ESIP Lab Project

**Geoweaver: a web-based system for managing compound geospatial workflows of large-scale distributed deep networks**

August Progress Report

Prepared by

Ziheng Sun, Liping Di

Center for Spatial Information Science and Systems

George Mason University

20180828

Contents

[Executive Summary 2](#_Toc523305779)

[Project Actions 3](#_Toc523305780)

[Project Kick-off 3](#_Toc523305781)

[System Architecture 3](#_Toc523305782)

[Development Environment Setup 3](#_Toc523305783)

[Web wrapper for deep learning and HPC library/software 3](#_Toc523305784)

[Geoweaver programming 3](#_Toc523305785)

[LSTM Experiment 4](#_Toc523305786)

[Geoweaver GitHub repository 5](#_Toc523305787)

[Evaluation 6](#_Toc523305788)

[Next Steps 7](#_Toc523305789)

Executive Summary

This month Geoweaver team has kicked off the project and started the development lifecycle as scheduled. Our progresses in August are listed in the following table:

|  |  |  |
| --- | --- | --- |
| **Milestones** | **Progress** | **Actions** |
| July 31: kick off, set up the development environment, develop web wrapper on top of open sourced deep learning/high performance computing library | 100% | * Kick off announcement in CSISS * Development environment discussion and setting up * Deep learning library selection (Keras/DL4j+Tensorflow) * Cloud instances started * High performance computing library (Spark) is installed on the instance VMs * A web wrapper is developed to invoke Spark and DL4j via the Web |
| Sep 30: develop workflow designer and data producer, complete bridge assembly between Geoweaver and data/function resources | 10% | * Create Geoweaver project in IDE * Download and configure client Javascript library D3 in Geoweaver |
| Nov 30: complete module integration, create and conduct LSTM experiment | 10% | * Move Landsat 5 and 8 images onto the created cloud VMs * Set up LSTM experiment environment on cloud |
| Jan 31: complete source code wrap-up, upload demonstration video, snapshot cloud instance, finish the GitHub final report and demonstrate it in ESIP winter 2019 | 10% | * Geoweaver repository on GitHub is created. |

Project Actions

Project Kick-off

We announced the kick-off of this project on July 30, 2018 in CSISS.

System Architecture

The architecture design is as shown in Fig. 1. The five components in Geoweaver will create a friendly web system for users to manage their deep learning experiments running on backend. Each experiment is managed as a processing workflow/VDP (virtual data product). We will stick to this design throughout the project.

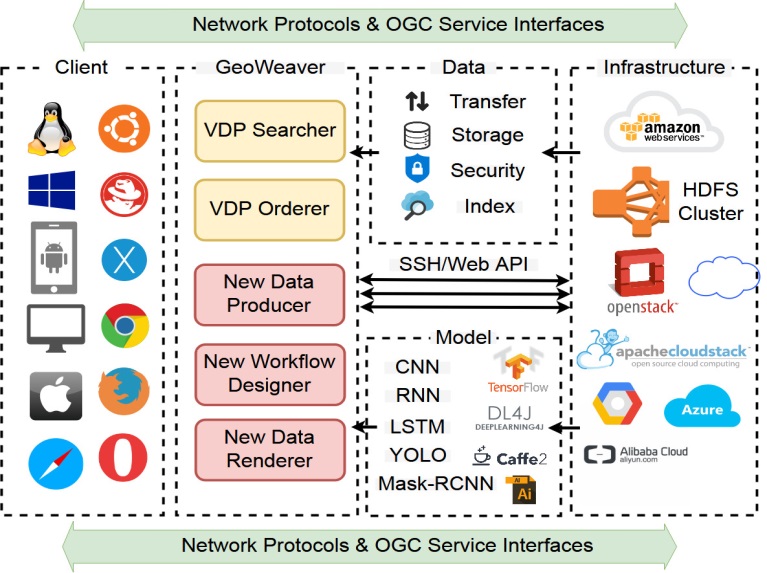


Figure 1. Architecutre design of Geweaver

Development Environment Setup

We choose Java (1.8) and Javascript as our main developing languages for the backend and frontend, respectively. Eclipse (Jee Oxygen) is used as our IDE. The deep learning library we use includes deeplearng4j and keras. Deeplearning4j can reuse keras. The cloud platform uses our private cloud: GeoBrain (<http://cloud.csiss.gmu.edu>). We have created ten instance VMs to run Spark and Deeplearning4j. Spark has been installed onto the VMs (Fig. 2).

Web wrapper for deep learning and HPC library/software

We reuse the backend of EarthCube CyberConnector to open the connection to the server resources including storage and computing power. We add new functions to the server application to enable the manipulation of deep learning and hpc software installed on the server. The real time logging is dynamically monitored and collected for further query from the client side.

Geoweaver programming

Geoweaver is created as a system backed by EarthCube CyberConnector (Fig. 3). D3.JS (<https://d3js.org/>) has been imported as renderring library. We are currently working on developing the interface.

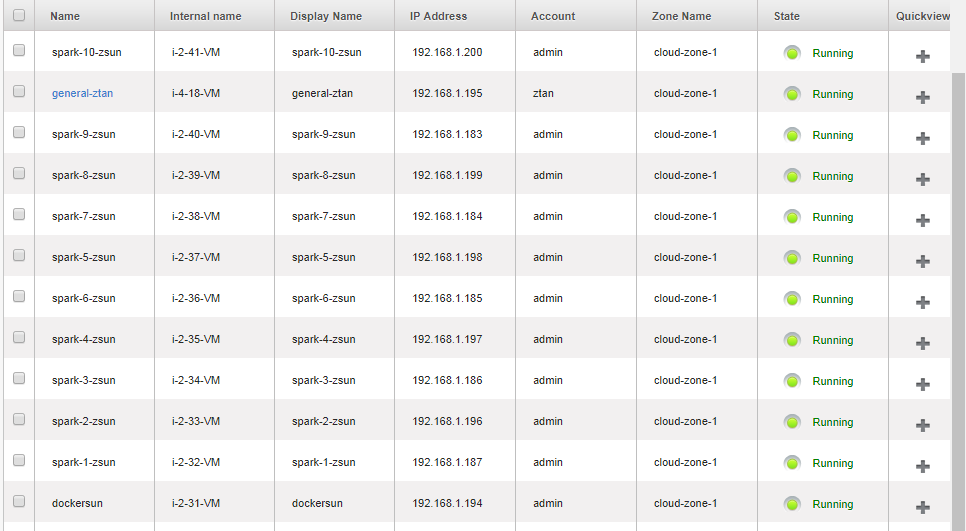


Figure 2. The list of created instance VMs

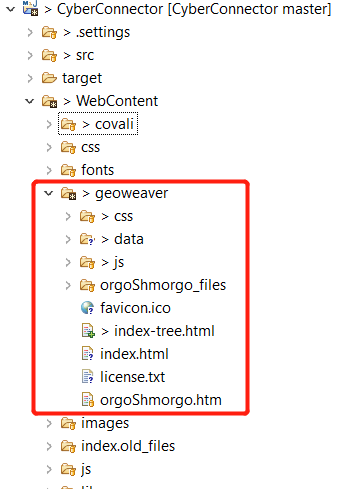


Figure 3. Geoweaver code

LSTM Experiment

We are transferring the data and scripts created in our LSTM experiment onto the created instance VMs. The experiment is trying to utilize the seasonal pattern of crops (Fig. 4) in improving the land cover land use classification in remote sensing.

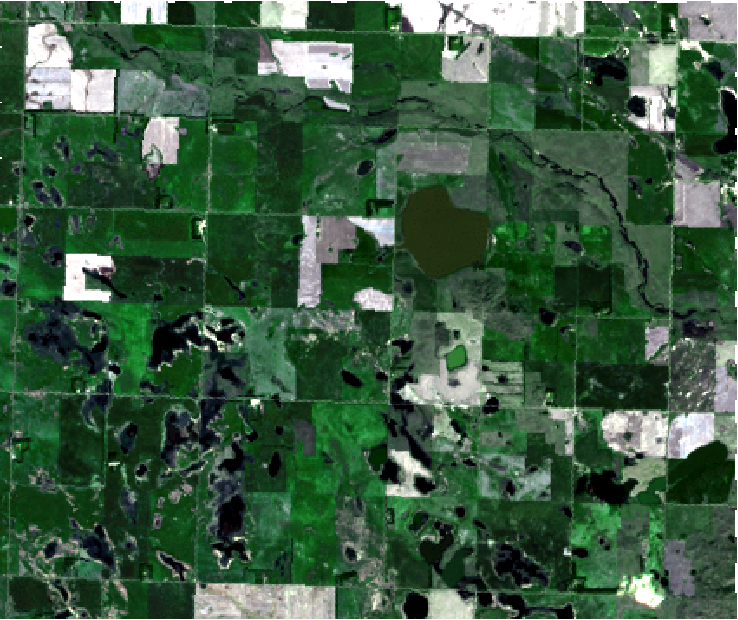
 

Figure 4. Left (26 May 2014) Right (15 Sep 2014)

Geoweaver GitHub repository

A repository has been created (<https://github.com/ZihengSun/Geoweaver>). Wondering how to add it to the ESIPFed GitHub organization (<https://github.com/ESIPFed>).

Evaluation

Not yet.

Next Steps

* Continue the development of Geoweaver user interface.
* Complete the development of Geoweaver backend VDP components.
* Use Geoweaver to manage our LSTM experiments.
* Evaluate Geoweaver.