FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

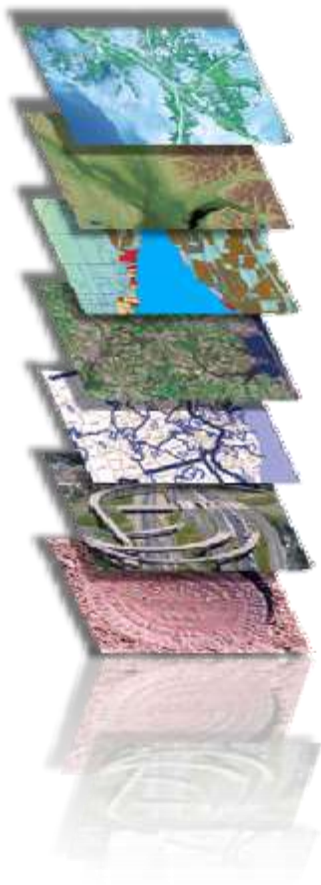
Advancing Geographic Place-Based Understanding *for* Decision-Making

FGDC GeoCloud Incubator

Initiative Outcomes, Lessons Learned and Best Practices

Eldrich “Rich” Frazier
Technical Advisor
Federal Geographic Data Committee
July 12, 2016

Background on GeoCloud Sandbox Initiative



- Initiated as an Architecture and Technology Working Group activity in December 2009
- Five-year project to test and monitor externally-hosted Cloud data and service solutions for the geospatial domain, to support the Geospatial Platform activity
- Initiative Sponsors:

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

U.S. DEPARTMENT OF HOMELAND SECURITY

GSA

Cloud Community Platform Initiative – *Mission Needs*

Business Drivers

- Growing pool of identified agency applications seeking cloud benefits
- Quantify savings, identify risks, ease migration

Platforms

- Construct a discrete subset of secure supporting platforms
- Develop processes for rapid application migration
- Leverage certifications and platform commonalities
- Develop chargeback mechanisms for self-sustaining funding

Best Practices

- Capture lessons learned
- Document techniques, issues and solutions
- Document cost models and benchmarks
- Build a cloud platform support community
- Maintain platforms, scripts and updates for client agencies

Results

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

GeoCloud Programmatic Goals, Objectives, Activities and Outcomes

Architectures

- Identify requirements-driven solution architectures and platforms for various sized deployments of geospatial data and services

Acquire, compose, **document, and deploy reference platform cloud that support Geospatial Platform standards**

Cost Models

- Document and assess cost models to support scalability, reliability, and redundancy

Monitor costs, loads, issues and options in support of OMB IT project document guidance

Certification

- Expedite FISMA (security) certification and accreditation for agency adoption of packaged solution architectures

Certify Geospatial Solution Packages to facilitate re-use

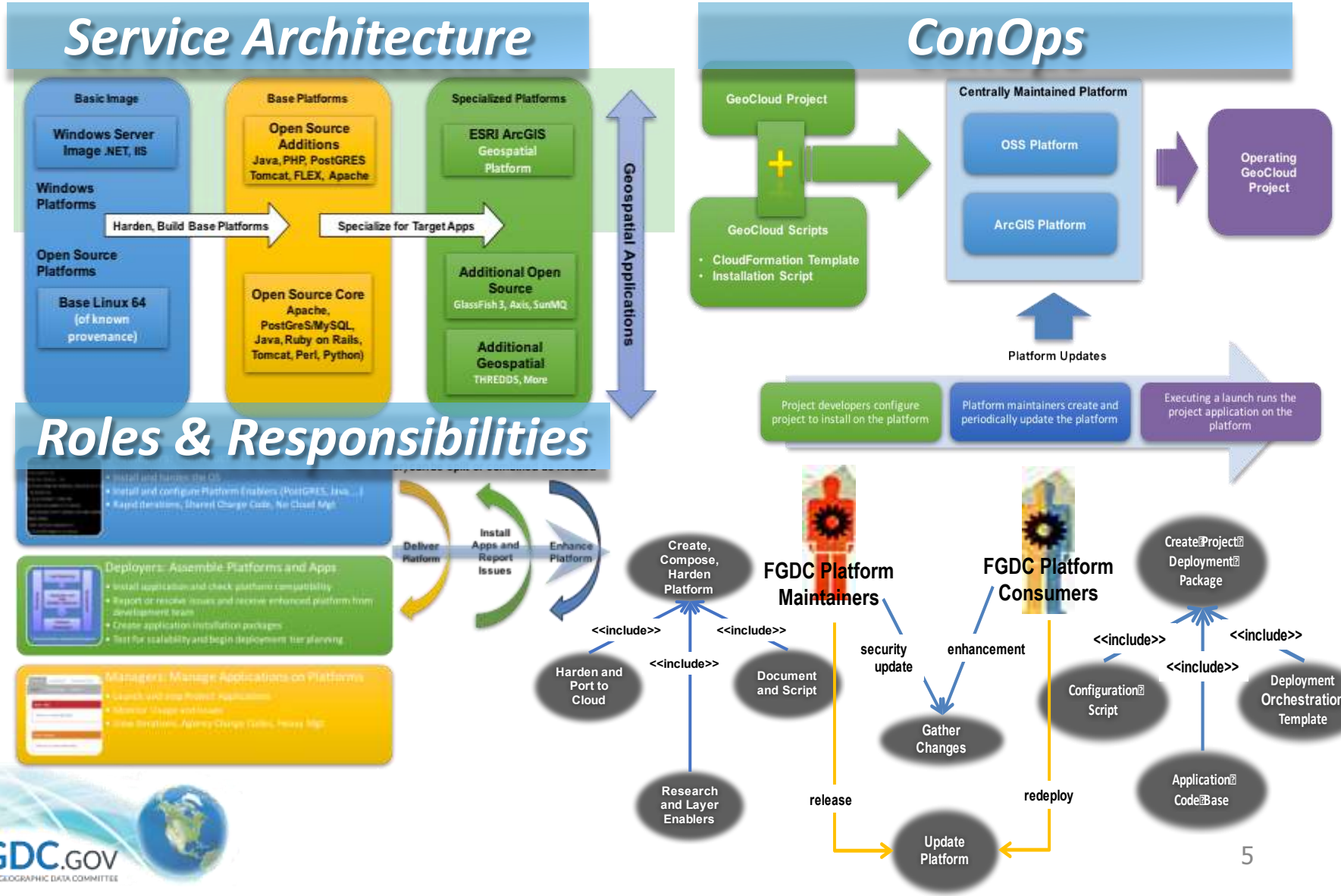
Comparisons

- Support and collect cost comparison information from agencies for existing and externally-hosted Cloud solutions
- Document lessons learned and best practices

Document Best Practices and guides to agencies on adoption of geospatial Cloud infrastructure

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

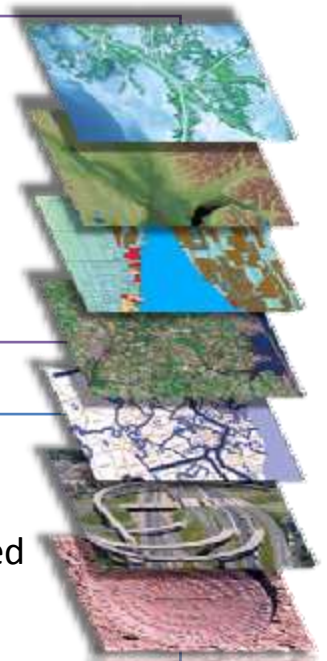
GeoCloud Outcomes – Development of a Cloud Services Framework



GeoCloud Technical Goals and Objectives – *Community Platform Sandbox*

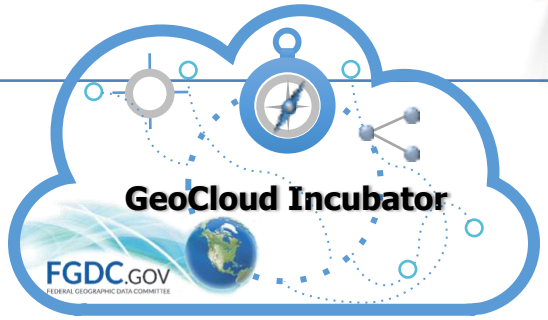
Goals

- Protype/Pilot applications for use on a GeoCloud Platform
- Determine the application's deployment and scalability requirements
- Move to production account, operate and document metrics and experience
- Graduate to full operation as appropriate



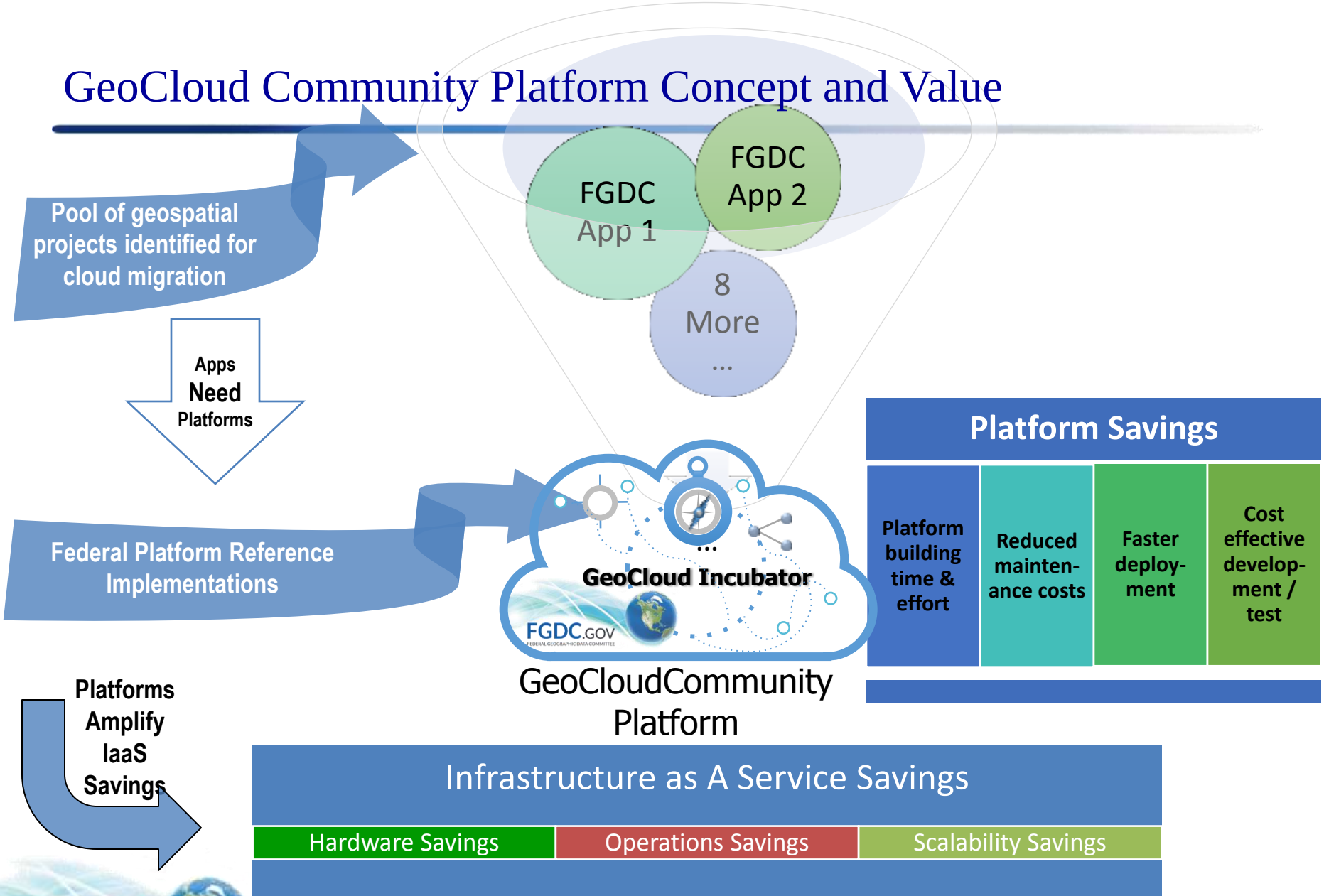
Objectives

- **Identify** and document the steps required to install the application
- **Document** any new application platform requirements and receive an enhanced platform from the platform development team
- **Create** and test application code bundles and installation scripts
- **Test** application for scalability and determine deployment strategy
- **Migrate** application to a production account and document

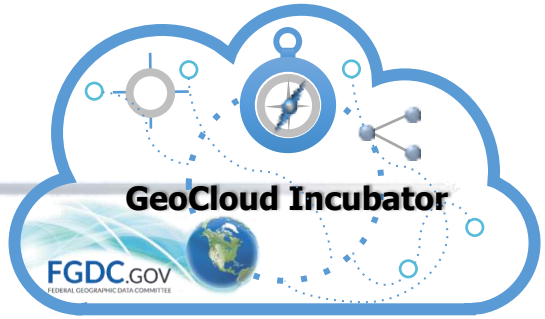


FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

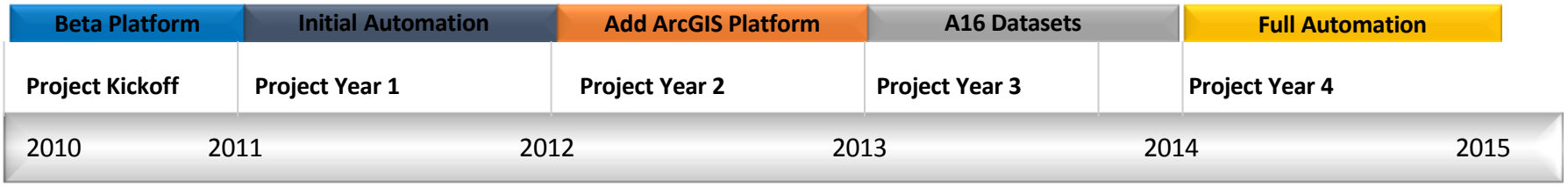
GeoCloud Community Platform Concept and Value



GeoCloud Sandbox History



Project Kickoff	Project Year 1 – General Purpose	Project Year 2 – Add ArcGIS	Project Year 3 – Focus on A16 Datasets	Final Project Year
• Build manual platform • Develop business processes • Build community portal • Find agency candidates	• Build limited platform automation • Add OpenGeo to OSS platform • 11 projects across 6 agencies	• Increase GeoSpatial focus • Add GeoNode to OSS platform • Initial hardened ArcGIS platform • 15 agency projects across 10 agencies	• Revise and update OSS and ArcGIS platforms • Initial CKAN prototype • 4 projects, 4 grandfathered A16 projects	• Introduce platform.mu with full automation • Add Ubuntu OS • Migrate 2 new services – GeoNode and GeoShape • 6 new projects



FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

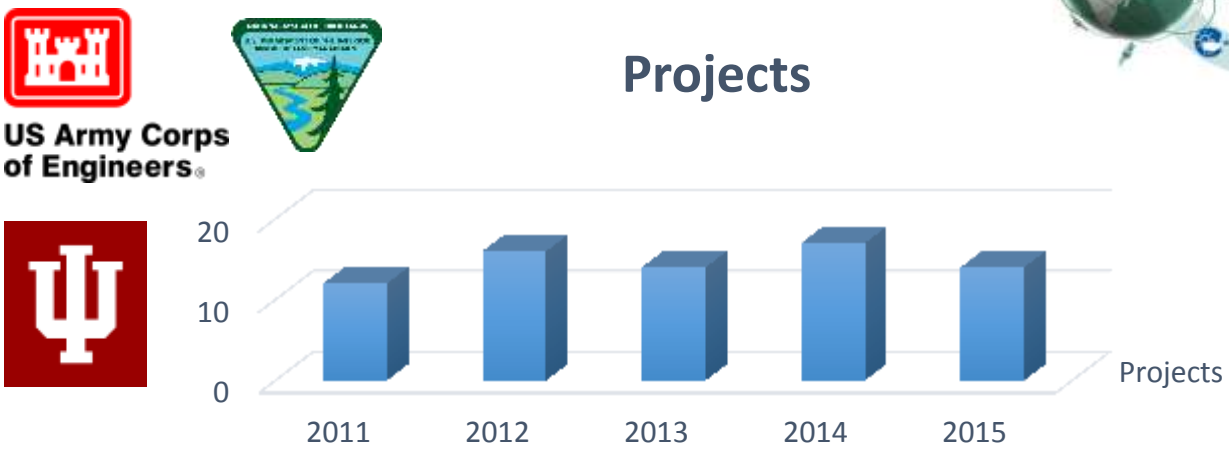
GeoCloud Community Platform – *Participating Organizations*

*Sample Participating Organizations

- Federal
- State
- Academia
- Tribal
- International

GeoCloud Incubator

US Army Corps of Engineers



Cloud Community - Sample Target Applications

~37 Projects

Platform	Based On	Readiness	Base Platform	Category	Most Suitable For
Wetlands Mapper	Windows/IIS/ESRI	Tier 1	Windows ESRI Platform with additions for ORACLE as needed	Normal	Public Cloud
National Elevation Data Set	Windows/TomCat/ORACLE /ESRI	Tier 3		Compute+ Data	Private/Community
National Map	Windows/TomCat/ORACLE /ESRI	Tier 3		Compute + Data	Private/Community
EPA Lakes & Ponds	Windows / PostGres/ESRI	Tier 2			
Semantic Wiki +	Linux	Tier 1	Linux Open Source Platform with Java/Ruby on Rails/Tomcat,/EJB3, PostGres, MySQL (alt) Python, PHP, Perl and Apache Good candidate for a single base platform	Normal	Public Cloud
NOAA Particles	Linux/Java/Axis/MQ	Tier 2		Normal	Public Cloud
GEOSS GeoNetwork w. additions	Linux/TomCat/PostGres	Tier 1		Normal	Public Cloud
TIGER Shapefiles	Linux / Apache / Perl/Python	Tier 2		Normal	Public Cloud
IOOS Registry +	Linux/TomCat/PostGres	Tier 2 ?		Normal	Public Cloud
ERDAP	Linux/TomCat/THREDDS	Tier 1		Data	TBD
Semantic Content Mgt - Drupal 7	Linux LAMP	Tier 3		Normal	Public Cloud
HHS Connect	Linux/EJB3 (GlassFish V3)	Tier 1		Normal	Public Cloud
Emergency Web Response Site	Linux LAMP	Tier 3		Normal	Public Cloud
OpenStreetMap	Linux/MySQL/Ruby On Rails	Tier 2		Normal	Public Cloud

GeoCloud – Pilot/Prototype Reports

National Center for Education Statistics

Amazon Hosted ESRI ArcGIS Servers
Project Final Report

NSDI Cooperative Agreements Program
Category 7: Geospatial Platform Cloud Service Test bed
Interim Report

Date: November 1, 2012

Agreement Number: G12AC20138

Project title: Ca
Implementation

Organization: Indiana University
620 Union Drive, R
Indianapolis, IN 46
http://www.iupui.ed

Principle Investigator: David J Bodenham

V. 317-274-2455
F. 317-278-1830
lntu100@iupui.edu

Table 7. Cost comparison between AWS services and local server

	Cost		Amazon EC2	Local Server
Fixed (one-time) cost			None	Around $2000 to purchase the computing instances
Monthly	Network cost		$1 for data transferring / month	None (Included in the maintenance fee)
	Storage		$22 /month	None (Included in the first purchase)
	Computing power	Not reserved	$255 / month	~$17 (Assuming the server can be used for 10 years, and then the each month cost would be (2000/(10*12)))
		Reserved	$76 / month	
	Maintenance cost (Cooling, system, sys admin maintenance, room etc)		None	$200 / per month (Assuming $100 for cooling, network, and room fee, and $100 for paying a sys admin to check and maintain the server)
Yearly	Total	Not reserved	$3300	$2604
		Reserved	$1104	
Note: The cloud computing cost is based on the standard large instance with 2 CPU cores of EC2 to host GEOSS Clearinghouse. If more advanced instance types are selected, the computing power cost would increase				

Figure1 August 2012 Indiana Harvest Parcels

Figure 1: AWS Console Monitoring Interface

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

GeoCloud Technical Services and Resources Tested

GeoSpatial

GeoShape*

ArcGIS

GeoServer

GeoNode *

ckan

Management

Confluence **

JIRA

X-WIKI

WordPress

Drupal

splunk> **

Jenkins

App Servers

Apache Tomcat

django

node.js *

RAILS

Microsoft .net

LAMP

sftp

Platform Tools

JAVA J2EE

Microsoft SQL Server

MySQL

ORACLE DATABASE

PostgreSQL

Gluster

mongoDB

Operating Systems

CentOS

ubuntu

Microsoft Windows Server System

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

Examples
– EPA How's My Waterway

GeoCloud Incubator

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Examples – USGS BISON Tool

GeoCloud Incubator

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

USGS
science for a changing world

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

Examples – HUD and EPA

GeoCloud Incubator

FGDC.GOV

NEPAnode provides a suite of geospatial enriched shared web services, software, and GIS tools to assist practitioners, resource managers and decision makers in reporting and understanding environmental assessments and permitting.

The HUD Storefront unifies public access to HUD's portfolio of applications and geospatial data resources, including the 11 National Geospatial Data Asset (NGDA) datasets that HUD is currently contributing to the Federal geospatial data portfolio.

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

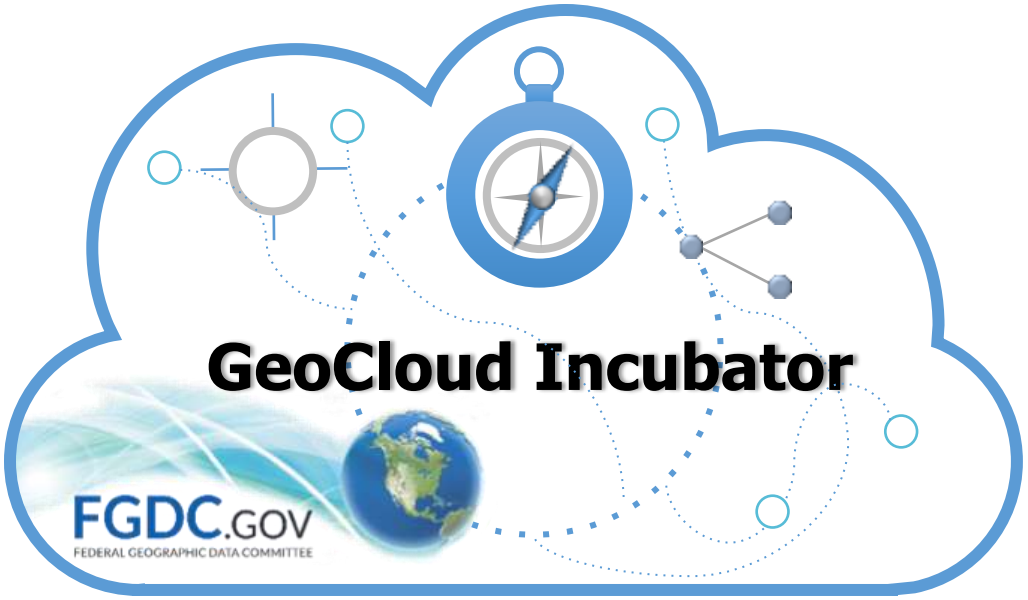
GeoCloud – Lessons Learned

- **Collaboration:** Community of Practice was a positive method for this effort
- **Cost Efficiency:** Cloud offers strong potential for operational savings
- **Architecture:** Application re-engineering generally necessary to take optimize use
- **Security:** Securing Cloud Services very challenging

GeoCloud – Recommendations

- Continue Collaborative R&D of Emerging Cloud Services
- Education and Awareness
- Share Lessons Learned
- Establish Resources via the GeoPlatform
- Security Architecture Development and Efficiency

Summary



Questions?

FGDC.GOV
FEDERAL GEOGRAPHIC DATA COMMITTEE

Rich Frazier
efrazier@usgs.gov