telescope operations, and for the transfer of produced data away from the telescope to intermediate and then to long-term data archiving facilities. These facilities shall allow for the re-processing and re-use of old data; in addition, they shall serve as dispenser of Science-ready data (VO-compliant) and processing tools. They shall also allow for the publication of data into VO and the interoperability with databases produced by other facilities.

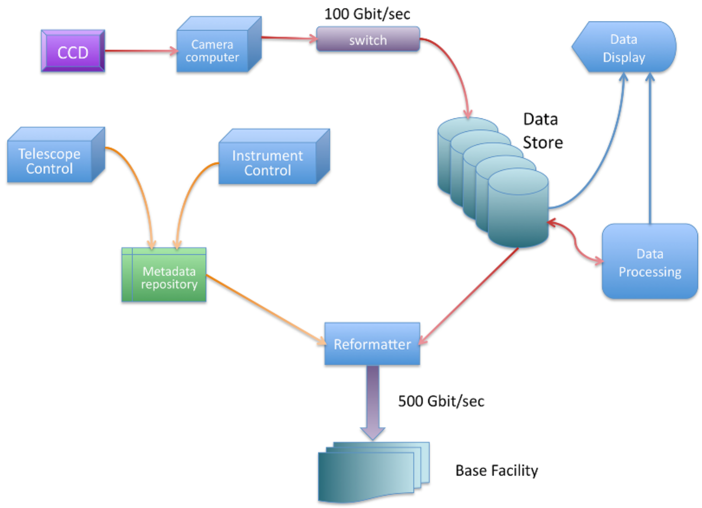

Figure 3. Block-diagram of the reference design for the on-site Data Handling Control System. Details are given in Sect. 3.2.

The three instruments planned at the Coudé focus of EST shall consist of various channels each and include 32 detectors to date. The science goals of EST and technical solutions identified to date point to use detectors with 4000 × 4000 pixels, 2 bytes/pixel depth and reading up to 100 frames per second. The data flux estimated from each instrument is 18.9 GB/s to 24.6 GB/s, depending on the instrument; the overall data flux from the three instruments exceeds 80 GB/s. Twelve hours of EST operation at maximum data rate translate in a data volume produced of over 3.5 PB per day. The data store at the telescope shall have a total capacity for several days of consecutive observations of order of 25 PB per week. In addition to the scientific data, many other details of the facility operation shall also need to be recorded. At the moment, it seems that only some high-energy and particle physics experiments (e.g., BaBar at SLAC or LHC at CERN) produce or shall produce data volumes similar to what can be expected from the EST.

The DHCS of EST shall consist of various sub-systems to support multiple different operational modes for the telescope, the running of different observing programs on the same day, the interaction with users at the telescope, as with remote users, the automated processing of data for both standard calibration (dark-current, flat-field, polarimetric calibration, image degradation correction) and global reduction of the data volume prior to moving it off the summit. The purposes of such processing might include 1) Data quality assurance, 2) Data volume reduction, 3) Evaluation of the solar structures being observed, 4) Standard data processing and data restoring. The data volume could be reduced to 2/3 with evaluation of observations quality and to 1/10-1/50, depending on the instrument, with on-site data processing. These two actions would allow to reduce the data volume expected from EST to that planned from other astronomical facilities, e.g. ATST and LSST. Among the various modules of DHCS, the data display sub-system shall provide for the visualization of the acquired data and metadata in an efficient manner. The primary use of this capability shall be to provide near real-time feedback to the operators and users about the telescope operations. The data display may have to combine the raw data stored on disk with metadata extracted from the repository in order to provide a coherent set of information to the users, also for engineering purposes.

3.2 Reference Design

Figure 3 shows the current design for the on-site DHCS. In this design, all the data from the instruments shall be written to a central data store as it is acquired. This central store shall utilize a shared pool, that shall be partitioned among experiments and instruments in a dynamic manner. By using a homogeneous storage facility, the data can be efficiently managed using a single set of applications.

In the current design, the central storage area, with a size of approximately 25 PB, is based on traditional hard-disk technology. Alternatives may be explored closer to the time of construction, depending on techno-

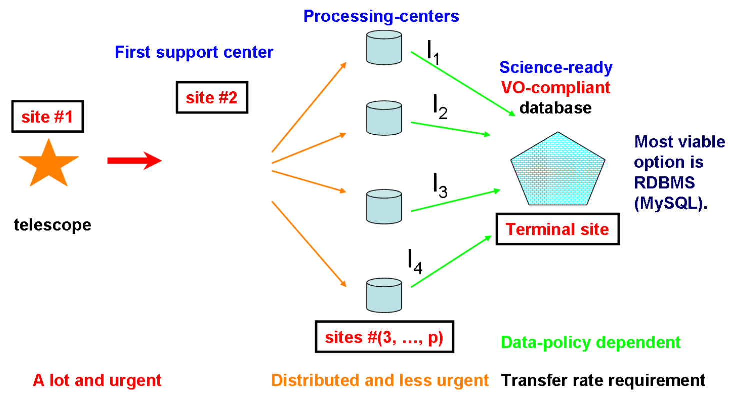

Figure 4. View of the data stream from the telescope to the base center (raw data archive), one or more processing centers (I_n), and the terminal VO-compliant database. Details are given in Sect. 3.2.

logical development of technologies. Assuming a hard disk size of 50 TB in 2018 and a fully redundant RAID configuration for speed and security, the data store shall require several racks and space for the associated electronics and cooling. The data store sketched above is comparable presently to the implementation of a 600 TB data store, complicated but not a serious challenge. In fact, Petabyte storage arrays are already available from multiple vendors.

The connection from the camera computers to the data store would presumably be through a switched network layer, to allow for a highly efficient point-to-point transmission of the camera data stream. To maintain the high data rates, enterprise class hardware shall be needed as well as possibly fiber optic connections. No current technology can support the transfer rate of up 100 Gbit/s from the detectors expected for the EST, but development of 100 Gbit Ethernet promises to reach these speeds in the near future at a moderate cost. Other technologies, such as PCI Express, Infiniband, Myrinet may offer similar transfer rates in the coming years. Therefore a single connection may suffice for almost any detectors designed to date. In addition, if technologies do not reach the desired speeds in the expected time-frame, multiple connections could be used for a single camera, streaming data from different taps on the chip over separate physical connections, as well as to connect the data repository to long-term storage units.

The processing needs at the telescope shall require processing nodes on-site. For instance, application of Speckle interferometry on the images acquired at the full rate of 100 frames per second would require a sizable computing cluster, comprising several 100 cores, with significant RAM and fast interconnects; these may be difficult to manage at the telescope.

All the metadata collected from the instrumentation and telescope systems shall be stored in real time in a central repository. Some metadata may be recorded at near real-time rates for each images (acquisition time, instrument configuration, AO seeing information). Other metadata from instruments or the observatory control systems may be recorded at slower rates. Receiving data from many subsystems at a variety of rates, the repository shall be strongly transaction driven. The most reliable system for recording such information at these rates shall be a database system, most likely a well tuned relational database. Such a system shall have to support the writing of information at a high rate into the database, while at the same time supporting queries of the database contents from other telescope systems in near real-time (perhaps with a lower priority). An in-memory database might support such demanding real-time uses, while this information could transit to a traditional disk-based database to maintain a historical record of the database contents (e.g. for engineering purposes). The entire database should probably be mirrored to the first support center or long-term archives, though not in real time.

The metadata repository would also be queried to produce the FITS or other metadata headers to attach to the raw Level-0 data in the data store in the process of converting those data to Level-1 formatted data. A timestamp and other essential image identifying information, written both in the metadata repository and as a small tag attached to each Level-0 image, would be used to associate the appropriate metadata to the raw data frame. The reformatting of the data may take place on a resource-available basis, possibly after the end of the daily observations, or on-demand for certain small data sets. This latter would facilitate the processing of limited data sets through a reduction pipeline for immediate data quality assurance purposes.

The data display system shall extract the necessary data from the data store and possibly the metadata repository. These data may need to be partially processed prior to display, so some computing power shall be incorporated in the display system. The algorithms to be applied may be fixed, with the possibility of incorporating new software as needed. In order to extract images from the data store and read them for display at rates of up to 10 frames per second, a high-speed connection to the data store shall be required. Proper access controls shall need to be implemented to avoid resource contention or the reduction of the data acquisition rate due to data display tasks. Multiple data displays may be operating simultaneously showing different data.

Figure 4 shows the general strategy for the data distribution beyond the telescope, down to one or several processing centers, and to the terminal Virtual Observatory Compliant Database of EST (hereafter referred to as VO-Compliant Database, VOCDB). The current design considers the transfer of the data away from the telescope to an intermediate, first support, data storage facility at the base of the mountain. This facility shall allow a more convenient and economical installation of a portion of the data handling facilities and better working conditions for the staff. The base center is likely to be somewhere in the Canary Islands due to location planned for the EST. This base center shall be connected with a very fast dedicated Ethernet connection to the telescope. Sufficient data transfer rates might be achieved with soon-to-available 100 Gbit Ethernet connections, or the 1 Terabit Ethernet expected to be fully available by the beginning of the EST commissioning period. Such a connection would probably require the installation of a special physical link, based on single mode fiber optics, between the telescope and the base facility. On the other hand, the logistic complexity in monitoring the flow of physical media up and down the mountain makes data move through physical media a less than desirable solution.

The base center shall comprise a large data store, with a pool capacity at least several times that necessary at the telescope, that is of the order of 25 PB. It may be easiest to use a hard-disk based storage system also for this purpose, although tape-based or other removable media might be considered especially for longer term storage. Presumably the data shall only be staged here temporarily before onward transfer to other long-term archives. Given the much slower rate of transfer away from the Canary Islands over normal telecommunications lines, it may be necessary to process and reduce the data volume at this stage. This may require significant processing capacity at the base facility.

After this first stage, the data would continue to flow to continental Europe to dedicated processing centers. This shall have the advantage that EST infrastructures at the Canary Islands take charge of telescope operations only, while production of science-ready data is in charge of dedicated resources. Another point is that network capabilities within mainland Europe may be still more favorable 10 years from now then the one available between the Canary Islands and the peninsula.

The processing centers shall be responsible for the data reduction and shall provide the EST-VOCDB with VO-compliant, science-ready data. Therefore, they shall be equipped with high performance computers able to handle several 10's of TB produced daily, on average. Then, the science-ready data shall be moved or duplicated from EST data centers to the EST-VOCDB responsible for the long-term data storage and VO-diffusion of EST data. This center shall take also charge of the interoperability with the VOCDB from other facilities, e.g. ATST. This could be achieved with the EST-VOCDB either serving as a duplication center for ATST science-ready data or using virtualization tools for accessing the data stored on the ATST-VOCDB.

In the current design, the long-term data storage utilizes the current standard technology of rotating magnetic disks. These hard disks have undergone consistent technological development and have a large commercial base driving future development. It might be reasonable to expect a factor of about 50 improvement in the hard-disk areal density in the coming decade. This would result in hard disks with upper capacities of 100 TB when EST becomes operational. On the other hand, there are several new data technologies being developed for high-density, persistent data storage. These include 3D optical of Holographic (HVD) storage and Magnetoresistive (non-volatile) RAMs (MRAM), that have started to come into the market (e.g. InPhase Technologies/DSM, Freescale). While these hold promise for future data storage needs, hard disks shall probably remain the storage technology of choice in the timescale of the data storage system design for EST. Magnetic tape recording is another option, though the operational simplicity and immediate data access of systems based on hard disks probably points to their dominant role for data storage at the telescope.

The EST-VOCDB shall run a significant amount of science-ready data available on-line. In order to allow for a comfortable time-response to the users, the data frequently accessed together could be co-located and re-arranged when moving from the processing centers to the EST-VOCDB. Indeed, the data access from users is driven by an object-scientific interest. Therefore, the science-ready data from the processing centers could be organized in storage units based on object classes from science goals.

At the moment, a relational database management system like MySQL seems to fit in with the requirements and reference design of the DHCS sketched above. It is a freeware, widely used in science and commercial databases. At the moment, iRODS is the state-of-the-art software for data storage visualization that would allow for the handling of the various storage units and the remote access to other units of the EST-VOCDB.

4. CONCLUSION

We have summarized the main functional and technical requirements for the definition of the Control and Data Handling systems of EST. We have also presented the technical alternatives identified to date and the current design of both systems. We plan completing the conceptual study of these systems with the selection of the general software architecture, the description of the interfaces between the various sub-systems handled by the CS and DHCS, the analysis of synchronization, alarm, and message requirements, as well as of data reduction and data visualization requirements and of relevant technologies.

ACKNOWLEDGMENTS

The conceptual design study of EST has been partially supported by the European Commission through the collaborative project Nr. 212482 "EST: the large aperture European Solar Telescope".

REFERENCES

[1] Collados, M., Bettonvil, F., Cavaller, L., et al., "European Solar Telescope: Progress status", Astron. Nachr., in press (2010)

[2] Collados, M., Bettonvil, F., Cavaller, L., et al., "European Solar Telescope: Project status", these proceedings (2010)