



# Data transfer from ALMA to North America

David Halstead, Mark Lacy  
National Radio Astronomy  
Observatory



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



# ALMA telescope

- Largest mm/submm telescope ever built.
- Interferometer – combines signals from multiple antennas to form an image.
- Inauguration occurred at the OSF on March 13<sup>th</sup> 2013
- All 66 antennas delivered, all at high site (except for maintenance).
- Multinational project with many partners, three ALMA Regional Centers (ARCs): US, EU and EA
- Operated “space mission” style, with pipeline data processing and a science archive at each ARC allowing data reuse.
- First PI projects released to public from the ARCs January 2013
- Cycle 3 observations began in October 2015.

# Data Transfer within Chile

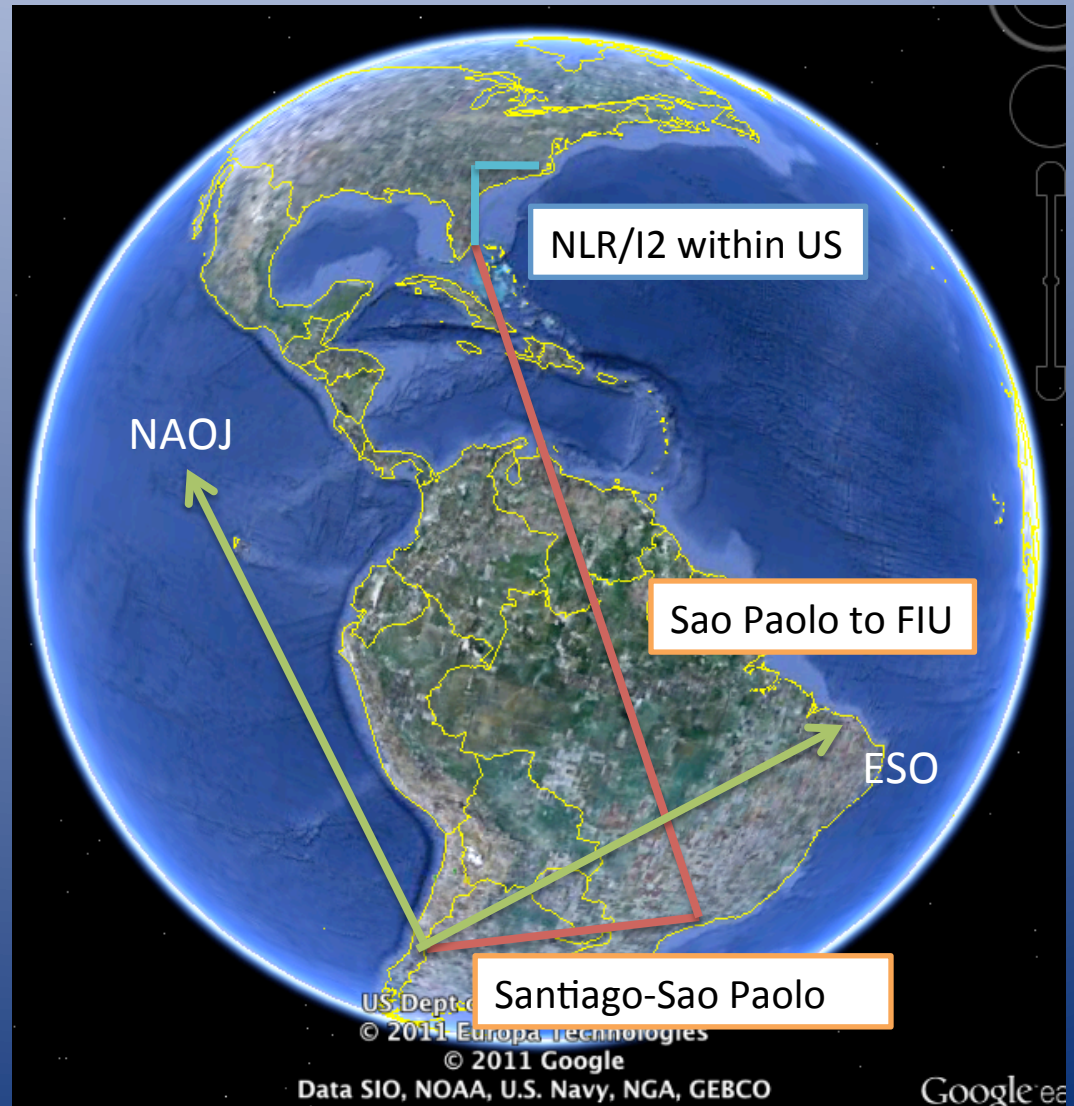
- AOS to OSF: 48 dedicated fibers
  - (1Gbp/s upgradable to 10Gb/s)
- OSF to Santiago currently 100Mb/s (part microwave)
- Upgrade project from ALMA development:
  - OSF to Calama fiber built 2014; waiting on revised environmental impact report before use.
  - Calama to Antofagasta provided by Telefonica
  - 2.5Gb/s from Antofagasta to SCO from EVALSO/REUNA
  - Redundant fiber loop via Argentina planned
- Pipeline run in Santiago to 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



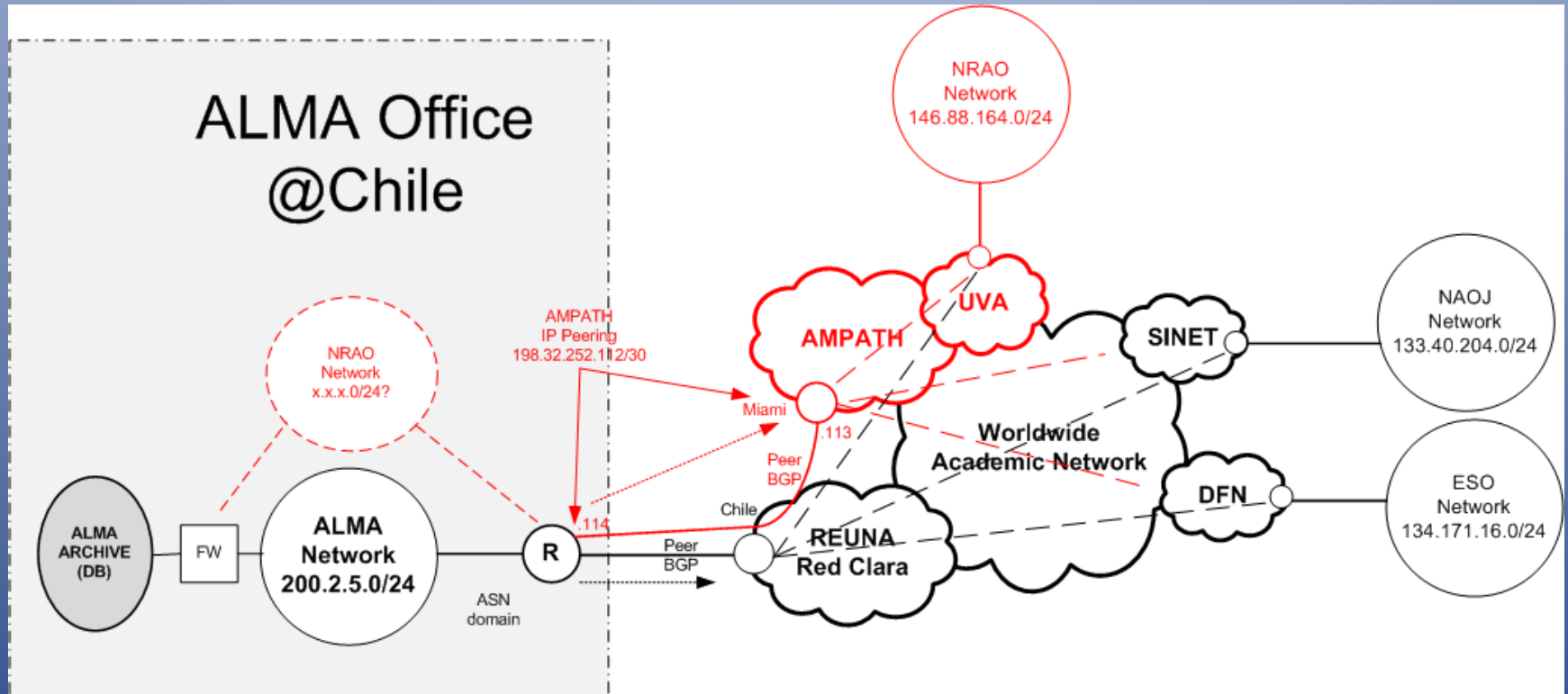


# Data transfer – Chile to NA

- 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.
- Typical rate obtained during peak data transfer periods is 2-300Mb/s, with bursts up to 600Mb/s.
  - 90% is bulk data with low QoS.
  - Remainder is database sync and telepresence
- MOUs in place between AURA/AUI and AUI/REUNA
- Opportunities for improving Chilean astronomer access under consideration



# Paths from ALMA



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

# ALMA Science data rate evolution

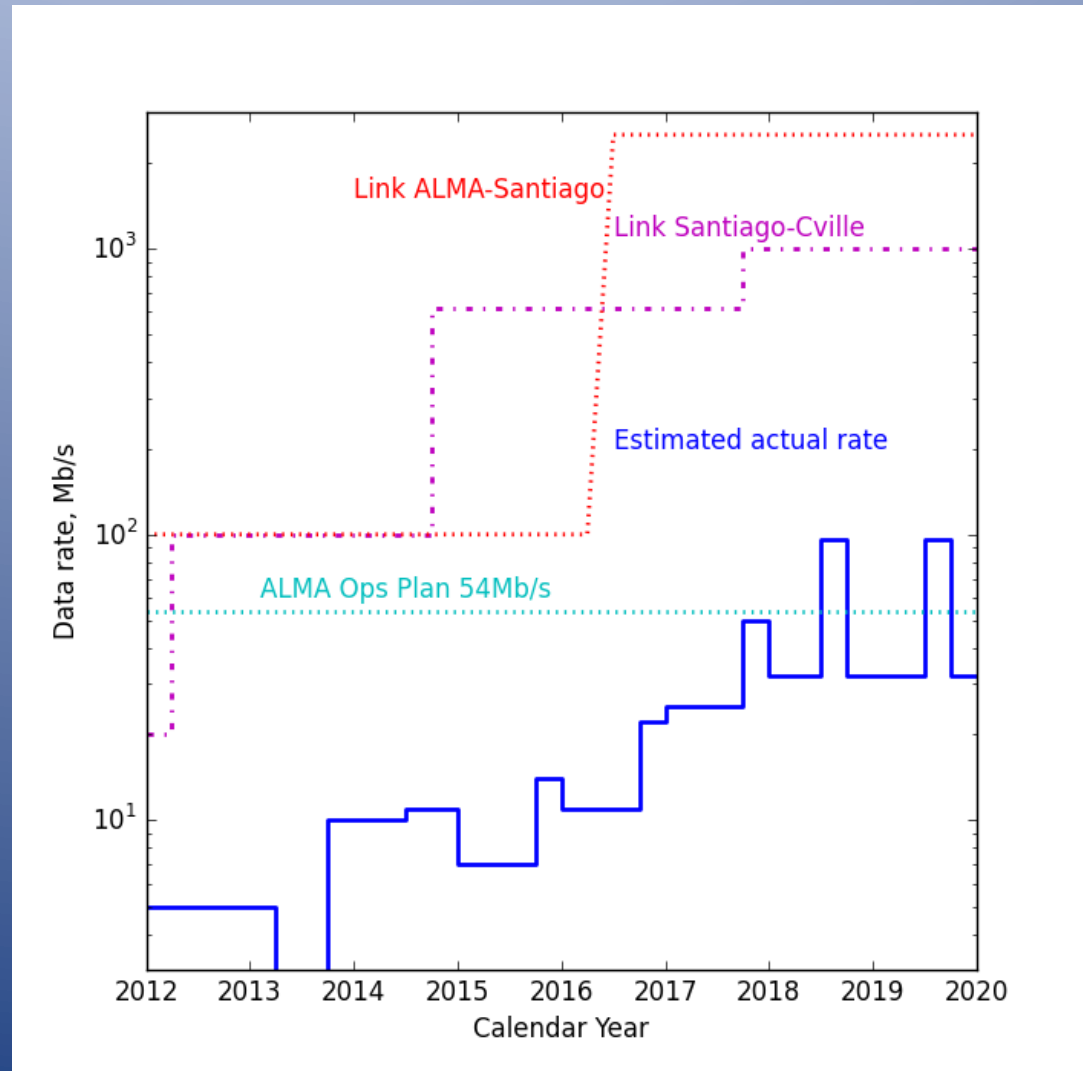
- ALMA Cycle 0 completed (Oct 2011-Jan2013)
  - 16-24/50 antennas used (data rate proportional to square of antenna number)
  - ~5-10% of array time for science
  - Data inflated to supply users with intermediate products
  - Total data volume was about 20TB
- ALMA Cycle 1 complete (~Aug 2013-Jun2014)
  - 32-40/50 antennas, plus 7/12 compact array
  - ~10% of array time for science
  - Users will not get intermediate products, better software means unnecessary data not taken.
  - 40TB over 1yr (ALMA archive hit 50TB in March 2014)
- ALMA Cycle 2 complete (June 2014-Sept 2015)
  - ~34 main array antennas, 10 compact array
  - ~15% of array time for science (but some carryover from Cycle 1)
  - 70TB in a 17 month Cycle.
- ALMA Cycle 3 started Oct 2015, runs until Sept 2016.
  - 36 main array antennas, 10 compact array
  - ~25% of array time for science
  - Estimate total of 80TB in 1 year.
  - Mean data rate of 40Mb/s during observations.

# Future Cycles

- Data rates will increase as we transition into full operations. Expect Full Science cycles (~2018 onwards) to have mean data rates ~100Mb/s during observations, but could be ~50% higher. “Duty cycle” of observations will also increase (by about a factor of two) as testing and maintenance procedures improve.
- Important to note that data rates vary through the configuration cycle. When long baseline configurations are scheduled the data rate goes up for two 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 comparable to the raw data in the largest configurations.
  - Large configurations tend to be scheduled June-November due to better weather conditions.

# 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

- In December last year the ALMA archive in North America reached 150TB, all transferred from Chile – our thanks to the SAACC and our colleagues at AmLight, REUNA and JAO!
- Ramp-up of ALMA data rate has been slower than anticipated, allowing us to stay ahead of the curve.