Current Data Practices & Issues

National Radio Astronomy Observatory

Brian Glendenning

Data Management and Software Department

Atacama Large Millimeter/submillimeter Array

Karl G. Jansky Very Large Array

Robert C. Byrd Green Bank Telescope

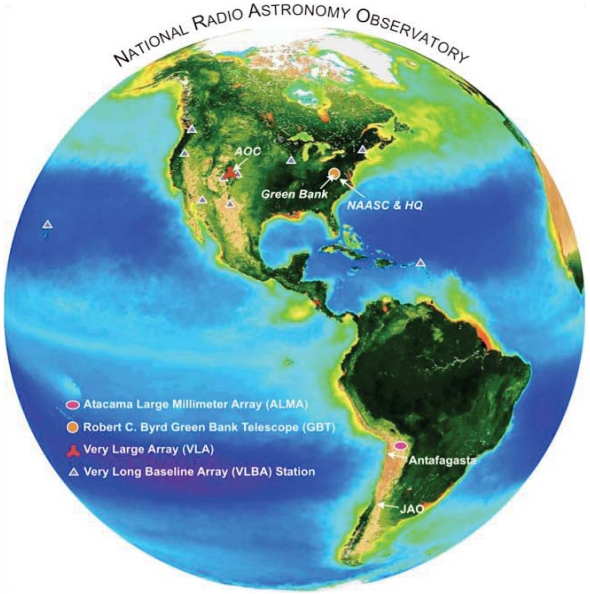
Very Long Baseline Array

NRAO: Four best-in-class radio telescopes

DMS

Jansky VLA

GBT

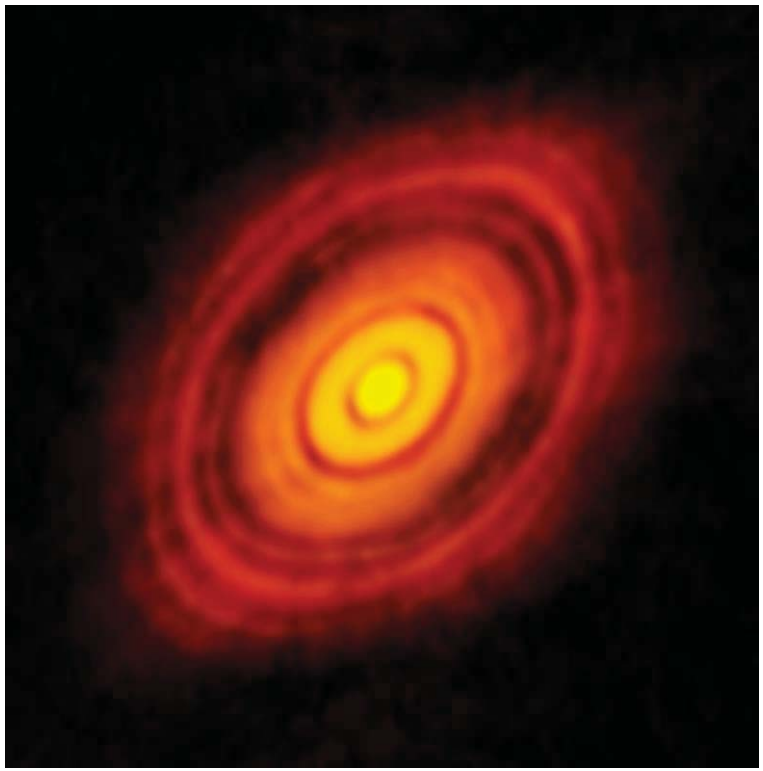


VLBA

ALMA

ALMA=ESO/NAOJ/NRAO/Chile

Images



ALMA: HL Tau Protoplanetary Disk

Estimated press circulation: 0.5B

VLA: Manatee Nebula (W50)

NRAO Telescope Suite

- Observing measurements
  - Premiere radio telescope system in the world
  - Tremendous breadth!
  - Wide Frequency Range: <1 – 1000 GHz
  - Wide Size Scale Range: 0.0002 arcsec – 10 arcmin
  - Continuum & Spectral (10,000 channels typical for spectral)
  - Measures all 4 polarizations
- Observing System
  - Hundreds of peer reviewed PI proposals (10 – 100 hours typical, VLA Sky Survey = ~5000 hours)
  - Oversubscribed, heavily for ALMA (pent-up demand)
  - Files in (proposals, programs), raw data out
    - Commissioning automated data processing

History: NRAO and NSF CI

- NSF Internet support has fundamentally enabled modern NRAO operations

- Relationship with HPC centers has been intermittent
  - NRAO data is too big to be easy but too small to be interesting
  - Occasional contacts going back to the late 1980s

Questions

2013 NSF Large Facilities Workshop

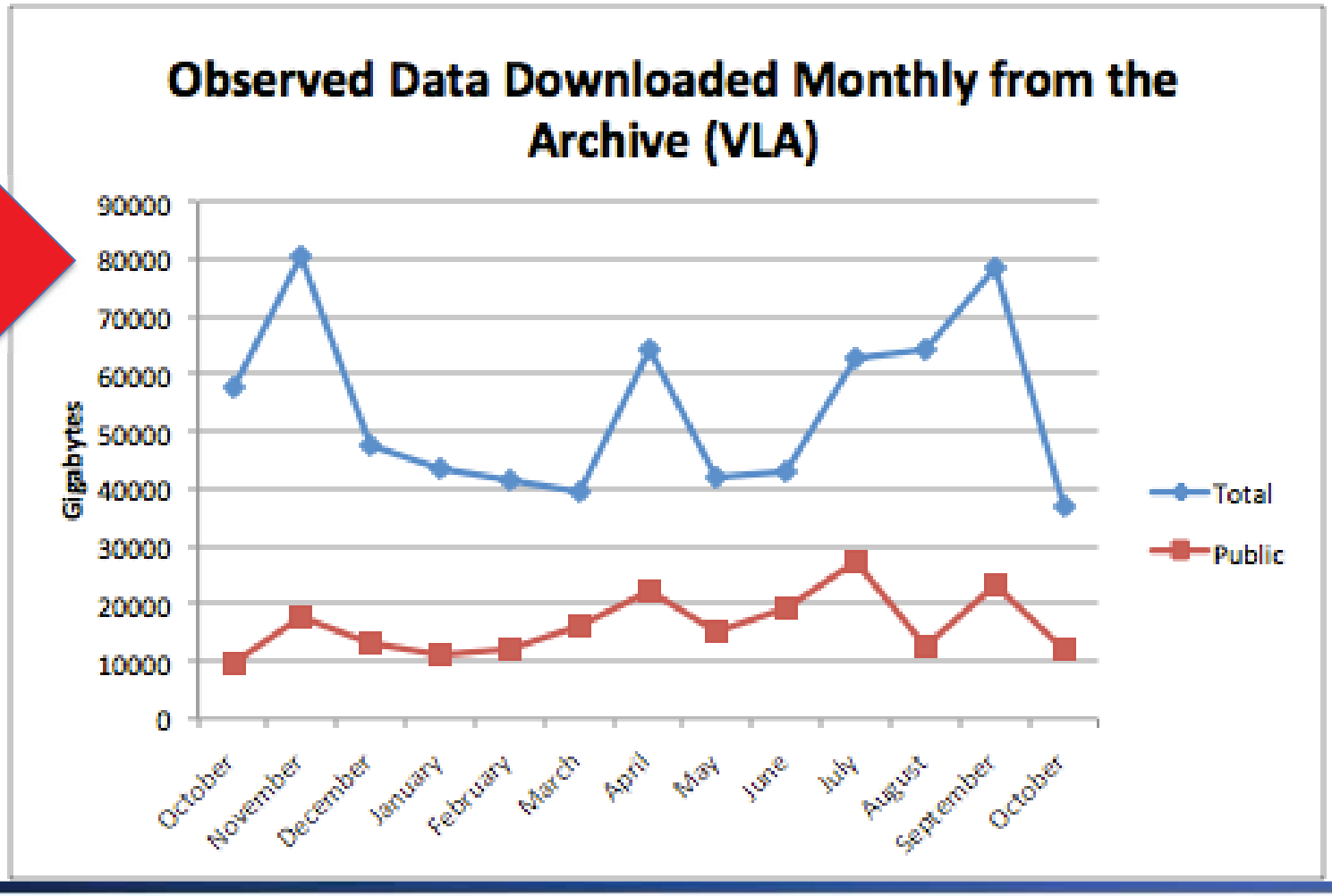
- What should NSF Facility "Data Management" best practices be?
  - Can/should this be formalized?
- How do we keep data management systems in construction project scope?
  - Often thrown out to obtain only modest cost savings
  - Construction projects often dominated by grizzled veterans
    - Data Management = chart recorder + HP-IHC
- Can the various national HPC centers/networks play more of a role?
  - Gap: big-ish data problems, hard for facility but not interesting for HPC research
- What metrics should we use?

Data Volumes - today

- Raw Data
  - Hardware capabilities: ALMA, GBT, VLBA ~1 GB/s = 30 PB/y
    VLA 16 GB/s = 0.5 EB/y
  - Actual data ingested into NRAO Archive (all telescopes) ~1 PB/y
  - Limited by: scientific need (e.g., most sources vary slowly), but also:
    - NRAO budget,
    - User pain of dealing with large data sets,
    - Size of in-house/user computational facilities,
    - Network pipe sizes, …
  - Possible game changer (if real): rapidly varying sources (ms timescales)
- Resulting Images
  - Typical: 1000^3 (Gpix) (x 1-4 polarizations)
  - Coming: > $10{,}000^3$ (Tpix),
  - VLA Sky Survey (if all images, channels kept): 6 Ppix

Data retrieval per month (VLA)

80 TB



DMS

Data Stewardship

Frank Press,
President Carter's
Science Advisor

- ***We have all the data for all the telescopes for all their existence***
  - And the data is still accessible
    - Need the bits, knowledge of the formats (80 character card images!, ½" tape record lengths!), and comprehension of the data models (software, humans)
- Had we given the raw data to a third party, would we still have it? How can we be confident we will still have it in 35+ years?
  - Feared path:
    - "This is really important, we'll help!"
    - "You don't need as much budget now!"
    - 10 years later: "Something else is really important!!"

Current In-house Computing

- We have several in-house computational/Lustre (I/O) cluster combinations use for various telescope and science operations purposes, as well as general community use ***(>200 user groups / year)***
  - Big users tend to have their own similar facilities, many small users rely on NRAO
  - ~150 16 core nodes, 3 PB Lustre storage
- Parallelization advancing, nearly through the entire processing chain
  - We are more high-throughput than high-performance computing at present
- Efficient, effective, …, but:
    - At the limit of what we can support
    - Limits data rate, reprocessing, hard for users to replicate

Example: VLA Sky Survey – So many pixels…

DMS

- ALL-SKY at 0.6" (36Mpix/deg^2) :
  - 34000 deg^2 : 1.22Tpix = 4.9TB per "image"
  - Continuum : 9 images = 44TB (plus 7 images, 3 epochs = 100TB)
  - Spectral Cubes : (1024ch, 5 images = 25000TB = 25PB)
    - 5.5 Ppix! This would be a lot of image pixels to sift through!
    - NRAO CANNOT AFFORD THIS! must compress <1%
      - Or calculate on demand (high compute or long latency)

- Computing estimate:
  - Coarse cubes (14 planes): 10-20,000 core-days per pass
    - Probably several passes per epoch, at least initially
    - ~OK
  - Fine cubes (900 planes ideally): x64
    - ***Not feasible for NRAO***

Computing comments

- Why do >200 user groups/year use our tiny facilities?
  - Because it's easy (software installed, logins, easy access to archive data, interactive use allowed)
  - Any alternate system has to be easy
- Our algorithms tend to be high in I/O / FLOP, HPC centers tend to be high in #cores / storage
- Lots of users each of whom need a relatively modest number of cores
  - Get data 1-2 times per year, "bursty"
  - Increasing automation, but still a lot of interactive use
- Far more buzz in observational astronomy about Amazon AWS than (e.g.) XSEDE
- Many international aspects to astronomy software; US only APIs are a problem
- Our software packages have long lifetimes (20+ years is the "new" one)
  - Good and bad – embeds a lot of knowledge, can be hard to adapt

We are not unusual (slide courtesy M. Norman, SDSC)

HPC for the 99%

Idea 1: *architect* an HPC system where 99% of the jobs run inside a single rack with full bisection BW

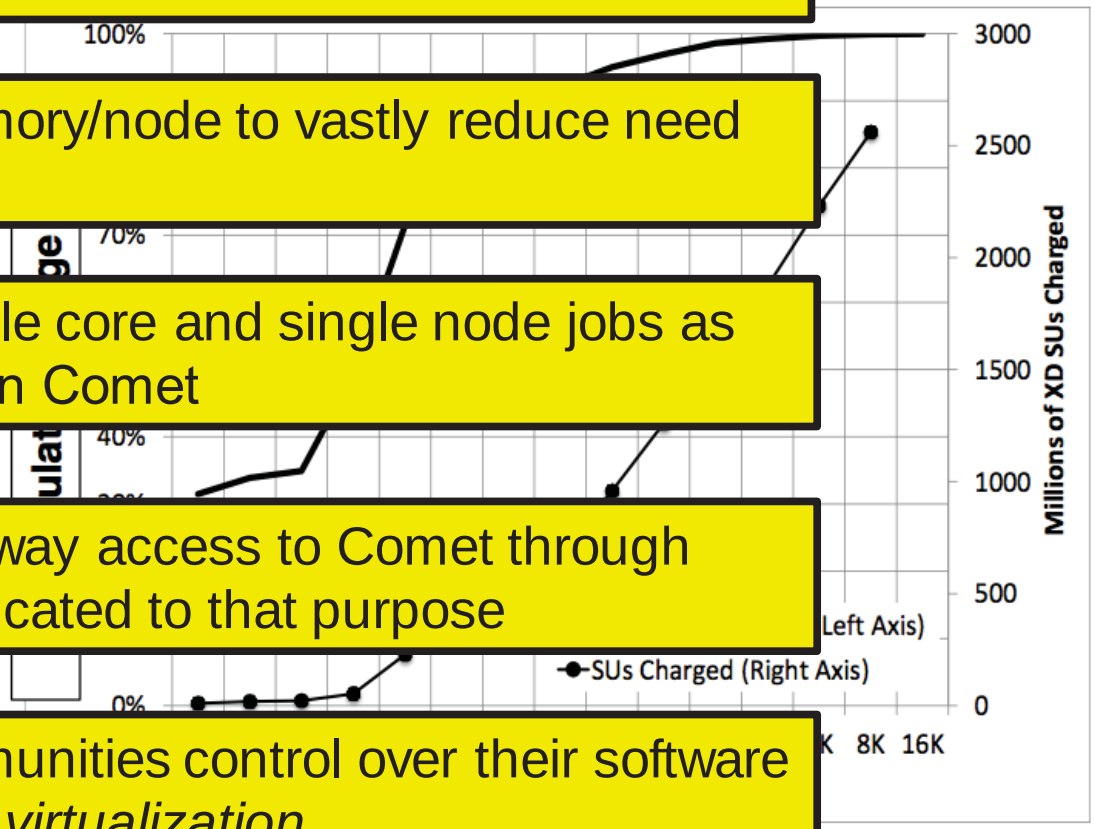
- **99% of jobs run on NSF resources in 2012 use**
- **And >50% of the total core NSF resources**

Idea 2: *increase* memory/node to vastly reduce need for multiple rack runs

Idea 3: *embrace* single core and single node jobs as "first class citizens" on Comet

Idea 4: *support* Gateway access to Comet through ancillary servers dedicated to that purpose

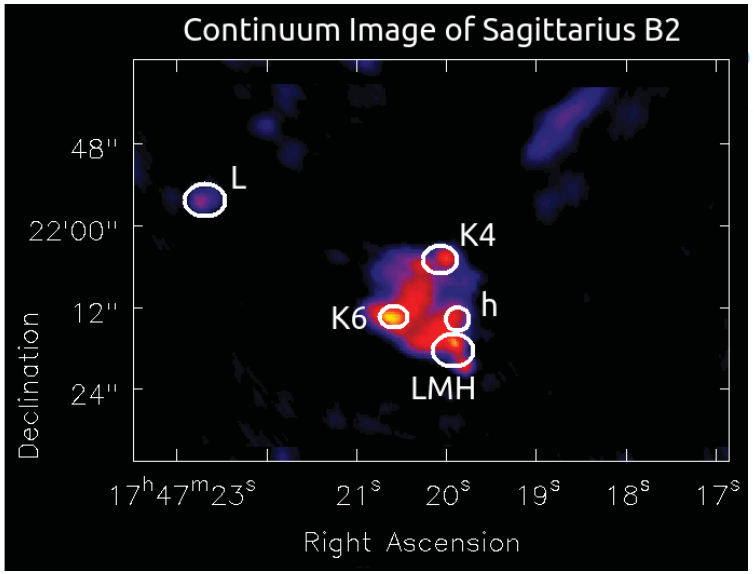
Idea 5: *provide* communities control over their software environment through *virtualization*

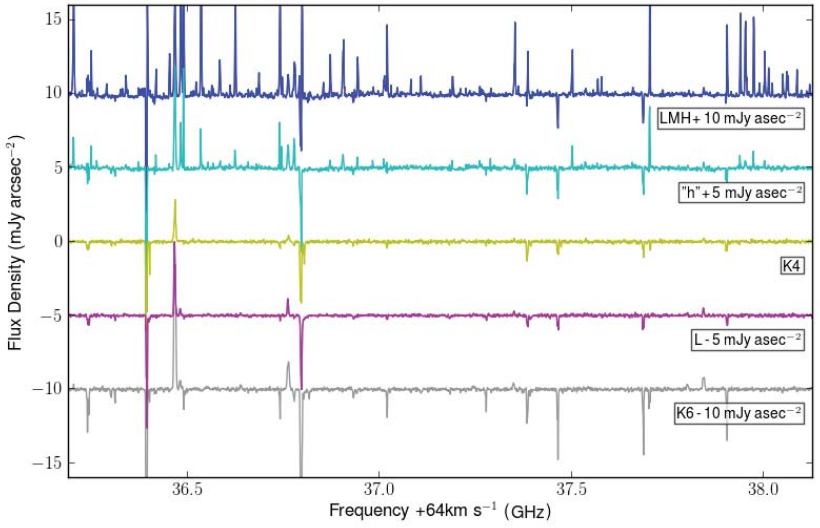


SDSC Comet seems well aligned with NRAO needs

Visualization

- We typically have 3+D images a spectrum at every pixel
- Complex physics, chemistry, …
- Terapixel images will become commonplace
  - All sky cubes = petapixel
- No good plan in place for visualizatioh, modeling
  - Both big data & visualization aspects
  - We have surveyed existing tools





Radio Astronomy - Future

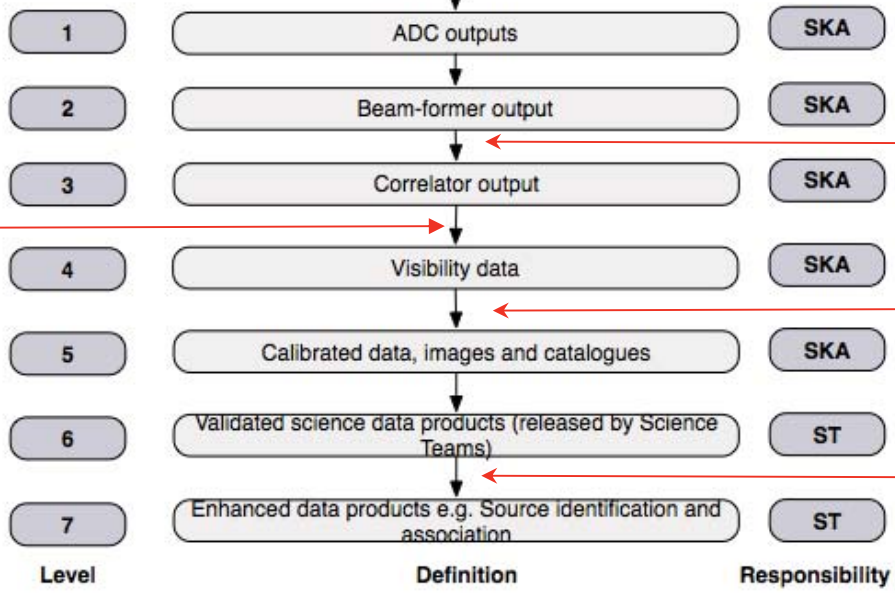
- Square Kilometer Array (and precursors) in various stages of construction, pre-construction
  - No formal US involvement
  - Australian Pawsey supercomputing centreaimed at the Australian *precursors*
- NRAO is participating in science/technical discussions of "next generation" VLA, (~10x antennas, ~10x antenna separation)
  - Comple,entary to SKA, ALMA
  - Idea is to propose to "Astro 2020" process for construction start next decade
  - ngVLA will require Exaflop computing
    - We will need partner(s) to develop a credible proposal and then build the telescope
    - Industrial scale; no possibility of building in-house

Data flow

Low frequency aperture array

Dish arrays

Slides courtesy SKA project



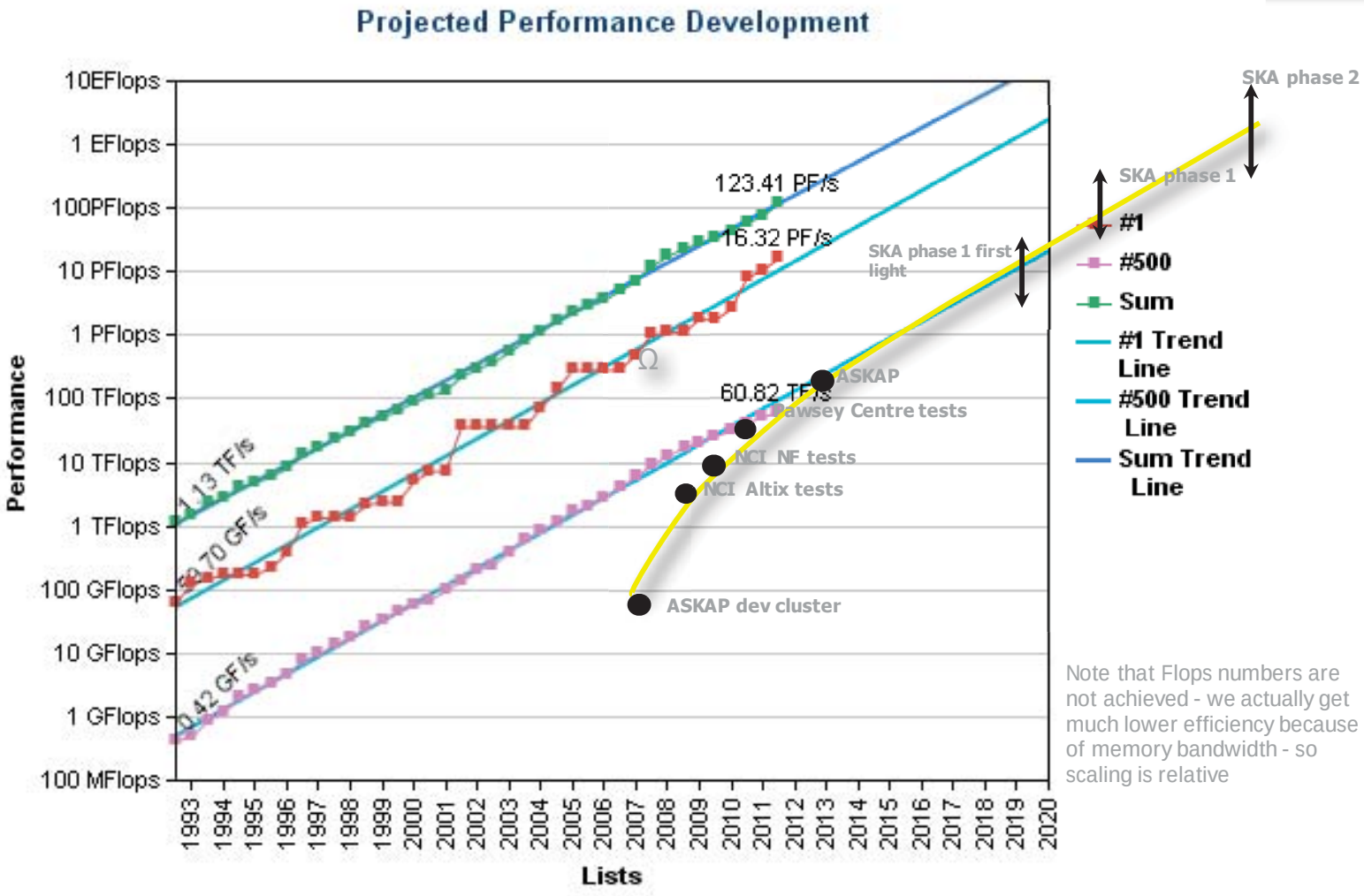
0.3 to 3 TB/s

10 - 500 TB/s

~ 100 PB data set read multiple times over several days

e.g. 1 year Redshifted Hydrogen survey ~ 4EB

SKA data processing rates



Note that Flops numbers are not achieved - we actually get much lower efficiency because of memory bandwidth - so scaling is relative

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc

www.nrao.edu