The ALMA Front-End Archive Setup and Performance

A. Wicenec[1], P. Burgos[2], A.Checcucci[1], A. Chen[1], B. Jeram[1], H. Meuss[1], A. Persson[1], R. Cirami[3]

[1] European Southern Observatory, Garching, Germany
[2] Joint Alma Office, Chile,
[3] INAF-OAT, Italy

ALMA SCIENCE ARCHIVE

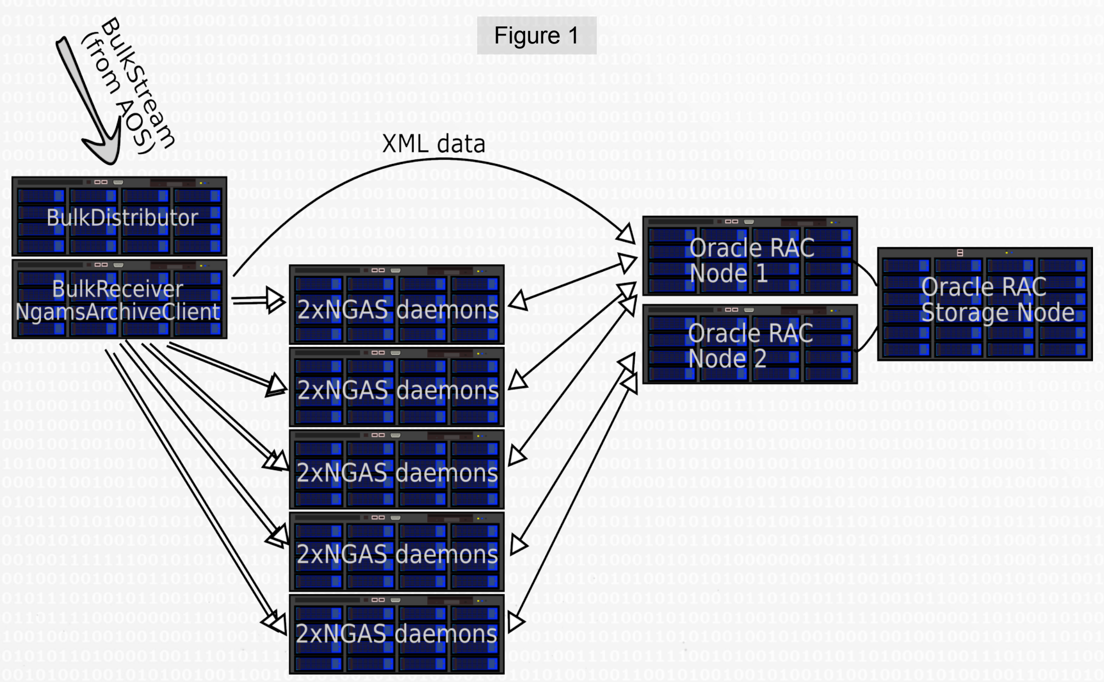
Abstract

The ALMA front-end archive system has to capture up to 65 MB/s for a period of several days plus the data of about 150,000 monitor points from all 66 antennas and the correlators. The main science data is delivered through Corba based audio/video streams and finally stored on SATA disk arrays hosted on 5 computers and controlled by 10 daemons. All data is collected by software components running on computers in the antennas and then sent through dedicated fiber links to the Array Operations Site (AOS) at 5000m and from there to the Operations Support Facility (OSF) at 3000m elevation. The various hardware and software components have been tuned and tested to be able to meet the performance requirements. This paper describes the setup and the various components in more detail and gives results of various test runs.

Science Data Archiving

Most of the ALMA science data will be produced by two independent correlators, one for the main array of 50 antennas and another one for the compact array (12 x 7m + 4 x 12m antennas). The data rate from the correlators is limited to 65 MB/s in total, but will only be reached for special projects, the average data rate is limited to 6.5 MB/s over one whole year, resulting in a total science data rate of about 200 TB/year. ALMA can observe 24 hours a day and thus there are no 'natural' time intervals where any data transfer, archiving or processing could catch-up with the observations. Since the full data rate of 65 MB/s could be delivered for several hours of even days the data capturing system has to be able to sustain this data rate as well. Given the requirement of a 24/7 operation the data capturing system has to be both highly reliable and it also needs to allow down-times of parts of the system without side effects on normal ALMA operations. In order to achieve these goals we have executed a long series of tests trying to optimize the deployment, configurations and inter-connects of the various parts of this vital part of the ALMA data flow. The deployment of the final system is depicted in figure 1 and shows a set of 7 machines, where one is the distributor, a second machine is running the archive receiver, which is placing the received data on a small solid state disk RAID array. From there it is picked up by a daemon, which is sending the data files using the HTTP protocol in a round robin fashion to 5 data storage nodes running a total of 10 NGAS [1] daemons. In addition the whole system needs the Oracle database in order to record the scientific meta-data and file meta data of the NGAS. Oracle is running as a load-balancing, fail-over Real Application Cluster (RAC)[2] on three machines.

Figure 1



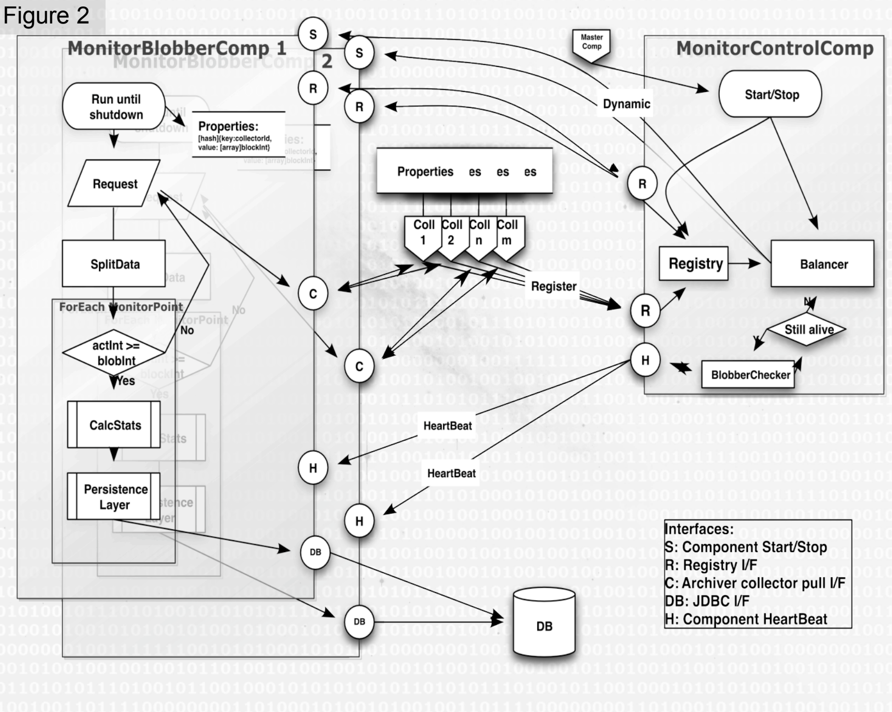
Performance

The backbone of the data transfer between the AOS and the OSF is based on an implementation of Corba[3] A/V streaming in TAO[4]. The final test had to prove that the system is able to cope with the requirements to store up to 18 files/second and up to 65 MB/second (whatever is first). In order to make this test as realistic as possible we have used the machines, which had in the meantime been sent to Chile for the final installation at the OSF. The results are summarized in the table below.

Number of files	Duration[days h:m:s]	Total volume [GB]	Average transfer rate [MB/s]	Average file size [MB]	Max. burst transfer rate [MB/s]
61387	5 21:00:20	3230.68	6.52	53.89	235

The maximum sustained data rate was measured to be 95.6 MB/s and the maximum frequency of small files archived was 36 Hz. These numbers are all well above the requirements of the fully deployed ALMA observatory.

Figure 2



Monitor data archiving

Archiving the monitor data poses a very different challenge on the data flow and in particular the database. There are many tens of thousands sensors installed on the hardware and data is collected at a frequency of up to 62.5 Hz. The figure on the left shows the conceptual design of the monitor archiver module. It consists of a controller component and one or multiple so-called blobber components. Designing a system, that would be able to store every single data point in a database row was regarded to be too expensive compared to storing the main science data. Thus in particular the high sampling rate monitor points are stored in binary large objects (BLOBs) and in addition for each of these BLOBs there are statistical properties calculated and stored in the same row. The 'integration' time and thus the size of the BLOBs is configurable, but will typically be of the order of 1 minute. In order to be able to scale the system up to the expected 150,000 monitor points, the number and deployment of the blobber components is also adjustable. The hardware related monitoring code and also the database properties population is supported by code generation technology, else it would be almost impossible to deal with the many thousand devices and the database initialization.

REFERENCES and further information:

[1] J. Knudstrup, A. Wicenec, and S. Johnston. NG/AMS: The Next Generation Archive Management System. In D. A. Bohlender, D. Durand, and T. H. Handley, editors, ADASS XI, volume 281 of ASP Conference Series, pages 181-+, 2002
[2] Oracle RAC: http://www.oracle.com/database/rac_home.html and http://en.wikipedia.org/wiki/Oracle_RAC
[3] CORBA: http://www.corba.org/ and http://en.wikipedia.org/wiki/Common_Object_Request_Broker_Architecture
[4] The ACE ORB (TAO): http://www.theaceorb.com/

Main ALMA web site: http://www.almaobservatory.org
ALMA Archive twiki: http://almasw.hq.eso.org/almasw/bin/view/Archive/WebHome

Atacama Large Millimeter/Submillimeter Array