THE EVOLUTION OF THE SIMULATION ENVIRONMENT IN ALMA

T. Shen, R. Soto, N. Saez, G. Velez, S. Fuica, T. Staig, N. Ovando, Joint ALMA Observatory, Santiago, Chile
J. Ibsen, European Southern Observatory, Santiago, Chile

Abstract

The Atacama Large Millimeter /sub millimeter Array (ALMA) has entered into operations transition phase since 2014. This transition changed the priorities within the observatory. Most of the available time will be dedicated to science observations instead of technical time for commissioning activities including software testing. The lack of the technical time surfaces one of the weakest points in the existent infrastructure available for software testing: the simulation environment of the ALMA software. The existent simulation focuses on the functionality aspects but not on the real operation scenarios with all the antennas. Therefore, scalability and performance problems introduced by new features or hidden in the current accepted software cannot be verified until the actual problem explodes during operations. Therefore, it was planned to design and implement a new simulation environment, which must be comparable, or at least, be representative of the production one. In this paper we will review experiences gained and lessons learnt during the design and implementation of the new simulated environment.

OVERVIEW

The Atacama Large Millimeter /sub millimeter Array (ALMA) started the operations phase transition in 2014. This transition changed the priorities within the observatory. Most of the available time will be dedicated to science observations instead of technical time for commissioning activities including software testing. (since software/hardware integration was the former priority during construction phase). The lack of the technical time surfaces one of the weakest points in the existent infrastructure available for software testing: the simulation environment of the ALMA software. In ALMA, simulation capabilities were initially developed to satisfy Control and Correlator subsystem needs, supplying them with virtual hardware devices to interface with the software components being developed. Additional simulation layers and capabilities were added during the years by different teams, but they were focused on the functionality aspect but not on the real operation scenarios with the whole array. Therefore, scalability and performance problems introduced by new features or hidden in the current accepted software cannot be verified until the actual problem explodes during operation. The lack of a representative testing environment will seriously impact the efficiency of the ALMA incremental software release process.

It was planned to design and implement a new simulation environment, which must be comparable, or at least, be representative of the production one. Duplicating the production environment was not an option, since it would be prohibitive from the point of view of the associated cost. Adjustments in the existent simulation architecture had to be introduced, but with special care on keeping the simulation environment comparable in terms of CPU load, network bandwidth throughput, memory usage and software configurations. The selected platform to provide computing power is based on blade technology of Cisco Unified Computing System (UCS). The new simulation platform will provide the required amount of time for testing purposes, at same time it allows us to maximize the efficiency of the reduced technical time available with operational hardware, which will be dedicated only for the final validation of a new release and small set of features that interact directly with the system. In this paper we will review experiences gained and lessons learnt during the design and implementation of a new simulation environment.

PRODUCTION ENVIRONMENT

The hardware related to the ALMA array (antennas, photonic references, correlators, etc) are controlled through an STE [1], which is a collection of servers, real time computers, network switches, storages and databases, that are configured accordingly as a single platform to support the execution of the ALMA observing software [2]. In the Fig. 1, the list of servers, and real time computers are shown.

EXISTENT SIMULATION ENVIRONMENT

The existing simulation platform is a reduced scale version of the STE dedicated for operation, which contains few servers are available to simulate an array of few antennas. This simulation environment has been very useful during the AIV stage [3] where the focus was the commissioning of newly assembled antennas. At that moment, no more than two antennas were usually configured into a single STE.

It became evident that the existent simulation environment needed to be upgraded since the pressure to use the production hardware for scientific observation is extremely high. Currently, the most critical deficiency in the current simulation environment is in the deployment aspects. It differs in terms of network layout, network throughput, computing power, available memory and disk I/O.

ISBN 978-3-95450-148-9

THE NEW SIMULATION ENVIRONMENT

The following requirements were defined for the new simulation environment:

- To be able to simulate up to 66 antennas.
- To be able to simulate 4 basebands of BL and ACA correlators.
- To use the same network design than production
- To be representative of the production environment in terms of available processing powers, memory, network bandwidth.
- To keep the same O.S., server hardware architecture.

As it is impossible to replicate the amount of servers involved in the operational environment due to the cost, the strategy presented here is to group components within a more powerful server in order to reduce the required number of servers, but at the same time to maintain the deployment characteristic/complexity of the production environment. In this context, the criteria to group components must make sure that critical resources such as network bandwidth, memory allocation, CPU load in the simulated environment, etc. must be representative and comparable with the production environment.

The virtualization approach was tested but with poor results, as the I/O performance required cannot be sustained by virtual machines. In the other hand, adding another layer of abstraction in the testing scenario doesn't add any value when one of the main requirements is to emulate the performance in the operational environment. Instead, a consolidation approach in more powerful servers was explored.

The methodology to find the right consolidation factor is to take a baseline performance measured in the existent hardware during end-to-end tests [4]. Then reconfigured in a more powerful servers with more components, the limit is found when the load stops to grow linearly. Investigation done in [5] allows us to estimate the required computing power in order to simulate an environment up to 66 antennas and the equivalent for the processing of the 4 basebands in the BL correlator and ACA correlator Within this investigation, the foundation to provide computing power is based on blade servers, and the following consolidation can be achieved in each server, as shown in the Fig. 1: i) 8 antenna real time computers (ABM) can be consolidated into a single server, ii) GNS, GAS type of servers will be map 1 by 1 between the production and simulation environment. So far, there are 8 servers of this kind, iii) 4 CDP nodes can be consolidated into a single server and iv) 4 ACA CDP nodes can be consolidated into 1 blade server. In total, the number of the required blade servers summed up to 32.

CISCO UCS as Source of Computing Power

Considering the tests performed, the most critical hardware resources are i) number of processors and cores, and ii) the available memory of the server. Based in these technical needs, a search was developed to find a suitable platform to provide such amount of processing capacity. Finally, CISCO UCS [6] was identified as the best solution. It comprises CISCO B22 servers with 48GB Ram, 2 six cores CPU, and CISCO UCS5108 Chassis, which allow us to have a scalable solution with 32 blades in modular chassis. The UCS5108 Server Chassis can accommodate a maximum of 8 blade servers and provides a total of 1.2 Terabits per second of available Ethernet throughput.

Networking

From the network point of view, the Cisco 6200 series Fabric Interconnect switches are fully compatible with the existing ALMA network backbone (Catalyst 6500 and Nexus 5000), ensuring that the new system will both be able to provide the maximum intended throughput and keep the same network layout than production environment [3]. The latter is a very important decision factor, because it will allow us to use the same configuration and administration tools developed for the production environment. As shown in [3], we just needed to add an additional instance of VRF (Virtual Routing and Forwarding) in the existent network domain. The new VRF instance contains the same Layer-3 network topology than the production VRF instance

Storage

High performance Storage Class provided by EMC VNX2 covered the need for storage in the simulation environment. Another advantage of Cisco UCS is that the same Fabric Interconnect also serves to provide 16 Gpbs redundant fibre channel connections to the storage.

USE CASES

Since the moment the new simulation environment was put into production, it has been the key platform for the verification phase of the incremental software release process [7]. The usual testing scenarios are: (i) regression tests, (ii) scalability tests, iii) new feature/functionality tests and (iv) bug fixes testing.

Besides the aforementioned scenarios, it is worth to mention that this simulation platform has been also very useful to verify the performance of key components of the ALMA software, such as concurrent access to the TMCDB database, bulk data transfer and system start-up parallelisation. It has also been useful to find long-standing integrity issues such as memory leaks and memory corruption that are seen after long observations, and to test new deployment strategies before going into production.

It has been possible to stress the access to the TMCDB database, which allows testing different solutions to alleviate the bottleneck situation. Applying the improvement has resulted in reducing the time spent during subsystems initialization and observations in the operation environment. Having the same network layout, the Bulk data transfer network loads can be tested in a similar way as in the operational environment, helping to find introduced throughput or timeout problems early in the testing process.

ISBN 978-3-95450-148-9

item	Description	type	# Equipment in production	Consolidation factor	# Of required blade server
1	Antenna Bus Master, Central LO RT computers	abm	68	8	8.5
2	CentralLO RT computers, DMCs	lo-x, cob-dmc	7	8	0.875
3	CDP nodes	cob-cdpn	16	4	4
4	CDPMaster, CCC, COJ-CDPMaster, COJ-CC	cob-cdpm, cob-cc, coj-cdpm, coj-cc	4	2	2
5	ACA CDP nodes	coj-cdpn	32	4	8
6	General Network Servers	gns	2	1	2
7	General Application Servers	gas	6	1	6
				Total # of blades	31.375

Figure 1: The different types of servers and real time computers configured the production environment and how they could be consolidated into more powerful servers.

It is now possible to test the system start-up parallelization scalability in simulation, to prevent the introduction of situations that may increase the initialization times in an unreasonable amount. Using the similarities to the production environment it has been possible to recreate real observing scenarios while analysing the processes for memory leaks and memory corruption, which has allowed increasing the continuous uptime of the system, by maintaining the resources usage low and reducing unexpected crashes.

This platform also allows us to prepare the upgrade of the operative system to RHEL 6.6 and to port ALMA software into 64 bit architecture.

FUTURE WORKS

Future works are located in the enrichment of simulation behaviour of hardware devices, such as antennas pointing, correlators modes, etc. We expect to incorporate concepts such as model in the loop or hardware in the loop, which gives us the advantage to use exactly the same software than production (currently the hardware drivers are not being tested, since they are being replaced by simulators [8]), therefore to achieve a better coverage with our testing process.

It is also considered to use the idle time of this simulation environment to run automatic testing. Our goal is to support continuous software integration, as part of the ALMA software delivery process [7], and this environment will be perfect to execute nightly builds.

CONCLUSION

A new simulation environment was designed and implemented. It fulfilled entirely the defined requirement, specially being a representative testing environment of the production environment. After its introduction, the amount of technical time requested on the production environment for software testing has been reduced considerably.

The testing environment has the same network configuration. Servers and devices (simulated) are deployed exactly in the same way as in production. Therefore the same configuration and tools are being used in both places.

Blade servers have demonstrated to be an excellent alternative to provide computing power, which scales, taking account that the observatory's data center is located in the Atacama Desert and providing power and space is not something trivial. Cisco UCS, as a complete solution, has the advantage that non-additional network devices are required to procure in order to put the blade servers chassis into production. It was important that this solution is fully compatible with our existent network design. Internally, Cisco UCS's Fabric Interconnect switches provide high availability and redundancy by design; therefore our former in house mechanism implemented by using network bonding at the O.S. level is not necessary anymore.

Finally, with the experience learnt in this area, it is planned to upgrade to production environment using the same technology in the next year.

ACKNOWLEDGMENT

The authors acknowledge the contribution of all the members of the ALMA Department of Computing to the success of this project.

REFERENCES

[1] T.C. Shen, R. Soto et al., "ALMA Operations Support Software and Infrastructure", Vol. 9149 91492D-6, proceedings of SPIE, SPIE' 12, Amsterdam, Netherlands (2012).

[2] A. Farris, R. Marson, J. Kern et al. "The ALMA Telescope Control System", Proceedings of ICALEPCS, ICALEPCS' 05, Geneva, Switzerland (2005).

ISBN 978-3-95450-148-9

[3] T.C. Shen, N. Ovando, et al, “Virtualization in network and servers infrastructure to support dynamic system reconfiguration in ALMA”, Vol. 8451, 84511N, proceedings of SPIE, SPIE’ 12, Amsterdam, Netherlands (2012).
[4] A. M. Chavan, B.E. Glendenning, J. Ibsen et al, “The last mile of the ALMA software development”, Vol. 8451 84510Q-1, proceedings of SPIE, SPIE’ 12, Amsterdam, Netherlands (2012).
[5] M. Mora, J. Avarias, A. Tejeda, et al, “ALMA software scalability experience with growing number of antennas”, Vol. 8451, 84510W, proceedings of SPIE, SPIE’ 12, Amsterdam, Netherlands (2012).
[6] Cisco Unified Computing System website: http://www.cisco.com
[7] R. Soto, T.C. Shen, N. Saez, and others, "ALMA release management: a practical approach”, proceedings of ICALEPCS, ICALEPCS’ 15, Sidney, Australia (2015).
[8] M. Mora, J. Ibsen, J. Kern, and others, “Hardware device simulation framework in ALMA control subsystem”, Vol. 411, p.462, proceedings of ADASS, ADASS XVIII, San Francisco, USA (2008).

ISBN 978-3-95450-148-9