Khushboo Singh Database Design and Management-Sec1 Discussion 1

MongoDB ENABLED FINANCIAL SECTOR

Abstract- Over 30 years' database selection trends have been changed drastically due to the application requirements. To select a database, it is required to understand the user requirements. Mobile, social and cloud based applications are using structured, semi-structured and unstructured data due to heavy volume of data which is doubling every 1.2 years.

INTRODUCTION

Organization needs to cope up with the scalability and performance**.** Relational databases are stored in rows and columns and can not be properly used for analytics. As well due to their monolithic scale-up design, consistent network and hardware quality-of-service. The requirement of storing complex data and the database which is query rich and provides good analytical capabilities pushed to use a new database. A database which provides agility to accommodate flexible and iterative development practices. The database which supports horizontal scale-out deployments coupled with fault- tolerance to meet up the business demanded SLAs due to the increase in servers, coupled with elastic and on-demand cloud computing which has increased the operating cost. Due to all this need MongoDB came into the picture.

MongoDB INCREASING BUSINESS VALUES

MongoDB helps application to move into production within weeks [1]. Its the fastest growing database which has been downloaded over 10 million times with 2000+ customers including Fortune 100 companies like ADP, Cisco, Bosch, CARFAX, eBay, MetLife, Salesforce.com, Telefonica and Viacom. MongoDB serves 10 top financial, electronics, media and entertainment, retail, telecommunication and top 8 technology and top 6 healthcare among the Fortune 500 and Global 500 companies.

MongoDB combines strong consistency, powerful query language and rich indexes of Relational database as well flexible data model, performance and scalability of NoSQL database. MongoDB features include JSON Document Model with Dynamic Schemas, Text Search Aggregation Framework, Auto-Sharding for Horizontal Scalability and MapReduce Full. It also provides flexible index support, rich queries built-in replication for high availability, advanced security, large media storage with GridFS.

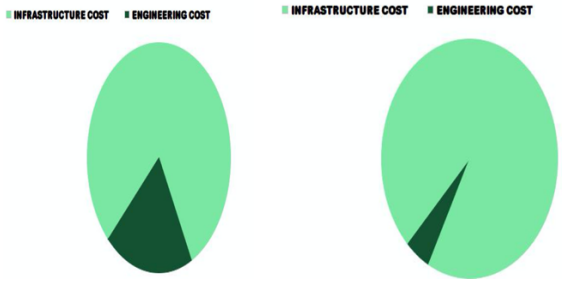

Fig1: Cost involved over the year

ADVANTAGES OF MongoDB:

MongoDB 3.0 has been equipped with the features which provides the below benefits:

1.*Storage Engine API*	It allows to plug-in different storage engine mixing storage engines on same replica set/sharded cluster which can be further integrated.
2.*Improved Concurrency Control*	Leads to more CPU usage due to higher level of concurrency.
3.*Compression*	Data is compressed on disk using two compression algorithms. Snappy is the default algorithm which provides good compression with relatively low overhead. Zlib is another algorithm which is better than Snappy, Prefixes are used to compress indexes.
4.*Consistency with Journaling*	Writes are written to journal as they come which is better for insert heavy workloads with guaranteed durability.

Advanced MongoDB helps in real-time analysis, log data, messaging, content management and enables IoT(Internet of Things) enhancing performance that is why different sectors started using it as one of the databases.

MongoDB IN FINANCIAL SECTOR

1.**Reference Data Management**:

Financial institution faces a lot of challenges in managing and distributing reference data globally. Managing database schema at the same time while integrating and replicating the data cross geographies is costly and time consuming [2]. MongoDB makes it simple to distribute and synchronize data efficiently across the globe by native replication and partitioned architecture. MongoDB reduces database maintenance for schema migration by using dynamic schema. Data is replicated for each change not in batches where changes are applied with no down time and without impacting the existing application. Both cache and database cache are always up-to-date ensuring a congruent view across all trading entities in a global organization.

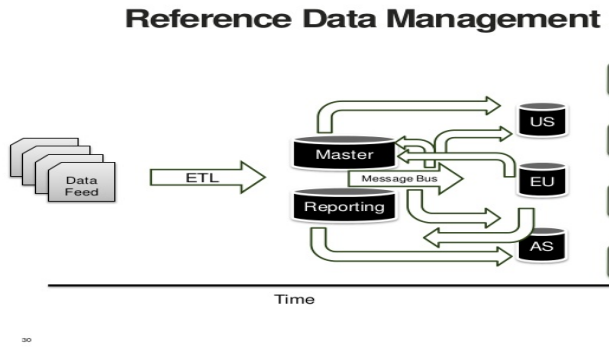

Fig 2: Reference Data Management

It is being used for an American Investment bank/Retail Bank.

Benefits: Reduced cost for infrastructure management. Data replication meets SLA with timely reference. Company saved around $40m as they were charged once for data from TR/Bloomberg.

2.**Single view of Customers**:

It is difficult to identify the need of the customer with isolated data across organization with no consistent view across customer. This can be resolved using single view of customer which helps to predict the customer requirements improving customer service.

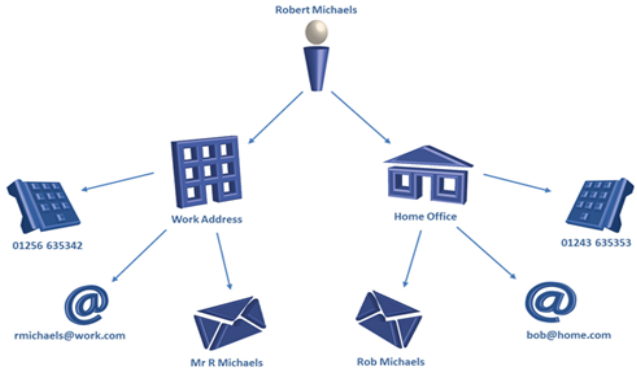

Fig 3: Single view of Customers [3]

MetLife is an example of the above. They blended data from 70 separate source systems and 20 screens within 3 months using MongoDB. Using flexible schema where the system is loosely coupled so that changes to upstream won't breakdown the downstream application. Sharding helps to add more data and users by globally scaling out the users.

3.Retail bank transaction log:

Fetching the data is quite costly whereas bank needs to maintain logs of transaction for five years. MongoDB helps to read requests from different systems which is further maintained.

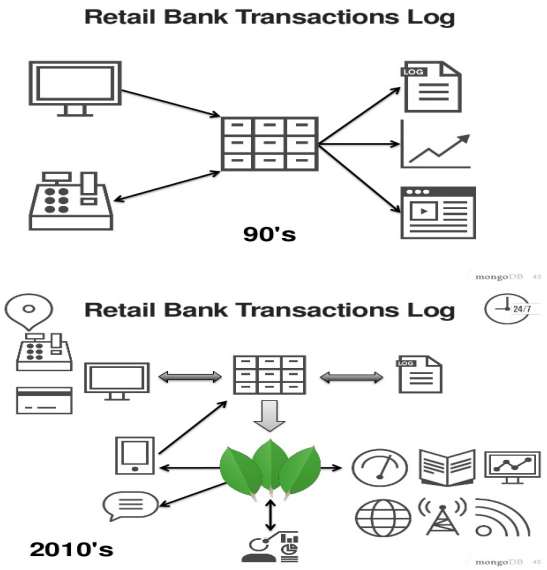

Fig 4: Retail bank Transaction

4.**Trade Repository**:
MongoDB's flexible schema helps to integrate diverse trades in a single database reducing cost by scaling storage on commodity hardware. Hence, Financial firms can store data for 7 years or more.

5.**Risk Aggregation and Reporting**:
Investment Bank needs a congruent view to effectively manage risks. They need to consolidate and analyze multiple risk by creating a single view across asset classes or counterparties. Allowing granular access to any data attribute using native aggregation framework provides a powerful reshaping of data at a very large scale for intraday analysis.

DISADVANTAGES OF MongoDB:
Usage of memory: As it stores key name along with every document hence consumes more memory and end up with slow query performance [4].
No Joins: Joins are not possible as compared to relational database results in the creation of multiple queries which is manually joined in the code.
Still under development: SQL was developed in 1980s whereas MongoDB came into picture in 2009 so lacks documentation and support.

Conclusion:
MongoDB is an open source, non-relational database based on three key features: queryable, schemaless and scalable. It has its own pros and cons but making the right choice of database helps to deliver the required business result [5]. If you need to manage the scalability and diversity of data by reducing the cost, enabling new types of applications, running on the latest hardware or in cloud then MongoDB should be a good choice.

References:
[1]https://www.mongodb.com/industries/financial-services
[2]https://www.mongodb.com/events/webinar/managing-reference-data
[3]http://bluesheep.bluegroupinc.com/
[4] http://blog.easylearning.guru/what-is-mongodb-the-advantages-disadvantages-of-mongodb/
[5] Ian White - http://www.businessinsider.com/how-we-use-mongodb-2009-11