1. ***Introduction***

This project aims at setting up a local cloud computing using ***Openstack*** and ***BDAS*** (Berkeley Data Analytic Stack).

With ***Openstack***, a local cloud computing environment could be deployed, so that web services like creating virtual machine on and creating cloud storage become possible in our local server.

Berkeley Data Analytic Stack is a software stack that provides an outline about what are the essential components for setting up a big data analysis cluster, components like Hadoop and Spark are necessary for Big Data analysis job.

So, we attempt to build a local cloud computing environment not only targeting at providing simple computational service, but also able to launch a big data analysis cluster that is with ***Hadoop*** and ***Spark*** installed and configured.

First, we start with ***XenServer*** as the hypervisor, which proved to be a very unclever choice. Though ***XenServer*** does very good in virtualization technology, it does not provide any Web Interface so that a pure ***XenServer*** could not be configured as a web service. ***Openstack*** could be installed together with ***XenServer***, so that a web service interface could be used to launch virtual machine on a ***XenServer***. However, there are many incompatibility between ***Openstack*** and ***XenServer***, so after putting much effort into this, we decide we should not continue investing into ***XenServer*** based ***Openstack***, which seems to be too costly.

Then, we switch to set up ***Openstack*** on ***Ubuntu Server***. Openstack has a relatively completed operational manual for using Ubuntu as the underlying hypervisor. But still it needs quite a lot extra configuration to have the cloud computing platform built with Hadoop, Spark, and also works with the existing network.

We have found out way to make Openstack works, but still there are problems that need more knowledge in Openstack to solve. In this project, we did learn a lot about how a cloud computing platform works, and how to set up one, it is a very precious experience for all the teammates. Setting up local cloud computing is not an easy job, but it really worth the effort, especially when Cloud computing becomes a trend of the new development in Computer Industry.

***2. Motivation***

There are several advantages of using a local cloud computing environment. It provides a limited accessibility like desktop computing, while also provides a combination of computing resource like cloud computing. Its security and convenience is what a public cloud service cannot match. There are three particularly important reasons why we want to build a local cloud computing environment for Carnegie Mellon University - SV campus.

**Local Resource can be conveniently and safely used**

There are many kinds of cloud computing services all around us these days, like Amazon Web Service (AWS) and Google Cloud Computing Services. They all provide highly reliable computing resources and user friendly UIs, but they all face one problem: If you want to do any kinds of data analyzing, you have to upload your original data up to the cloud. This is considered time consuming and inconvenient, especially when the amount of data needs to be processed is extremely large. As we are in the age of big data, situations like this are inevitable. On the other hand, although the public cloud services all provides storage services like Amazon Simple Storage Service (S3), we should also consider the security issue. These storage systems are considered reliable, but not totally trustworthy, especially when you are storing data with great importance. With a local cloud environment, however, these issue no longer exist. All CMU faculties and students can use the local cloud just like using desktop computer. The data can be placed safely and conveniently on the cloud side. Users can access their data conveniently and no longer need to waste time waiting for the data uploading to the cloud. Besides, security is no longer a problem. As the local cloud only grants accessibility to specific users, in this case, CMU students and faculties, the data stored in the cloud is highly secure.

**Pricing**

Local cloud is a low pricing cloud environment. After the universities have purchased the servers, everything else is free except for the charge of electricity. This is why a local cloud is so necessary, it is much more cheaper than using a public cloud service. For example, the AWS charges users in an interesting way, because it has so many charging areas. Many users may just notice the price of the instances, like the on demand price of m1.large is $0.175 and m1.small is $0.077, but they failed to realize the pricing of other services. For example, when you are using the AWS cluster to do a mapreduce job, several more charges besides the price of the instance should be considered, including the price of the AWS EMR (Elastic Mapreduce) cluster, the price of the I/O data, the price of the Auto Scaling Group, and the users may even have to pay for the S3. The combination of all these prices can be considerably large. However, when using the resource of the local cloud computing environment, everything is totally free. All we have to do is to purchase several servers and the cloud environment is ready to service for several years. Besides, as the universities already got these servers, there’s no reason to not make use of them.

**Local Cloud Service can provide customized service**

Public cloud services are convenient and fast, but compared to local cloud computing environment, it cannot provide so much customized service. Using a local cloud service, we can adjust almost everything from the operating system to the software installed in it. The public cloud service, like AWS, do provide many kind of different customized AMIs (Amazon Machine Image). But to satisfy the need of the most, most of the AMIs only installs the most basic tools and software, it is hard for a user to customize the image. If a user had to make a customized version of AMI, he or she had to create a snapshot, which would be charged quite a high price. With local cloud computing environment, however, users are allowed to modified the image or upload his own customized image in order to acquire a customized service. This is considered to be a marked advantage of setting up a local cloud environment.