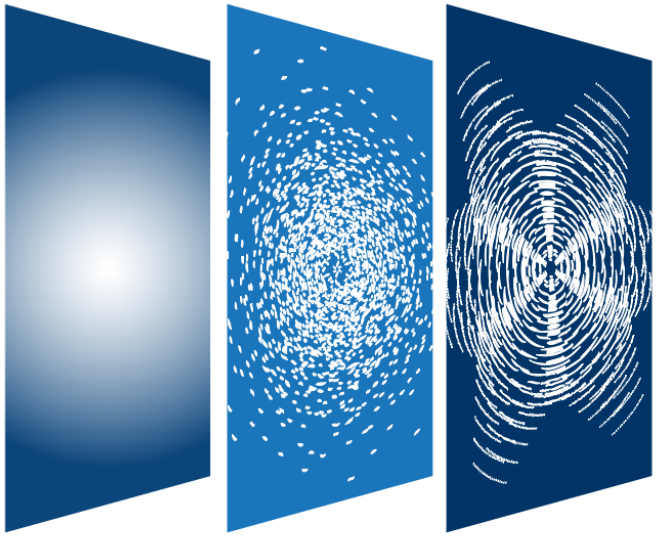


CASA

Common Astronomy Software Applications

AWS/XSEDE

Alternative Compute Resources

James Robnett
DMSD

NAOJ National Astronomical Observatory of Japan

ESO

NRAO

ASIAA

AWS and XSEDE

- AWS: Amazon Web Services
  - Amazon's cloud computing platform
  - Largest collection of compute resources in the world
  - Use based cost structure
- XSEDE: Extreme Science and Engineering Discovery Environment
  - NSF Funded consortium of Supercomputing facilities
  - Replaces NSF TeraGrid
  - Free to use but only available to United States PIs

AWS Overview

- Why Amazon (versus Microsoft Azure or Google)
  - Dominating market, richest set of tools and resources
  - Resource price is roughly the same across vendors so porting labor dominates cost
  - May examine other options in the future
- What is AWS
  - Collection of physical resources and software tools which enable ad hoc creation of compute environments.
  - Create single workstations or arbitrarily large clusters to match processing.
  - Create shareable storage pools ranging from GBytes to PBytes
  - Pay as you go cost model

AWS Resources and terms

- Regions (separated by 100s of miles) and zones (10s of miles)
  - Define boundaries for resource interaction
- Amazon Machine Image (AMI): essentially an OS image
  - Canned and user generated images supported (e.g. NRAO CASA images)
  - Some directives wired into image (VM type, root device size, etc)
- Instance (specification of cpu and memory)
  - 1 to 100s of cores, 1 to 1000s of GB (2TB)
  - GPUs available
  - Can be shared or dedicated hardware
- Elastic Block Storage (EBS)
  - Local SSD for some instance types
  - SAN-esque SSD or HDD can be mapped to instances
- Elastic File System (EFS): pNFS based distributed filesystem
- Simple Storage Service (S3) and Glacier for long term

XSEDE

- 16 Major facilities nation wide
  - Resources granted through quarterly proposal/review process
    - Million core hour, Multi-Tbyte requests common
  - Primarily large scale general purpose Linux clusters
  - Individual facilities supply roughly 50x the processing and 5x the storage of existing NRAO clusters, architecturally very similar.
  - Typically memory and storage limited compared to NRAO clusters
- Primarily focused on broad parallel large jobs (1000's of cores)
- Some systems (e.g. SDSC Comet) focused more on HTC, many small job or narrowly parallel job submissions
  - Shared compute resource model can trigger contention
- NRAO still investigating for operational fit

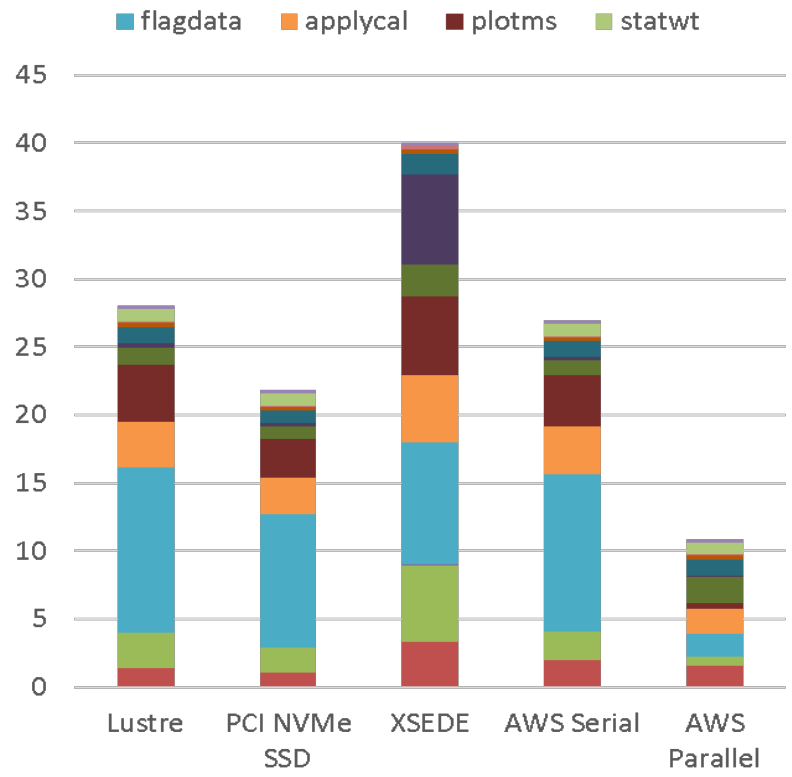
CASA on AWS/XSEDE

- Installation effectively just 'tar –xz <tarball>', a few secondary packages may be necessary.
  - Most labor is in automation, state tracking, data transfer
- NRAO supported CASA installations
  - NRAO generated AWS public machine images exist
  - NRAO supported installation on SDSC/Comet (XSEDE)
- Growing documentation repository, tutorial planned for this year
- DMSD HPC group will assist community with both environments
  - Driven by interest
  - Limited by competing projects
- Highest utility for both AWS/XSEDE is in automated, batch processing, neither is terribly effective for interactive processing.

Processing Comparison

- TSKY0001.sb32295801.eb32296475.57549.31722762731
- Observed: 2016-06-10T07:36:49
- ASDM: 348 GB
- Processed MS: 1018 GB

- Lustre: AOC 1PB Lustre system
  - Operational 9/23/16
- PCI NVMe SSD
  - Intel P3608 1.2 TB NVMe SSE
- XSEDE: COMET SDSC
  - Failed due to memory limitations
- AWS Serial
  - 1.5 TB provisioned storage
    - 20K I/O Ops
- AWS Parallel
  - Same as Serial
  - 8 Processing Engines + 1 Management



Useful Links

- AWS
  - https://aws.amazon.com/documentation/
  - Compute: https://aws.amazon.com/documentation/ec2/
  - Storage: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Storage.html
  - Pricing: https://aws.amazon.com/ec2/pricing/
  - NRAO https://casa.nrao.edu/casadocs/casa-5.1.1/casa-on-amazon-web-services
- XSEDE
  - https://www.xsede.org/overview
  - https://www.xsede.org/using-xsede
  - Facility specific user guides: https://www.xsede.org/userguides
  - https://www.xsede.org/web/guest/usage-policies
  - Writing successfully proposal webinar: https://www.xsede.org/web/xup/course-calendar
  - NRAO: Planned for late spring