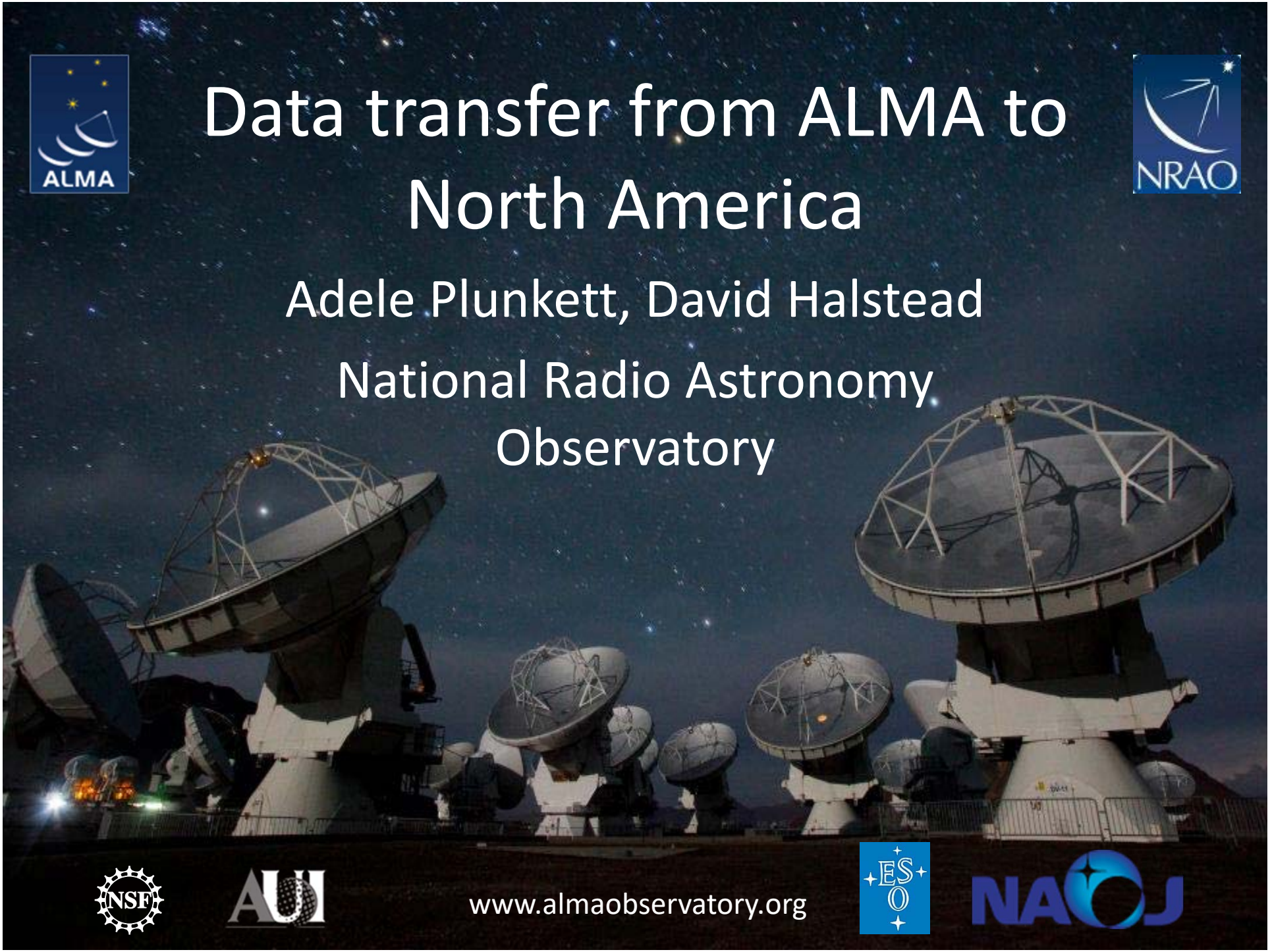


Data transfer from ALMA to North America

Adele Plunkett, David Halstead

National Radio Astronomy Observatory

www.almaobservatory.org

NRAO and ALMA

ALMA is a multinational project with many partners, and three ALMA Regional Centers (ARCs):

- NA: NRAO, Charlottesville, VA, USA
- EU: ESO, Garching (Munich), Germany
- EA: NAOJ, Mitaka (Tokyo), Japan

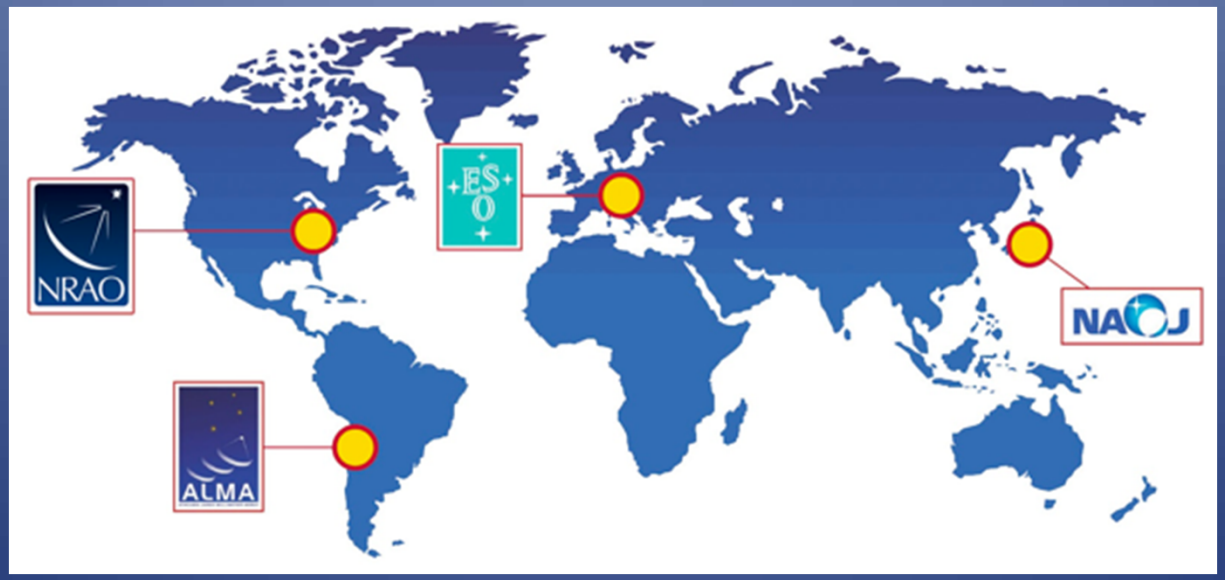


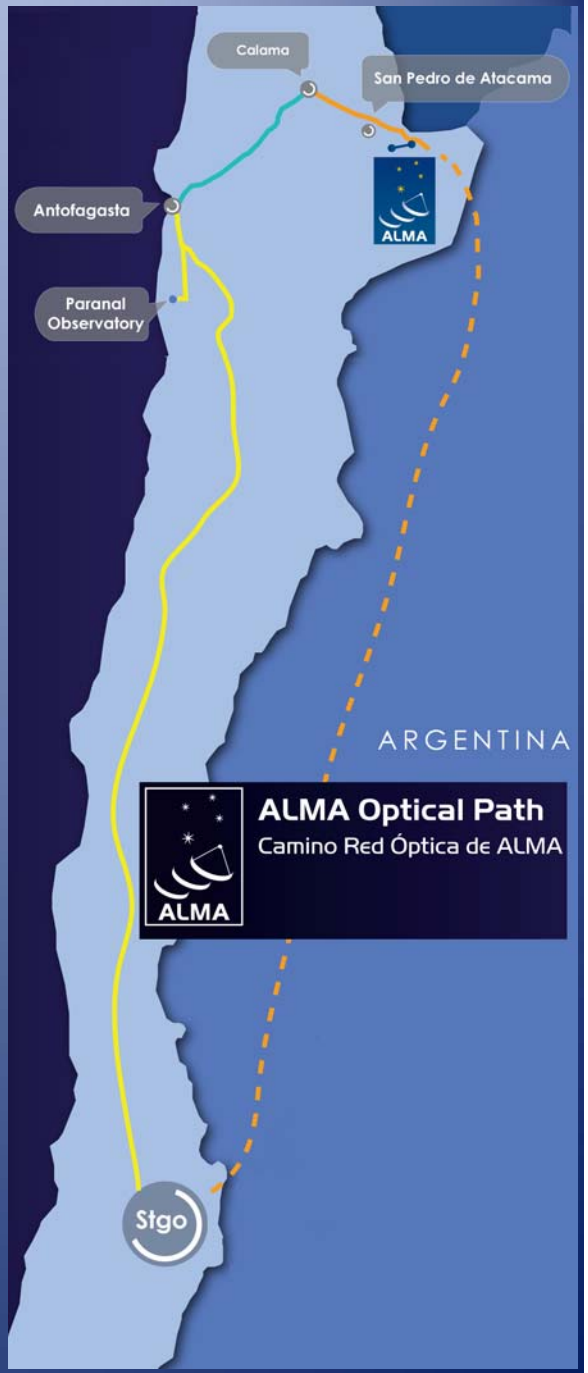
Image: https://science.nrao.edu/facilities/alma/images/arcs.jpg

ALMA telescope

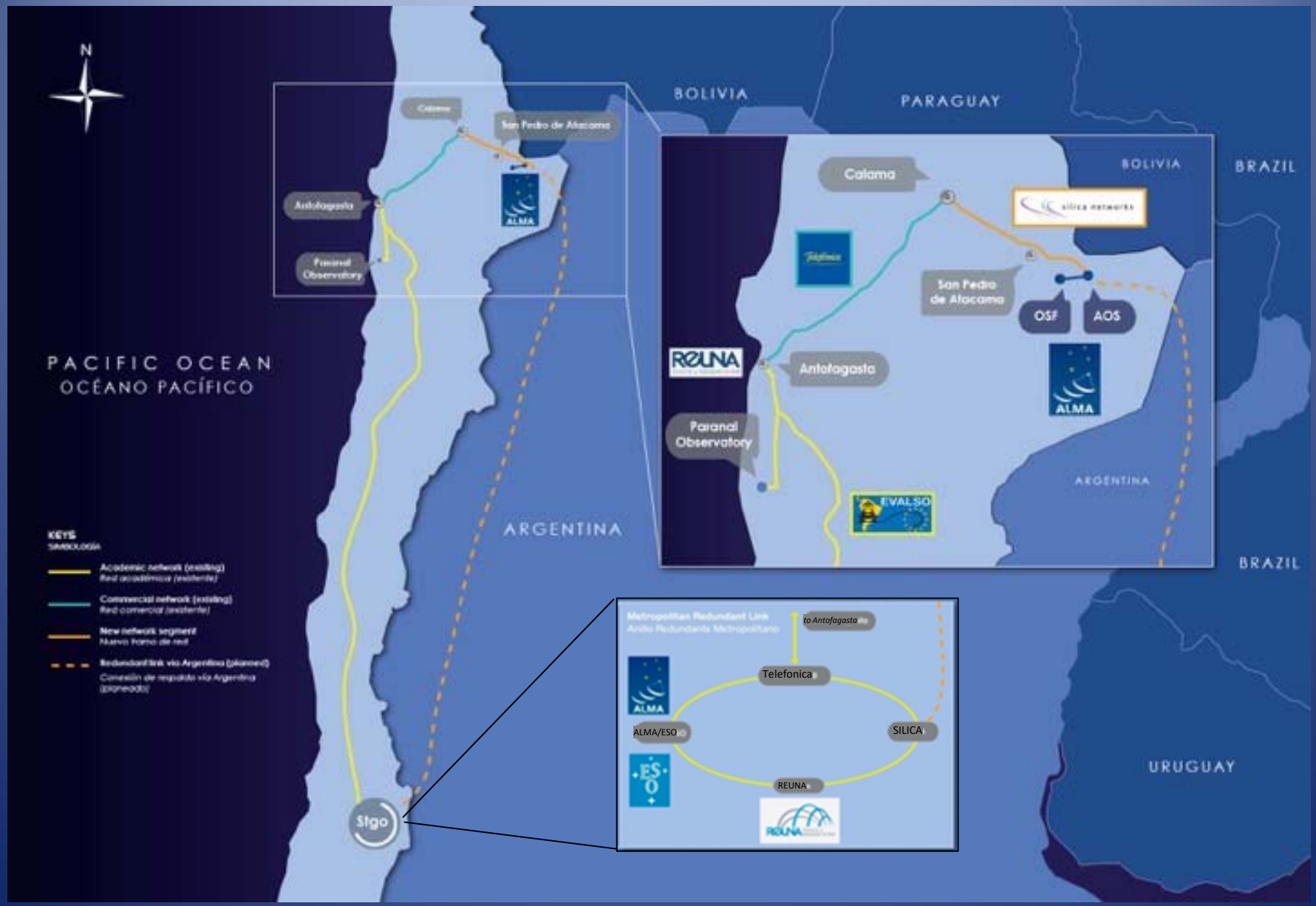
- Largest mm/submm telescope ever built. As an interferometer, it combines signals from multiple antennas to form an image.
- All 66 antennas operational at high site (except for maintenance)
- Sub-arrays possible, and generally 3 projects observed at once:
  - 50 x 12m-antennas (main array)
  - 12 x 7m-antennas
  - 4 x 12m-antennas, observing in "Single Dish" mode ("Total Power")
- Operated "space mission" style, with pipeline data processing and a science archive at each ARC (and Chile).
- First PI projects released to public from the ARCs January 2013.
- Annual calls for proposals in March-April (currently for Cycle 8)
- "Control Room Extension" at Santiago Central Office.
- Cycle 7 observations began in October 2019, but halted due to pandemic shutdown in March 2020
- *Cycle 7 restarted in March 2021, currently continuing with ~40 antennae*

Data transport within Chile
From the telescope

- Operations Support Facility to Santiago 2.5Gb/s
  - fiber to Calama,
  - commercial fiber Calama to Antofagasta,
  - REUNA from Antofagasta to Santiago
- Redundant fiber loop via Argentina planned



Filippi, G., Ibsen, J., Jaque, S., et al. 2014, SPIE, 91520V
https://doi.org/10.1117/12.2054306

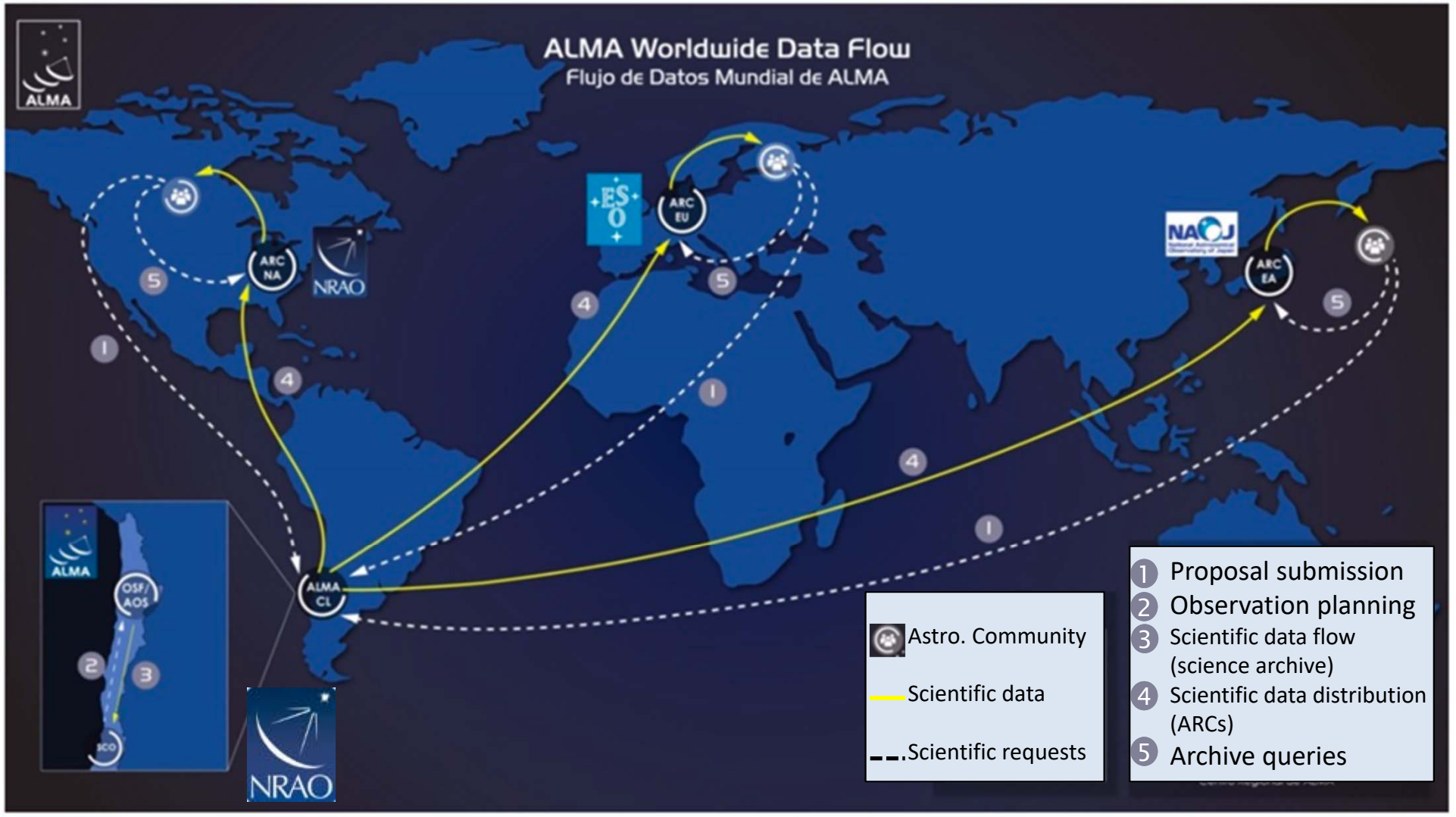


Filippi, G., Ibsen, J., Jaque, S., et al. 2014, SPIE, 91520V
https://doi.org/10.1117/12.2054306

Data transport
Chile to Charlottesville (NAASC)

- MOU signed between AUI/REUNA for local link to SCO
- Santiago to ARCs: individual ARC contracts with REUNA and NRENs.
  - NA: Joint AURA-AUI agreement for 1Gbps committed (burstable to 10 Gb/s capacity) from Santiago to US NREN via Latin America's Nautilus Point of Presence
  - Link from NRAO to Internet2 through UVa is 10Gb/s
- Typical rate obtained during peak data transfer periods is 2-300Mb/s, with bursts up to 600Mb/s
- Currently working on establishing network monitoring, and improving our understanding of how the link performs in typical load conditions (~1TB/day)
- Recent transition to most data processing being done in Santiago. Some manual processing at ARCs.
- The North American ALMA Science Center (NAASC) hosts the ALMA Archive, and computing for NA users.

Data transport (cont'd)



Cycle 7 (interrupted)

From: Sean Dougherty
Subject: ALMA Operations shutdown
Date: March 18, 2020

Dear colleagues:

With the current developments in the country related to coronavirus, the CMT has decided to start preparing for a total shutdown of the Observatory over the coming days. This will enable JAO staff to comply with request of the authorities to reduce contact with others to reduce the spread of the virus and to protect yourselves and your families.

The plan is:

- All SCO-based staff should be working from home, unless critically required to be at SCO e.g. an AoD in the SCO Control Room until Thursday evening and a member of the CMT to act as SCO MoD each day. Please coordinate your priorities your work from home with your supervisor and do not hesitate to ask for any technical support you may need to make remote work effective.
- The number of staff and contractors at the OSF is gradually being reduced. Any travel changes, including those for self-arranged travel, will be duly reimbursed. The flight situation is very dynamic, so please be patient and be prepared for changes in flight schedules.
- By Friday March 20th, only those essential staff needed to support the safe shutdown of the observatory will be onsite, led by Stuartt Corder. The focus will be the AOS shutdown and to bring the transporter down to the OSF.
- By Saturday, we expect to complete the AOS shut down. A small power generator will keep the maser powered up in order to protect its stability. All other equipment will be powered down.
- On Sunday, the only people onsite will be those required to complete the IT and Power System shutdown tasks and to lock down all buildings and equipment to ensure all assets are protected. A small generator will provide power to the fuel station and to the water treatment plant. All other systems will be off.
- By Sunday noon the last staff will leave the site.

Cycle 7 (restarted)

From: Sean Dougherty
Subject: Return to Operations Update
Date: March 24, 2021

Dear colleagues,

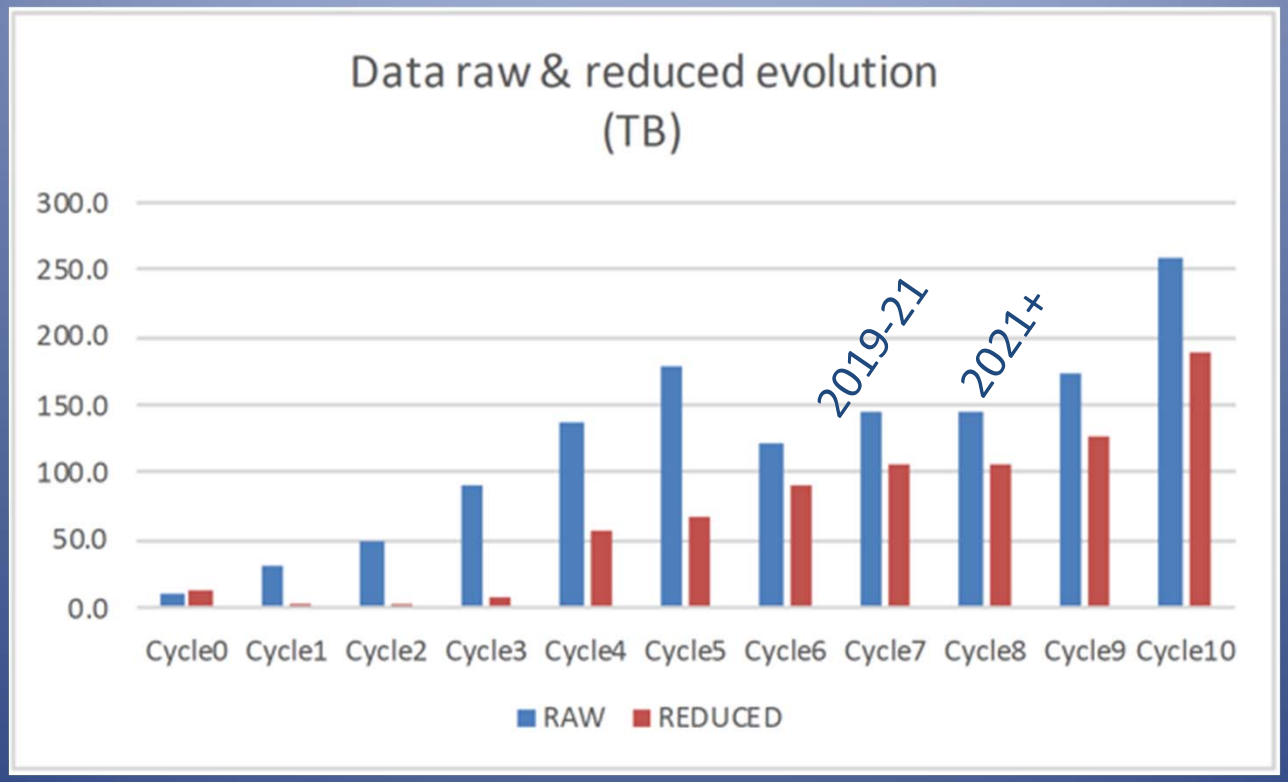
I hope that you and your families are all continuing to stay safe and healthy.

Over the past year, from the moment the decision to close the ALMA observatory was taken, everyone at the JAO has participated in some way in helping to work our way through the myriad of challenges presented by the pandemic, from the caretaker team activities, preparing and returning to the OSF, going back to 5,000 metres to restart the systems in the AOS TB and antennas, until this past week when we celebrated the milestone of starting PI science observations – 364 days after the last science observation. Thanks to all. You are such a wonderful and extraordinary team, that made it possible to once again have ALMA exploring the wonders of our Universe!

This week we have seen a dramatic spike in the new daily cases numbers and on the ICU burden. In response, many comunas in the Metropolitan Region have moved to Paso 1 of the Paso-a-Paso program to try and contain the spread of the virus. The SCO remains open currently for those staff considered "essential" or those who do not live in a Paso 1 comuna and who have arranged to work at the SCO. For staff members that live in a Paso 1 comuna and have not been identified as essential, please follow the quarentena guidelines and continue working from home until further notice.

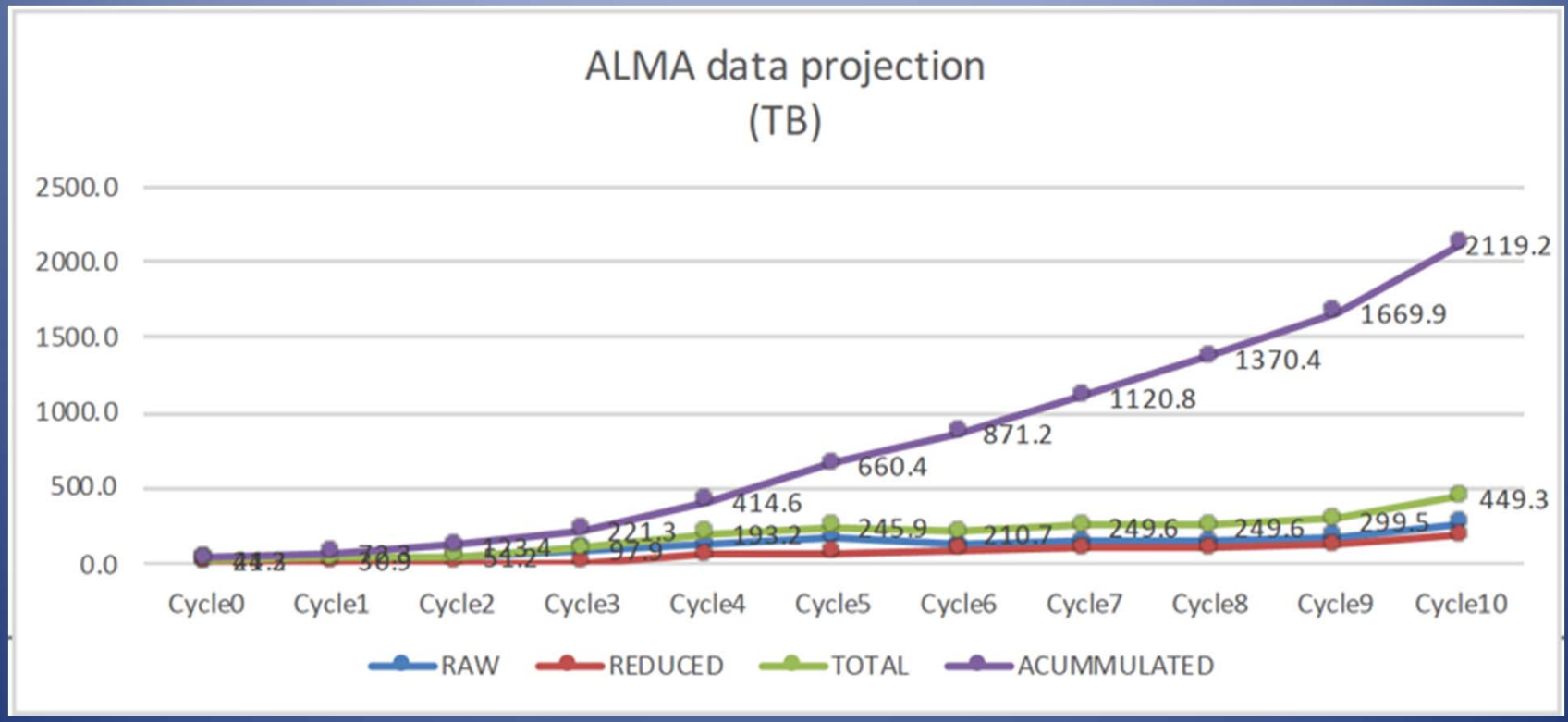
Following the framework put in place in May 2020, the CMT has continued to monitor closely the status of the pandemic to inform the Return to Operations decision-making process. Notwithstanding the high number of new cases that remains a concern, there are a number of indicators that show the current situation has not reached the levels seen in June 2020. The impact of the vaccination campaign progress, both nationally and for our staff, is now also being

ALMA Science data storage



Source: ALMA data rates and archiving at the NAASC NAASC Memo 110 (v5.1)
ALMA Archive Review, 2019

ALMA Science data projection



Source: ALMA data rates and archiving at the NAASC NAASC Memo 110 (v5.1)
ALMA Archive Review, 2019

Cycle 6-7

- No longer taking data in 2 streams (Water Vapor Radiometer duplication eliminated)
- ~55% of array time for science
- 40 antennas as minimum; sometimes achieved 66
- Total volume will be ~220TB, including image products (which will constitute about 30% of the total data volume)
- Shifting the majority of data processing at JAO in Santiago *before* distribution to the three ALMA Regional Centers

Future Cycles

- Now running in “Full Science” state, with mean data rates ~100Mb/s during observations
  - “Duty cycle” of observations will slowly increase as testing and maintenance procedures improve.
- Best guess estimate for the next 3 years (including product size mitigation) is around 200-300TB/yr (raw data and products roughly equal)
- Important to note that data rates vary through the configuration cycle. When long baseline configurations are scheduled the data rate goes up for these reasons:
  - Data sampling needs to be faster to prevent beam smearing at the field edges
  - The data products, which are also mirrored from Santiago, also increase in size, to become larger than the raw data in the largest configurations
  - So far, long baseline campaigns have tended to have low observing efficiencies, however this may change

“Next ALMA Correlator”

- A correlator upgrade is being developed
  - In line with ALMA Development Roadmap [1], and ALMA2030 vision
  - Likely completion ~2028, as part of ALMA2030 upgrade
  - Specifications available in draft by working group [2]
- Deployable location physically separate from BLC (potentially OSF)
- Expected data rate increase is about a factor of four, corresponding to a data rate of ~1PB/yr (not all projects will need the extra channels)
- Details to be defined along with correlator partner

[1] https://www.almaobservatory.org/en/publications/the-alma-development-roadmap
[2] https://science.nrao.edu/facilities/alma/science_sustainability/Specifications2ndGenCorrelatorV2.pdf

Summary

- Ramp-up of the ALMA data rate has been slower than anticipated, allowing us to stay ahead of the curve.
- Data flow mostly from JAO to ALMA Regional Centers, with data processing mostly at JAO.
- Still monitoring how the network performs when transferring ~10TB/day in multiple parallel streams.
- Would like to establish a link with 10Gb/s of dedicated bandwidth within the next 1-2 years to improve transfer speed to and from Chile for bulk reprocessing, and to help with occasional large data and metadata transports (e.g. a DB export)
- Most new developments (e.g. next generation correlator) on ~5yr timescale can probably be accommodated without increasing the data rate by more than a factor ~4