Solving the Software Legacy Problem with RISA

Remote Interface to Science Analysis Software: XMM-Newton SAS + EXOSAT interactive analysis

A. Ibarra and C. Gabriel
XMM-Newton SOC, European Space Astronomy Centre (ESAC) / ESA, Madrid, Spain

ABSTRACT

Nowadays hardware and system infrastructure evolve on time scales much shorter than the typical duration of space astronomy missions. Data processing software capabilities have to evolve to preserve the scientific return during the entire experiment life time. Software preservation is a key issue that has to be tackled before the end of the project to keep the data usable over many years.

We show RISA (Remote Interface to Science Analysis) as a solution to decouple data processing software and infrastructure life-cycles. RISA is based on the client/server paradigm (using axis and tomcat server to deploy the application) and it has been developed to be used in GRID environments, but it can be easily adapted to any other system architecture such as cluster or cloud computing. The system allows the user to search for any XMM-Newton data (pointing or slew observation) using SIAP protocols and name resolver services. The application has been designed as a mission independent analysis tool, being possible to implement tasks and workflows from different missions.

RISA architecture wraps the existing code, encapsulating the software in a virtual machine. The Client contains all the functionality that will enable the access to the data. The Server contains the logic to process the requests and send the job information to the final execution nodes. Finally, the Data layer is in charge of gathering the data from the archives and to send it to the execution nodes.

We will also present the first studies to reactivate the data processing software of the EXOSAT mission, the first ESA X-ray astronomy mission launched in 1983, using the generic RISA approach.

RISA Handling XMM-Newton SAS + EXOSAT IA Data Processing Tools

User Interface

XMM-Newton & EXOSAT

- Multiple searching options.
- Vizier query to resolve names.
- SIAP query to retrieve XMM-Newton and EXOSAT information.
- Load previously saved sessions.

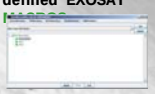
XML Parser

XMM-Newton

- GUI converter of SAS Tasks and XML Parameter file.
- Select your Tasks from the SAS Tasks window.
- Configure your parameter values.
- Pre-defined SAS threads.

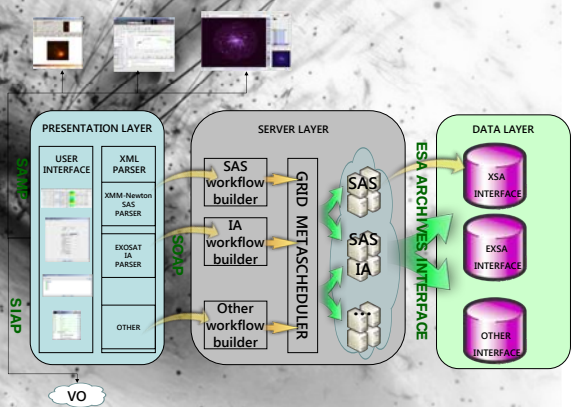
EXOSAT

- Parse EXOSAT MACRO files.
- GUI converter of EXOSAT MACROS
- Select the pre-defined EXOSAT



EXOSAT MACROS are recipes that contains much of the knowledge and expertise concerning the extraction of scientific information

RISA ARCHITECURAL DESIGN



VO Interoperability I

XMM-Newton & EXOSAT

- RISA makes use of VO technologies such as SAMP
- SAMP allows RISA to send and receive information from other applications such as VoSpec, ds9 or Aladin.

Workflow Builder

XMM-Newton & EXOSAT

- The Server processes the request.
- Creates the workflow corresponding to each mission
- The Metascheduler creates job templates and take care of the job submission and execution.

Data Processing

XMM-Newton & EXOSAT

- Each software is built and stored in a Virtual Machine infrastructure.
- Job submission automatically detect the execution nodes available for each mission.
- Each execution node has direct access to the Archive Interfaces.

VO Interoperability I I

XMM-Newton & EXOSAT

- The system is able to work with the results using VO apps. making use of the SAMP VO protocol and retrieving products.

XMM-Newton SAS within RISA

XMM-Newton Science Analysis Software

HISTORY

- SAS is the package used for interactive and pipeline data reduction of all XMM-Newton instruments.
- Highly portable, optimizing its usefulness to a large scientific community.
- Written and maintained by a large group of scientific software developers and instrument calibrations specialists.
- XMM-Newton data will not be superseded by new or better observations for many years and their full scientific exploitation could require repeatedly more refined analysis compared to that stored in the archive during the next decade.

Software Preservation

Enable Continued Access to Data

Trying to make your software project portable to multiple platforms will help to make your code sustainable in the long term.

Several different scenarios can be identified to enable continued access to data such us:

- Repeating and verifying research results (using the same or similar setup)
- Reanalyzing data in light of new theories.
- Reusing data in combination with future data.
- Identifying new use cases from new questions.
- Verifying data integrity.
- "Squeezing" additional value from data.

Each set of data is unique: once lost means lost forever

Some of the benefits of working thinking in software preservations are:

- New insight and knowledge.
- Improve research outcomes.
- Greater return on investment

The most important reasons for preservation are the ones we do not see now

EXOSAT Interactive Analysis Test Case

EXOSAT Interactive Analysis

HISTORY

- EXOSAT IA was developed by specialized scientist in the 80's.
- Developed to run on HP1000 computer running RTE operating system.
- The use of 16-bit machine and the limited capabilities of the operating system did not make for easily portable software.
- In 1990 ESA decided to port the IA to UNIX operating systems transcribed all 8340 Telemetry tapes onto optical disks.
- HEASARC converted the EXOSAT scientific products to FITS.

PORTING

- Recently, it was decided to port IA software to Linux and Mac OSX platforms.
- Encapsulate EXOSAT IA software in a Virtual Machine, to preserve the source code, compiler and third party libraries.
- Convert the existing EXOSAT IA MACROS to XML files.
- Plug-in the EXOSAT XML Parser into RISA Presentation Layer.
- Plug-in the IA workflow builder into RISA Server Layer.
- Analyze EXOSAT data using RISA!!!

The legacy of this original platform remains even today.

SUMMARY: Software preservation is a key issue in projects where data are the final product of a project. Astronomy missions, where data have a intrinsic value, need to preserve access to the data and in most cases the data may only be used if the the software is also kept alive.

In this paper, we have demonstrated that RISA architecture can be used to preserve the data access by enabling the sustainability of the software through virtual machines, and enabling software access trough web services.

Although software preservation is difficult to consider during the operational phase of a projects, while the software is in use, we need to be aware that the most important reasons for preservation are the ones we do not see now.

Contact

Aitor Ibarra Ibaibarriaga (aibarra@sciops.esa.int)
Carlos Gabriel (cgabriel@sciops.esa.int)

Astronomical Data Analysis Software and Systems XXI. Paris, France. Nov 6-10, 2011