

Spring 2015 - Big Data Open Source Software Project (BUEX-V 594)
May-01,2015

PROJECT PROPOSAL

Dev-ops for Apache Storm on
Openstack

I.	TEAM.....	4
II.	DESCRIPTION	4
III.	DELIVERABLES	5
IV.	TIMETABLE	5
V.	NEXT STEPS	5
VI.	ACKNOWLEDGEMENTS	6
VII.	APPENDIX.....	6

I. Team: Working individually

- Oliver Lewis
- Email: oliverlewis7@gmail.com
- Github username: oliverlewis
- Portal name: oliveral

II. Description

The project is about using a Dev-ops tool to install Apache Storm on an Openstack cluster. For this project after evaluating various configuration management and orchestration tools like Chef, Puppet, Ansible.

Reason to use Ansible: I've decided to use Ansible as a dev-ops tool for this project because of its simplicity. Ansible is rich with built-in modules that help you do anything from trivial tasks such as installing a package to more complex tasks. Ansible uses playbooks that are an equivalent to chef recipes. Ansible is also agentless, which means that there's no need to install anything on your servers in order to use it. Playbook semantics are so intuitive and easy to write that we no longer bother connecting to servers and running commands manually through the shell

Ansible is an IT automation tool which deploys software and configures systems on multiple servers using SSH protocol. Python based Ansible stores information about deployment and configuration in a yaml format file, named Ansible Playbook. Tasks defined in the playbook allows you to have identical configurations and software across multiple machines in your infrastructure.

[http://cloudmesh.github.io/introduction_to_cloud_computing/class/lesson/devops/ansible.html#ref-class-lesson-devops-ansible]

Apache storm uses Zookeeper for handling its distributed configuration service. So basically I have to setup zookeeper and then configure it to work on a setup of Apache Storm. This can be done on a single machine but for this project I will try a more practical and realistic scenario of using multiple machines.

I'll be using a 3 node Openstack network.

Node 1: Will contain the Zookeeper server and Storm Nimbus node

Node 2 & 3: Will be the supervisor nodes.

III. Deliverables:

1. I will write 2 Ansible playbooks to first install all the prerequisites required for Apache Storm. The first playbook will install Apache Zookeeper and take care of the prerequisites.
2. The second Ansible playbook will install Apache storm on the Nimbus node and start the Storm nimbus command.
3. It will also install Apache Storm on the supervisor nodes and start the Storm supervisor command.

IV. Timetable

Provide detailed information on the expected timetable for the project. Break the project into phases, and provide a schedule for each phase.

	Description of Work	Start and End Dates
Phase One	Setup Openstack 3 node environment.	April 13 th , 2015
Phase Two	Learn to use Dev-ops tool Ansible.	April 17 th , 2015
Phase Three	Install Zookeeper using Ansible.	April 19 th , 2015
Phase Four	Install Apache Storm using Ansible.	April 24 th , 2015
Phase Five	Test Ansible script on a new environment.	May 7 th , 2015

V. Next Steps

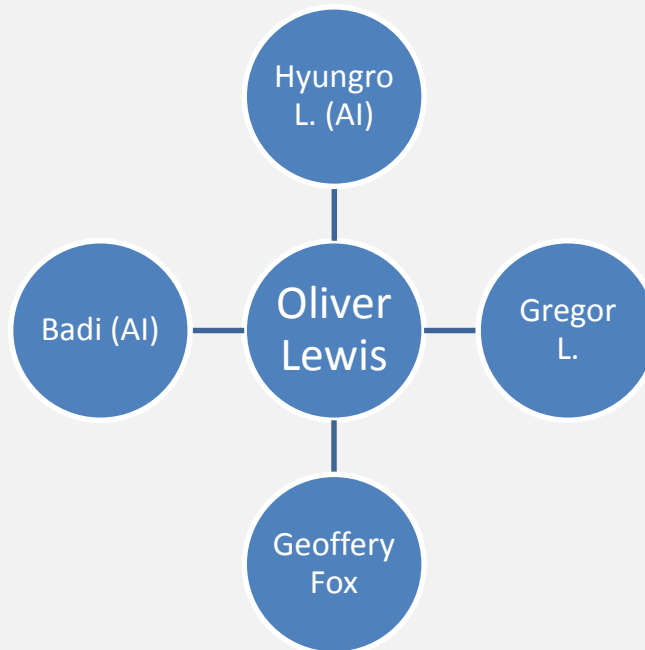
Specify the actions required of the readers of this document.

- Be able to install any apache component by reusing the Ansible template created for the installation of Apache Storm.
- Try using other dev-ops tools to create and manage a similar deployment. Example: Tools like Puppet or Chef.

VI. Acknowledgements:

I would like to thank Prof. Geoffery Fox for giving me this topic as my independent study.

I'd also like to thank all the instructors of the course who I have collaborated directly with.



XI. Appendix

Documentation on Github: <https://github.com/futuresystems/465-oliverlewis/project>

http://cloudmesh.github.io/introduction_to_cloud_computing/classes/bdosp_sp15/week_plan.html