

Private Cloud Environment within OpenStack

Jingchao Yang (presenter)

Keith Clarke, Wendy Guan, Mengchao Xu, Fei Hu, Chaowei Yang, and others

NSF Spatiotemporal Innovation Center

George Mason University, UC-Santa Barbara, and Harvard University





Outline

- Project Overview
- Cloud Infrastructure and Network Configuration
- Service Communities
- Accomplishments and Sample Projects
- Impact
- Plans for the Next Reporting Period



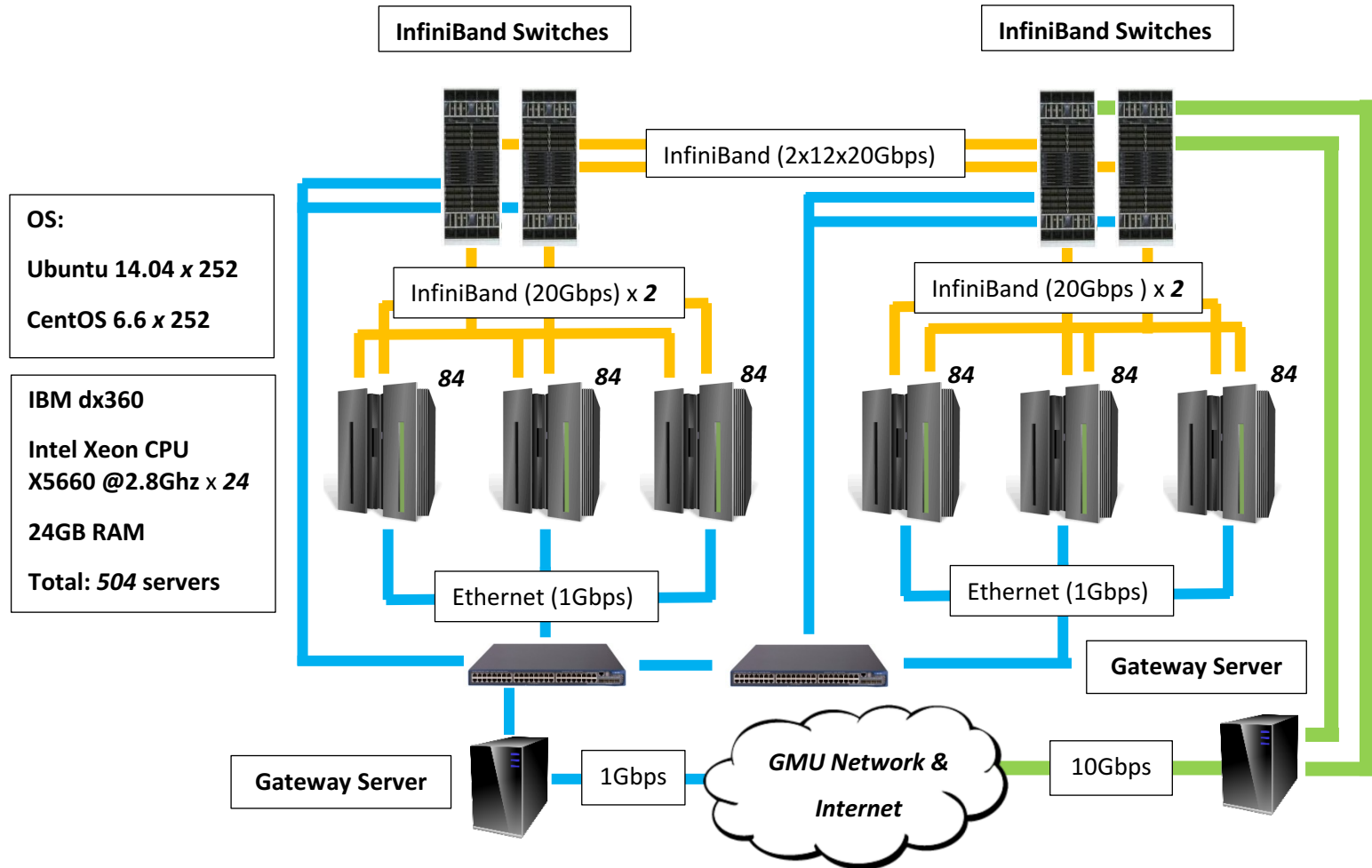


Project Overview

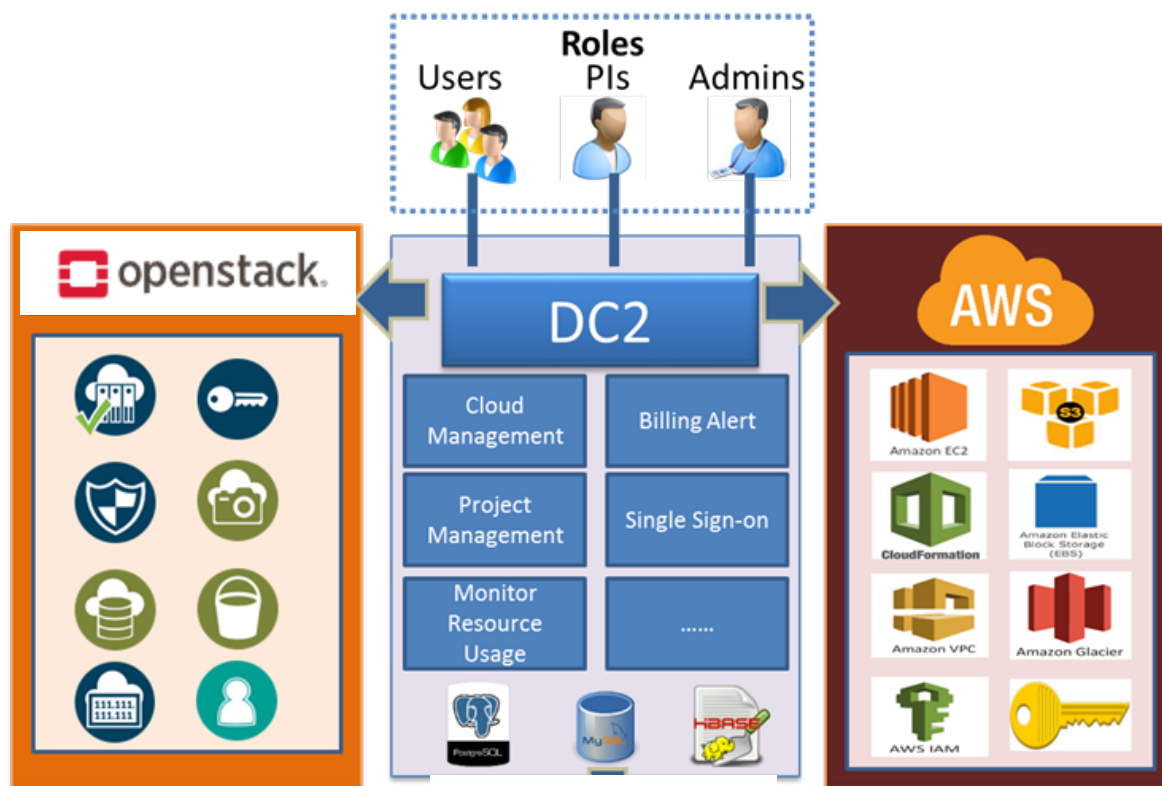
- Big data analytics requires significant computing support for model simulations, e.g. dust, flooding, fire, storm, require High Performance Computing (HPC)
- Most center projects need computing support
- The community suggested the building of a cross site center wide computing infrastructure to remove duplication and improve computing capability to meet center advancement needs.
- Develop and maintain a spatiotemporal computing infrastructure.
- Provide spatiotemporal computing coordination to all projects that need computing by
 - Acquiring a high performance computing facility
 - Developing data mining components to detect security, usage and other patterns of interest
 - Maintaining a highly capable research staff support for computing optimization



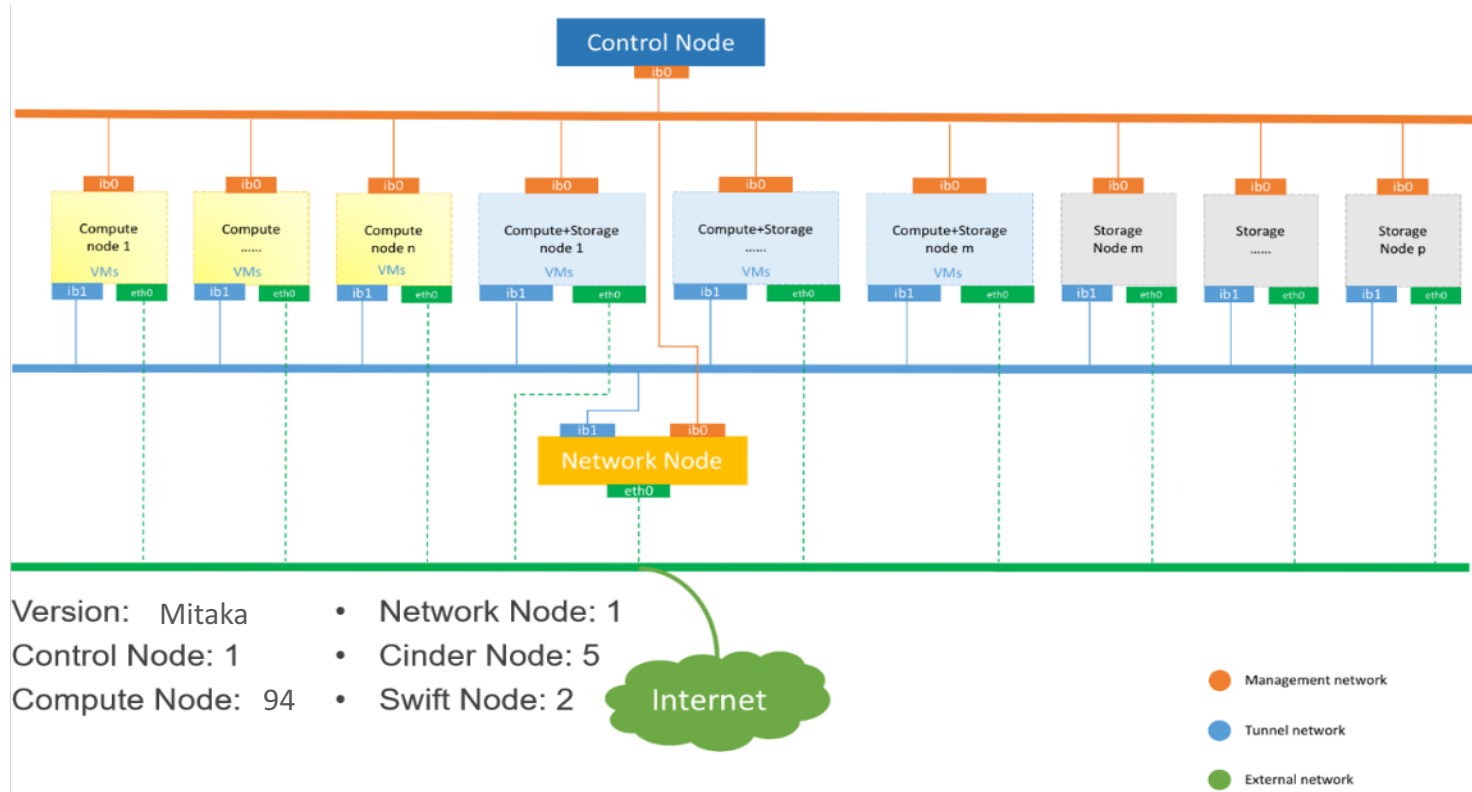
Infrastructure setup and network configuration



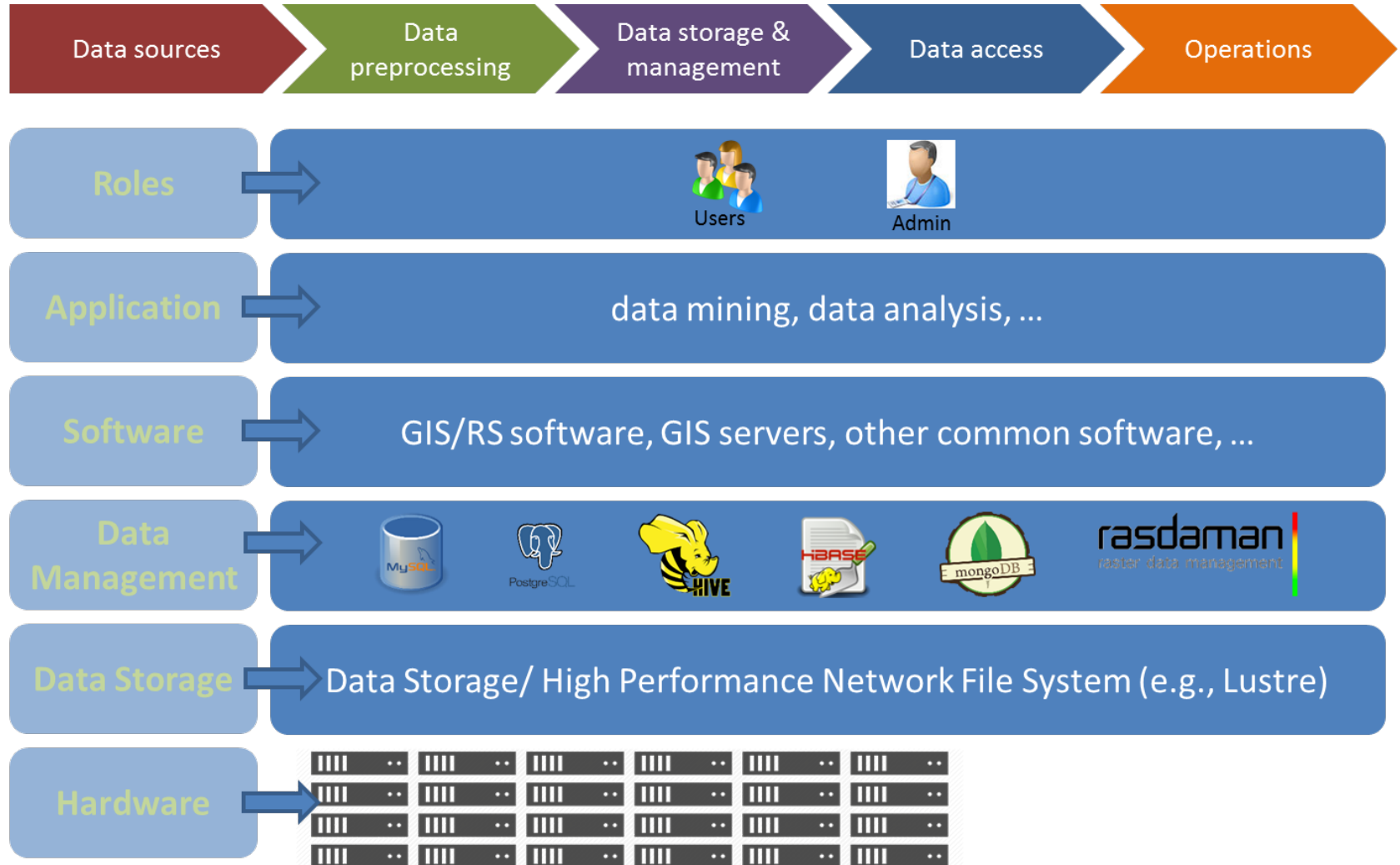
Cloud software management architecture



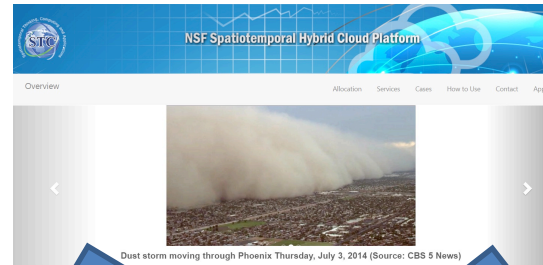
Cloud configuration example - Openstack



Planned Software/data infrastructure



Service Communities



EARTH CUBE

*NSF EarthCube Integration
and Testing Environment (ECITE):*



Dashboard : Project Management > [Project](#)

Create Project Operate Refresh Delete Project Tag

| <input type="checkbox"/> | Action | Project title | Description | Date Created | Cloud Account | Allocated Funding (\$) |
|--------------------------|--------|------------------|---|--------------------------|-------------------|------------------------|
| <input type="checkbox"/> | | BCube | BCube in ECITE | Apr 22, 2016 2:19:06 PM | aws-ecite-operate | 2000.00 |
| <input type="checkbox"/> | | emily-project | For testing | Apr 19, 2016 5:52:13 PM | aws-ecite-operate | 500.00 |
| <input type="checkbox"/> | | STCTestbed | STCTestbed | Apr 11, 2016 2:27:28 PM | aws-ecite-operate | 2000.00 |
| <input type="checkbox"/> | | EarthCube-DASHER | NSF EarthCube DASHER CD | Apr 7, 2016 4:27:49 PM | aws-ecite-operate | 500.00 |
| <input type="checkbox"/> | | CINERGI | Community Inventory of EarthCube Resources for Geosciences Interoperability | Feb 29, 2016 10:26:16 AM | aws-ecite-operate | 2000.00 |
| <input type="checkbox"/> | | CHORDS | Cloud-Hosted Real-time Data Services for the Geosciences | Feb 22, 2016 11:01:28 AM | aws-ecite-operate | 2000.00 |





Accomplishments

- Hardware: an NCCS (NASA Center for Climate Simulation), donation of 504-node computer cluster with two 20Gbps internal connections and a 1Gbps management connection
- Websites:
 - Information portal: cloud.gmu.edu
- Documents:
 - Maintenance documents
 - User manuals
 - Operation specifications

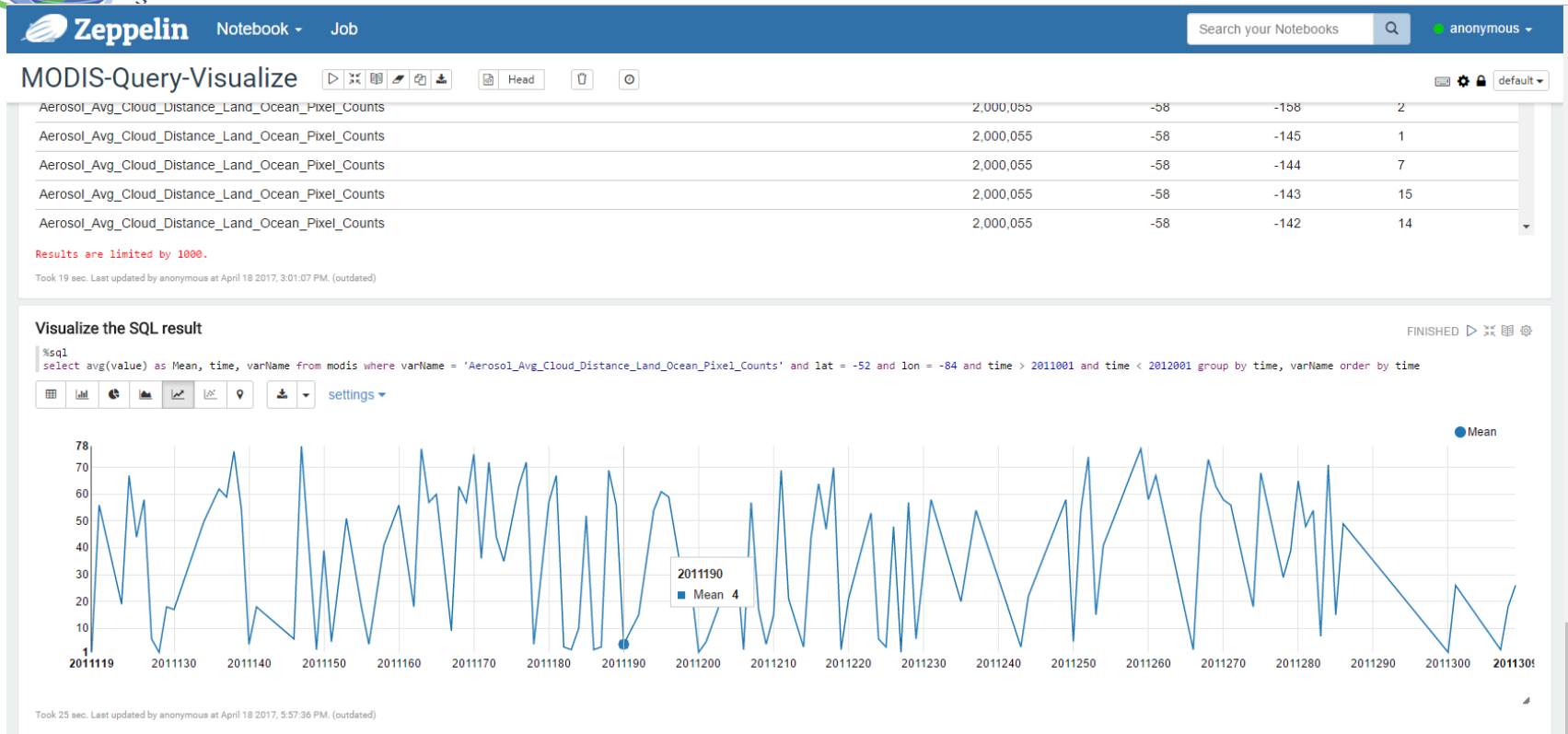


- Consolidate and reduce cost spent on computing resources at the center
- Provide a computing testbed for center projects and potential IAB directed research
- Accumulate advanced computing operation experiences
- Serve NSF EarthCube program through ECITE
- Serve ESIP through its cloud computing cluster
- Serve the community through AAG, Geo4All, etc.

Sample projects on the cloud



GMU-15-01 ClimateSpark: An In-memory Distributed Computing Framework for Big Climate Data Analytics



A web portal to ease the remote interaction between climatologists, climate data, analytic operations, and computing resources (e.g. using SQL query and Scala/Python notebook)

AWS:

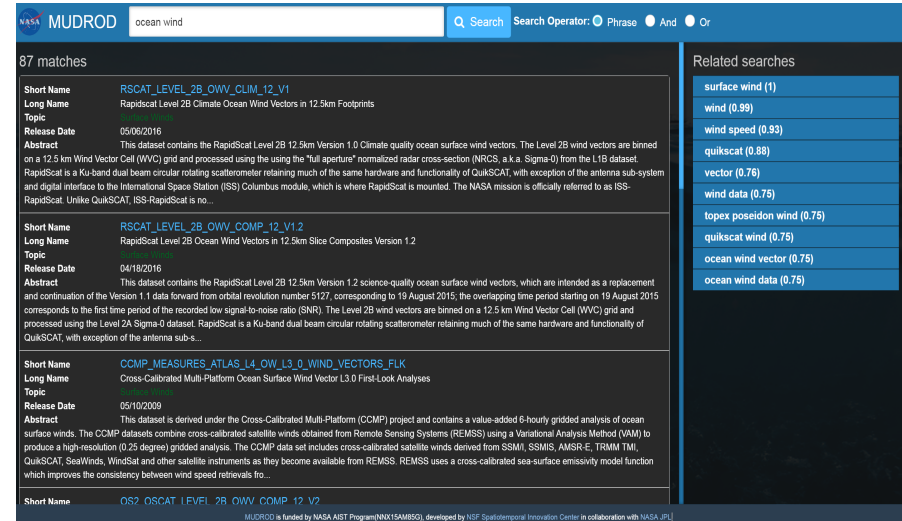
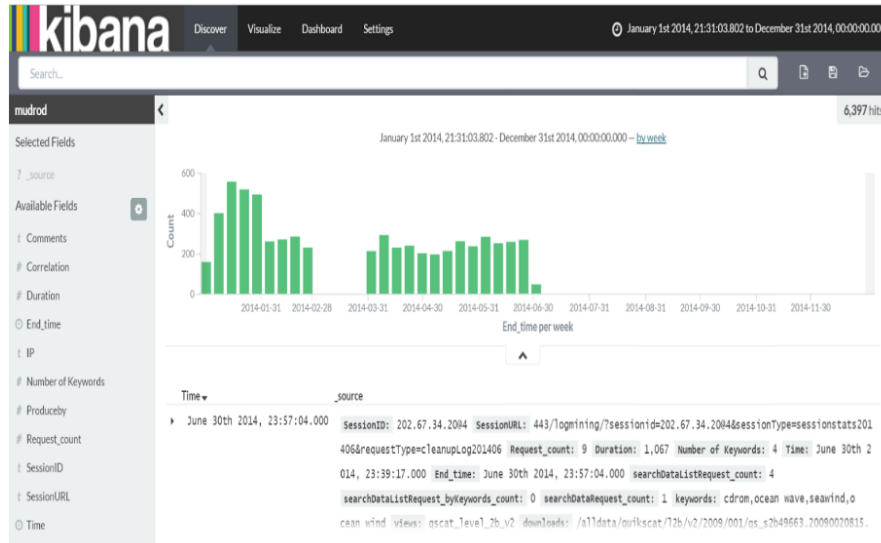
<http://10.192.20.172:8080/#/notebook/2C172AC58>

Supported by NCCS (NASA Center for Climate Simulations/Daniel Duffy)





MUDROD, NASA AIST 2014, Managed by Mike Little



- Analyze web logs to discover user knowledge
- Data management
- Search engine
- **OpenStack:**
- <http://199.26.254.151:8080/mudrod-service/>





GMU-15-09 Planetary Defense Funded by NASA PDCO

PLANETARY DEFENSE (PD) FRAMEWORK GATEWAY

[HOME](#) | [FILEDEPOT](#) | [VOCABULARY LIST](#) | [CRAWLER DB](#) | [HOW TO USE](#)

INTRODUCTION

The NASA Planetary Defense Coordination Office (PDCO) was established in 2016 to study the mitigation of potential Near-Earth Object (NEO) impacts to our home planet. NASA Goddard and National Nuclear Security Administration (NNSA) established a collaboration to study the short time response options to potentially hazardous objects (PHOS). The motivation of designing this architectural framework is to develop an integrated architecture for the process of detecting, characterizing and mitigating NEO threats. The project is currently a conglomeration of individual facilities conducting separate research and this framework is meant to help define a collaborative system based on data reporting and sharing across various elements of the collaboration project. With this architecture, the NEO project will increase the efficiency, accuracy, and timeliness of its assessments. The benefits that come from—1) maximizing the linkage between different organizations, scientists, and amateurs that are conducting individual NEO research; 2) leveraging the current available computing resources to conduct computation intensive tasks, including data storage, data analysis and modeling processes; 3) developing a sustainable architecture for the NEO defense system—to facilitate faster information sharing, more efficient computing capability, and the creation of a more effective mitigation design; 4) extending flexibly to future additions or evolutions of assets from the broader PD community. By employing this architecture, the NEO efforts can achieve objectives more efficiently, while further exploring the possible mitigation methods in order to protect our home planet.

The architectural framework is designed to accommodate the process of detecting, characterizing and mitigating Near-Earth Object (NEO) threats. The architecture organizes current data and resources into useful information and correlates that information with the goals of the project. The architectural framework will enable scientists, organizations, and decision makers to locate, identify and resolve definitions, properties, facts, constraints and issues with potentially hazardous asteroids. Our major focus is to design the data and information flow that models the complete process from NEO detection, to the design of mitigation strategies.

The data and information flow depicts what kind of information will be input to and output from each subsystem within the whole process. There are mainly three subsystems, NEO detection, characterization, and mitigation. During the NEO detection process, a diversity of observation systems are utilized to detect the potential NEO. These observations are submitted to databases such as IAU's Minor Planet Center for cataloging. Based on the remote observations, attributes of a NEO are calculated through trajectory analysis, such as diameter, spin state, mass, and orbit. The result of the analysis will trigger NEO characterization process through meteoritic samples or precursor work to obtain the physical and chemical properties. During the NEO characterization process, the calculated NEO properties are fed into physics-based models to calculate the energy deposition, hydrodynamics, and earth impacts, such as earthquakes and tsunamis. The model results are used for mitigation analysis and designing multiple reference missions. Each mission will be assessed for its effectiveness and risks.

The system architecture is designed to describe the supporting infrastructure for the framework. There are four levels of infrastructures that support the mitigation system. Firstly, remote observation systems are needed to detect and track potential NEOs. Secondly, databases and catalogs are needed for the effective management of NEO records with easy data access and sharing between organizations and agencies. Thirdly, computing infrastructure from Goddard is needed for efficient and accurate calculations for trajectory analysis, effectiveness assessment, risk analysis, and a visualization platform. Finally, HPC clusters at NNSA and NASA are needed to handle the intensive computation during physical based modeling. In addition, the promotion of interoperability, interagency information sharing, and improved computing capabilities can largely facilitate NEO defense decisions. As a strategic information asset, the NEO defense architecture framework will serve as an enabling tool to: support the alignment of resources with NEO defense actions, facilitate capability management, and ensure that operational goals and strategies are met.

USER LOGIN

Username *

Password *

[Create new account](#)
[Request new password](#)

[Log in](#)

This system is funded by NASA Goddard and NSF (NNG16PU001), and supported by NASA AIST (NNX15AM85G). Developed and hosted by NSF Spatiotemporal Innovation Center on the Hybrid Cloud Service.

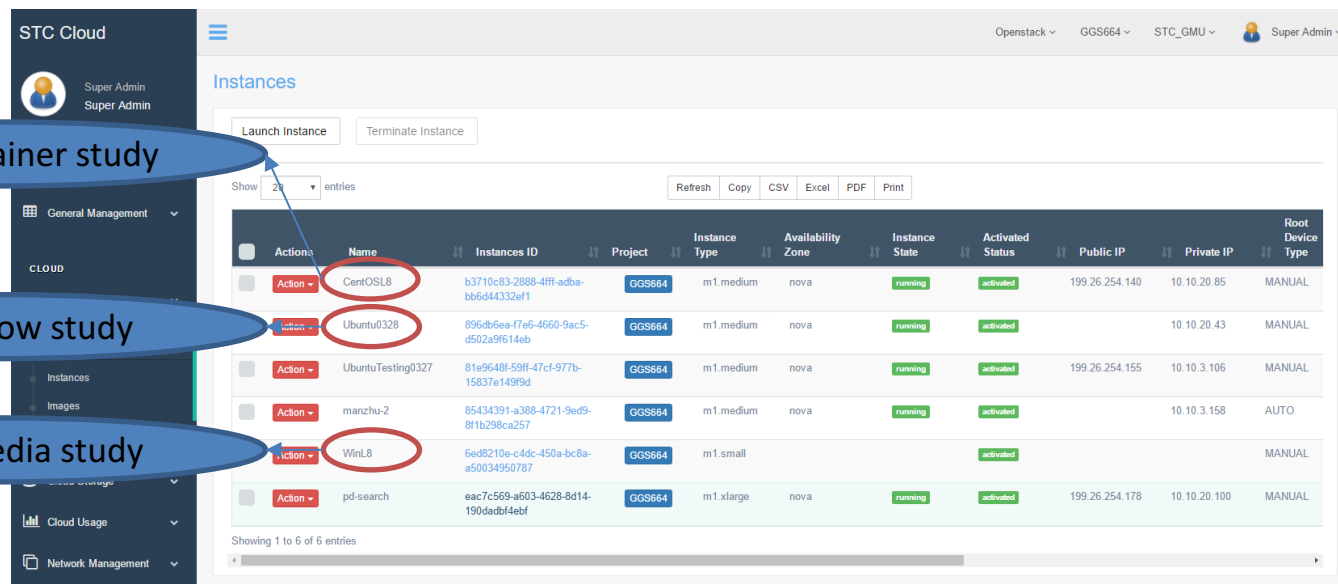
- Domain Crawler
- Search Ranking
- Thanks to Myra Bambacus and Bernie Serry

OpenStack:

— <http://pd.cloud.gmu.edu/>



- GGS 664 Spatiotemporal Data Structure
 - Easy the teaching processes by using platform to customize instances according to needs.



STC Cloud

Super Admin
Super Admin

General Management

CLOUD

Instances

Images

Cloud Usage

Network Management

Instances

Launch Instance Terminate Instance

Show 20 entries

Refresh Copy CSV Excel PDF Print

| | Actions | Name | Instances ID | Project | Instance Type | Availability Zone | Instance State | Activated Status | Public IP | Private IP | Root Device Type |
|--|----------|-------------------|--------------------------------------|---------|---------------|-------------------|----------------|------------------|----------------|--------------|------------------|
| | Action + | CentOS8 | b3710c83-2888-4fff-adba-bb6d44332ef1 | GGS664 | m1.medium | nova | running | activated | 199.26.254.140 | 10.10.20.85 | MANUAL |
| | Action + | Ubuntu0328 | 896db6ea-f7e6-4660-9ac5-d502a9f614eb | GGS664 | m1.medium | nova | running | activated | | 10.10.20.43 | MANUAL |
| | Action + | UbuntuTesting0327 | 81e9648f-59ff-47cf-977b-15837e149f9d | GGS664 | m1.medium | nova | running | activated | 199.26.254.155 | 10.10.3.106 | MANUAL |
| | Action + | manzhu-2 | 85434391-a388-4721-9ed9-8f1b298ca257 | GGS664 | m1.medium | nova | running | activated | | 10.10.3.158 | AUTO |
| | Action + | WinL8 | 6ed8210e-c4dc-450a-bc8a-a50034950787 | GGS664 | m1.small | | | activated | | | MANUAL |
| | Action + | pd-search | eac7c569-a603-4628-8d14-190dadbf4ebf | GGS664 | m1.xlarge | nova | running | activated | 199.26.254.178 | 10.10.20.100 | MANUAL |

Showing 1 to 6 of 6 entries

- Different instances assigned to students, providing different operation system environments and preinstalled with different software



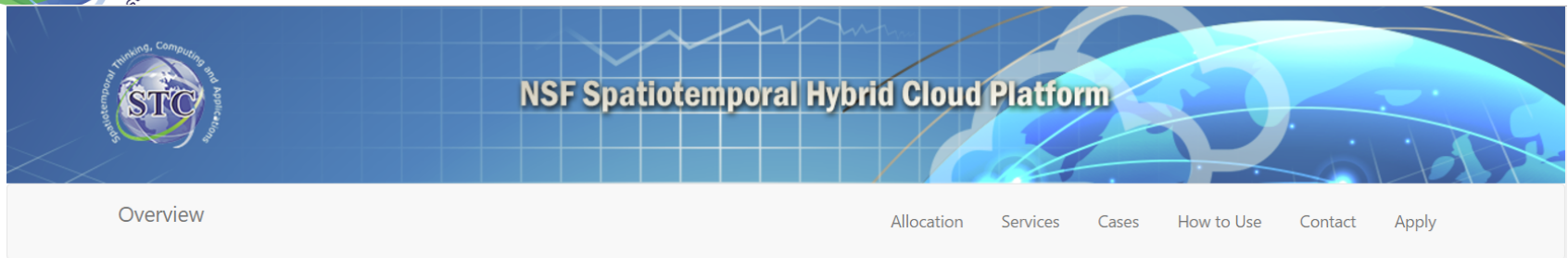
Functionalities being added

- One-button Hadoop
- Auto Scaling
- Load Balancing
- Develop big data management platform





Apply for Platform Resources



Note: Please fill out **all** the sections/fields. If you have any question or additional requirements, please contact us: stc@gmu.edu

* You will receive a confirmation email while we are processing your application

Overview

Project name:

Enter project name

Project participant role(s):

Enter role(s) (e.g. Sarah - project manager; Peterson - priviledge user...)
Please list your names and relevant roles (project manager*, priviledge user or end user). Your account(s) will be created based on the role(s)

Contact name:

Enter name

Contact email:

Enter email

Organization(s):

Enter applicants' organization(s) (e.g. Sarah - gmu stc center; Peterson - nasa...)

Project description and purpose:

- <http://sites.cloud.gmu.edu/sthcp/form.php>





Contacts

- Jingchao Yang, jyang43@gmu.edu
- Phil Yang, cyang3@gmu.edu

