Astronomical Data Analysis Software and Systems VIII
ASP Conference Series, Vol. 172, 1999
D. M. Mehringer, R. L. Plante, and D. A. Roberts, eds.

Archive Storage Subsystem for the ESO VLT

Jennifer Dunn, Steve Cockayne, Séverin Gaudet, Shannon Jaeger

National Research Council Canada/Herzberg Institute of Astrophysics, 5071 West Saanich Rd., Victoria, BC, Canada V8X 4M6

Miguel Albrecht

European Southern Observatory, D-85748 Garching bei Munchen, Germany

Abstract. The Archive Storage Subsystem (ASTO), developed by the Canadian Astronomy Data Center, is part of the On-Line Archive Facility in use by the European Southern Observatory (ESO) Very Large Telescopes (VLT) and New Technology Telescope (NTT) Archive System. It is composed of several programs that allow for the preparation and writing of media in parallel. Its function is to respond to requests to create archive and user media on magnetic tape, CD-ROM, or DVD. Parts of this system have also been integrated into the Gemini Telescopes Data Handling System. The design and function of this system are presented.

1. System Overview

The Archive Storage Subsystem (ASTO), developed by the Canadian Astronomy Data Center, is part of the On-Line Archive Facility in use by the European Southern Observatory (ESO) Very Large Telescopes (VLT) and New Technology Telescope (NTT) Archive System. It is composed of several programs that allow for the preparation and writing of media in parallel. Its function is to respond to requests to create archive and user media on magnetic tape, CD-ROM, or DVD. Parts of this system have also been integrated into the Gemini Telescopes Data Handling System. The design and function of this system are presented. The ASTO system supports writing of two types of requests:

Archive data are for long term use and will generally be available for on-line retrieval. Archive data can be grouped together based on attributes such as data type (science data, calibration, etc.). A directory structure is supported for archive requests. Typically, archive data are constantly received and they are stored when there are enough data to fill a media volume (e.g., CD-ROM). The media volume is then ingested into an on-line data system where multiple users can view and/or retrieve the data.

A user, perhaps an astronomer or observer, may have a collection of files that they wish to have written to media. This system provides the user with a means of producing a unit of media (e.g., tape or CD-ROM) containing the data which they can keep. The user request can specify the type of media, the files to write, and if the files are to be compressed. The specification of a directory structure is also supported. A user can request the media to be made at any time, even many months after the observation has been carried out. A GUI has been created but is not discussed here.

Initially data are made available to ASTO by placing them into ASTO's configured magnetic disk area called the Data Stage. Information about the files are also put into ASTO's database. The SYBASE database is also used as persistent store for the status of a request and to provide the ability to rollback actions if a problem occurs.

2. Data Stage Design

Three Data Stage areas are required on magnetic disk:

1. The Archive Data Stage is the magnetic disk area which contains files destined for archive media.
2. The User Data Stage is the magnetic disk area which contains files destined for user media.
3. The Media Data Stage is the disk area used to prepare a media volume for writing.

The Data Stage areas are distinct and can be placed on different disk partitions. This will ensure that, for example, archive data do not consume all disk which would prohibit the preparation and writing of a media volume. The number of media devices that can be simultaneously written is typically constrained by the disk space available in the Media Data Stage and number of devices available for writing.

3. Database Design

ASTO maintains a database which it uses and accesses. The table names are configurable and contain information such as

1. detailed information about files available locally (name, file size, status, compression information, and format),
2. and information about the creation of media (type of request, source of the request, date requested, number of copies, type of media to write, etc.).

4. Media Creation

The various components of the ASTO system manage the writing of media. Each request to write media must proceed through a series of steps called actions. Figure 1 illustrates the actions and rollbacks that are performed when writing media system. It is possible to rollback at any stage prior to the cleanup stage.

The actions are defined as follows:

1. Ingest. This involves the placement of files into the ASTO Data Stage and ingestion of these files into the ASTO storage database. This action is performed either by ASTO or by the VLT On-Line Archive System

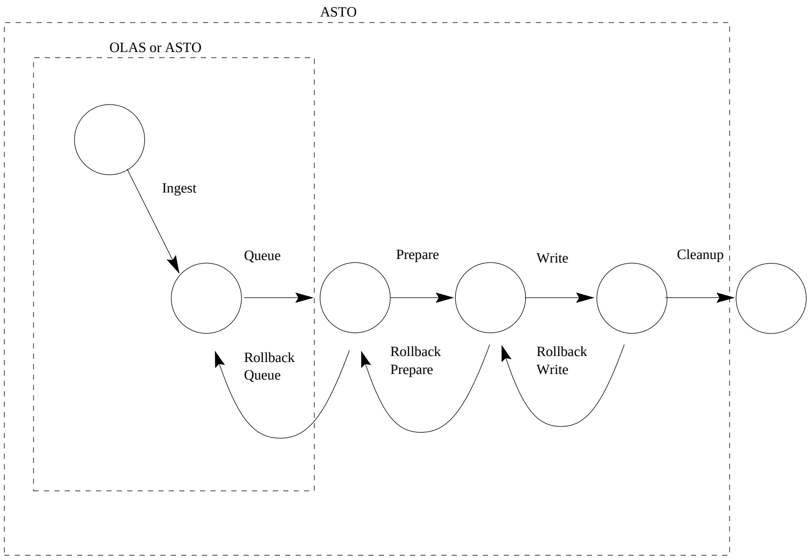

Figure 1. ASTO actions.

(OLAS). At this point, a request will have a name and be assigned a unique identifier.

2. Queue. Files are grouped together which are to be written onto the same physical unit of media, whether the media be CD-ROM, DVD, or tape. For example, a single request may contain enough data to fill three CD-ROMs. Each of the CD-ROM's worth of data is considered to be a unit, and each unit is assigned a unit name that is a combination of the media name and a unit counter. This grouping allows the preparing and writing of the CD-ROMs to be run in parallel, rather than sequentially. An algorithm is used to ensure that, if possible, directories are not split between CD-ROMs. At the completion of each of the following steps, the database is updated to reflect the current state of a request. This action is performed either by ASTO or by OLAS.
3. Prepare. The preparation stage involves whatever is necessary to write the media. For example, the burning of a CD-ROM involves the creation of the physical image before writing. This can be a time consuming process; thus, being able to run this step in parallel with other preparations streamlines this stage. The software which performs the preparation is configurable. For example, a CD-ROM physical image could be created using either the commercially available `Gear`, or the public domain `mkisofs` CD-ROM authoring software.
4. Write. This step performs the actual writing of the medium. For a CD-ROM, this would involve the burning of the prepared physical image from the Data Stage to the physical medium. The script that performs the writing is configurable. For example, `Gear` or `cdrecord` can be used to write the previously prepared physical volume to a CD-ROM. This system supports the writing of multiple copies to multiple devices. A logical device

name, which corresponds to one or more actual devices, is specified. The device is reserved at the start and released at the end of the writing process.

5. Cleanup. This involves the removal of any prepared physical image and any other information created during the preparation stage as well as the removal of the files from the Data Stage. It also releases the reserved device(s).

Once the creation of user media has completed, no further actions need to be taken by ASTO. The media is simply taken by the user for personal use. However, for archive media the media needs to be added to an archive. For example, archive media would then be made available for on-line retrievals. For CD-ROMs, this may involve placement into a CD-ROM jukebox. ASTO includes several executables which perform this task. In general, different executables are used for different types of media.

5. Configuration

ASTO is a highly configurable system. Examples of the information that can be configured are:

1. paths to the different Data Stage areas,
2. the database server name, database name, and database tables names,
3. names of scripts to use for the preparation and writing of a given media type,
4. media unit names, which are a combination of the media name plus a unit counter, which is a number padded with zeros. The default number of zeros can be configured,
5. recognized types of compression,
6. and recognized file formats and their associated extensions.

6. Conclusion

ASTO's ability to perform archive or user media requests has been proven. Parts of this system have been integrated into the Gemini Telescopes Data Handling System. The flexibility of the system allows for many different configurations. Because the actions to produce the media are sufficiently distinct, recovery due to a file error is easily handled. In addition, the ability to write and prepare the media in parallel produces a more time efficient system. A prototype version has been delivered and is in use at the NTT and VLT telescopes as well as ESO headquarters in Garching.

Acknowledgments. To my family, Rodger, Jade and Marina Dunn, for keeping me sane and to the intelligent people I work with.