OpenStack Capstone Sprint Report

Justin Anderson
John Nagel
Cesar Ramirez
John Rensberger

Table of Contents

1. Sprint Planning Meeting - 03/15/2013

1.1 Sprint Backlog

Our backlog can be found in the Microsoft Project plan located in our Git repository, which is accessible from Trac and Github. We also have been using Trac's ticket feature for small tasks and we are striving to actively update them accordingly. As we're doing weekly sprints, the most pertinent information can be found in the roadmap section of Trac. We update this weekly after meeting with our mentor to reflect the goals we set for each sprint. This also provides a nice historical view of what we've done and any future plans that we have.

2. Sprint Review Meeting - 03/22/2013

2.1 Customer Demo

We have been working on implementing Hadoop as requested by department faculty. We haven't reached the point where we can demonstrate anything new. Most of our work and effort has been dealing with debugging Quantum (the networking component of OpenStack), and working on Hadoop installation on top of OpenStack. Once we have working installation we can work on a working demo.

2.2 Stakeholder Involvement Review

We have decided to prioritize our use cases and primarily focus on Hadoop as this is the most widely requested service. For the remaining use cases we will address them as time permits, but we want to ensure that we will deliver Hadoop and thorough deployment documentation before continuing on the remaining use cases. We are staying very involved with our stakeholders through meetings. This week we have had three meetings, two with IBM and one with CS faculty.

2.3 Data Management Review

We've been very meticulous on keeping Trac as current and relevant as possible. We use tickets to track tasks and keep a weekly log of our biggest goals in the roadmap section. We also have improved how all project data is stored as Trac and Git are not our only data management resources. We maintain all documents on a shared folder on Google drive that our sponsor has access to which allows us to provide them any documents or files that we're working on in real time.

2.4 Requirements Review

Our progress is still mostly in line with the original requirements that we specified at the beginning of the semester. Most of our resources have been poured into discussing uses with faculty and our sponsor as well as setting up an operational development environment. Along the way we've been learning much more about how OpenStack works and each of the services that make it up. A large amount of our requirements lie in documentation and this will likely continue through the remained of the semester. The other requirements are the same from Sprint Report 2. We are working on get Quantum fully functional. Once we have Quantum working, we will be working on setting Hadoop cluster on OpenStack, and documenting the whole process.

2.5 Progress Review

2.5.1 Work Completed

We were able to secure a location and some fairly decent hardware. We first tried installing DevStack on this hardware but repeatedly ran into issues that we were unable to solve since we were not familiar with OpenStack. We decided to instead only use DevStack in single-node configurations on virtual machines as we could easily revert the environment with snapshots if there were any issues.

We then rebuilt the Ubuntu operating systems on all of our physical hardware and began installing Grizzly release candidates using Ubuntu cloud packages. We are still working on this installation process and hope that it will provide us with a stable and up-to-date development environment.

2.5.2 Work Not Completed

We are still setting up our multi node environment since DevStack did not work correctly for us. We hope to also get Quantum fully working after the environment is setup so we can begin using networking on our instances. Once we have Quantum working we can work on setting Hadoop on top of OpenStack using Heat for automation. Once we have competed this successfully, we will document our work.

3. Sprint Retrospective Meeting - 04/02/2013

3.1 Top Highlights

- Focused the scope of the project
    - With the time we have spent deploying and debugging OpenStack as well as talking with IBM, we have narrowed the scope of the project. Our primary goal for the next sprint is to fully implement the Hadoop use case.
- Developed a more thorough understanding of Hadoop
    - We have done a significant amount of research and met with interested department faculty to gain a firm understand of Hadoop and the department's intentions. We will use this knowledge to provide Hadoop deployment services on top of OpenStack.
- Completed a manual deployment of OpenStack Folsom
    - We rebuilt Ubuntu on the physical hardware (again) after doing testing on Grizzly. We now have four nodes running, one hosts the controller and networking services. The other three then each host compute services, of those three one also is hosting swift (object storage) and another is hosting cinder (volume storage).
- Began implementing Savanna, a hosted Hadoop solution for OpenStack.
    - Savanna is a new project that aims to provide a service, similar to Amazon Elastic MapReduce, that integrates with OpenStack components. It provides much of the functionality that we hoping to achieve, but it is very young and under-documented, so there is significant work to be done.

3.2 Top Lowlights

- Grizzly was buggy and poorly documented
    - We encountered several bugs in our install of Grizzly. We were using packages from the Ubuntu cloud repository and they were not up-to-date with bug releases. We spent a large amount of time debugging issues to come up with fixes and then find documented bugs. For this reason we dropped Grizzly and reverted to using Folsom.
- Still working on getting all aspects of Quantum functioning.
    - We still are working on Quantum. We have made some progress, but not enough to meet our requirements.

3.3 Reflection on Improvements

We wish we were further along with Quantum. We have been spending more hours than originally planned on this single component. We wish we had done more research on Quantum and developed more elaborate questions for IBM. This could have sped the process up a significant amount. It would also allow us to work on other use cases in more detail.