Atacama Large Millimeter Array

Technical Specifications for the Initial Purchase of Linux servers for the ALMA Archive NGAS system and the ALMA Archive relational databases

COMP-70.02.00.10-001-A-RFP

Version: A
Status: Released

2007-10-31

Prepared By:	**Organization**	**Date**
Beate Marx	European Southern Observatory	2007-10-31
IPT Leader Approvals:	**Organization**	**Date**
G. Raffi Gianni Raffi Digitally signed by Gianni Raffi DN: cn=Gianni Raffi, c=DE, o=ESO, ou=ALMA Date: 2007.11.05 16:55:42 +01'00' B. Glendenning Brian Glendenning Digitally signed by Brian Glendenning DN: cn=Brian Glendenning, c=US, o=ALMA/NRAO, ou=Computing IPT, email=bglenden@nrao.edu Reason: I am approving this document Date: 2007.11.05 09:22:34 -07'00'	European Southern Observatory National Radio Astronomy Observatory	
System Engineering Approvals:	**Organization**	**Date**
ALMA Project Manager Approval:	**Organization**	**Date**
T. Beasley Tony Beasley Digitally signed by Tony Beasley DN: cn=Tony Beasley, c=US, o=ALMA, ou=JAO, email=tbeasley@nrao.edu Date: 2007.11.05 12:02:03 -03'00'	National Radio Astronomy Observatory	
Release Authorization:	**Organization**	**Date**
Hans H. Rykaczewski Hans Rykaczewski Digitally signed by Hans Rykaczewski DN: cn=Hans Rykaczewski, c=DE, o=ESO, ou=ALMA Division, email=hrykacze@eso.org Date: 2007.11.06 12:12:37 +01'00'	European Southern Observatory	

Change Record

Version	Date	Affected Section(s)	Change Request #	Reason/Initiation/Remarks
A	2007-31-10	All		First Issue

Table of Contents

ACRONYMS & DEFINITIONS

ALMA	Atacama Large Millimeter/Submillimeter Array
ALMA AOS	Array Operations Site (Llano de Chajnantor)
ALMA OSF	Operations Support Facility (San Pedro, Chile)
ALMA SCO	Santiago Central Office (Santiago, Chile)
ARC	ALMA Regional Center
ATF	ALMA Test Facility (Socorro, U.S.)
CPU	Central Processing Unit A quad-core CPU is a chip-level device containing four independent microprocessors
ESO	European Organisation for Astronomical Research in the Southern Hemisphere
NAOJ	National Astronomical Observatory of Japan
NRAO	National Radio Astronomy Observatory
NGAS	Next Generation Archive System
Oracle RAC	Oracle Real Application Cluster
RAID	Redundant Array of Independent/Inexpensive Disks
STE	Standard Test Environment
XML	Extensible Markup Language

1 Purpose and Scope

This document provides the technical specification for the purchase of hardware for the ALMA Archive system, namely:

- Linux servers for the NGAS system (bulk storage),
- Linux servers for the relational databases (XML metadata, TMC, miscellaneous science support),
- Additional servers for other purposes, and
- Network components and computer room space (racks/floor space).

These servers will be located at the various ALMA sites, namely:

- at the various production environments located at the Operations Support Facility (OSF) in San Pedro de Atacama, the Santiago Central Office (SCO) in Santiago de Chile, and at the ALMA Regional Centers (ARC) in Charlottesville (U.S.), Mitaka (Japan), and Garching (Germany),
- at various Standard Test Environments (STE) in Charlottesville, Socorro and at the ATF, in Mitaka and in Garching,
- at a development environment in Garching.

No ALMA Archive servers will be located at the high Array Operations Site (AOS).

2 Overview

ESO is the intergovernmental European Organisation for Astronomical Research in the Southern Hemisphere. On behalf of its thirteen member states ESO operates a suite of the world's most advanced ground-based astronomical telescopes located at the La Silla Paranal Observatory in the Atacama desert in Chile. The ESO Headquarters are situated in Garching near Munich, Germany.
The Atacama Large Millimeter/Submillimeter Array (ALMA) project is a partnership between Europe, Japan and North America in cooperation with the Republic of Chile. ALMA is funded in Europe by ESO, in Japan by the National Institutes of Natural Sciences in cooperation with the Academia Sinica in Taiwan and in North America by the U.S. National Science Foundation in cooperation with the National Research Council of Canada. ALMA construction and operations are led on behalf of Europe by ESO, on behalf of Japan by the National Astronomical Observatory of Japan (NAOJ) and on behalf of North America by the National Radio Astronomy Observatory (NRAO), which is managed by Associated Universities, Inc.

ALMA is located on the high-altitude Llano de Chajnantor (5000 m elevation), east of the village of San Pedro de Atacama in Chile. ALMA's location in the Atacama Desert is one of the highest, driest places on Earth, making it ideal for astronomical research at millimeter/submillimeter wavelengths, which are absorbed by atmospheric moisture. When completed (in 2011), ALMA will be the largest and most capable imaging array of millimeter/submillimeter telescopes in the world.

3 Planned ALMA Archive Environment (production)

The purpose of the ALMA archiving subsystem is to provide services for persistent archiving and retrieval of observational data, observation descriptors, images and technical and environmental data (configuration, monitoring, and logging). A variety of different types of data needs to be supported, which can be grouped into the following categories:

- Bulk data: Observational data and images.
- Metadata: All kinds of data entities describing observations and the bulk data. These kinds of data are formatted in XML.
- TMC data: Configuration, monitoring and logging data.
- Miscellaneous science support data.
- Shift logs.

Once the ALMA site is fully operational (from 2012 onwards) the bulk data is estimated to have a size of 0.5 TeraBytes per day; it will be stored in the next generation archive system (NGAS) developed by ESO's data management division. The remainder of the types of data listed above will be stored in a relational database system (Oracle); its size is expected to be around 15% of the bulk data size.
The bulk data and also the vast majority of the relational data will be produced at the AOS/OSF; however, at the OSF there will be only limited storage capacity for bulk data, which will be kept for a limited amount of time (> 8 weeks) and transferred to the SCO at daily intervals. The SCO will hold a full copy of the bulk data and will act as the master node towards the ARCs, which will also hold full copies of bulk data. The relational data, which will consume much less disk space, will be fully held at the OSF, the SCO and the ARCs.
Figure 1 shows a high level description of the hardware needed at the OSF and the SCO up to 2012. In addition servers will be needed for antenna testing (limited time) and for disk handling. The ARCs will be clones of the SCO.

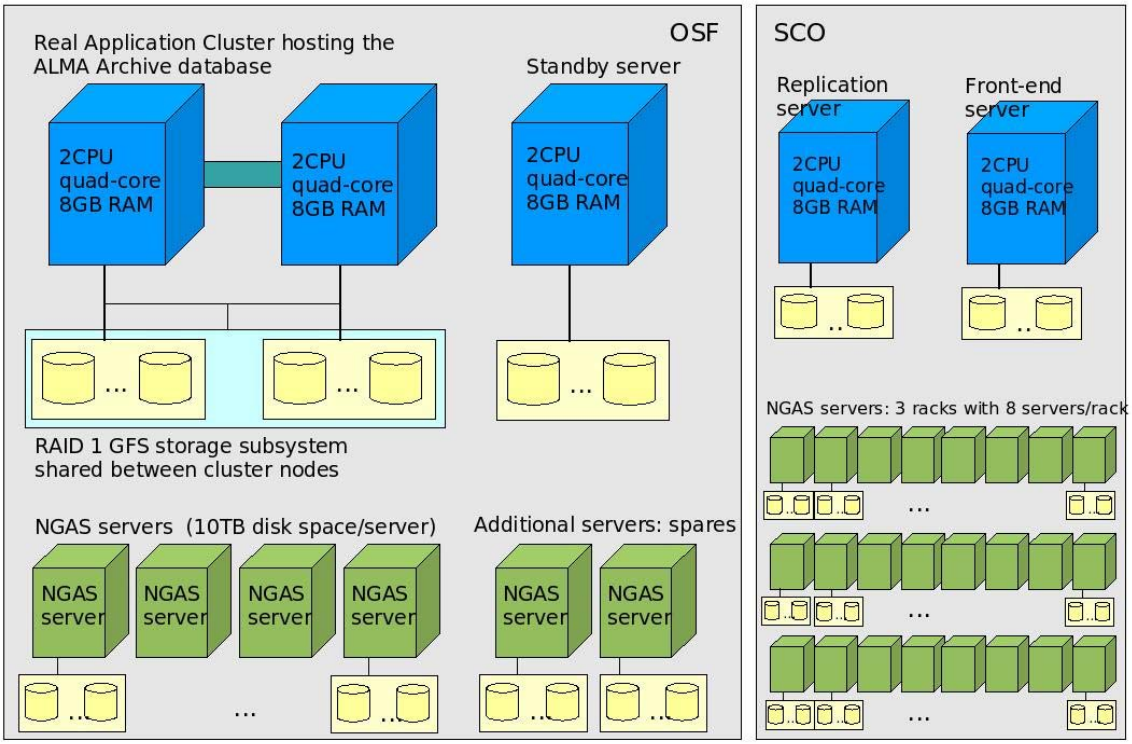

Figure 1: Servers needed at the OSF and the SOC.

All servers shown in the picture above shall be configured as described in chapter 5. They will be running Linux x86 Red Hat Enterprise. The next sections will go into more detail on the ALMA hardware needs covering the first year of purchases.
For that time period only the servers needed at the OSF as shown in figure 1 will be purchased (as neither the SCO nor the ARCs will be operational before 2010).
Additionally the first rack of SCO NGAS servers will be purchased with a limited number of disksto be installed initially at the OSF..

3.1 Operations Support Facility (OSF), Chile

Servers needed at the OSF fall into three different categories:

- Servers dedicated to relational database systems (Oracle)
- NGAS servers
- Additional/spare servers

Additionally the first rack of SCO NGAS servers will be installed at the OSF.

3.1.1 OSF Oracle servers

There will be three Oracle servers at the OSF. Two of them will form a cluster sharing a RAID 1 (mirrored disks) GFS storage system; the third server will host the OSF standby database. All three servers will have 2 quad-core CPU, 8G RAM and 12x750GB disks[1]. More servers will be added to the cluster if it becomes necessary, e.g. for performance reasons.

3.1.2 OSF NGAS servers

There will be four NGAS servers at the OSF (four are needed to cope with the incoming data), each of them having 2 quad-core CPU, 2GB RAM and 12x1TB disks. Should one of them fail then it will be replaced by one of the two special purpose servers..

.

3.1.3 OSF Special Purpose servers

Two additional servers will be held at the OSF. During normal operation they will serve as disk handling units; they may also be used to replace a broken NGAS server. These servers will have 2 quad-core CPU, 2GB RAM, a system disk and a 3U chassis with 12 slots for disks.

3.1.4 First rack of SCO NGAS servers

Initially, while the SCO is not yet operational there will be one rack of SCO NGAS servers located at the OSF holding 8 servers, each having 2 quad-core CPU, 2GB RAM and slots for 24 disks. For the first year each of these servers will hold 6x1T disks; more disks will be added when the need arises. Once the SCO is operational this rack will be moved there. It is estimated that the first NGAS rack with the full set of 24x1T disks will be able to cope with the data collected up to the start of full ALMA operations in 2012.

4 Planned ALMA Archive Environment (test and development)

[1] If at the time of purchase there are bigger/cheaper disks, different CPU or memory types available that are as fast as or faster than the ones mentioned here, then this other, newer hardware should be purchased.

4.1 Enhanced Test Environment (enhanced STE)

One "enhanced" STE (located in Garching) will be setup to match the OSF environment. This STE will be used to reproduce and debug OSF problems and will thus simplify remote maintenance. This STE will consist of four servers, two of them meant to mirror the database cluster at the OSF each having 2 quad-core CPU, 8GB RAM and 4x750GB disk space. Having the same amount of RAM available on these servers will help tuning the OSF environment, e.g. by running tests that will allow finding the best combinations of tables that should be pinned into memory. For the same reason (tuning) it would be very useful having the same kind of processors and disks/controllers as the OSF servers. The third server will play the role of standby server as well as master node in a replication environment having 2 quad-core CPU, 2GB RAM and 4x400GB disk space. The fourth server will be the NGAS server, also having 2 quad-core CPU, 2GB RAM and 4x400GB disk space.

4.2 Development Environment

Also in Garching a development environment matching the OSF setup will be installed allowing to define database related configuration and procedures. It will be used to test the Oracle cluster, multimaster/streams replication procedures, switchover/failover scenarios and the interaction between the three. It basically has the same hardware requirements as the enhanced STE. This environment is already in place, except for one server that should be used for software development and nightly builds.

5 Detailed Description of Hardware

The current servers are assembled by a vendor based on a component hardware specification provided by us. This allows us to control the components of the servers to a much higher degree than by specifying just a model. The specifications can be modified if certain parts are no longer available or to follow technology advances. Any modification has to be discussed between us and the provider. More substantial modifications like an exchange of a mainboard or a different type of controller has to be tested and verified before they go in production. There are three different types of machines, but the difference between them is kept to a minimum in order to ease maintenance and exchangeability. The differences are the necessary number of disk slots, amount of memory and CPU speed. A typical set of specifications looks like the following:

- Two Processor Intel Quad-Core Xeon
- 2GB RAM, registered DDR2 400 with ECC

  - Intel S5000PSL motherboard with dual GB network interfaces
  - 2 x Gigabit Ethernet Controller (supported by Linux)
- 3.5" slimline floppy drive
- Slimline DVD-ROM IDE
- Raid Controller: 3Ware 9550SX-12ML: 12 Channel Serial-ATA II RAID Controller, 256 MB ECC Cache Memory, with BBU (Battery Backup Unit)
- 12 x 500 GB Western Digital Serial-ATA Hard disks 500 GB, 7200 RPM
- Chenbro RM31212B 19" 3 HU Chassis with redundant power supplies
- 12 x S-ATA HotSwap Trays, Rack Rails for 19" Rack

These specifications are for a 3HU standard server as they are currently used for stand-alone archive machines running all archive software, including the data base, NGAS and the archive front-end software. The same kind of hardware will also be used later for NGAS front-end machines. The Oracle database servers will be very similar except that for this application we need at least 8GB of memory. The servers in the NGAS cluster will be 24 slots and thus a Chenbro 5 HU Chassis and two 3Ware controllers instead of one will be used. This application will benefit from quad core CPUs. 2GB of memory are sufficient in this case and the power supply has to adequate for this kind of load. As mentioned above, the exact specifications can be modified by the bidders, but only for well specified reasons, e.g. availability or technological advantages and within as small as possible boundaries. In particular changes of the chassis and the controllers are critical as this would affect the interoperability of the NGAS system. After the selection of the vendor we will also specify details like cabling order between the controllers and the disk slots.
The SATA disks specified above can be exchanged with different ones, but we need to know the exact type and will check for higher than average failure reports. In case the offered disk is not already known to us we would like to get a few samples for tests. All servers will run Linux and thus in particular the availability of drivers for the used network chips and the overall compliance of the mainboard with Linux are essential.

6 Scope Of Supply

6.1 Procurement of servers and disk space

New servers will be needed for setting up the ALMA Archive production environment (for the first year this is restricted to the OSF). Test and development servers need to be purchased according to chapter 4 (also see Appendix).
The contract shall cover the purchases for one year starting 2008 and include the servers listed in the Appendix. It is further foreseen to purchase disks for production and test and development as outlined in the Appendix.
The offer should quote the price for all the requested servers and disks, to be delivered to the Headquarters of ESO in Garching.

Additionally every server shall be tested at the vendor's site before delivery to ESO. The expected test is that the computer is run for one day, exercising some load on the CPU and on disk access. ESO will make available the application software to be used for this. Offers shall quote a separate price for this test and a short description of it.

ESO commits to purchase the listed equipment in steps. The offer should be valid also for a second year in case all the items would not have been bought yet by ESO in the first year.
After that time ESO will intend to proceed with a new request for quotation, to take into account the new servers and disks available by then.

6.2 Maintenance and Support

The contract will include Maintenance and Support.

The following level of maintenance is defined:

Upon notification, replacements for defective parts shall be delivered to the ESO Headquartres in Garching within the next business day (Mo-Fr, 8.00-18.00 local time). ALMA has spares on-site in Chile to perform fault isolation and replacement.

Access to telephone and hotline support during normal working hours (Mo-Fr, 8.00-18.00 local time) should also be included.

For each level the maintenance service shall be priced as a percentage of the price charged for the hardware. The date when the maintenance period starts for purchased items should be specified in the offer, taking warranty periods into account.

At the end of the year ESO will provide, formatted as Excel file, an updated list of hardware purchased for the maintenance cost calculation.

6.3 Technical Partnership

It is expected that the selected service provider will also support ESO with respect to technology issues such as:

- being the main information source concerning new technologies and products
- providing technical support and advice with respect to new items
- subject to availability, providing loaned equipment for product evaluation

Bidders are requested to provide an indication of their capability to be effective on the above mentioned areas.

7 Appendix

7.1 Production and Test and Development Servers

7.1.1 By Location

Environments	Number of environments	Number of servers / environment	Number of single/dual-core CPU / server	Memory / server	Disk slots / server	Total number of servers
OSF	1					
OSF cluster		2	2 quad-core	8G	12	
OSF standby		1	2 quad-core	8G	12	
OSF NGAS		4	2 quad-core	2G	12	
OSF spare/DiskHandling Units		2	2 quad-core	2G	12	
SCO NGAS		8	2 quad-core	2G	24	
						17
Enhanced STE (Garching)	1					
STE cluster		2	2 quad-core	8G	8	
STE standby/replication		1	2 quad-core	8G	8	
STE NGAS		1	2 quad-core	2G	8	
Nightly build / NGAS		1	2 single-core	2G	8	**5**
						22

Colour Coding: Yellow = production, Blue = test and development

7.1.2 By server type

Type of Server/Hardware	Description	Count
Oracle server OSF	2 quad-core CPU, 8G RAM, Chassis 3U with 12 disk slots	3
NGAS server OSF/disk handling unit	2 quad-core CPU, 2G RAM, Chassis 3U with 12 disk slots	6
NGAS server SCO/ARC	2 quad-core CPU, 2G RAM, Chassis 5U with 24 disk slots	8
Disk	1T disk	100
Oracle server (enhanced STE)	2 quad-core CPU, 8G RAM, Chassis 2U with 8 disk slots	3
NGAS server (enhanced STE,Dev)	2 quad-core CPU, 2G RAM, Chassis 2U with 8 disk slots	2
Disk	450G disk	20
TOTAL number of servers		22