Twitter with Cassandra

Adhish Singla 201403004
Saurabh Jain 201301128
Gaurav Singh 201264059
Shantanu Parashar 201202022

Abstract

We aim at building an online social networking service that enables users to send and read short messages using a distributed DBMS – **Apache Cassandra** which is an open source distributed database management system designed to handle large amounts of data across many commodity servers, providing high availability with no single point of failure. Cassandra offers robust support for clusters spanning multiple datacenters, with asynchronous masterless replication allowing low latency operations for all clients.

We are using Cassandra because of the following reasons:

1. Decentralized

2. Supports replication and multi data center replication

3. Scalability

4. Fault-tolerant

5. Tunable consistency

6. MapReduce support

7. Query language – CQL

Features:

- Tweets

- Messaging

- Grouping Posts Together

- Trending Topics

- Adding and Following Content

- Email Verified Accounts

Keywords: Twitter, Apache Cassandra – Distributed DBMS, Microblogging, Cloud