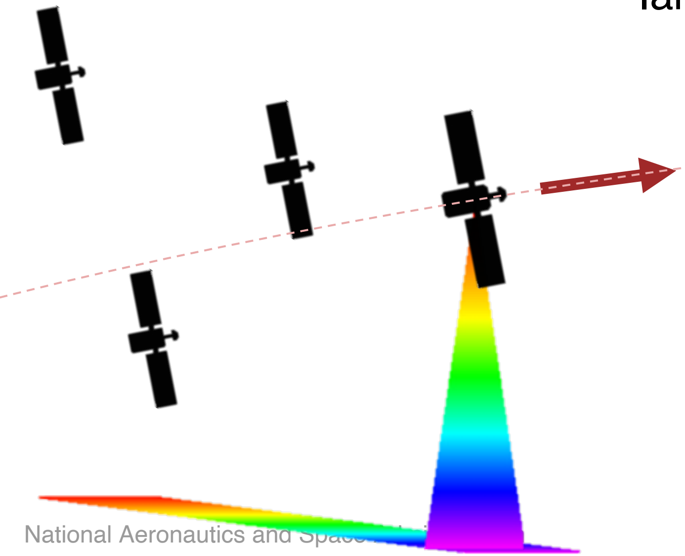
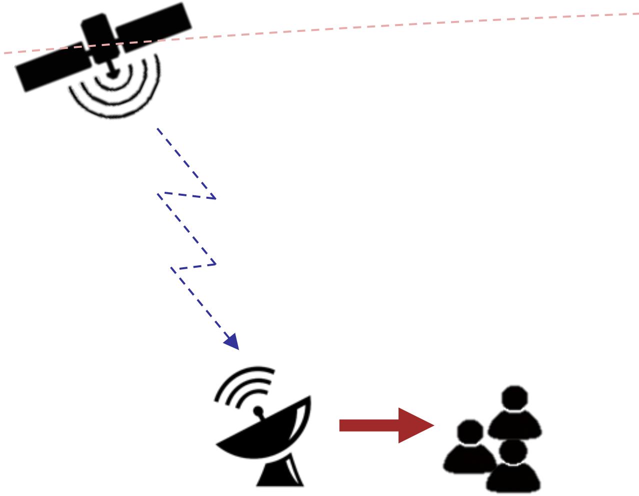
NASA's High-End Computing Capability: Growing to Support Science Data Processing for the Earth System Observatory Missions

Jon M. Jenkins, Bob Ciotti, Andy Michaelis, Peter Tenenbaum,
Ian Brosnan, and Piyush Mehrotra

4 March 2022





HECC Overview

NASA Ames' High-End Computing Capability (FY21 Budget: $47M)

- NASA's only cloud-scale private HPC cloud infrastructure
- Similar economics of scale to CSPs who run multiple hyperscale data center
- Robust power infrastructure from aero/windtunnel heritage (each >100MW)
- Current Pad/Power infrastructure can scale out to ~1 Billion SBUs (7.5x)
  - straight forward to double again
- > 620,000 CPU cores > 614,000 GPU cores
- > 17,000 compute nodes - (delivered 110 million SBUs FY21)
- > 100 PB of on-line data storage
- > 350 PB of off-line tape data
- Supplementary analysis systems
- Scientific Consulting for Optimization and Help (significant code speed up)
- Data Analysis and Visualization – as a service

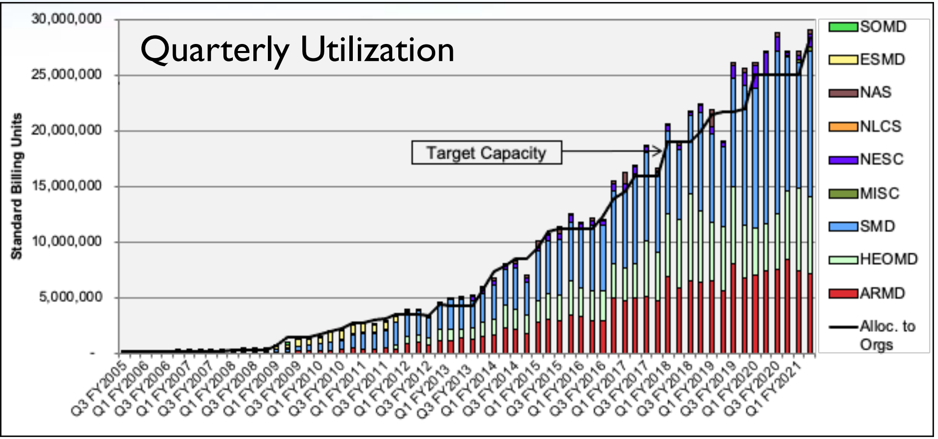
Traditional focus on modeling and simulation (Data Producer)

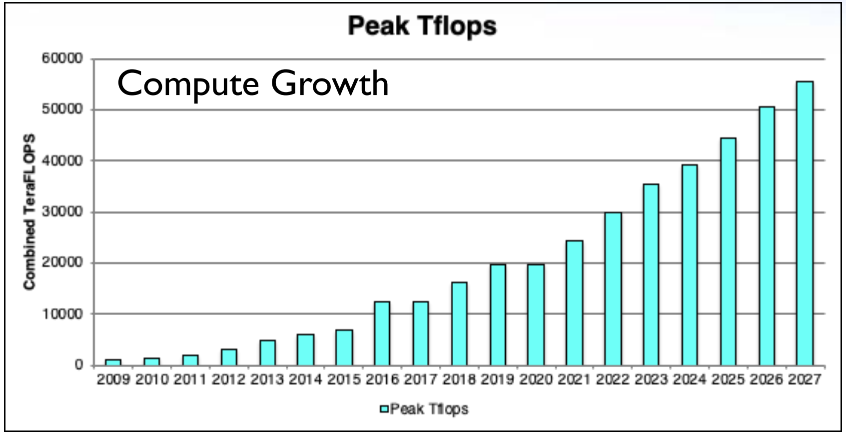
- Evolving Support for Hybrid Computing (On-prem/Cloud)
- Improvements around Latency (reservations and dedicated systems)

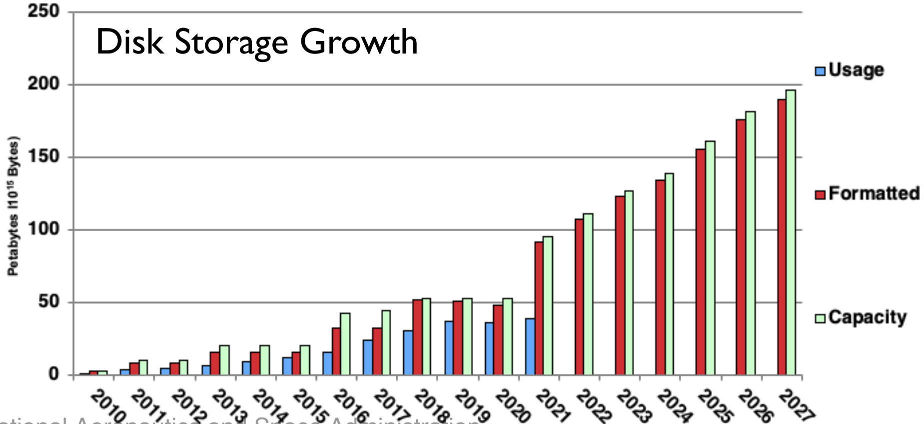
Growth in the size and nature of SMD data sets and SPD-41 motivating changes in HECC

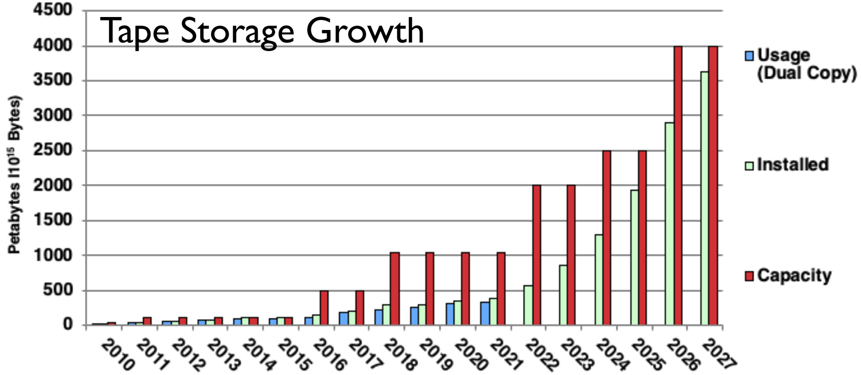
- ECCO (Ocean Only Model Outputs) – 4 PB data set accessible through the NAS Data Portal
- GEOS Coupled – 3 PB of scratch space (https://gmao.gsfc.nasa.gov/GEOS_systems/)
- ECCO with GEOS5 – increased simulation output ~10 – 20 PB
  - Ocean Biology (e.g., predict whale migration/feeding patterns based on food sources estimated from global models)
- NASA's Earth Exchange (NEX) - 5.9 PB of on-line storage, used for data cache based on projects requirements for a given funding cycle.
- SBG is expected to collect 10-15 TB day^{-1} and produce ~75 TB day^{-1} of data products

Historic and Projected Capability Growth









* 1 SBU (Standard Billing Unit) represents the work that can be performed on a dual-socket Broadwell node in one hour.

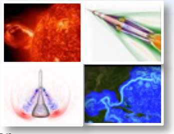
HECC: Meeting Tomorrow’s Computational Challenges

SMD/ESO compute/data requirements motivate change

- **Hybrid Computation between HECC and Cloud**
  - Early successes demonstrate capability (e.g., HySDS, NEX)
- **Tailoring file systems to large observational data sets**
  - Many small files, rather than data streams from simulations
  - More random file i/o
  - Increase use of SDD based storage: 7 PB of SDD
- **Factor Compute node needs of SMD/ESO into hardware selection**
  - Variety of multigenerational systems available to optimize compute
  - Larger memory vs. smaller number of required cores
- **Node scheduling to support different use cases**
  - Calendar based scheduling – e.g., 100 nodes at 10 am
  - Dedicated resources – special queues for high priority work
  - Enable autoscaling up from 1 to N nodes for embarrassingly parallel jobs
  - Preemption for priority work
- **High Availability Compute/Data Module**
  - Fault tolerant network to peer providers
  - Uninterruptable power for critical infrastructure
  - Improved support for lower latency applications

HECC Responds to SMD Requirements

Earth Science on HECC



Subseasonal to Decadal Climate Forecasts
HEC Resources for ECCO-IcES-Ext.01-Ext.02
GEOS/ECCO Coupled Model And Data Assimilation
SWOT KaRIn Data Simulation
Weather Systems Over North Africa And Tropical Atlantic Using Lidar Wind O
U.S.GODAE: Sustained Global Ocean State Estimation For Scientific and Practical Application
NNH20ZDA001N-MAP-BARAHONA
Aerosol Impacts on the Microphysical and Dynamical Processes of Tropical Convection
Coupled Predictions of Ozone and Climate: 1950-2100
OCO-3
NASA EARTH EXCHANGE
JPL OSSE Support
Improving Glacial Melt Rate Estimates Using ECCO And NASA OMG
Global Cloud Thermodynamic Phase And Ice Cloud Microphysics for Aqua AIRS, Suomi NPP, And JPSS
Multi-scale Modeling Simulations And Satellite Observations For Cloud-Precipitation
Drivers And Uncertainty Of Ocean Carbon Sources And Sinks
Orbiting Carbon Observatory 2
Fusion of AIRS/AMSU and CrIS/ATMS Observations via Radiometric Consistent Climate Fingerprinting
MISR Wind Assimilation with GEOSS DAS
Data Access for GEOS-5
Megacities Carbon Project
NASA Airborne Observations With The GEOS-Chem Model with Complementary Satellite And Ground-based Data F
Understanding Predictability and Improving Prediction of Atmospheric Blocking and Associated Extreme Weather
CloudSat ROSES 2018
Management Of Social-ecological Grazing Systems In The Altai Mountain Transboundary Zone
Models Of Regional And Global Mesoscale Dynamics
Aerosol Observations From Terra, Aqua, Suomi NPP, And JPSS Series Satellite Sensors
From Grounding Lines to Coastlines: An Integrated Approach to Barystatic Sea-Level Projections
An Observational and Modeling Study of High-Latitude Clouds, Precipitation, and Radiation Extratropical Cyclones
Atmospheric chemistry in the GEOS Earth System Model (ESM) and Data Assimilation System (DAS) at GM
Understanding Rapid Geomagnetic Secular Variation via Data Assimilation
Addressing Advanced InSAR Time Series Processing And Tropospheric Noise - NISAR Solid Earth Cal/val
A Reanalysis Of The Greenland Ice Sheet
Salinity And Stratification At The Sea Ice Edge
VSLS Halogen In GEOSCCM-GMI
Revealing The Mystery Of African Carbon Cycle
Toward disentangling causes for the substantial increase of CO2 seasonal amplitude in the Arctic
Tropical Cyclones in the GEOS-S2S System
Carbon-climate Interactions From Interannual To Long-term Carbon-climate Feedbacks
Predictability of Earth's Climate
Carbon Monitoring Study: Flux Pilot Project
Simulating Earth System Processes
Multi-model simulations of CO2 and CH4 for the Atmospheric Carbon and Transport (ACT) America project
Exploring the dynamic response of West Antarctica to Ocean Variability
Coupling high-resolution cloud-scale simulations and observations
Feedbacks Between Wind---Driven Surface Fluxes and Cloud Population Evolution During MJO:

HPC Cost Terminology vs Cloud Costs

- ## SBU* – Standard Billing Unit
  - SBU – Standard Billing Unit – Work completed in 1 hour on a dual socket Broadwell Node
  - SBU Cost == 75% of all HECC Compute capacity available day 1 of FY / FY Projected Budget
  - $0.56 in FY22 vs $0.47 FY21? -> Planned budget increase in FY22 (CR - No Actual Budget)

Cost is predictable
Fixed Budget

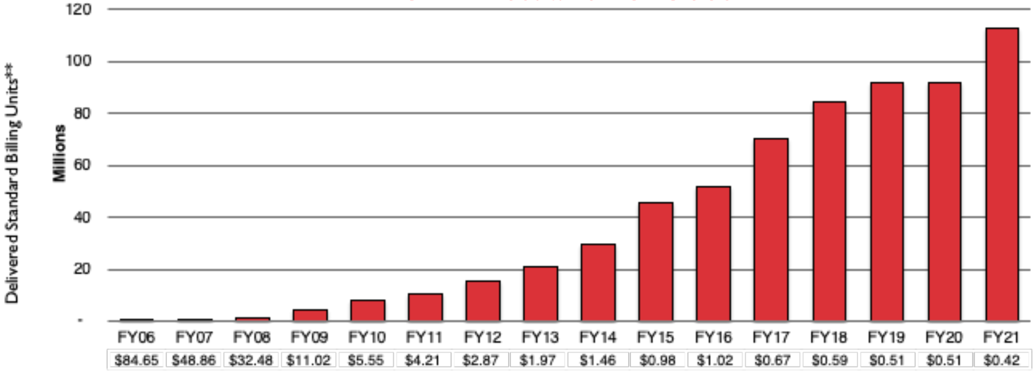
Costs are all-inclusive: facilities, power, personnel, hardware, networking, maintenance, and storage

- ## ROI – Actual SBU cost
  - Number of SBUs delivered / Actual FY Budget (typically ~85% - 110M in FY21)
  - FY21 SBU == $0.47 vs. ROI == $0.42
- ## Marginal SBU Cost
  - SBU cost for last programs who added funding to HECC planed procurement FY21 (@75% utilization)
    - » $0.20 / SBU - 3 years
    - » $0.12 / SBU - 5 years
    - » $0.09 / SBU - 7 years
- ## Marginal ROI
  - Likely - 10% better than marginal SBU cost
- ## 3-year payback time HECC vs cloud*

65 days on X nodes in cloud on-demand
is equivalent to
1095 days X nodes in HECC (# of days in 3 years)
(Substitute X for some number of nodes - e.g., 128)

  - HECC needs 6-12 months lead time
  - Based on recent CFD runs in cloud (AWS)
  - Comparable performance – no scaling advantage – 17x the cost vs. 3 year purchase at margin

ROI – Actual SBU Cost



Planned HECC Enhancements

- Improved support for containers
  - Singularity now – still security concens: set UID root,squashFS
  - Podman – hopefully can be path forward

- Support for HySDS, Ziggy, and other hybrid processing systems that need to span HECC and cloud resources
  - Looking at how to best address networking and egress
  - Improved connectivity to internet2, AWS
  - Research into how we best to reduce challenges across the HECC/Cloud interface

- Modernized Modular Facility – eliminating dependence on main facility built in 1980s
  - Reduce/Eliminate need for full facility maintenance
  - UPS powered storage and networking

- Rolling updates
  - Method to roll in all security updates without need to take entire system down
  - NASA security policy requires expedient patching - (e.g., < 7 days)

- Data "re-exporting"
  - Data sets can be accessed performantly inside HECC, but also externally with http range requests
  - May be used to support for SPD-41 data publication requirements

Hybrid Computing & Storage Architectures (HCSA) Evolving To Meet The Agency's Needs

- Explore & support HCSA, optimizing resource allocation as well as time to solution for modeling & simulation, general data processing, machine learning & collaboration.

- Leverage the unique features and capabilities of public cloud platforms and on prem private cloud High Performance Computing & Storage

Cloud

- Data, resilient control infrastructure, services capabilities, community engagement & consumers

HPC

- Large-scale lower* cost sustained computing capabilities (modeling & simulation, ML training, etc.)

Current Pilot Activities

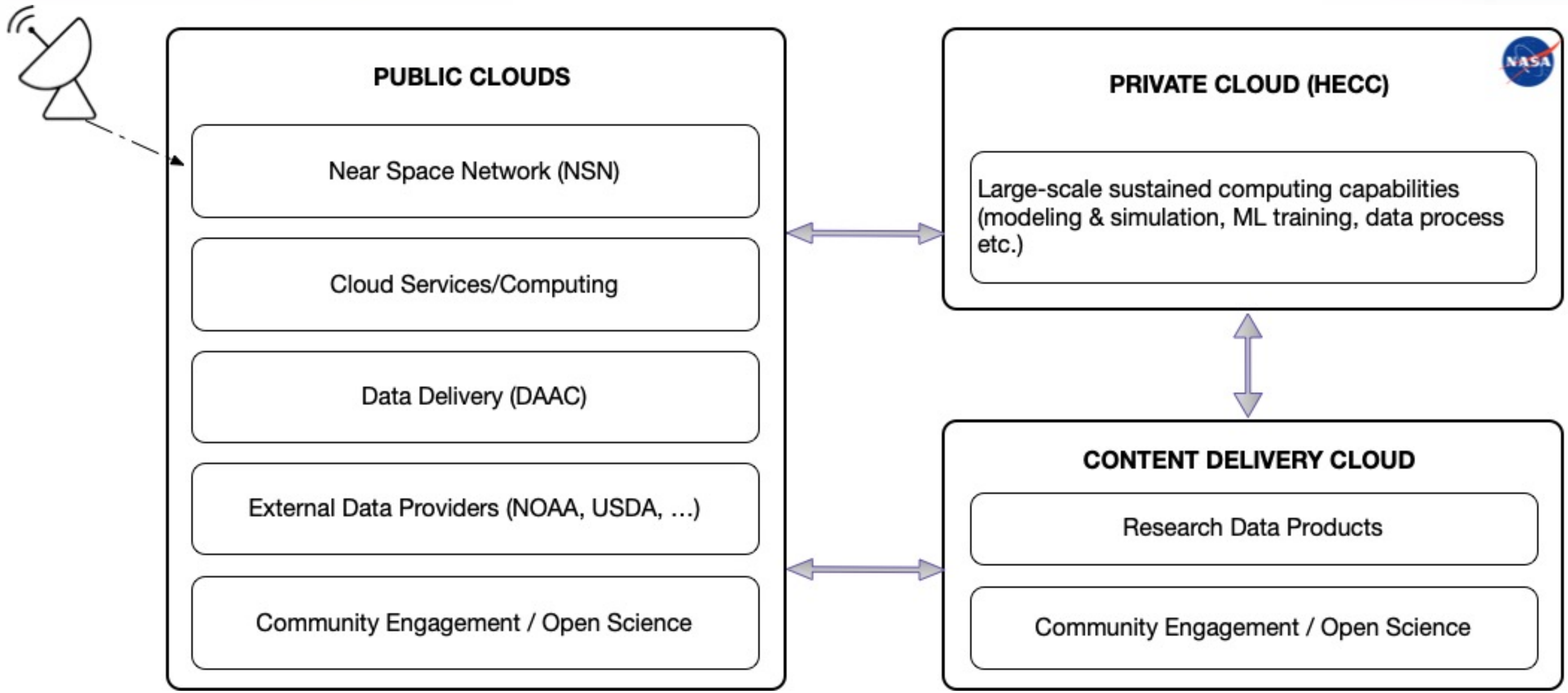
- NASA-ISRO Synthetic Aperture Radar (NISAR) HySDS (JPL) - https://hysds.github.io/

- Multi-Mission Algorithm and Analysis Platform (MAAP) - https://earthdata.nasa.gov/esds/maap

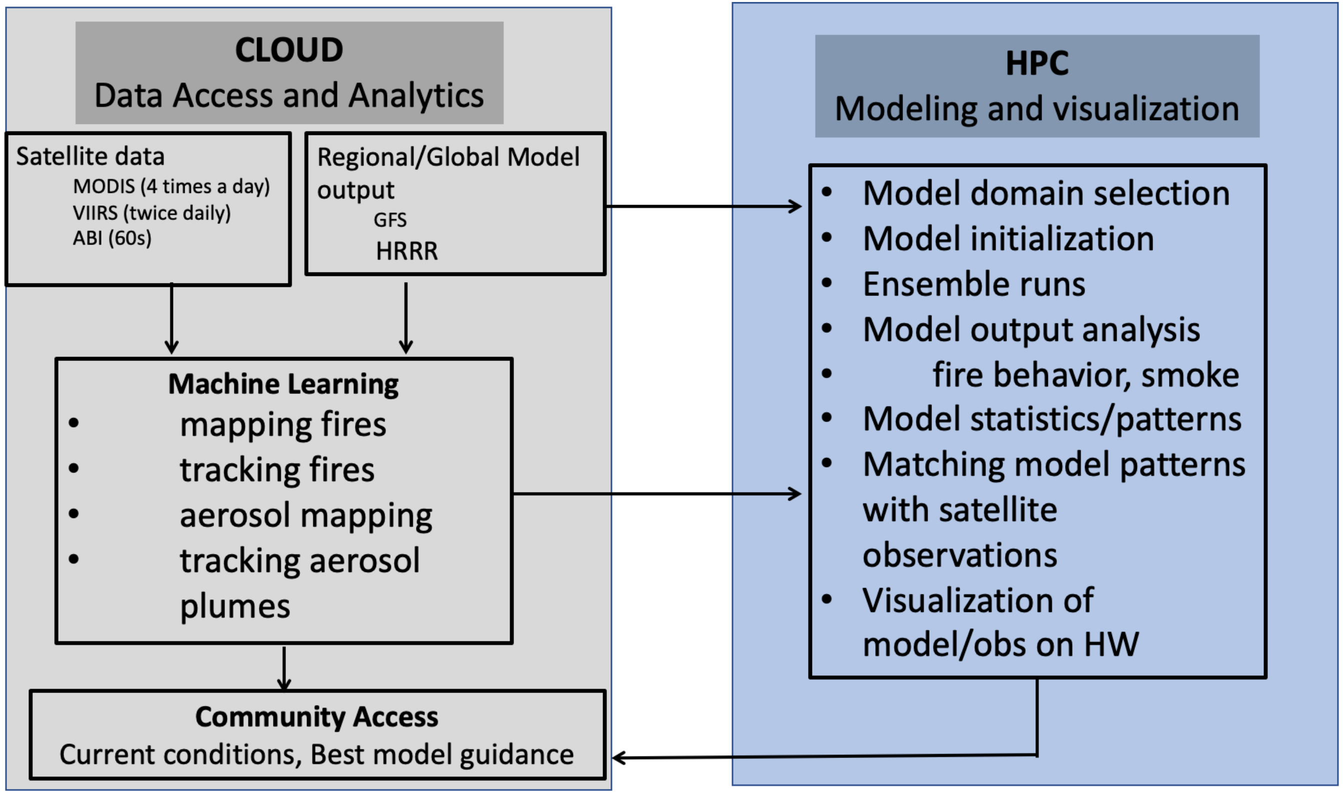
- Urgent computing using a hybrid Cloud-HPC architecture for streaming satellite data analysis and event-driven modeling (wildfire, Q3 2022)

* Relative to Public Cloud

Hybrid Computing Architectures (HCA)



HCA for Wildfire Monitoring and Modeling



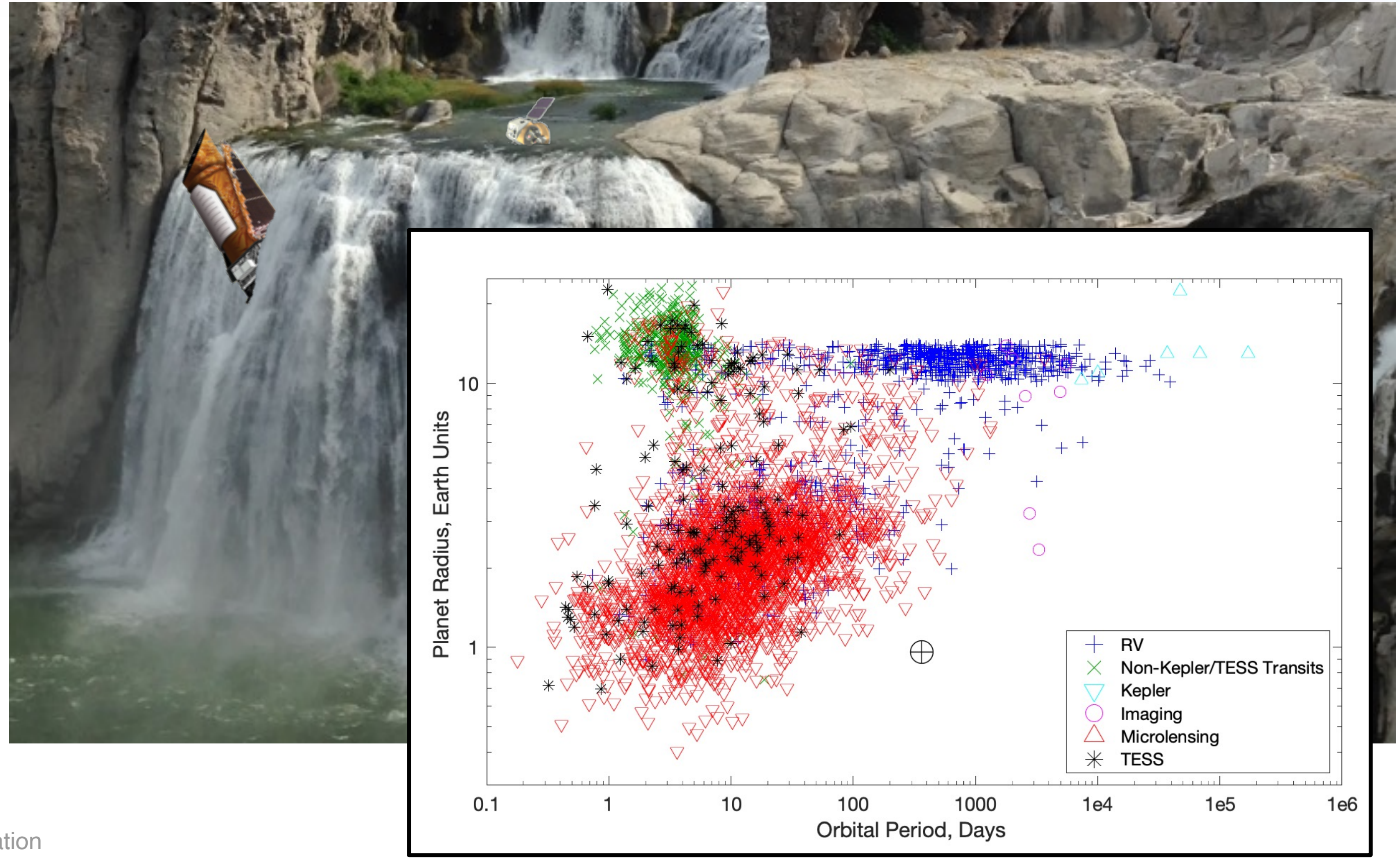
Science Pipelines at NAS

We've built and operated several major science pipelines for Astrophysics missions.

Science Pipelines at NAS

We've built and operated several major science pipelines for Astrophysics missions.

Kepler Mission:
Discovered 3186 of the 4935 known exoplanets to date
1 GB day^{-1}

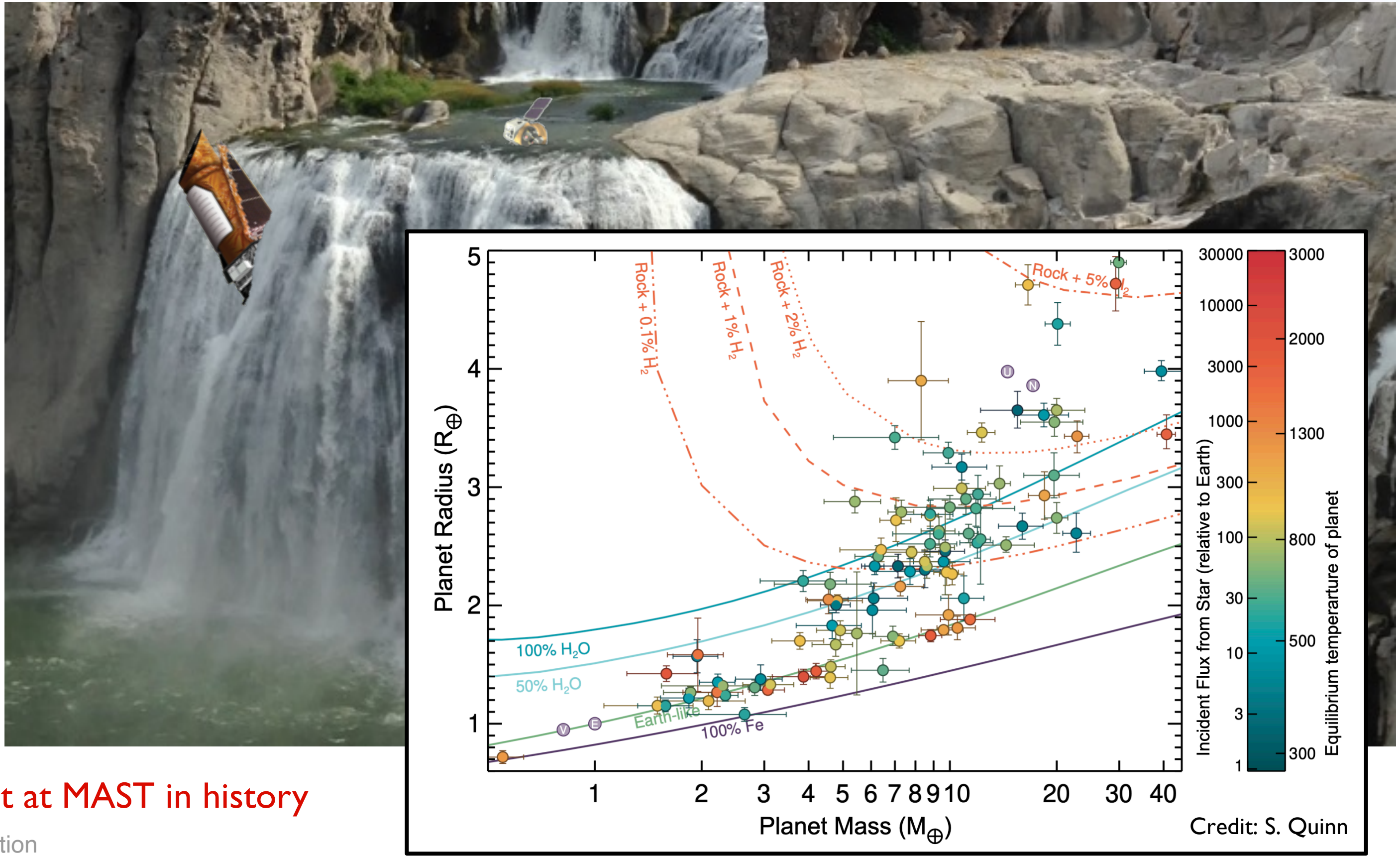


Science Pipelines at NAS

We've built and operated several major science pipelines for Astrophysics missions.

Kepler Mission:
Discovered 3186 of the 4935 known exoplanets to date
1 GB/day^{-1}

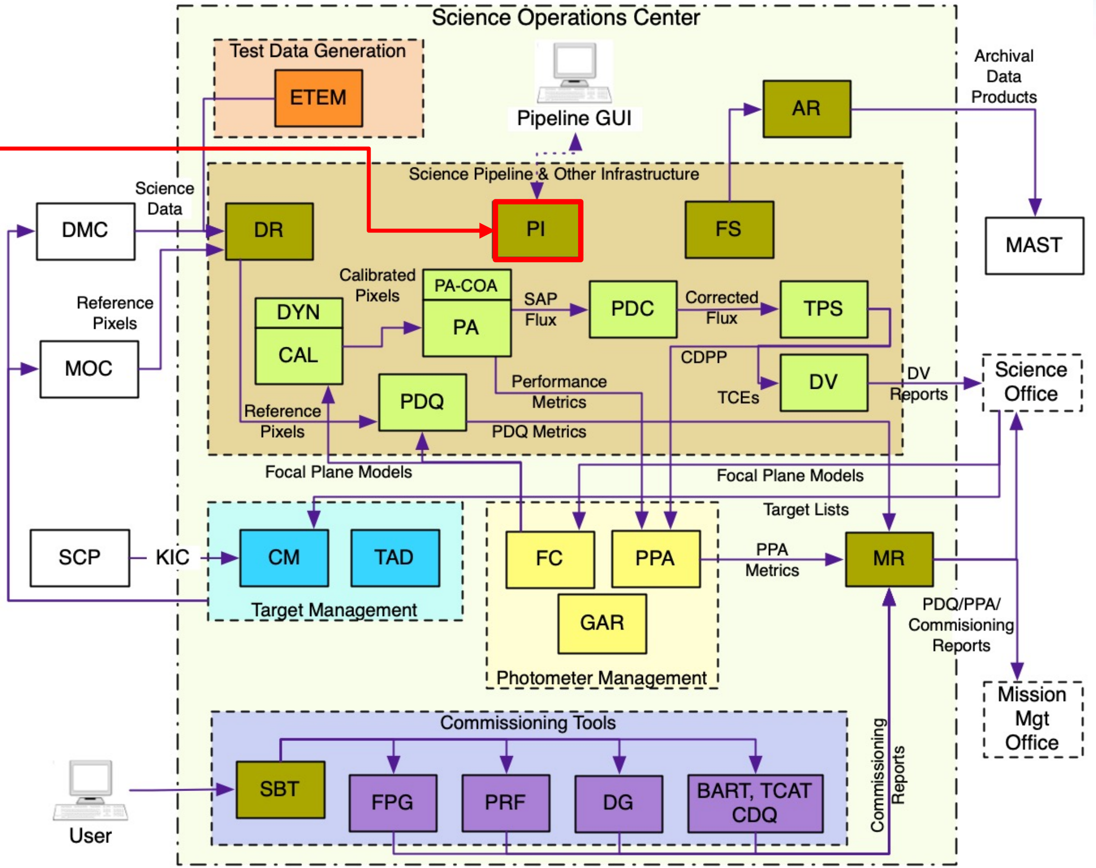
Transiting Exoplanet Survey Satellite:
199 Planets
72 < 4 R_{earth} with measured masses
13-50 GB day^{-1}



Most downloaded data set at MAST in history

Kepler Science Operations Center Architecture

Ziggy's Predecessor



The Kepler SOC Pipeline consisted of 24 Computer Software Configuration Items, including the Pipeline Infrastructure (PI), which has been split off as its own software project, *Ziggy*

Ziggy is capable of running complex science algorithms on large data sets on a laptop, workstation or on the NAS Pleiades supercomputer.

We've demonstrated running Ziggy in AWS

TESS SOC directly connected to NAS Pleiades backbone network

Performant Access to Compute and Filesystems

ziggy A Flexible, Scalable Pipeline Controller

Highly automated processing of large volumes of data

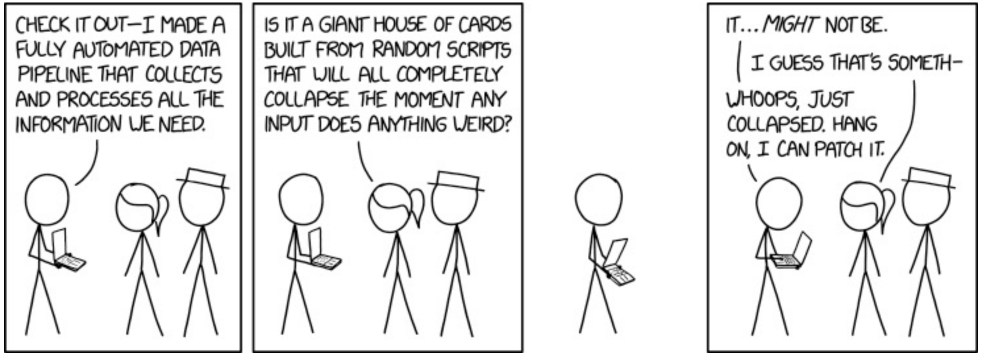
Automated dispatching to Pleiades for large to extremely large tasks

Provides permanent storage for all version changes, linkage between parameter values and pipeline tasks

Excellent logging and diagnostics for those rare occasions when something goes wrong

Extensively tested by Kepler and TESS processing missions

- Codebase: Java (some MATLAB, Python C, C++, Perl, Shell scripts)
- Relational database: mainly for data accountability
- Datastore: main storage of mission data and products
- Pipeline definitions: xml files that sequence algorithms, specify parameters, etc.
- Messaging: Remote Method Invocation (a Java API)
- Ziggy is being released under the NASA Open Software Initiative

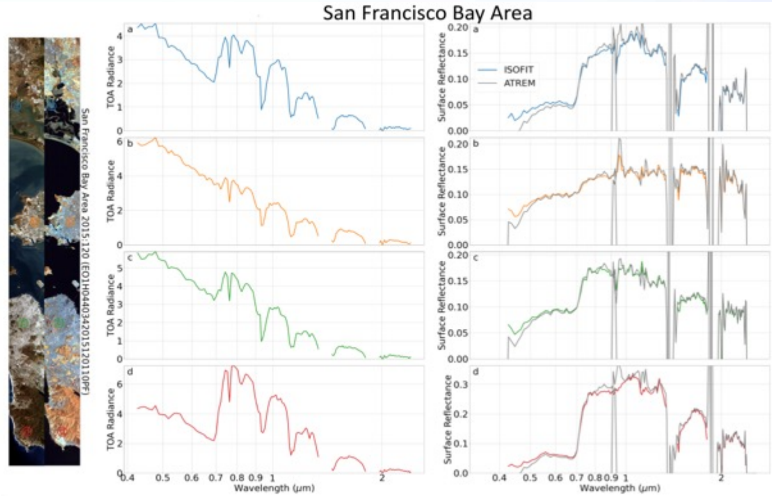


Ziggy is TRL 7 and a class C software under NPR 7150.2C

Surface Biology & Geology Pathfinder Studies

Hyperion Pipeline:

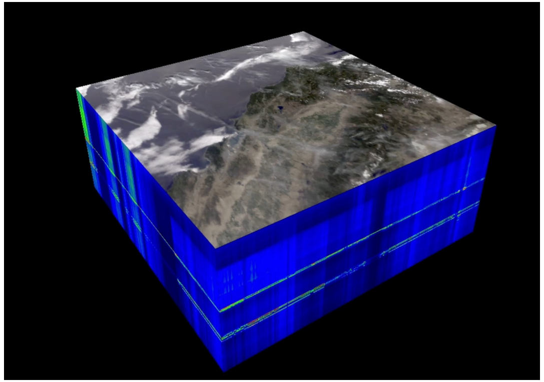
- Completed processing the 55-TB Hyperion data to top-of-the-atmosphere radiances (L1)
- Currently checking consistency of Hyperion surface reflectance results (L2) (RadCalNet and AVIRIS/Hyperion Comparisons)
- Future work: incorporate L3 algorithms for vegetative traits and/or aquatic studies



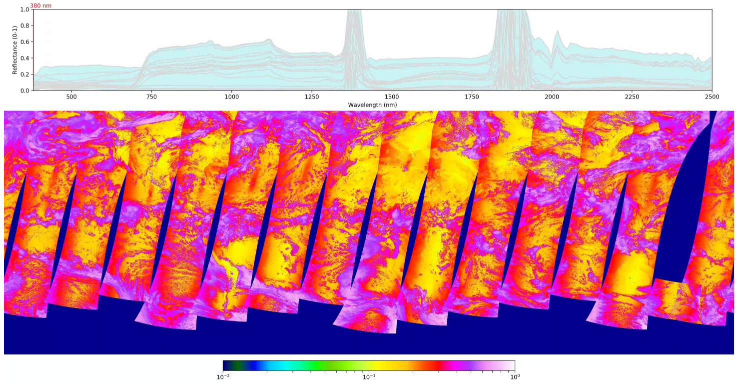
Credit: Y. Shinozuko

Ames Global Hyperspectral Synthetic Data (AGSD)

- Created using MODIS data together with an AVIRIS hyperspectral library and a linear unmixing model
- 1st year of AGSD data available on NAS Data Portal



Credit: A. Michaelis



Credit: A. Michaelis, W. Wang

National Aeronautics and Space Administration

https://data.nas.nasa.gov/aghsd

HECC Forward and Reprocessing Cost Est

- Costs Basis Est from last customer to buy HPC on Margin

  $0.10/SBU - 7 year

- Scaling Hyperion up to SBG
- 55 TB L0 ⇔ ~5 days of SBG
- 8520 SBUs for Hyperion is ~621,960 SBUs/year for SBG
- 7-year purchase, 256 nodes, 5 PB storage
- 4 PB of tape each year for L0 (2 copies)
- Reprocessing every two years
- 2 hours compute/2 hours transfer @30gb/sec

	Forward (KiloSBU)	Reprocessing (KiloSBU)	Annual (KiloSBU)	Compute Storage Tape	Maint Water Power	Projected HECC Cost (K $)
Year						
0				$ 4,827		$ 4,827
1	622.0		622.0	$ 168		$ 168
2	622.0		622.0	$ 168		$ 168
3	622.0	1,243.9	1,865.9	$ 168		$ 168
4	622.0		622.0	$ 168	$ 173	$ 341
5	622.0	2,487.8	3,109.8	$ 168	$ 173	$ 341
6	622.0		622.0	$ 168	$ 173	$ 341
7	622.0	3.731.8	4,353.7	$ 168	$ 173	$ 341
						$ 6,695
					percent utilization	25%

- 75% of capacity available: algorithm development, >L2 processing, HySDS hybrid post DAAC analysis, more frequent re-processing, etc.

Caveats on HECC Forward (re)Processing

- Pipeline Control Management and Persistent Services run HA in cloud
- Redundant capability to run pipeline in cloud during HECC unavailability
- HECC needs upgraded WAN connectivity – 100gb to DAAC, maybe NSN
- No cost included for egress from NSN or other
- No pipeline operational staff included
- Only saving L0 data to tape for later reprocessing
  - No substantial L1R or L2 stored locally – Costed about 5 PB in HW budget
  - No tape media migration costed
- Possibility to perform >L2 forward processing or HySDS type analysis
  - System @25% utilization
  - Preemptive and Calendar scheduling supported in PBS for high priority workloads (think spot for HECC)

Conclusions

- Opportunity for HECC to assist ESO data processing needs
  - Most impact likely in offloading 'heavy lift' processing requirements
  - Consolidation of requirements critical to developing economies of scale
- Collaborative benchmarking critical to understanding potential impact on ESO MDPS
- Dedicated Mission Support Systems can be sited (TESS SOC) under MOU
- Network connectivity/peering to remote systems such as AWS and NCCS are requirements driven
- Ideally, Processing pipelines transparently span Cloud and HECC, leveraging Cloud High Availability & lower compute cost on HECC
- Opportunity to publish data through HECC via the DataPortal or the cloud
- Compression Techniques may prove valuable for exponential data growth
- HECC is open sourcing Ziggy, a flexible, scalable science pipeline control infrastructure

Early Engagement and planning with HECC is essential to maximize impact