MACE TELESCOPE

Archival System & Remote Operation

A. Jain, *D. Sarkar, D.D. Sonvane, N. Chouhan,
S. Padmini, K. Rajesh, S. Srivastava, U.K. Gaur & K. Vaibhav

Large amount of data at varied rates is generated by various subsystems of the MACE telescope. The camera and data acquisition system generate maximum data at an estimated rate of 1kHz with an average throughput of ~ 20 MB/sec leading to the generation of ~ 500GB of data per observational night. All of the MACE data needs to be archived and processed for further analysis. MACE Data Archival Storage System (DASS) provides safe and secure storage for MACE data.

MACE DASS comprises off-the shelf commodity hardware and software to manage the storage system. The system is housed in the control room at telescope site. It serves the crucial function of archiving and retrieval of the raw data generated by the camera during observation, telemetry data generated from different subsystems as well as the processed data generated by various data analysis tools.

After data collection, the software provides for data staging, replication, backup and archiving of data at the telescope site for two years. Additionally, a copy of data is transferred to BARC, Mumbai for long term archival. As, it holds all the experiment data, APIs and mechanisms for data retrieval with proper authentication and authorization are also part of MACE-DASS.

Due to its remote location and harsh environment of the telescope site, there was a need for a remote facility for control and operation of the telescope as well for the remote availability of the data generated for detailed analysis later.

Data Archival System - Hardware Perspective

MACE-DASS is designed for having data ingestion at maximum rate of 20MB/sec at full resolution and frequency. With raw data rate of 40GB/hr and 1000 hrs of operations in a year the capacity required is 40TB/year. To retain data for 2 years, the on-site DASS has capacity of 80TB. Older data will be available at long-term permanent storage at BARC, Mumbai.

Design Challenges

A single server with 20 HDDs would have been sufficient for fulfilling the capacity. But the need for making a resilient storage against the adverse site environment raised several challenges. First, on high altitudes, hard-disk drives have higher failure rate. Second, space and power constraints (3kW) at the site, severely restrict the choices in redundancy. Third, due to constrained site accessibility, getting vendor support for spares and routine maintenance was a challenge for proprietary options. Instead, the DASS is developed using standard off-the-shelf commodity hardware, open source software components and in-house developed software for easy maintainability.

Hardware Design & Architecture

Storage system consists of a combination of Solid State Drives (SSDs), Hard Disk Drives (HDDs) and tape drives. SSDs are deployed only at critical stages and Bulk storage is realized using tape-based storage and hard disk-based storage in a redundant manner. For achieving higher fault tolerance, redundancy and diversity has been incorporated in the design.

The system is designed with 2 redundant storage controllers. Each storage controller hosts tape library of 60TB native capacities and hard disks of 40TB aggregate capacity in Storage Expansion Unit (Just Bunch of Disks-JBOD) connected using redundant path. Each JBOD houses 12 drives of 4TB capacities. Two RAID-6 arrays of 5 drives are configured in each JBOD and remaining 2 drives are maintained as hot spares. RAID-6 provides dual drive protection against failure of up to 2 drives in an array. All the Computer Servers have identical configuration. Under normal working conditions the Data Acquisition Server and the Data Analysis server perform their pre-assigned tasks, but if any of them fails the tasks of the failed server can be performed on the other server.

Both the storage Servers export the storage arrays over NFS/iSCSI. In addition, one Storage Expansion Unit, 4 hard disk drives, and a 24-port gigabit Ethernet switch are kept as spares on site.

Software Perspective

MACE telescope generates large volume of data that needs to be archived and analyzed. The event data generated by the telescope undergoes data analysis steps like pre-processing, calibration, image cleaning, parameterization, gamma-hadron separation, energy spectrum formation, light curve generation, and sky map production. Apart from this, telemetry data is collected to carry out data quality checks on the event data and evaluate system health. This analysis chain demands a data storage system capable of handling large data generated at varied rates from different

subsystems of the telescope.

Software Design

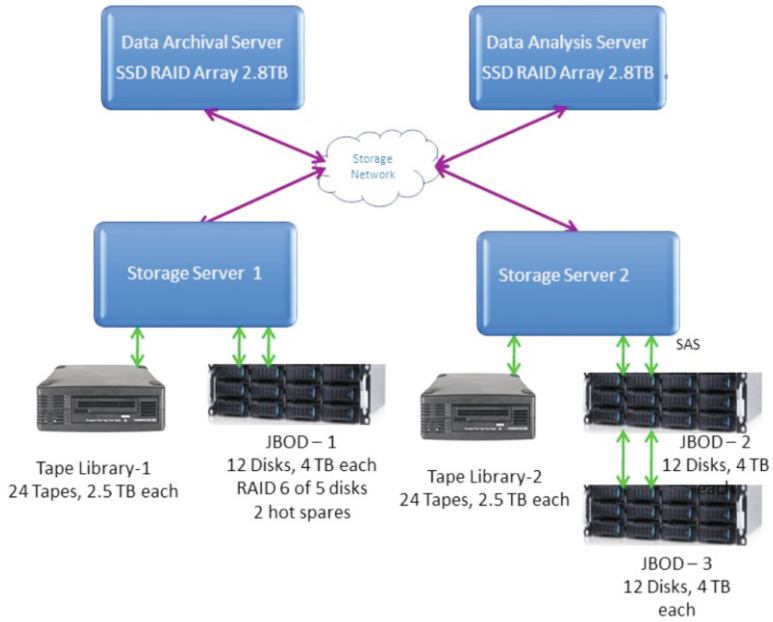
MACE Data Archival Software binds all the diverse hardware systems of DASS and abstracts its complexity from the end-user. The Archival software needs to meet functional requirements like data handling, archiving, meta-data mapping, data staging and backup, online and offline data analysis, and remote backup. To meet these requirements, software for Data Archival System is realized by a set of software – Data Archiving Software (DArS) and Data Staging Software with Space Management Module and Health Monitoring Module.

The Data Archiving Software (DArS) is responsible for data handling, archiving, and meta-data mapping. This DArS software has been designed following a layered architecture, consisting of five layers- the Data layer, the Data access layer, the Communication layer, the Service layer, and the Presentation layer; each layer having its well- defined functionality. The five-layered architecture has been designed with an object-oriented paradigm, with classes been strictly designed with a single responsibility using various design patterns. The interaction between the classes is guided by an event-driven architecture using the Event Listener and Notification scheme based on the Publish-Subscribe pattern. This supports loose coupling among the classes and hence provides scalability. To handle data rates at high speed with full utilization of the memory cores, Data archival software follows a multithreaded approach with each thread being assigned a single dedicated task.

The software architecture of DArS has been designed on an open-closed principle, keeping it loosely coupled for ease of maintenance and re-usability. The DArS software provides following major features:

- TCP based robust communication link with MACE Operator Console (OC) for command interpretation and event snapshot update.
- Dual redundant communication link with Camera.
- Acquisition & Archival of high rate Event Data, and Telemetry from different subsystems.
- File management.
- Archival of Run information and Meta-data mapping.
- Health Monitoring and Error handling.

Data Staging Software has been designed and developed to automate the process of replication of data, space management, health monitoring of storage system and RAID management. The DASS software logically visualizes the storage in 3 tiers: Immediate tier, comprising of SSDs, HDDs pool as Primary tier and Tape storage as Secondary tier. The raw data arriving from the MACE Telescope Camera is first stored on the Immediate tier of the Data Acquisition Server of around 2TB, sufficient to store raw data arriving over a fortnight period.



▲ DATA Archival System.

The data from immediate tier is simultaneously staged to primary and secondary tier of storage. Effectively two copies of data are maintained at the telescope site, one on the tape based archival system and the other on the disk based archival system.

During the staging operation, data is replicated and data integrity on both the peers is verified. A hash-based checksum of all the files transferred are calculated and matched. These hashes are also stored to check the integrity of the data stored in future. It allows seamless movement of data between the three tiers of storage system- Immediate (SSDs), Primary (HDDs) and Secondary (LTO Tapes).

Space Management Module is responsible for cleaning up the Data Acquisition Server data space regularly to accommodate MACE data in the SSD. Clean-up is based on First In First Out (FIFO) mode. It deletes the old data files to create space for new experiment data, ensuring that data is successfully staged to next levels. It ensures that the data redundancy is maintained and enough space is available for the upcoming experiment data. Health monitoring module, provides detailed as well as overall health of the DASS.

This module exploits the h/w specific utilities to accumulate the status information of the components like JBODs, Tape libraries, Storage servers, Network etc. and publishes it to standard format. This can be extended further in case the DAS system is augmented in future by h/w of another vendor. It is also responsible for Regular Check of Disk Space and migration of data and Logging of information and errors of RAID management.

THE camera and data acquisition system generate maximum data @ 1kHz with an average throughput of ~ 20 MB/sec. This leads to the generation of ~500 GB of data per observational night.

Software Implementation & Testing

DArS software has been implemented following an iterative and incremental life cycle model, deployed on the Linux operating system, using C++ as the programming language and following POSIX standards. Three-level functional testing was carried out consisting of unit testing of individual classes followed by integration testing of the packages developed, system testing, and performance testing; followed by system validation by the users.

Remote Operation

The facility of remote operation of the telescope from BARC, Mumbai alleviates the need to work in high altitude and harsh environmental conditions of site. This facility is made available from BARC Mumbai using ANUNET network over satellite link.

ANUNET: Infrastructure

ANUNET is a private wide area network inter-connecting all important units/locations of DAE to provide communication infrastructure for data, voice and video communication. ANUNET has robust four layers connectivity architecture *viz*. National Knowledge Network (NKN) links, Leased Links, Dedicated Fiber and VSAT links. At present, MACE system at Hanle is connected to ANUNET through a VSAT link of 3 Mbps throughput.

Run Control Architecture

The data acquisition and run control software for MACE telescope has been designed considering both local and remote setups.

THE data of events generated by the telescope undergoes data analysis, including pre-processing, calibration, image cleaning, parameterization, gamma-hadron separation, energy spectrum formation, light curve generation and sky map production.

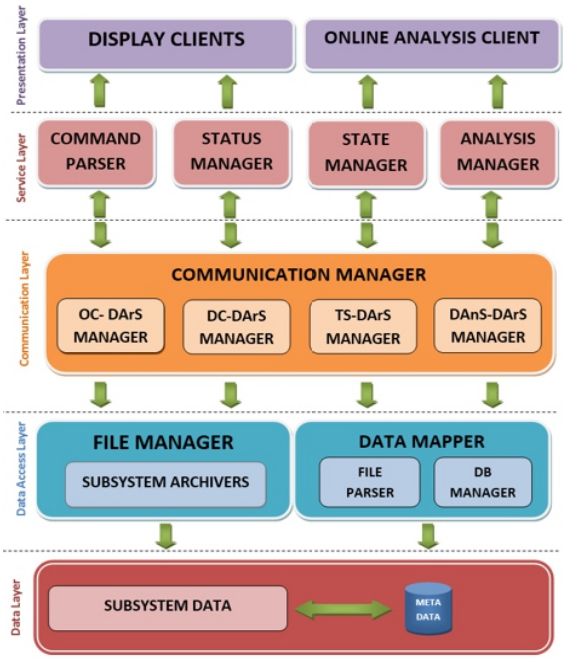
The software for MACE OC has been designed as a client-server architecture which consists of MACE command server, data server and MACE OC client. With Command & Data servers running at Hanle, OC provides both local as well as Remote Run Console Clients. It provides no conceptual difference between software working on site and remotely. MACE OC Client can be run on any of the geographically distributed site with Anunet network.

Lab Setup

For operation of the MACE telescope remotely, multiple labs are setup at BARC, Mumbai. An easy to use Remote Control room has been setup consisting of dedicated high-end workstation with multiple monitors for simultaneous control and monitoring of system parameters. It provides access to only the authorised users, limited by double authorisations, first at the Anunet network level, secondly for MACE OC Client. The authorised users can operate the telescope as per the observation schedule.

Multiple Controller & Observer

MACE OC also provides the feature of multiple OC Clients with one being the controller while rest being the observers. Based on the desired role, Controller and Observers have specified privileges which are granted by a centralized server. The controller will be able to orchestrate the observation run while the status is updated to all the observers. There is also a Master controller, which runs at Hanle site, with special privilege to override any remote controller at any point of time, in case of emergency.



▲ DArS SOFTWARE Architecture.

Data Optimization

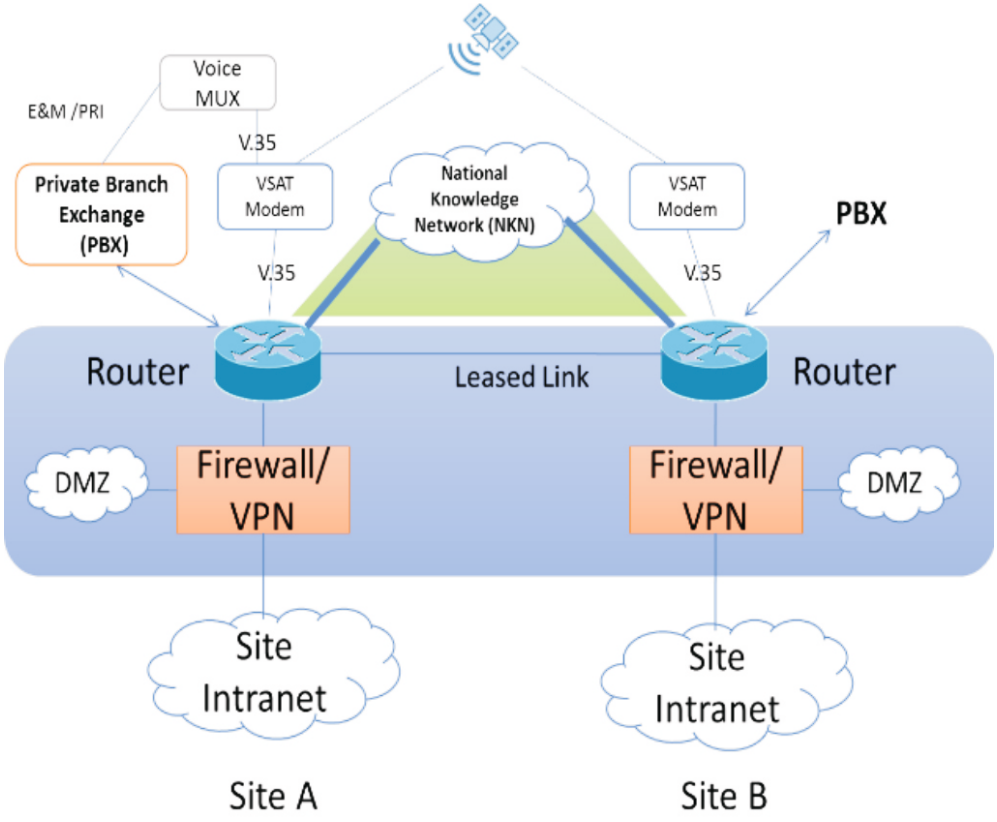
The remote operation must take place over shared satellite network with a limited allotted bandwidth of 256 kbps. To accommodate for network limitations on Anunet, optimization was carried out to make system responsive during remote operation by segregating the desired event and telemetry data to be displayed online and data that can be saved on site and retrieved later for data processing. Large amount of telemetry data is generated from each of the subsystems of MACE for monitoring purpose; the periodicity was reduced for remote setup to meet network bandwidth limitations. In case of Event data, only the integrated charge data was made available remotely while full profile data was available at local setup only. The periodicity is kept adequate enough for monitoring purpose. Meanwhile, complete telemetry and event data is archived locally and transferred offline to be analyzed later.

Remote Software Upgrade

During the maintenance, the software running on different embedded nodes as well run control software may need to be upgraded. Provision has been made to remotely install and upgrade the software at the telescope site over Anunet link by authorised developers.

MACE Data Retrieval

DASS also provides facility for the data retrieval. For in-situ use of data, at Hanle, all the data available in Immediate and Primary tier are available over Network



▲ ANUNET Architecture.

File System (NFS). However, access to this data is available only on the servers which are part of Archival system. For systems, which are external to DASS, data is available for download via MACE-Explorer or at AMBAR-Project share.

AMBAR is a centralized storage facility, using which data can be accessed from any machine on BARC Intranet. AMBAR service provides dedicated storage area for projects having huge storage requirement, in range of TBs. It is best suited for group of users to archive enormous amount of experimental data. MACE data is synched from Data Archival Server (Hanle) to Ambar projects share (BARC) over ANUNET, supporting the feature of Automatic backup and sharing of data among project team members.

The MACE data project share is available as a network drive that can be mapped to a Windows or Linux Desktop.

MACE Explorer is a web-based application developed for querying and retrieving observation event data based on various parameters. Users can search experiment data over tagged fields, like Run number, Source name or coordinates, Observation dates etc. and retrieve associated data files of interest. MACE Explorer provides interface for querying and downloading purposes only. However, it does not store or tag the data or allow any upload.

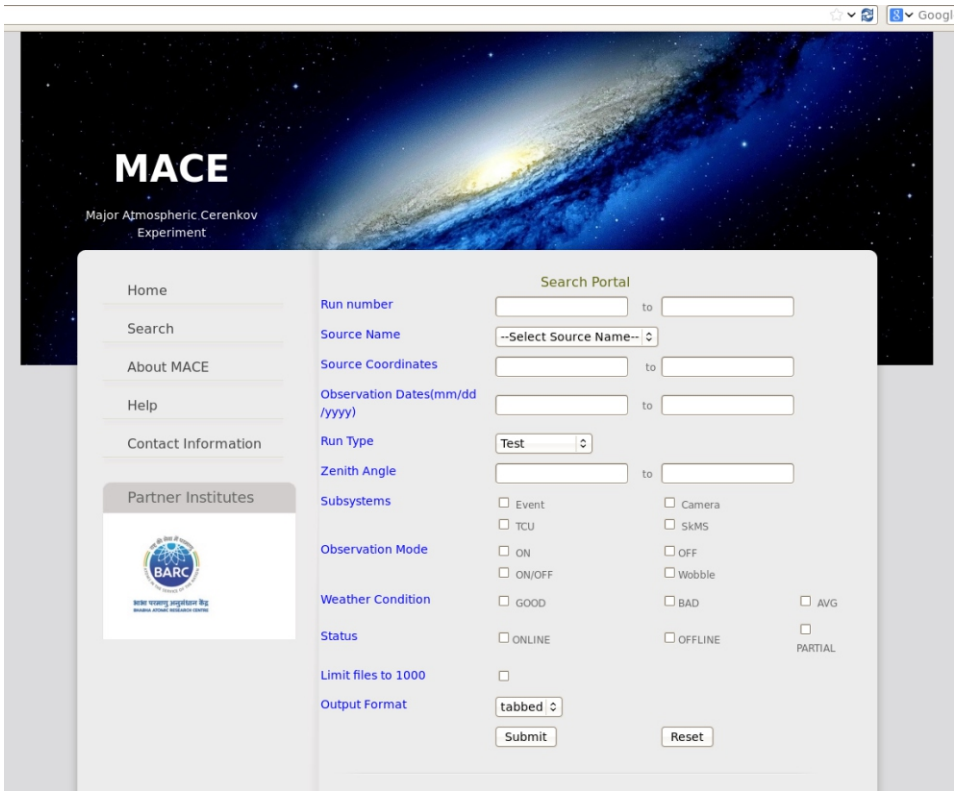
Results

DArS software could successfully handle the required rate of 1kHz. Approximately 840GB data was collected from 2018 to 2019 during the testing phase with 4 CIM Set up at the lab and approximately 1TB of Level '0' data was generated from the test runs at telescope site Hanle during the installation phase of the telescope from 2017 to 2021. DArS software was tested, commissioned and is currently operational at site.

Conclusion

MACE DASS has been successfully deployed at telescope site, Hanle and is operational with the desired performance. The MACE observation data is being archived in redundant manner at Hanle and is synced to BARC, Mumbai for long-term archival. Currently, around 4TB data of interest is available at Ambar for analysis and further processing. Considering the harsh environmental conditions of the site, the Remote Control Room Setups at BARC, Mumbai proved highly helpful to the developers and the users allowing operation, configuration, testing, troubleshooting and upgradation of the system from a more convenient environment during installation and commissioning phase of the telescope. Multiple researchers from different geographical locations can conduct or observe the

USING the remote operation facility, researchers from different locations can conduct their experiments simultaneously.



▲ MACE Explorer - Homepage.

experiment simultaneously. Remote operation of the telescope will be more efficient in near future as a high-speed link between BARC, Mumbai to Hanle of 200 Mbps bandwidth is in process to be commissioned very soon.

Acknowledgment

We are thankful to our colleagues from Electronics Division, Computer Division and Astrophysical Sciences Division for their inputs and suggestions during this work. Sincere thanks to our colleagues working at the telescope site, Hanle who contributed significantly during testing and commissioning phase. Special thanks to Anunet team for providing multi-medium reliable network for Data transfer and control of the telescope. Authors are grateful to Anita Behere, Director, E&IG and K.K. Yadav, Head, ApSD for their continuous support and encouragement. Apart from this, we are grateful to everyone involved in the MACE Project for their kind support.

The authors are from Electronics Division, Astrophysical Sciences Division and Computer Division of Bhabha Atomic Research Centre Trombay, Mumbai-400085.

debangana@barc.gov.in*