Community for Data Integration
Technical Stack Working Group

Cloud Computing Sub-Group

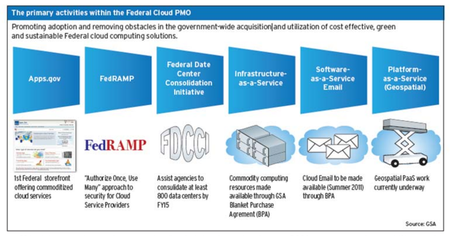
Is your Group looking into Cloud Computing?
Do you want to keep up to speed with other USGS/DOI Cloud Efforts?
Do you want to share coordination points?

This Cloud Birds of Feather is a new TSWG Sub-Group started in Spring by joining NGP, USGS Water to share information and coordinate

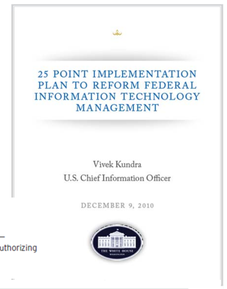
Example Outcome of Cloud BOF: Cloud BOF coordinated multiple sessions with ESRI and Amazon Architects with technical questions. One session was informal during TNM User Conference attended by 10 and another was a formal technical exchange meeting held during the ESRI User Conference with detailed review of pilot issues encountered and plan question on deploying ArcGIS on Amazon Web Services

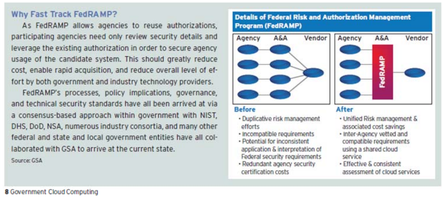
Learn More:
Go to CDI Wiki,
Tech Stack Working Group,
Link to the Sharepoint area
- Cloud Ramp-up Information
- Geospatial on Cloud Ramp-up
- Links to some USGS efforts

GSA Cloud Leadership
(Dave McClure (OCS), Bajinder Paul (OCSIT))



What *is* new about cloud computing is the speed with which the Federal Government has moved to adopt a "cloud first" policy as both a strategic and tactical solution to achieve operational efficiency. The implementation of the cloud first policy will achieve substantial cost savings, allow agencies to optimize spending, and allow agencies to reinvest in their most critical mission needs.





GSA and FGDC GeoCloud Pilot
(Doug Nebert)

Background

Cloud computing and data storage capabilities represent a great promise and potential benefit to federal agencies interested in providing scalable access to government geospatial data and services. At the same time, there are significant unknowns in adopting such technology or in performing cost-comparisons with in-house solutions. These issues are related to operating costs, particularly in an on-demand, burstable environment, software deployment options and configurations, computer security, certification and accreditation, and deployment options that involve data hosting and/or compute hosting.

Geospatial data, applications, and systems interfaces present unique requirements and capabilities that should be addressed systematically in the development of a deployed federal service architecture. Such solutions patterns need to be consistent with the recommended standards, services, data and metadata, and business practices called out in the FEA Geospatial Profile. The data lifecycle for geospatial data is currently being formalized for nationally significant datasets. Realization of the data lifecycle in a broad federal context will require the establishment of reliable geospatial data access services for all types of users.

Proposal

The Geospatial Line of Business, Technology and Architecture WG is seeking participants with data from federal agencies who would like to prototype deployment of public geospatial data and service offerings in a Cloud environment for up to a year, with initial operating costs underwritten by the Federal Geographic Data Committee (FGDC).

Outcomes

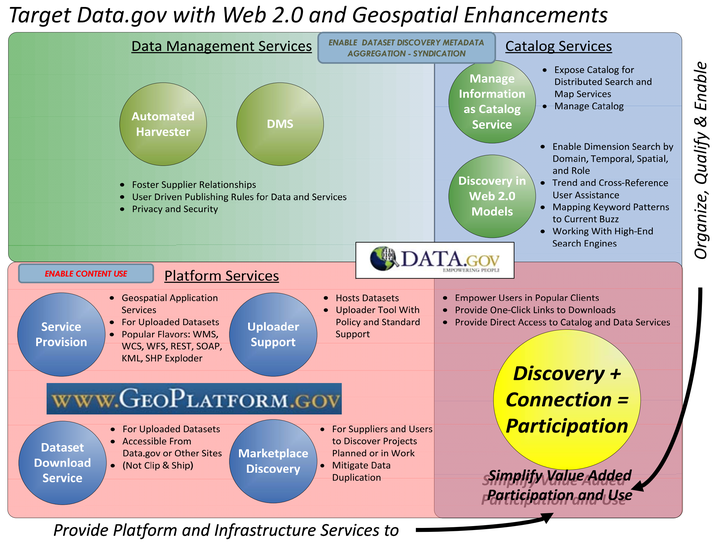
The intent of the GeoCloud Sandbox Initiative is to provide technical assistance, external-hosting, monitoring, and reporting on configuration options and results, usage, and costs for mission-critical geospatial data and software within a federal cloud computing environment. Key items to be addressed include:

- Identifying requirements-driven solution architectures for various sized deployments of geospatial data and services
- Documentation and costing to support scalability, reliability, and redundancy
- Addressing system security concerns to expedite FISMA certification and accreditation for agency adoption
- Supporting and collecting cost comparison information from agencies for existing and externally-hosted Cloud solutions

GeoCloud Project Candidates
Wetlands Mapper
US Census TIGER Shapefiles
IOOS Registry, Catalog, Viewer
NOAA ERDAP
USGS National Elevation Dataset
USGS National Map
EPA Lakes and Ponds
NOAA Particles in the Cloud
GEOSS GeoNetwork
Land Cover Analysis Tool

GSA and FGDC Geo Integration (data.gov, geoplatform.gov)
(Doug Nebert, Jerry Johnston (EPA), Marion Royal (GSA), Tony Lavoi (NOAA), Stephen Lowe (USDA), Rob Dollison (USGS))

Target Data.gov with Web 2.0 and Geospatial Enhancements



USGS
National Geospatial Program
Cloud Pilot Effort (Kevin Hope, Paul Wiese, Matt Tricomi)

- **Overall Goals:**
  - **Assess current state of Cloud capabilities**
  - **Assess NGP capabilities and propose a set that might be viable for initial cloud testing**
  - **Develop evaluation criteria**
  - **Conduct initial Cloud testing**
  - **Develop recommendations that can be used to inform FY12 planning decisions regarding cloud migration**
- **Data and Service Tests:**
  - **Move selected data files supporting overlay and basemap services to ArcGIS on a public cloud**
  - **Evaluate time needed to set up overlays, including WMS and REST services, in the cloud**
  - **Run caching tests to evaluate improved reliability and speed**
  - **Compare basemap startup times in the cloud versus internal**
  - **Test WFS services in the cloud**
  - **Explore any possibilities for testing ArcGIS 10 Image Server capabilities within the Cloud**
- **Bottom Line: Evaluate the utility of public cloud for data storage, service delivery, caching, etc**
  - **Evaluate cost, performance, risk, etc., versus current internal capabilities**
- **Download Tests:**
  - **Upload selected data to the public cloud and compare upload and download times versus current internal capability. Targets may include:**
    - **GeoPDF's**
    - **Vector downloads**
    - **Raster downloads**
    - **Bulk file GDB's**
    - **Staged GNIS and NHD Vector Download Files**
    - **Archived Raster tiles**
- **Bottom Line: Evaluate the capability of the cloud for upload and download of different data types versus internal**
  - **Evaluate speed, efficiency, reliability, cost, risk, etc...**

Mgmt Analysis	Procurement
	Security
	Workforce
	Cost Model
	Suitably/Selection
	Policies
	Communication
	Role Definition
	Project Plan Development

- Vector Pre-staged
  - NHD, SubRegion, Few Formats
  - Names, State
  - US Topo as PDF (not ZIP)
- New Pre-Staged
  - 1 meter NAIP 10:1
  - Transportation
  - Structures
  - Boundaries
  - HQ Scanned Maps PDF
- Atlas
  - Load Existing Zips

- Raster Pre-staged
  - NLCD 92,01,06
  - Elevation multi-res
  - 1 foot Nga
- New
  - Other 1 Foot

2.2 Million files

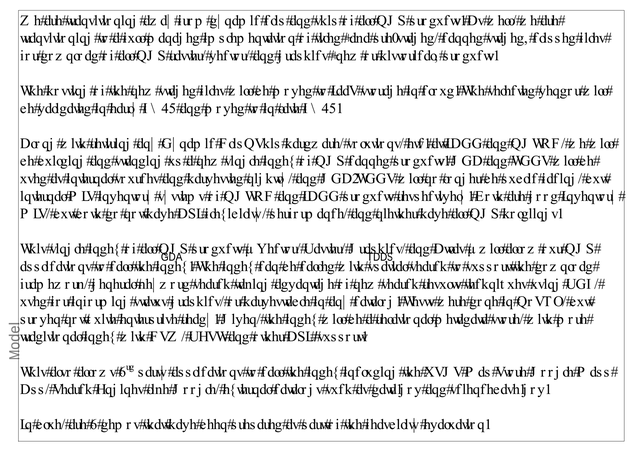
5 million features

Covering 365 million square miles

Reposted Monthly with thousand of features updated or as needed for events

Alternate Analysis discussions on 1 meter and 1 foot imagery are near final decision now. This assumes certain outcomes with such.

FY12 NGP Cloud Staged Product Hosting and Index -Solution Architecture Summary



USGS
Wisconsin Internet Mapping Group
(Gary Latzke, Jon Baier, Nick Estes)

- **Cloud Goals**
- **WIM has dozens of small applications for specific water analysis and exploration use cases**
- **Application traffic is niche and targetted**
- **Trying to lower IT Footprint and look for ways to ease deployment**
-
- **Heavily interested in PaaS or SaaS solutions like ArcGIS.com on AWS for quick deployment of these applications**

Other "Birds of a Feather" Cloud Activities

What else is going on in cloud?
USGS Biology/CBI looking into ArcGIS Service hosting on Cloud (Mulligan/Mancuso)
Hazards has been testing generating basemaps on Cloud, but hosting tiles in their private cloud (Smoczyk/Fee)
EROS building investment and business case for USGS Private Cloud (Binnie)

What else.... (write down, what, who, organization, interested in joining?)