

AVES: Cluster Computer system for INTEGRAL Scientific Analysis

Memmo Federici¹

INAF-IASF: Istituto di Astrofisica Spaziale e Fisica Cosmica Roma
Via Fosso del Cavaliere 100, 00133 Roma, Italy
E-mail: memmo.federici@iasf-roma.inaf.it

Bruno Luigi Martino

CNR-IASI: Istituto di Analisi dei Sistemi ed Informatica
Viale Manzoni 30 00185 Roma, Italy
E-mail: bruno.martino@iasi.cnr.it

Lorenzo Natalucci

INAF-IASF: Istituto di Astrofisica Spaziale e Fisica Cosmica Roma
Via Fosso del Cavaliere 100, 00133 Roma, Italy
E-mail: lorenzo.natalucci@iasf-roma.inaf.it

The AVES computing system, based on an “Cluster” architecture is a fully integrated, low cost computing facility dedicated to the archiving and analysis of INTEGRAL data. AVES is a modular system managed by a software resource manager (SLURM) that allows unlimited expandability (65,536 nodes and hundreds of thousands of processors); actually is composed by 30 Personal Computers with Quad-Cores CPU able to reach the computing power of 300 Giga Flops (300×10^9 Floating point Operations Per Second), with 120 GB of RAM and 7.5 Tera Bytes (TB) of storage memory in MHDFS configuration. AVES was designed and built to solve growing problems raised from the analysis of the big data amount accumulated by INTEGRAL mission (actually about 8 TB) and due to increase every year. To this purpose we have developed a software interface able to split the required analysis process in N processes automatically sent to N cores. Thus the whole computing time, compared to that needed by a Personal Computer with single processor, has been enhanced up to a factor 70.

The Extreme sky: Sampling the Universe above 10 keV - extremesky2009
Otranto (Lecce) Italy
October 13-15, 2009

¹ Speaker

Introduction

The INTEGRAL mission, [1] operational since 2002, has already generated an archive of pre-processed data of about 8 TB size up to this summer. Given the current data rate it is foreseen that the archive storage needs to be incremented by about 1.5 TB/year, to be supported at least until 2012 following the approved mission extension. In the meanwhile, more scientific analysis applications require to span over time period of years with thousands of Science Windows (SCW is a single pointing lasting 2000-3000 sec as part of a dither pattern for basic observation) to be performed in a single run. On basic, conventional PC systems this results in an overburden especially in terms of computing time that can be of the order of months. We have then approached a strategy aimed at increasing of computing power while keeping the system costs into a relatively low range. The developed AVES system allows to pull down the required computation time for the analysis from a factor 40 to 70, resulting in a strong increment of the rate of scientific production. Now that 7 years of scientific observations have been acquired.

1.1.Overview

The drivers considered for the design of AVES were (chosen by) low cost and high performance computing. In order to achieve the first goal it was decided to use only conventional and readily available components. The mechanical structure that houses the individual components of the cluster is composed of a common shelving, very robust and yet extremely economical. This is suitable to accommodate up to 60 units/nodes (240 CPU).

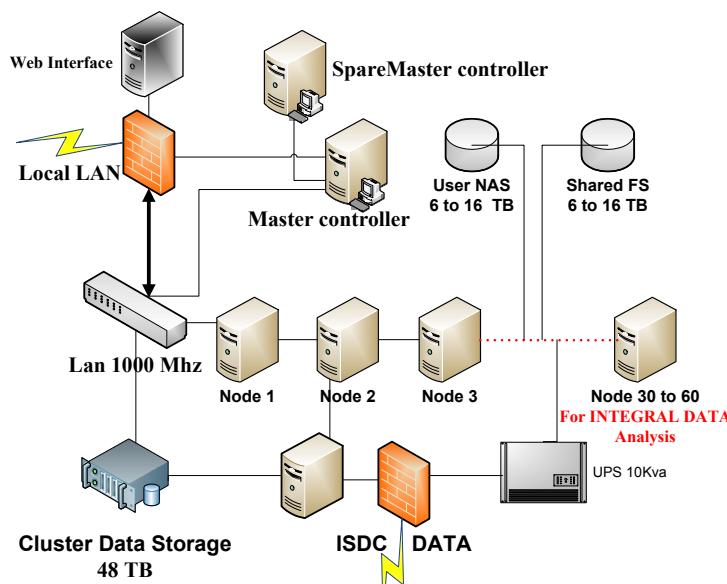


Figure 1: block diagram of the hardware structure of AVES

It is housed in a cool environment that guarantees the correct working temperature. The individual nodes that make up the cluster are of a commercial personal computer, then low-cost hardware with good quality, but conventional. The diagram (see Figure 1) shows the current configuration of AVES. It is at present composed of 30 nodes, each equipped with Intel Quad Core processor, 4 GB of DDR2 RAM with access to 1066 Mhz, and 250 GB of Hard Disk. Overall AVES develops,

with its 120 GB of RAM and 120 CPU computing power, close to 300 Giga Flops (300×10^9)

Floating point Operations Per Second). The cluster is managed by two Master units in redundant configuration, installed on both the resource manager SLURM (Simple Linux Utility for Resource Management) [2]. The INTEGRAL data reside on a server with NAS DATA-RAID-6 storage capacity of 48 TB and with redundancy built by an identical unit. From DATA-STORAGE the data are copied to an internal structure made by the Cluster Technology UFS (Union File System) that allows the merger and use free space on hard disk with a strong increase in the speed of scientific analysis. A further increase of the speed of analysis is determined by data handshaking provided through the interconnection network of intra-cluster full-duplex with a speed of 1Gb/s. The result of scientific analysis are then stored on a NAS RAID-5 of the current capacity of 6TB. Access to AVES is via a secure connection in the ssh protocol which allows users to export graphical windows. A system of user access through a web interface is currently under study. AVES has an individually power supply (UPS) controlled via WEB. The power of 10kVA can guarantee, in the absence of the main power line, an endurance of several hours for a load of up to 60 nodes.

2. Software strategy

AVES is a computing system designed to perform only the scientific analysis of data obtained from instrumentation provided with the INTEGRAL satellite. These scientific tests are carried out using the package developed from ISDC: the Off-Line Scientific Analysis (OSA) [3]. The OSA architecture does not allow easily splitting a single, time consuming process into multiple runs running in parallel. However, one can subdivide the analysis into multiple instances of calculation well ordered in time, with the result of splitting the computing load into the different units in a controlled way. To efficiently use the features of the calculation in clustering and to facilitate their use by all users a series of graphic interfaces have been produced that implement the following functions:

- Allocation of resources
- Run of specific tools
- Synchronization instances subjected to the calculation queue

2.1 Software Flow Charts

The AVES handling software is composed of about 50 scripts making use of the bash shell language. Actually it is possible to group the different tasks performed by these programs into 4 procedures which are described below by

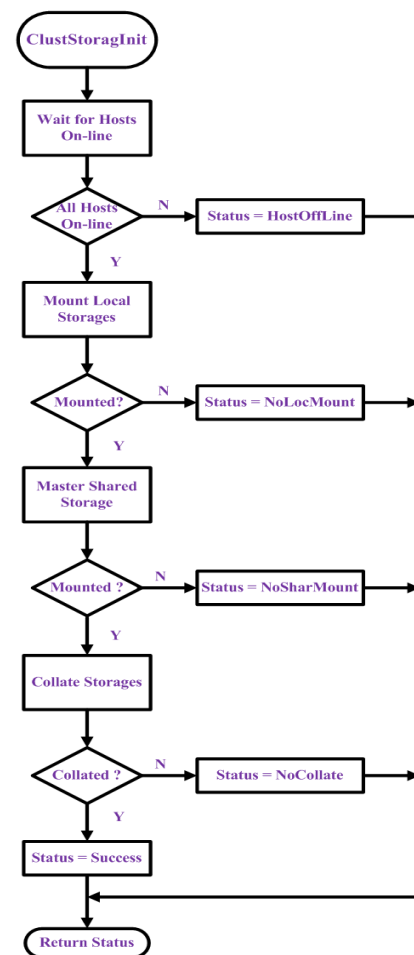


Figure 2: ClustStorageInit procedure

means of flow diagrams (see Figures 2,3,4,5)

2.1.1 ClustStorageInit

The ClustStorageInit procedure (see Figure 2), allows to create a shared UFS (Union File System) in which is available the free storage space usable as single folder. Furthermore it is able to dynamically add in the folder the free space from the active nodes of the cluster. Then, in case of success, enable the essential services for the communication between nodes and the resource manager SLURM.

2.1.2 Aves Login

This procedure (see Figures 3) manages the access of users, the setting of the needed environment variables and provides information about the available software and objects already created or analyzed. Then it activates the graphical interfaces (par. 2.1.5) needed to manage the entry of parameters.

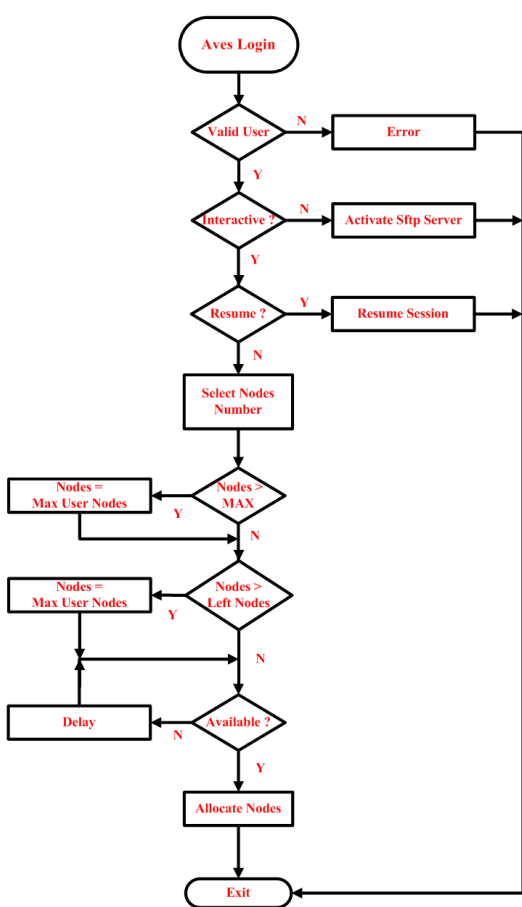


Figure 3: Aves Login procedure

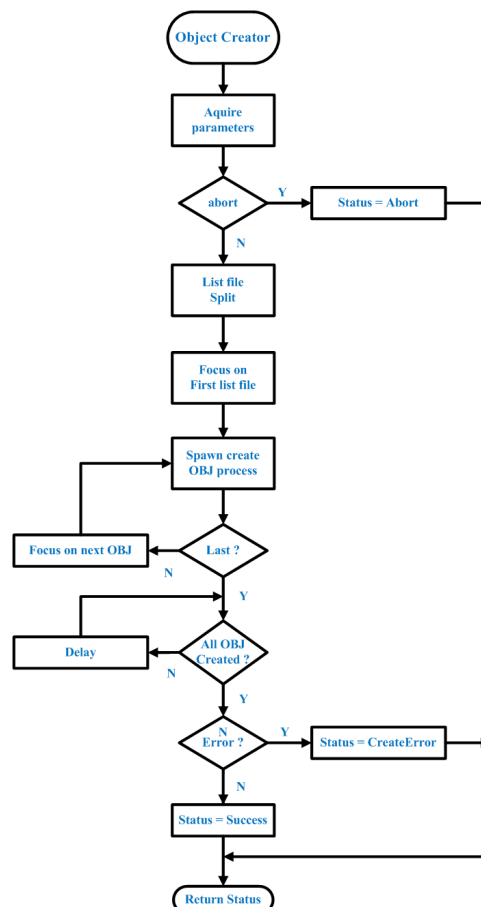


Figure 4: Object Creator procedure

2.1.3 Object Creator

This procedure (see Figure 4) splits the SCW file list into partitions that match the number of nodes allocated to the user and creates objects for further analysis. Current parameters are stored in profiles files ready to be used or edited during the next analysis steps.

2.1.4 Science Analysis

This procedure (see Figure 5) reads the collection of file lists prepared by the previous step and provides the scientific software with the string of chosen parameters. It is possible to save or retrieve the entire set of parameters for future activities.

2.1.5 Graphical User Interface (GUI)

The use of AVES is facilitated by a series of graphical interfaces (see Figure 6) that allow an easy selection of the parameters for the analysis and the saving of the used settings for future reuse. It is also scheduled, for the next software release, to implement a system of quality control for the data stored in the database. This will allow a further decrease in the overall time of analysis, and a desired improvement in the quality of scientific results.

Besides the tasks/procedures described in the previous sections, the AVES software contains utilities allowing performing statistics on the computing load associated to each user, including the used disk storage. This allows optimizing the storage areas of the scientific products that necessitate of important memory resources.

Furthermore within this project, a dedicated system for the reading/writing of intermediate/final data products has been finalized. Even though using the RAID-5 technology to guarantee the stability of the data, this system is capable of maintaining high speed performance. Thanks to this procedures the overall computation time on a big analysis run has been improved by a factor ~3.

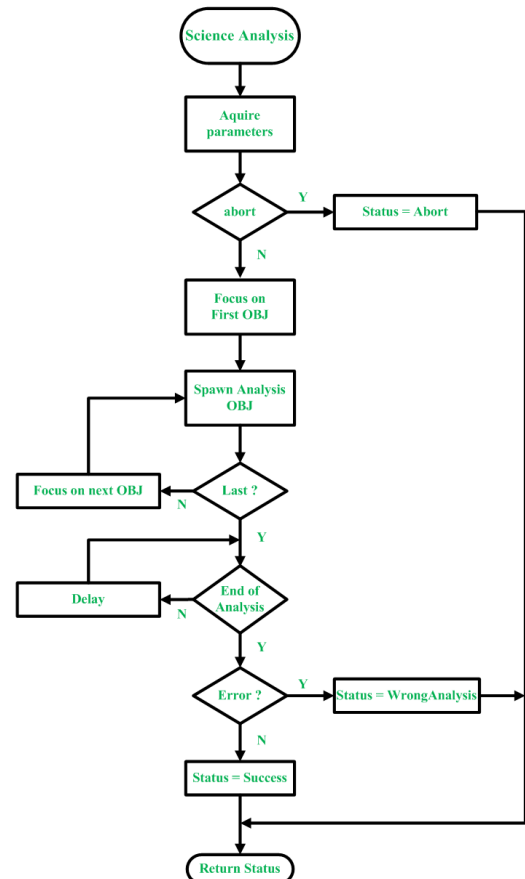


Figure 5: Science Analysis procedure

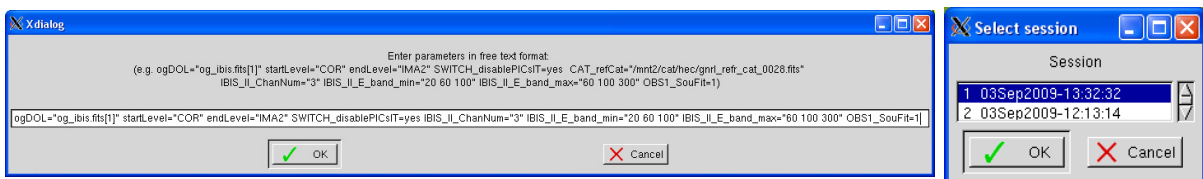


Figure 6: Two examples of GUI of AVES

POS (extremesky2009) 092

3. Scientific Analysis Example

Using AVES we were able to create, from raw (pre-processed) data files, higher level science products such as light curves, images and spectra for each Science windows using the standard OSA software. In Figure 7 it is shown the total spectrum in the 3–200 keV energy band, performed with IBIS [4] and JEM-X [5] instruments (exposure of 20 Ks), and a mosaic of images in 20-40 keV energy band of the Low mass X-ray Binary 4U 1705-440.

This timing and spectral analysis, performed with a single core PC, would require a time of the order of few months. Only a few days are needed to carry out this job with AVES.

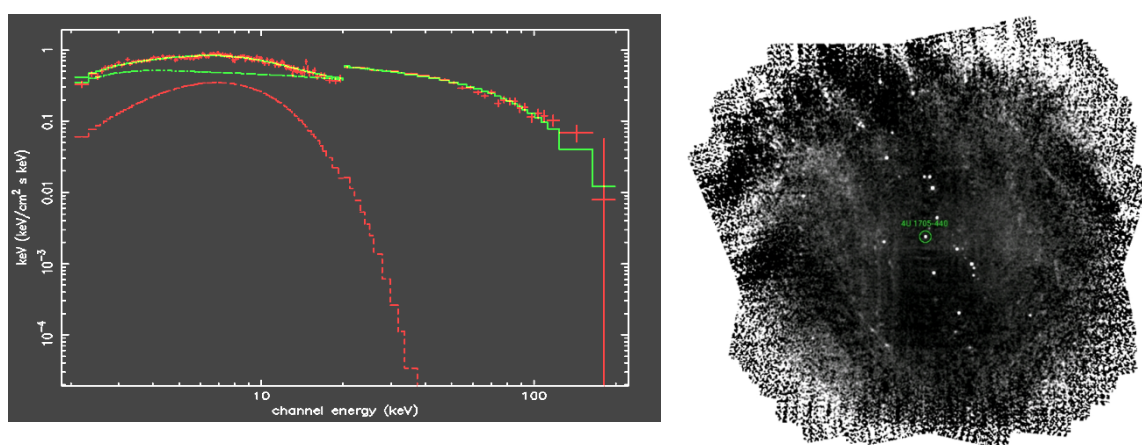


Figure 7: The spectrum and the mosaic of images (right) of the Low mass X-ray Binary 4U 1705-440, obtained with AVES.

References

- [1] C. Winkler et al., *Astronomy and Astrophysics* 411, L1, 2003
- [2] A. Yoo, M. Jette, and M. Grondona, *Job Scheduling Strategies for Parallel Processing*, volume 2862 of *Lecture Notes in Computer Science*, pages 44-60, Springer-Verlag, 2003.
- [3] A. Goldwurm et al., *Astronomy and Astrophysics* 411, L223, 2003
- [4] P. Ubertini et al, *Astronomy and Astrophysics*, 411, L131, 2003
- [5] N. Lund et al, *Astronomy and Astrophysics*, 411, L231, 2003