Technical Challenges and Lessons from the Migration of the GLOBE Data and Information System to Utilize Cloud Computing Service

John F Moses, NASA Goddard Space Flight Center, Greenbelt, MD, USA;
Brian Littlefield, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Global Learning and Observations to Benefit the Environment (GLOBE)
- Provides training and material on the use of GLOBE, including measurement protocols
- Maintains records and data across time periods
- Began Earth Day 1995 as a partnership between NASA, NOAA, NSF
- Grew to more than 48,000 trained teachers, 24,000 schools, 1.5 million students

GLOBE Data and Information System (DIS)
- Top Functional Requirements
  1. Provide training and material on the use of GLOBE, including measurement protocols
  2. Receive, check, accept member observation entries with protocols
  3. Ingest data from automated weather stations and other data sources
  4. Maintain records across protocols (over 127 Million since 1995)
  5. Visualize observations on a map
  6. Provide data analysis tools, graphs and data access tools via search and filtering
  7. Export observations recorded across protocols
  8. Enable data entry from mobile Apps
  9. Recognize contributions from schools and students
  10. Administer workshops, science Blog, Help Desk, Email, Mass Mailing

GLOBE DIS Modernization
- In 2010, NASA Goddard was asked to lead the evolution of the GLOBE DIS to a new architecture, evolving software and internet technologies to achieve program goals.
- The Goddard GLOBE DIS Team was formed and partnered with UCAR to develop a new enterprise portal, web application framework, and modern visualization and graphing features.
- Migration to new system completed July, 2012

GLOBE Team and Organization circa 2013
- UCAR GLOBE System Development and Operations
- GSFC GLOBE Software Engineering Services Task

GLOBE Transition Objective, Tasks, Milestones
- In 2013, NASA Goddard Science Data Systems was asked to transition the GLOBE DIS from UCAR to ensure its long-term stability, integrity and continued improvement.

Infrastructure Architecture Trade, Estimates, Drivers
- The UCAR GLOBE DIS hardware was near end-of-life and a trade study was prepared to evaluate alternative infrastructures
- 480 Hours of estimated labor for both solutions involves similar tasks
- Flash and Local Facility
  - Initial Cloud Configuration and Testing
  - Configuration of VMs and image transfer protocol
  - VM Update and maintenance (cloud maintenance)
  - Data configuration and testing
  - Initial SQL database validation
  - Redundant planning, testing and validation
- Amazon Web Services
  - Elastic server instance - logical setup & configuration
  - Configuration of VMs and image transfer protocol
  - AWS cloud maintenance and administration
  - AWS cloud configuration and testing
  - Initial AWS database validation
  - Failure planning, testing and validation

Final Production Architecture

Launch Readiness Checklist and Responsibilities

Amazon Specific Findings
- IPSV - Classic Elasti Load Balancer (ELB) entry; no support for newer Virtual Private Cloud architecture
- Load balancer properties are set by Amazon
- Autoscaling limits to handle large queues
- Customization limited – receptive by Amazon
- CNS domain cost requires static IP: Amazon ELB - IP may change at any time – impact to hand-coded social media external systems
- Difficult to trace
- HTTPS/SSL traffic gets redirected to www.google.com; however, HTTPS/SSL does not, require separate server redirect
- Handshakes are restricted normally at any time
- 1 to 2 week notice, takes a few hours
- Requires reconfiguring instances based on image
- Includes IP changes and DNS changes
- Database instance takes longer because of size
- Amazon has default limits dependent on service level
- Number of instances, servers, IP addresses
- Not always aware you’re up against a limit

Changes from the Original Plan
- 30% over original labor estimate (480 hours grew to 625)
  - System upgrades, additional test tasks, API from prototype to production
- Changes to System Environment
  - More instances available for database, application servers in production, smaller instances in staging (Med 5K/Week)
  - Amazon DNS domain cost requires static IP: Amazon ELB - IP may change at any time – impact to hand-coded social media external systems
  - Difficult to trace
  - HTTPS/SSL traffic gets redirected to www.google.com; however, HTTPS/SSL does not, require separate server redirect
  - Handshakes are restricted normally at any time

Availability and Uptime
- No quarterly 1 day power cycle outage (previous UCARI facilities)
- Hourly uptime including planned code refactoring outages: 
  - Jul 2015 - Aug 2016: 99.999% uptime
  - Aug 2016 - Aug 2017: 99.999% uptime
  - Aug 2016 - Aug 2017: 99.999% uptime

Staff Skills & Level
- UCARI support existed Feb 2014
- No significant change in GSFC staffing due to transition

Skills change in transition to GSFC
- Developers
  - Maintain knowledge of AWS Services
  - Focus on Postgres and lifespan capability
  - Focus on UCAR on Rails and Data Entry
- System Administrator
  - CNS services on post GSFC

Skills change in transition to AWS
- System Administration
  - Responsible to Amazon instance reutilization
  - Monitor and analyze instance performance
  - Respond to IP and instance limits
  - Plan architecture changes in anticipation of growing loads

Sponsor: NASA Science Mission Directorate