



# Data transfer from ALMA to North America

David Halstead, Mark Lacy  
National Radio Astronomy  
Observatory

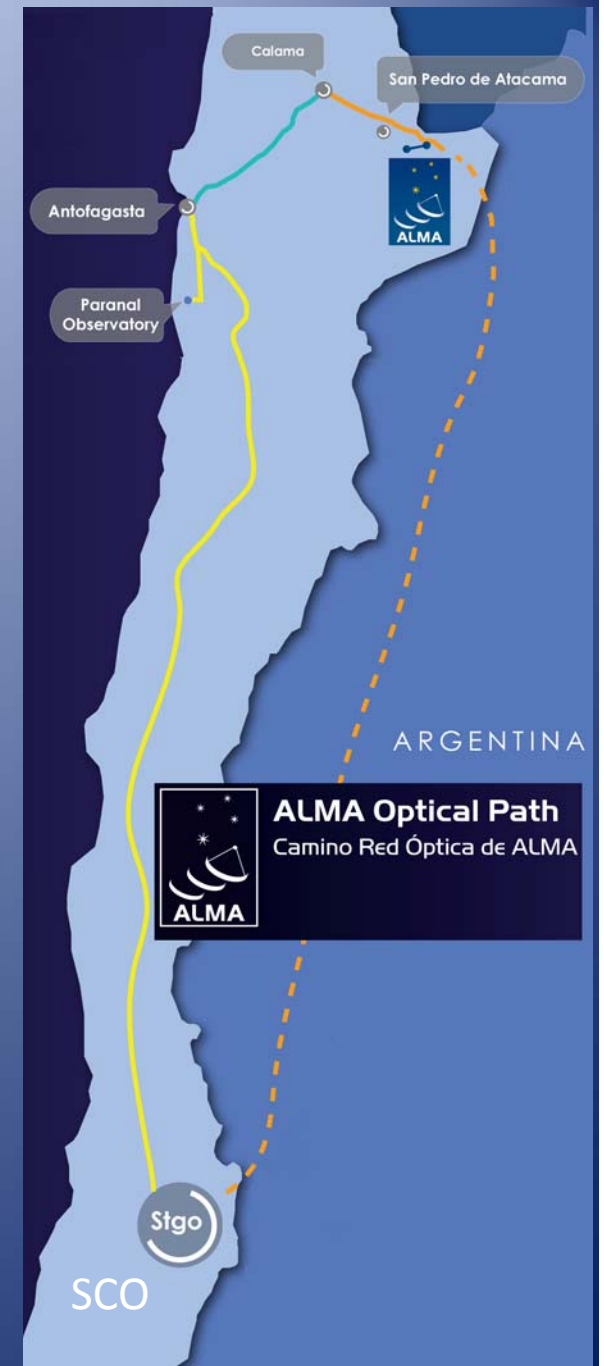


[www.almaobservatory.org](http://www.almaobservatory.org)



# Data Transfer within Chile

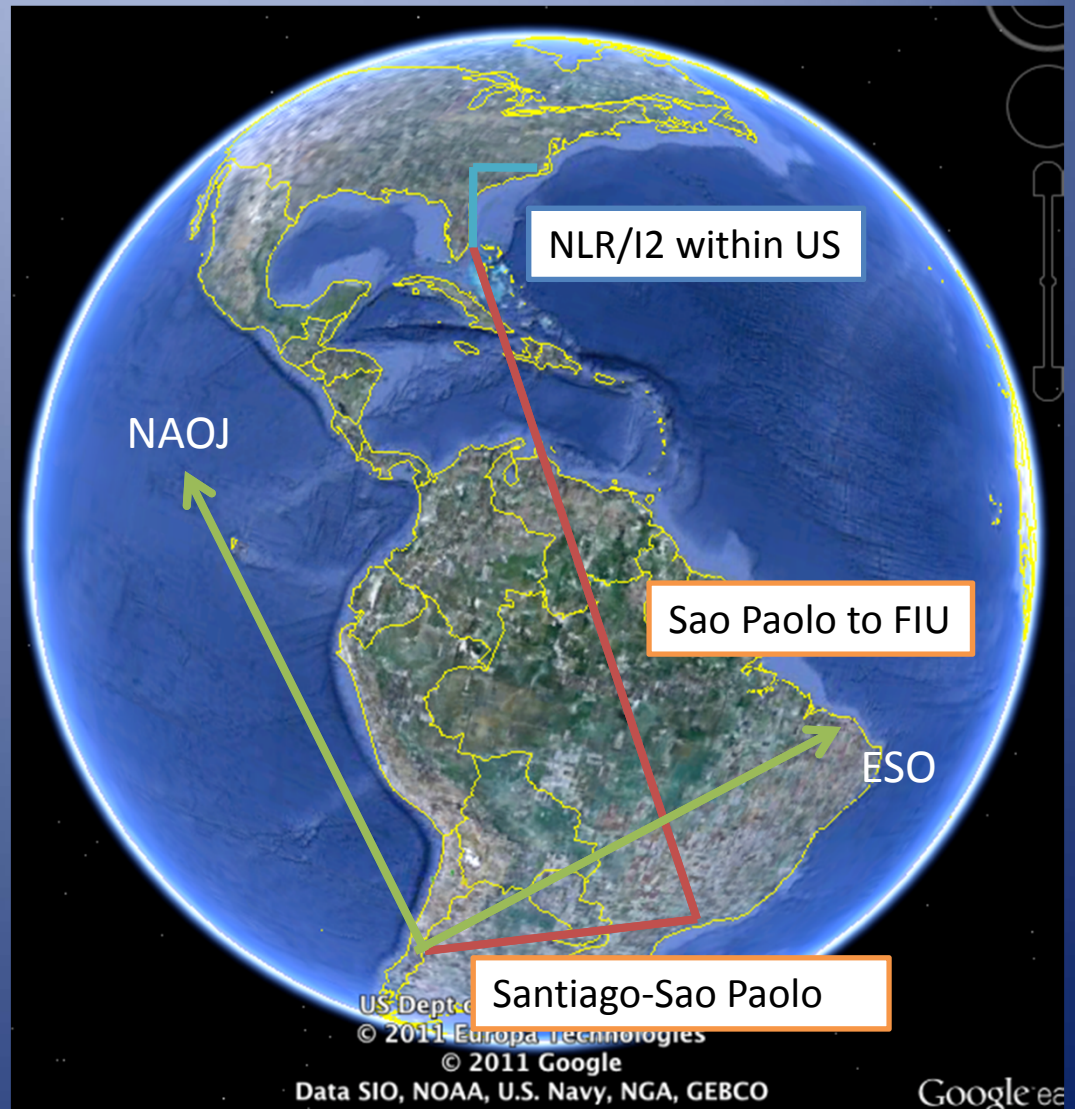
- AOS to OSF: 48 dedicated fibers
  - (1Gbp/s upgradable to 10Gb/s)
- AOS to Santiago upgraded in Jan 2015 to 2.5Gb/s with a new fiber link to Calama.
  - Calama to Antofagasta provided by Telefonica
  - 2.5Gb/s from Antofagasta to SCO from EVALSO/REUNA
  - Should start operations shortly
  - Redundant fiber loop via Argentina planned
- Pipeline run in Santiago, will ultimately produce L2/L3 data products, increases data size by 10-30%.
- Primary ALMA archive in Santiago (SCO)
- Santiago to ARCs: individual ARC contracts with REUNA and NRENs



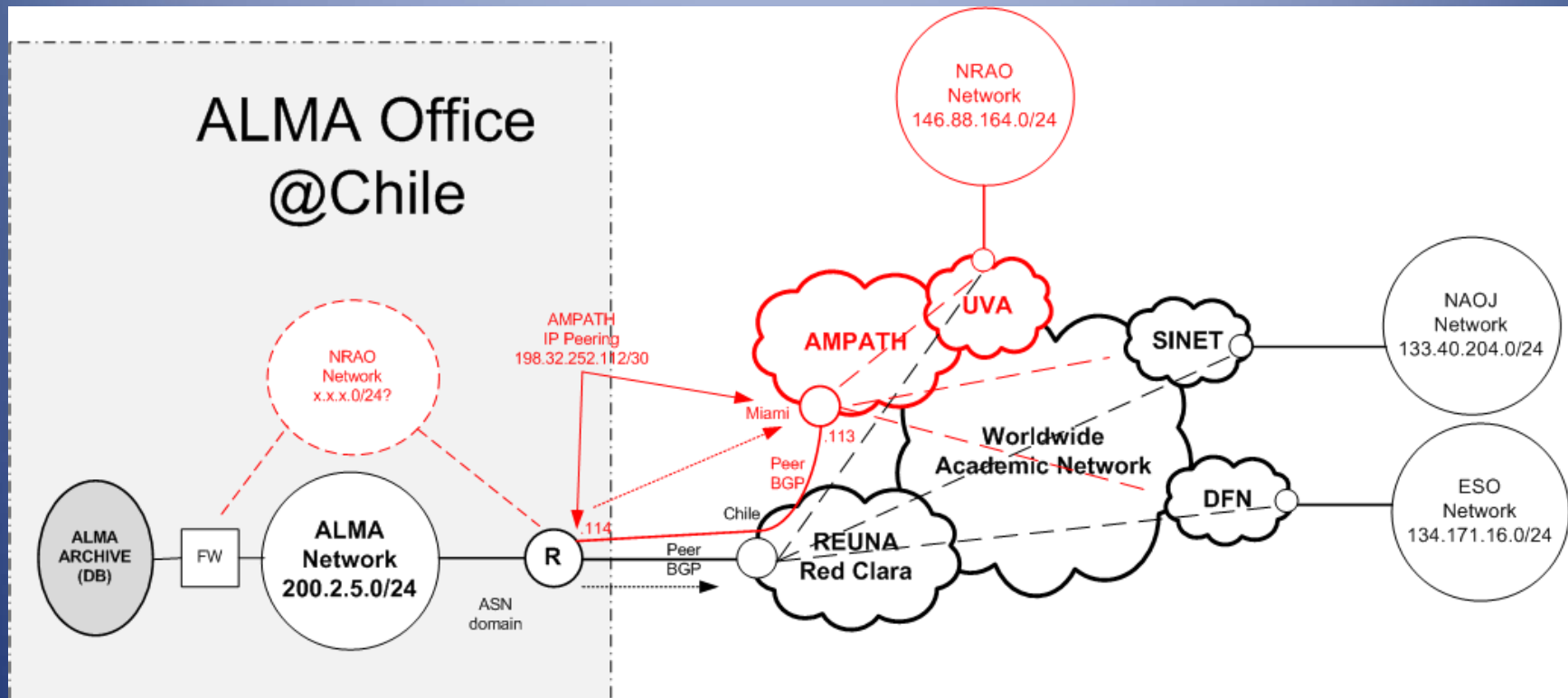


# Data transfer – Chile to NA

- Joint AURA-AUI agreement for NRAO to have 100Mb/s committed (burstable to capacity) of AURA's 622Mb/s link to Chile through Sao Paolo and Miami (FIU/AmLight) to the US research network backbone (NREN).
- MOU signed between AUI/REUNA
  - local link to SCO.
  - implementing international links.
- Will need 1Gb/s to the ARCS by 2016.
  - Large data volumes now being seen.
  - High rates  $\gg$  100Mb/s needed for specific operations over  $\sim$ 10hr periods.
    - e.g. time critical science verification data ( $\sim$ 1TB).
    - VLBI data ( $\sim$ 2.5Gb/s)?
- **100Mbps from SCO to REUNA for NRAO**
  - Requested upgrade to 1Gb/s with JAO/REUNA
- MOUs in place between AURA/AUI and AUI/REUNA.



# Paths from ALMA



Note: NRAO will be abandoning 146.88.x.x IP space in the next month

# ALMA Science data rate evolution

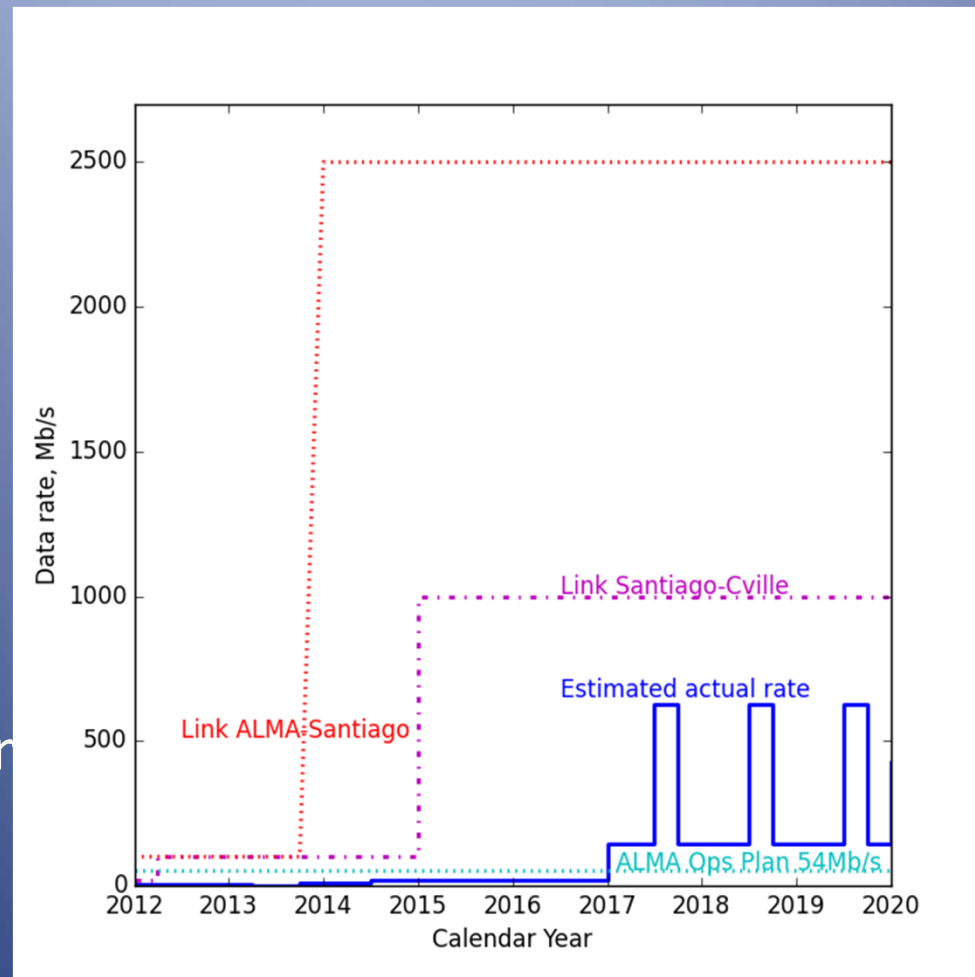
- ALMA Cycle 0 completed (Oct 2011-Jan2013)
  - 16-24/50 antennas used (data rate proportional to square of antenna number)
  - ~5-10% of array time for science
  - Data inflated to supply users with intermediate products
  - Total data volume was about 20TB
- ALMA Cycle 1/2 in progress (~Aug 2013-Nov 2015)
  - 32-40/50 antennas, plus 7-12/12 compact array
  - ~20% of array time for science
  - No intermediate products stored
  - Data rate about 40TB/yr (ALMA archive hit 80TB in Dec 2014)
- ALMA Cycle 3 starts in Nov 2015, ends November 2016
  - ~36 main array antennas, all 12 compact array
  - ~25% of array time for science
  - Best guess data volume is around 50TB for the 1 year cycle.

# Future Cycles

- Data rates will increase as we transition into full operations over next few cycles.
- Automated data pipeline will be fully implemented over next ~1-2 yr - needed to replace “human pipeline” for imaging and for calibration of the ~25% non-standard datasets (calibration pipeline for standard datasets is now in operation).
- Unthrottled, we expect data rates driven by a typical mix of science use cases to be ~100Mb/s, with seasonal peaks ~600Mb/s in the Austral winter/spring.

# Current data rate projections

- Assumes no imposed limit on data rate (cyan line is current Operations Plan rate).
- Blue line is for data generation
- Data transmission is per ARC



# Summary

- New fiber link from the telescope to Calama will soon allow 2.5Gb/s transfer rates to/from Santiago for both data flow and remote ops.
- Ramp-up of ALMA data rate has been slower than anticipated, allowing us to stay ahead of the curve overall.
- However, we have identified need for bursts of high data rate (>1Gb/s) even with the current observational parameters.
- Data rate expected to grow steadily to design value of 200 TB/yr over the next 2-3 years. Will need bursts of high data rates to deal with some projects.