Documenting Cassandra

Johan Hilding, johanhil@kth.se
Joel Bohman, jbohman@kth.se

Supervisor
Mads Dam, mfd@kth.se

Abstract

This is the project specification for our Bachelor's thesis project. In this document we have outlined the problem we are going to solve, which is to document the Cassandra data storage system. We have also included a relevant history to put Cassandra in perspective. For administrative purposes there is also a tentative timetable included.

Background

For 40 years relational database management systems (RDBMS) such as PostgreSQL have been the standard for storing data. However, the last decade has seen a huge increase in the amount of data that needs to be stored. Data sets of many terabytes, or even petabytes, are not unusual anymore. The RDBMS were not designed with these sizes in mind, and have trouble handling them.

Instead of fixing what is broken, several companies started to develop their own data storage systems without the strict relations and complexities RDBMS had. One of the first companies to do this was Google, developing BigTable[1] to help scale their search engine which stores billions of webpages. Another company was Amazon that developed Dynamo[2] to suit their needs. Both BigTable and Dynamo are proprietary software but as the word spread about their success and theory, a number of open source projects sprung to life. In 2009 the term NoSQL was coined to group these non-relational databases.

During 2008 Facebook faced issues trying to scale their inbox search system which contained a few terabytes of data and had a steady growth rate. To solve this issue they developed the database Cassandra[3][4] and later released it as open source software. Cassandra's main design goals were:

- Easy to scale it by just adding new hardware to keep up with the growth rate.
- Reliability, high performance and no single point of failure as its used by a high traffic site.
- A data model that could fit the data from the inbox search system.

To achieve these goals the developers looked at Dynamo for its distributed design and BigTable for its data model. The killer feature of Cassandra compared to RDBMS are its scaling capabilities.

During 2009 Cassandra has steadily gained momentum and has matured as a software project. Digg, a social news site, had troubles with a feature that required them to intersect two huge sets of data. Running the intersect could take up to 14 second on their RDBMS, which for a website is unacceptable. They did not see a future in tweaking their RDBMS[1] and instead they decided to switch to Cassandra, and can now run the intersect in under one second[5].

Problem

We are going to document the Cassandra architecture, producing a report that contains:

- An overview of Cassandra and its data model.
- Details about the algorithms that Cassandra use.
- Survey of real life projects that use Cassandra.

If time allows and our research brings up a significant problem we might try to contribute a patch or help with documenting Cassandra's design.

---

1. Through techniques such as sharding/partitioning

Project plan

The project will be divided into two tasks, research and writing.

The research part will consist of the following subtasks:

A. Getting familiar with the Cassandra community through mailing lists[6] and IRC[7]. This is not essential but will come in handy when we need to ask questions.
B. Learning about Cassandra using the slides and videos linked from the official website. This will help us to decide which headings that will be in our report and give us an overview of Cassandra.
C. Read the scientific papers that have been published about Cassandra and its algorithms[8].

The writing part will consist of these subtasks:

A. Figuring out the headings for the report.
B. Writing the overview of Cassandra.
C. Explaining the data model, perhaps producing some sort of a cheat sheet or real life examples. To be able to do this, we have to do research task B.
D. After finishing research task C, we will write about the algorithms that are used in Cassandra, detailing how, when and why they are being used.

Timetable

When	What
Beginning of February	Research task A starts.
Middle of February	Research A done. Research B starts.
End of February	Writing A & B starts. Research B half-way.
Middle of March	Writing A & B together with Research B done. Writing C starts.
Beginning of April	Writing C done. Writing D and Research C begins.
End of April	Only minor fixes are left before we hand the report in. We begin practicing for the presentation.

References

1. Bigtable: A Distributed Storage System for Structured Data, http://labs.google.com/papers/bigtable.html
2. Dynamo: Amazon's Highly Available Key-value Store, http://www.allthingsdistributed.com/files/amazon-dynamo-sosp2007.pdf
3. Cassandra, http://incubator.apache.org/cassandra/
4. Cassandra – A structured storage system on a P2P Network, http://www.facebook.com/note.php?note_id=24413138919
5. Looking to the future with Cassandra, http://about.digg.com/blog/looking-future-cassandra
6. Cassandra Users mailing list, http://www.mail-archive.com/cassandra-user@incubator.apache.org/
7. #cassandra @ Freenode IRC Network

8. Cassandra - A Decentralized Structured Storage System, http://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf