High-Resiliency and Auto-Scaling of Large-Scale Cloud Computing for NASA's OCO-2 L2 Full Physics Processing
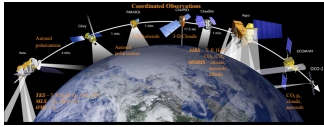
HySDS Team: Hook Hua, Gerald Manipon, Paul Ramirez, Phillip Southam, Michael Starch, Lan Dang, Brian Wilson
OCO-2 SDOS Team: Charlie Avis, Albert Chang, Cecilia Cheng, Lan Dang, James McDuffie, Mike Smyth
Jet Propulsion Laboratory, California Institute of Technology
Poster Number: IN43B-1733

Background

- OCO-2 launched on July 2nd, 2014, at the head of the A-Train
  - Collect global measurements of atmospheric carbon dioxide with the precision, resolution, and coverage needed to characterize sources and sinks in order to improve our understanding of the global carbon cycle
- NASA OCO-2 Science Data Operations System (SDOS)
  - Forward (1X) and bulk processing (4X)
- L2 bulk processing ported to NASA AMES Pleiades Supercomputer
  - L2 full physics processing on ~200 nodes (15X)
  - Running 48 x PGE processors on each compute node

Motivation

- Frequent science requirement changes
  - Needed agile science data system approach
- Increase in science computing needs
- Pleiades Supercomputer scheduled downtimes conflict with science processing requirements
- Need elastic and large-scale processing capability

Rapid Hybrid Cloud Enablement

- **Day 1**
  - Team formed to bring in cloud capabilities via **HySDS** (Hybrid Science Data System)
  - Introduction to HySDS, bursting out to cloud, current cloud computing strategies
- **Day 8**
  - HySDS team successfully integrated Level-2 Full Physics (L2FP) executable into HySDS and demonstrated parallel runs on **Amazon Web Services (AWS)**
  - L2FP developers validated outputs from AWS
- **Day 17**
  - End-to-end testing with SDOS (at JPL) HySDS (in AWS)
  - Benchmarking on AWS to **optimize cost-effectiveness**
- **Day 21**
  - SDOS gets ownership of PCS-HySDS subsystem
  - Technology transfer training
    - AWS basic and HySDS operations
  - Access to HySDS source control
- **Day 37**
  - HySDS official delivery and operational handoff to SDOS.
  - Introduction of on-premise HySDS workers on OCO-2 cluster to handle uploads and downloads.
  - **High-resiliency operations on AWS "spot market"**
    - **90% cheaper than "on-demand"**

Hybrid-cloud Science Data System

- Leveraging the Hybrid-cloud Science Data System (HySDS)
- A set of loosely-coupled science data system cloud services. Data Management, Data Processing, Data Access and Discovery, Operations, and Analytics
- Data system fabric over heterogeneous computing infrastructure
- Large-Scale, High-Resiliency, and Cost-effective hybrid approach
- AWS spot market support
- Faceted Navigation for all SDS interfaces
- Real-time metrics
- Crowd-sourcing and Collaboration
- Real-time Analytics
- Provenance for Earth Science

HySDS Architecture

- SDS can be dynamically provisioned at on-premise at JPL, DAAC, and/or at public cloud providers
- Scalable workers enables control over data movement and monitoring of data movement between data centers
- Compute and Object Storage are horizontally scalable
- Data products in cloud object stores
  - Redundancy and replication policies
  - Scalable high-performance storage
  - Accessible as URLs
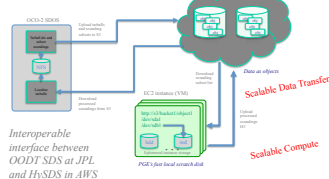- Use of AWS GovCloud addresses export controlled issues
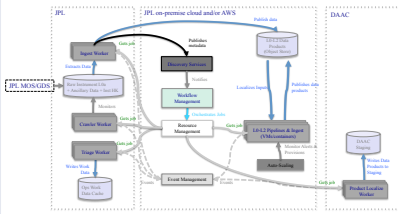
Partitioning Jobs in AWS

- Workers processed subset of soundings per granule
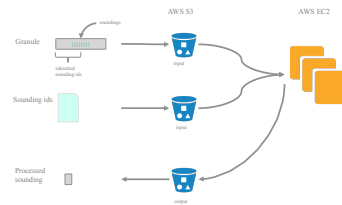
Compute and Storage Usage

- *Compute* instances can scale up to demand
- *Object storage* can scale up *data volume* and *aggregate data throughput* to demand

Interoperable interface between OODT SDS at JPL and HySDS in AWS
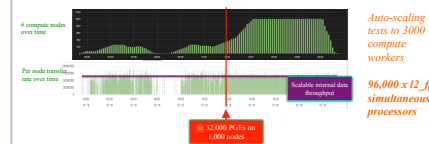
AWS Transport and Compute

- Entry points into AWS are different
  - JPL to US-West-1 on 10Gbps
  - JPL to US-West-2 over internet
- Transport approach
  - Stream data from JPL to S3/US-West-1
  - Maximize JPL-AWS network speeds
- Compute in other AWS regions
  - E.g. EC2/US-West-2
  - Move data from S3/US-West-1
  - Results moved back to S3/US-West-1
- Asynchronously localize results back to JPL from S3/US-West-1

Auto-Scaling Science Data System

- The size of the science data system compute nodes can automatically grow/shrink based on processing demand

Auto-scaling tests to 3000 compute workers

96,000 x l2_fp simultaneous processors

@ 32,000 PGEs on 1,000 nodes

High-Resiliency & Spot Market

- Major cost savings (~10X)…if can use spot instances
- On spot market, AWS will terminate compute instances if market prices exceed bid threshold
- HySDS able to self-recover from spot instance terminations
- Running in spot market forces data system to be more resilient to compute failures

Large-Scaling "Market Maker"

This OCO-2 production run of 1000 x 32vCPUs affected the market prices
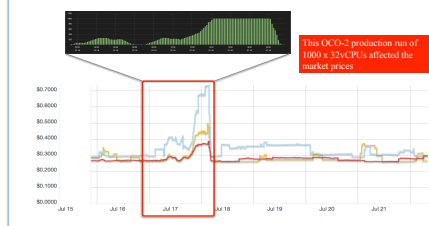
Large-Scale Considerations

- **Spot market terminations**
  - If market prices passes tolerance bid threshold, then terminations
- Availability Zone (AZ) **load rebalancing**
  - Terminations of nodes for balancing across AZs
- **Instance failures**
  - 99.9% reliability means 1 hardware failure per 1000 nodes
  - E.g. "EC2 has detected degradation of the underlying hardware hosting one or more of your Amazon EC2 instances in the us-west-2 region. Due to this degradation, your instance(s) could already be unreachable."
  - E.g. disk failures
- "Job drain"
  - Addressing failures leading to job drain from work queues
- "**Thundering herd**"
  - API rate limit exceeded
- "**Market Maker**"
  - You affecting spot market prices
- **S3 object store** performance optimizations needed
- **Instance** startup with data caching
- **Auto-scaling**
  - slow scale-up needs AWS tweaks
  - scale-down group vs self-terminating instances

Provenance Support

- Data production compliant to Provenance for Earth Science (PROV-ES) specification
- Near real-time view of provenance from production
- Faceted search of provenance
- Visualization of provenance
- NASA ESDSWG PROV-ES Working Group
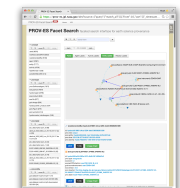
HySDS for OCO-2 L2 Full Physics

- Scalability
  - Scaled up to 3000 x 32 vCPUs = 96,000 vCPUs
- Operability
  - Flight operational hardening
- Visibility
  - Faceted search interface for everything
- Cost
  - Running on AWS spot market is ~10X cheaper
- Accepted by OCO-2 as usable and cost-effective for reprocessing scenario
- Interfaced with existing SDSes at JPL (e.g. OCO2 SDOS)
- Passing along lessons learned