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

Distribution of the OPUS Data Processing System

Daryl A. Swade, James F. Rose

Computer Sciences Corporation, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218

Abstract. The OPUS data processing software system was developed for converting raw telemetry from the HST into standard FITS format data files. The OPUS process control and monitoring components, which are applicable to many pipeline projects, are being distributed on CD-ROM along with associated documentation.

1. Introduction

The Space Telescope Science Institute (STScI) has developed OPUS as a data processing software system for converting raw telemetry into standard Flexible Image Transport System (FITS) format data files. OPUS is a dynamic, event-driven, distributed processing system that provides an environment designed to handle a large number of observations processed through many steps across a network of computers (Rose 1998). OPUS is also an automated system that monitors processing and provides facilities for error identification and repair.

The Hubble Space Telescope (HST) OPUS pipeline has been operational at STScI since December 1995, and the OPUS baseline system has now been packaged so that other spacecraft missions and observatories can take advantage of this flexible system for data processing projects. OPUS has been recently ported to the UNIX operating system and can currently run on Sparc/Solaris, ALPHA/DIGITAL UNIX, PC/LINUX, VAX/VMS, and ALPHA/VMS platforms. OPUS also supports a pipeline running on any mix of these platforms.

2. OPUS Components

OPUS can be considered as having three components. First, the baseline OPUS system is a distributed processing platform, independent of the applications that it runs and monitors. Second, the Application Programming Interface (API) is a library of software packages that supports the blackboard architecture for the distributed platform as well as for the standard applications. In addition, OPUS consists of a collection of instrument specific pipeline applications that, until recently, have been developed primarily for the instruments aboard the HST. STScI is currently also developing applications for the Far-Ultraviolet Spectroscopic Explorer (FUSE) mission (Rose et al. 1998).

2.1. Baseline System

OPUS has adopted a blackboard architecture using the standard file system of the native operating system for interprocess communication (Rose et al. 1995). In this model processes do not communicate with one another, but simply read and write to a common 'blackboard' instead of having a single 'controller' process which must be continuously aware of the activity of other processes in the system (Nii 1989). This technique effectively decouples the interprocess communication from the individual processes that comprise the data processing system.

Within the OPUS distributed processing system, a variety of independent processes are run sequentially as processing steps in the pipeline. The system also allows multiple instances of any single process to be run simultaneously without interfering with each another. Multiple pipelines, or independent paths, are also supported. Any pipeline can be configured to run its processes across multiple nodes on a network of computers.

In addition to several copies of pipelines with identical processing steps, OPUS supports any number of different pipelines all running on the same network. Thus, in addition to the science pipelines, OPUS can, for example, accommodate an engineering data pipeline.

The OPUS environment is configured through a set of simple ASCII text resource files that describe the command line arguments, pipeline triggers, how steps get triggered, and other control information (Boyer & Choo 1997). The pipeline path file defines a set of cluster-visible directories on the available disks.

In order to monitor the system, OPUS provides two Motif pipeline managers (Rose, Choo, & Rose 1996). The process manager assists with the task of configuring the system and monitors the status of each process. The observation manager views the pipeline activities, monitoring the progress of datasets through the pipeline, and flagging observations that are unable to complete pipeline processing. Multiple managers can each monitor separate pipelines without interference from one another.

OPUS provides facilities for handling data processing problems such as missing data, absent calibration files, or other unexpected situations. The OPUS pipeline provides convenient ways to investigate problems: examine process log files, list data file headers, view observation processing history (trailer) files, and finally restart the troubled exposure at any step in the pipeline.

2.2. API

An OPUS Application Programming Interface is currently under development. See Miller (1999) for details.

2.3. Pipeline Applications

The OPUS applications are programs or scripts that tend to be specific to individual missions or science instruments. Processes (or scripts) can be triggered in three ways: the most common is to allow the completion of one or more previous pipeline steps to act as the process trigger. Another useful technique is to use the existence of a file of a certain class as a trigger. Alternatively, one can use a timing device to trigger an OPUS process (e.g., wake up once an hour).

With OPUS, mission software developers can write specific applications that use the OPUS architecture and API. Experienced STScI OPUS developers are available for consultation and help (opushelp@stsci.edu). For examples of OPUS pipeline applications see Rose (1997), Schultz et al. (1999) and Swam & Swade (1999).

3. Other Missions Using OPUS

While the OPUS system was developed at the STScI, the blackboard system and the OPUS API are independent of the HST mission. The following missions are now using OPUS as their science data processing pipeline system: FUSE (Far Ultraviolet Spectrographic Explorer), INTEGRAL (International Gamma-Ray Astrophysics Laboratory), SIRTF (Space Infra-Red Telescope Facility), AXAF (Advanced X-Ray Astrophysics Facility), and the MSSSO (Mount Stromlo and Siding Spring Observatories) MOSAIC project.

In addition, many groups are considering the OPUS platform for their data processing projects. Over 40 copies of OPUS CD-ROM and FAQ have been distributed to potential users.

4. OPUS CD-ROM

The OPUS baseline system is currently distributed on CD-ROM to help other institutions with their own pipeline management. The OPUS CD-ROM comes complete with the process manager, the observation manager, a set of sample applications, and all the resource files required to get a sample pipeline running.

The sample pipeline is distributed with the OPUS environment to demonstrate some of the major capabilities of the OPUS system. This example simply converts a stack of public GIF images from the HST archives into standard FITS files. In the sample pipeline, the GIF images provide the 'raw telemetry' files that would normally feed a real data reduction pipeline.

More importantly, the sample pipeline is a working illustration of how to put together your own production pipeline. Each of the tasks in the pipeline demonstrates a different variety of 'trigger' that activates the task. For each task there is a separate text resource file showing the variety of switches and parameters available to the system. In addition, the pipeline resource files and path files for the sample pipeline are provided in simple text files that you can both examine and modify.

5. OPUS FAQ

The OPUS CD-ROM and the sample pipeline are fully documented in the OPUS Frequently Asked Questions (FAQ). The FAQ is available on the OPUS CD-ROM or at `http://www.stsci.edu/software/OPUS/opusfaq.html`. This document explains how to install the system on your computers, how to run the sample pipeline, how each of the managers works, and what the different resource files are about. As more projects get experience with the OPUS environment, this document expands with further clarifications, and is revised with each new release of the OPUS CD-ROM.

6. Summary

While the OPUS system was developed at the STScI, the blackboard system and the API are independent of the HST mission. HST mission specific applications are not portable, but the experience of the OPUS team in developing complete pipelines for the HST, FUSE, and other potential missions is available. Packages such as OPUS provide a true resource for NASA and ESA projects in a cost-conscious era where the software development cycle can and should be better controlled. The OPUS platform and the OPUS software libraries can be reused, forming the basis for the rapid development of robust data processing applications.

For further information see the OPUS web page at `http://www.stsci.edu/software/OPUS` or send e-mail to opushelp@stsci.edu.

References

Boyer, C. & Choo, T. H. 1997, in ASP Conf. Ser., Vol. 125, Astronomical Data Analysis Software and Systems VI, ed. G. Hunt & H. E. Payne (San Francisco: ASP), 42

Miller, W. W., III 1999, this volume, 195

Nii, H. P. 1989, in Blackboard Architectures and Applications, ed. V. Jagannathan, R. Dodhiawala, & L. Baum (San Diego: Academic Press), xix

Rose, J. 1998, in ASP Conf. Ser., Vol. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 344

Rose, J., Heller-Boyer, C., Rose, M. A., Swam, M., Miller, W. W., III, Kriss G. A., & Oegerle, W. 1998, in SPIE Proc., Vol. 3349, Observatory Operations to Optimize Scientific Return, (Bellingham: SPIE), 410

Rose, J., et al. 1995, in ASP Conf. Ser., Vol. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 429

Rose, J. F. 1997, in ASP Conf. Ser., Vol. 125, Astronomical Data Analysis Software and Systems VI, ed. G. Hunt & H. E. Payne (San Francisco: ASP), 38

Rose, J. F., Choo, T. H., & Rose M. A. 1996, in ASP Conf. Ser., Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes (San Francisco: ASP), 311

Schultz, J. J., Goldstein, P., Hyde, P., Rose, M. A., Steuerman, K., Baum, J., Perrine, R., & Swade, D. A. 1999, this volume, 199

Swam, M. S. & Swade, D. A. 1999, this volume, 203