Volume Title
ASP Conference Series, Vol. **Volume Number**
Author

Closing the observing loop across continents: data transfer between Chile and Europe

R. Hanuschik[1], S. Zampieri[1], M. Romaniello[1], C. Cerón[2], A. Wright[2], C. Ledoux[2], and F. Comerón[1]

[1]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei München, Germany*

[2]*European Southern Observatory, Alonso de Cŕdova 3107, Vitacura - Santiago, Chile*

Abstract. The different tasks supporting the operations of ESO telescopes in Chile are carried out at widely separated locations, with some taking place at the observatory site and others in Europe at ESO's headquarters. A fundamental requirement to make such an operations scheme viable is the need for a stable way of transferring the large amounts of data generated by the telescopes to Europe on the shortest possible timescale. We review technical progress that has allowed ESO in the last years to move from the transfer of data on physical media, with delays of up to two weeks, to the current transfer of most of the data stream through the internet within hours. We also describe the possibilities that will be open in the near future with EVALSO, an European Union co-funded project to provide full-fiber connectivity from the ESO observatory site on Paranal, and the nearby Cerro Armazones hosting the telescopes of the University of Bochum, all the way to Europe. Given the recent choice of Cerro Armazones as the future location of the European Extremely Large Telescope, upgrades in the communications infrastructure of Paranal and Cerro Armazones are very relevant to the ability to operate the current and future facilities there.

1. Introduction

The operation of world-class ground-based astronomical facilities is nowadays a global enterprise, in which operations processes can take place at locations thousands of kilometers away from the telescope. Underpinning the capability of managing operations in such geographically distributed manner is the ability to quickly transfer the vast amounts of data generated by current astronomical instrumentation across the operations sites, so that the various processes composing the operations scheme of the facility can feed back on each other in a timely manner.

The ESO observatory of Cerro Paranal (Chile), hosting the Very Large Telescope (VLT), the VLT Interferometer (VLTI), the Visible and Infrared Survey Telescope for Astronomy (VISTA) and, in the near future, the VLT Survey Telescope (VST) is no exception to these demands. The operations facilities on Paranal are complemented in the ESO Headquarters in Garching near Munich (Germany), where user support, data processing and permanent storage takes place. During the initial 10 years of operations, limitations of the available bandwidth coupled with the growing overall data rates coming from the increasing number of telescopes and instruments on Paranal forced ESO to

transfer the data from Chile to Europe using physical media, with the subsequent time lag of up to two weeks between the generation of the data and their detailed quality analysis. Recently it has become possible to use the Internet to transfer this data stream within hours or less of it being produced. However, qualitative leaps lie ahead with the arrival to Paranal in the period 2009-2014 of the VLT/VLTI 2nd generation instruments and of the VST. Furthermore, the communications infrastructure on Paranal will be used to transfer the data from the future European Extremely Large Telescope (E-ELT), to start operations in the nearby Cerro Armazones around 2020.

2. VLT end-to-end operations and its communication needs

Since its initial design in the mid 1990s (Quinn et al. 1998), VLT operations were conceived as an integrated system (Péron 2008) strongly relying on flexible scheduling. Underlying this end-to-end model is a system with tools to manage the flow of data and information. A central element in this scheme is the ESO archive, located at the ESO Headquarters in Garching, where all observations collected at the observatory are stored and made available to operations groups in charge of quality control, instrument trending and health check, and data package preparation and distribution (Eglitis & Suchar 2010), as well as to the users. The essential role played in operations by the archive thus makes the operations workflow critically dependent on the speed with which the data obtained at the observatory can be ingested in it. The end-to-end model has been validated by its successful implementation for over 10 years (Kaufer & Comerón 2010), and is now applied with minor variations to the operation of the La Silla facilities and to the other facilities on Cerro Paranal. Many elements of this model have been adopted by the planning of operations of the ALMA millimeter and submillimeter array (Andreani & Zwaan 2008). Conveniently evolved in the coming years, it will also form the basis of the operation of the E-ELT (Spyromilio et al. 2008).

3. Current demands

The transfer of data through the Internet (Zampieri et al. 2009) from Paranal to Garching became operational in late 2009. At present, the transfer of data from Paranal (and also from the ESO La Silla observatory near La Serena, Chile) to the ESO Archive in Garching takes place online. Data are also copied at the observatories onto USB disks, which are shipped to Garching on a weekly basis. The double channel is designed to ensure the prompt online transfer of high priority data without saturating the available bandwidth, in which case data can be made to overflow to the disk channel.

3.1. Why fast data transfer?

Having the data available in the ESO archive in Garching within a short time from acquisition has positively impacted both the user experience and the observatory operations. Investigators of programs executed in Service Mode can access their proprietary data in almost real time to speed up their scientific exploitation. More than 100 datasets are downloaded per month on the average. In addition, data handling operations are simplified on both sides of the Atlantic by eliminating most of the logistics associated with the transfer and handling of physical media. Finally, the Data Processing and Quality Control group in Garching can perform in-depth quality control (Hanuschik

2007) in quasi real time using the most recent files and feed the results back to Paranal typically within one hour from when the suitable frames are acquired. This removes most of the needs for similar activities on the mountain.

3.2. Current data volumes and data transfer capabilities

The current average production rate of VLT-VLTI instruments is 20-25 GB of data (all data volumes given in this paper are after compression) per 24-hour period. The bandwidth currently reserved for data transfer from Paranal is currently 9.12 Mbit/s, out of a total available bandwidth of 11.2 Mbit/s, corresponding to about 100 Gb over 24 hours. The total bandwidth from Santiago de Chile to Garching is 50 Mbits/s. The indication from the first year of operations of the Scientific Data Transfer is that the nominal available bandwidth should be over dimensioned by at least a factor of two with respect to the average data rate to be transferred.

4. VLT/VLTI second generation, survey telescopes, and beyond

4.1. VLT/VLTI second generation instruments and survey telescopes

With the new telescopes and instruments scheduled to start operating at Paranal in the next few years Moorwood (2009), the amount of data that will be produced, and that will have to be transported to Garching, will significantly increase. The most significant contributors will be VISTA, with an average production of 75 GB/night of compressed data; VST (2010, 35 GB/night compressed), SPHERE (2011, 40 GB/night) and MUSE (2012, 30 GB/night). The expected data volumes of VLTI 2nd generation instruments are smaller, adding another 20 GB/night to the data volume. The total average data production of Paranal in 2014+ will thus be of about 225 GB/night, requiring a sustained bandwidth of about 21 Mbits/s for their transfer to Garching within 24 hours. The current bandwidth from Paranal to Garching will thus fall short by a factor 2 from that needed to transfer the data produced on Paranal.

4.2. The E-ELT

The nightly data volumes to be produced by E-ELT instrumentation is very difficult to predict one decade in advance, but the instrument concepts that have been developed have provided already a crude estimate of the average nightly data production to be set at the level of 1-2 TB (uncompressed) per night, with large variations depending on the instruments and modes actually used on a given night. A rough estimate of the data transfer requirements of the Paranal observatory, including the E-ELT on Armazones, is thus 0.5-1 TB/night (compressed). The quantity may increase further through the upgrade of planned Paranal instruments with larger format detectors.

5. Dealing with future demands: EVALSO

The daily transfer of the Paranal and Armazones data production over the internet will require a bandwidth increase of roughly one order of magnitude with respect to the capabilities of the current setup over the next decade. Furthermore, new applications such as the implementation of a remote interaction mode involving the transfer of data just acquired from the observatory to a user located on another continent, would require

Gbit/s-level capabilities. The EVALSO project (for Enabling Virtual Acces to Latin South American Observatories) (Filippi et al. 2010), funded by the European Union under its Framework Program 7, is expected to provide such capabilities for Paranal and Armazones already in the very near future.

The main goal of EVALSO is to create the missing parts of the physical infrastructure to connect the Paranal and Cerro Armazones observatories to Europe with a high capacity link. To this end, a consortium was formed in 2007 by seven European institutions (the GARR consortium, the University of Trieste, and the Astronomical Observatory of Trieste in Italy, Queen Mary University of London, NOVA in the Netherlands, the Astronomical Observatory of the Ruhr University of Bochum, and ESO, plus the REUNA and RedCLARA networks in Chile). The project will use the ALICE/ALICE2 research network infrastructure within South America and transatlantic connection to European National Research Networks (NREN) via GEANT2. ESO has procured the infrastructure needed to connect Paranal to the existing networks linking Santiago with Europe, and in particular a 75 km-long fiber link between Paranal and the access point to the Chilean backbone. The EVALSO infrastructure is now completed and undergoing commissioning. In the EVALSO implementation currently being put in place, the capacity of the path between Paranal and Armazones is limited by that of the transatlantic link, which with the planned ALICE2 upgrade is expected to exceed 1 Gbit/s. Tests on this segment using the current ALICE infrastructure have already achieved a sustained transfer rate exceeding 100 Mbit/s between Santiago and Garching using UDP-based file transfer tools. Considering that all existing links have higher nominal capacity and that the planned upgrade within ALICE and the trans-Atlantic link will even increase such limits, even faster transfer rates are foreseen in the near future.

The new infrastructure should thus provide sufficient data transfer capacity for at least the next decade on Paranal and Armazones, enabling data files produced by the scientific instruments to be stored in the Garching archive within seconds from being produced and making data processing to start in a practically instantaneous manner. The high capacity of already existing NREN would make it possible the extension of fast data transfer to many other locations in Europe, which may be used in the future to enable remote interaction with the facility as described above (Comerón et al. 2008).

References

Andreani, P., & Zwaan, M. 2008, SPIE, 7016, 31
Comerón, F., Filippi, G., & Emerson, J. 2008, SPIE, 7016, 59
Eglitis, P., & Suchar, D. 2010, in Ensuring long-term preservation and adding value to scientific and technical data, ESA, in press
Filippi, G., Jaque, S., Liello, F., Chini, R., Utreras, F., Wright, A., Lemke, R., & Heissenhuber, F. 2010, SPIE, 7740, 49
Hanuschik, R. 2007, ASP Conf. Ser., 376, 383
Kaufer, A., & Comerón, F. 2010, SPIE, 7737, in press
Moorwood, A. 2009, in Science with the VLT in the ELT era, Springer-Verlag
Péron, M. 2008, in The 2007 ESO Instrument Calibration Workshop, Springer-Verlag
Quinn, P. J., Albrecht, M. A., Ballester, P., Banse, K., Chavan, A. M., Grosbol, P., Péron, M., & Silva, D. R. 1998, SPIE, 3349, 2
Spyromilio, J., Comerón, F., D'Odorico, S., Kissler-Patig, M., & Gilmozzi, R. 2008, The Messenger, 133, 2
Zampieri, S., Forchi, V., Gebbinck, M., Moins, C., & Padovan, M. 2009, ASP Conf. Ser., 411, 540