Data collected for TEAM Presentation on Cassandra!

Cassandra Introduction

Apache Cassandra is a database management system which is distributed, high performing and open source.

It is one of the NoSQL databases.

Previously, traditional database systems were used, but as the time passed by amount of data to be stored and maintained increased and hence increased the need of highly scalable storage.

When such a huge amount of data was stored using traditional database, huge breakdowns were witnessed.

Loss of data , unable to maintain the data was a common problem.

Thus, giving rise to the invention of NoSQL databases.

Hence, Cassandra is one of many commodity servers which provides highly scalable, large storage and high availability of database without a single point of failure.

Cassandra is one of the most used and trusted databases today. Following are the reasons:

1. Fault Tolerant: Cassandra is a distributed database, hence replication of data among all servers is one of the main feature of this database.

Thus proving it to be fault tolerant that is there remains no single point of failure. As data is replicated so even if one server

fails, other are available to server the purpose of user.

2. Performant: It proves to be consistent in performance as compared to other NoSQL databases because of its architectureal choices.

3. Decentralized: No centralization of processes avoiding single pint of failure. Also avoids any network bottlenecks makind every node identical.

4. Durable: It can be used for applications which do not afford loosing any data and sustains the data and load for a long time, which means maintains persistance.

5. Elastic: Easy read and write of applications.

6. Professionally Supported: Supports all third party platforms.

References:

https://en.wikipedia.org/wiki/Apache\_Cassandra

http://cassandra.apache.org/

Book: Cassandra The Definitive Guide.pdf