



Data transfer from ALMA to North America

David Halstead, Mark Lacy
National Radio Astronomy
Observatory



www.almaobservatory.org

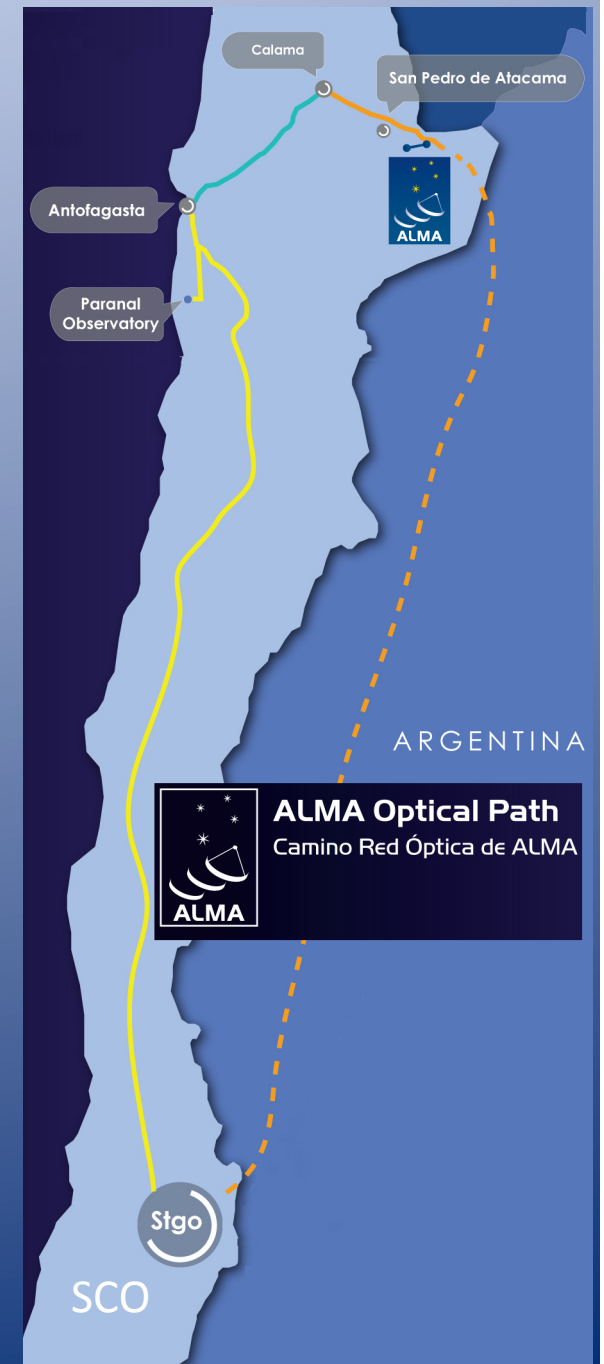


ALMA telescope

- Largest mm/submm telescope ever built.
- Interferometer – combines signals from multiple antennas to form an image.
- Inauguration occurred at the OSF on March 13th 2013
- All 66 antennas delivered, ~50 accepted
- Completion scheduled for late 2013
- 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.
- Cycle 0 PI projects released to public at ARCs January 2013
- Now moving to Cycle 1 observing

Data Transfer within Chile

- AOS to OSF: 48 dedicated fibres
 - (1Gbp/s upgradable to 10Gb/s)
- OSF to Santiago: 100Mb/s
 - Upgrade to 2,500Mb/s 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/NRENS
- 100Mbps from SCO to REUNA for NRAO
- MOU signed between AUI/REUNA



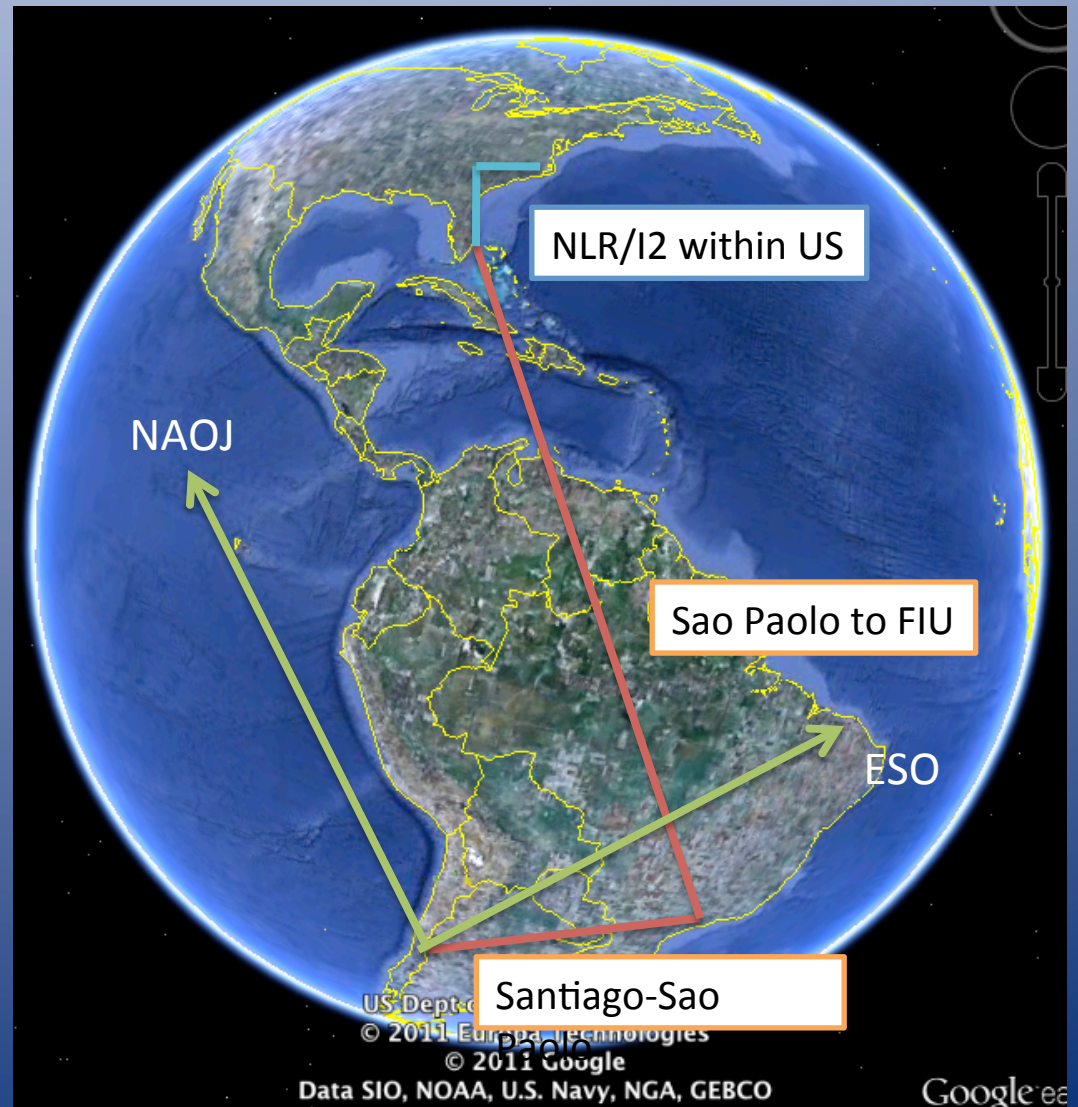
Data transfer – Chile to NA

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) to the US research network backbone (NREN).

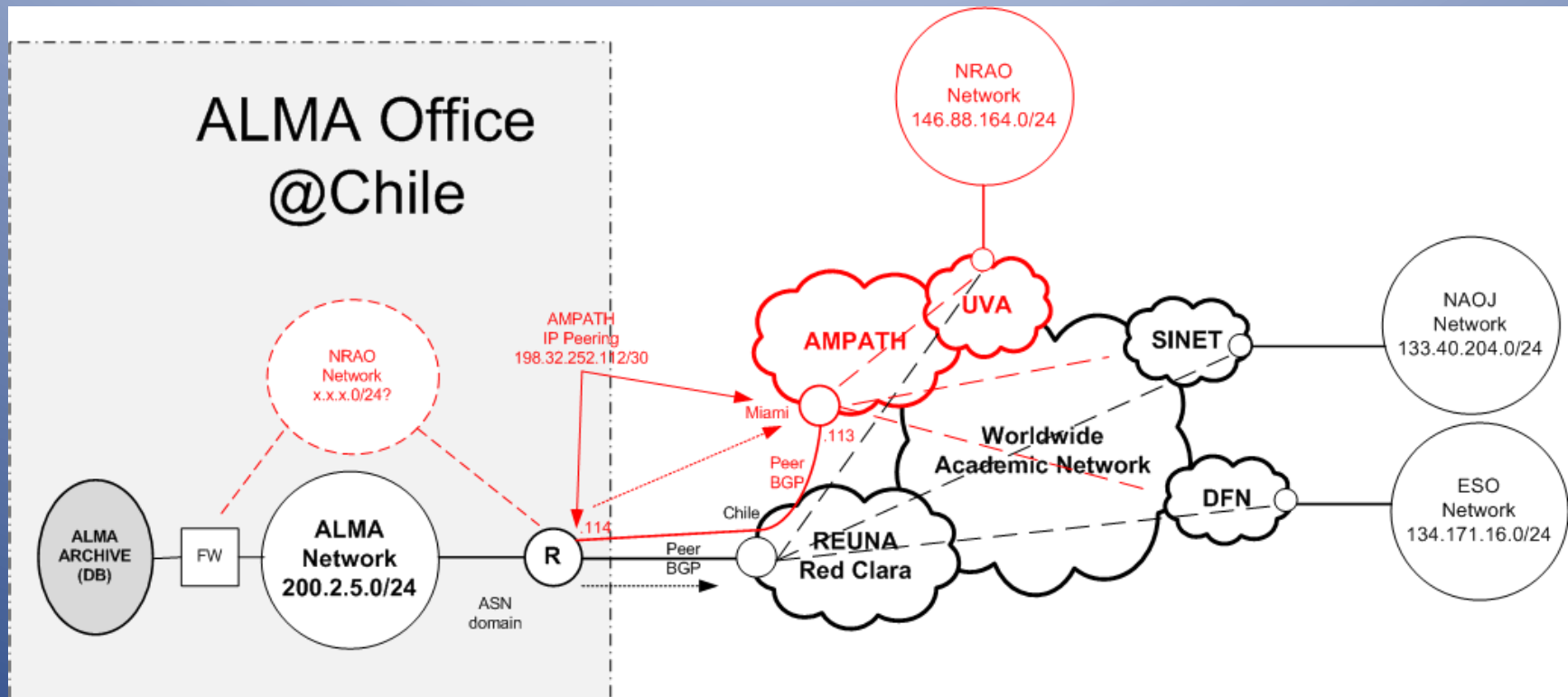
MOU signed between AUI/REUNA

Will likely need ~1Gb/s (10% of 10Gb/s) by ~2016.

90% is bulk data with low QoS. Remainder is database sync and telepresence



Paths from ALMA



Note: NRAO will be abandoning 146.88.x.x IP space in the next 4-6 months

ALMA Science data rate evolution

- ALMA Cycle 0 just completed (Oct 2011-Jan2013)
 - 16-24/50 antennas used (data rate proportional to square of antenna number)
 - ~5-10% of time for science
 - Data inflated to supply users with intermediate products
 - Total data volume was about 20TB
- ALMA Cycle 1 starting (~Apr 2013-Apr2014)
 - 32-40/50 antennas, plus 7/12 compact array
 - ~10% of time for science
 - Users will not get intermediate products, better software means unnecessary data not taken.
 - Best guess is about 40TB over 1yr
- ALMA Cycle 2 (~2015)
 - ~40 main array antennas, all 12 compact array
 - ~25% of the time for science
 - Data rates limited to ~50TB in 1 year.

Future Cycles

- Current Operations Plan uses an ALMA data rate estimated a decade ago, before the scientific capabilities of ALMA were properly understood.
- Effort underway to treat data as a resource, similar to observing time.
 - This will result in impacts on the scientific capability of ALMA
 - Time averaging will decrease the calibration quality
 - Frequency averaging will reduce the accuracy of continuum estimation
 - Gathering data over less bandwidth will reduce the ability to make serendipitous discoveries, and thus the archival value of the data.

Alleviation plans

- Need to ensure that data transfer and storage costs are kept as low as possible to allow ALMA to reach its full potential.
- Pooling of data transfer from Santiago to the Northern hemisphere and efficient distribution between the ALMA regional centers an important part of this plan.

Current data rate projections

- Assumes no imposed limit on data rate (cyan line is current Operations Plan rate)

