NoSQL for non-relational databases with Hadoop

Need to have ability to very quickly answer questions. Needs to handle 10s of thousands of queries.

What are the movie recommendations for this customers and etc.

NoSQL. Scale horizontally forever.

Tools in NoSQL.

* HBase. Built on top of HDFS for Web application and etc.

NoSQL When you need to operate very quickly.

When RDBMS didn’t cut it

Random Access to Planet-Size Data need to live on Distributed data. Horizontally capable. Always gets bigger and bigger. Billions of Data. Website like Amazon. At a Large Scale.

**PREVIOUSLY,**

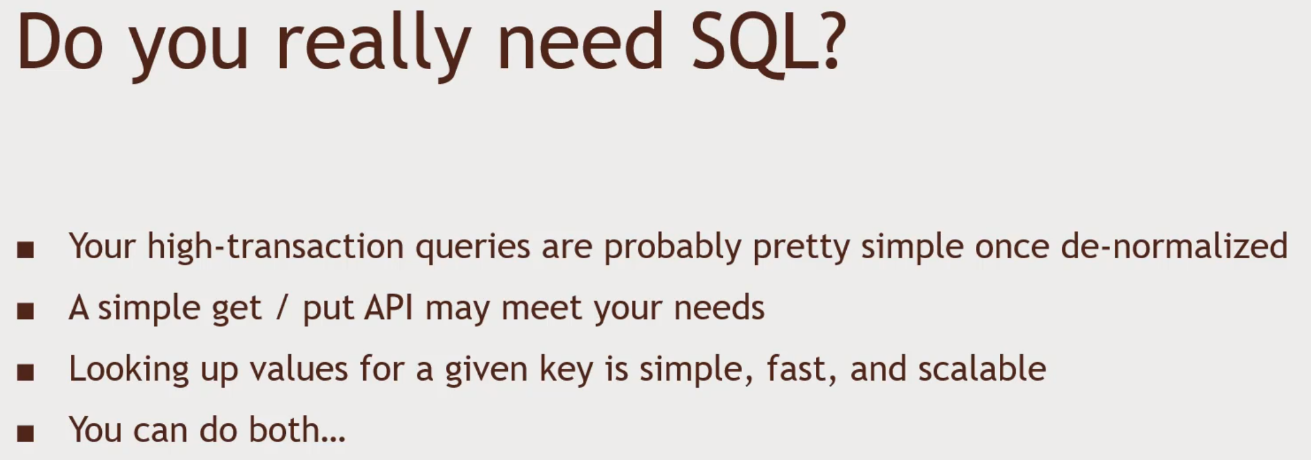
Denormalize data- LOTS OF LOTS OF TABLES. Rematerialize the views

Caching Layer – Giant CACHE.

Master/slave setups Multiple databases, writes, reads, Maintaining require a lot of time

Sharding

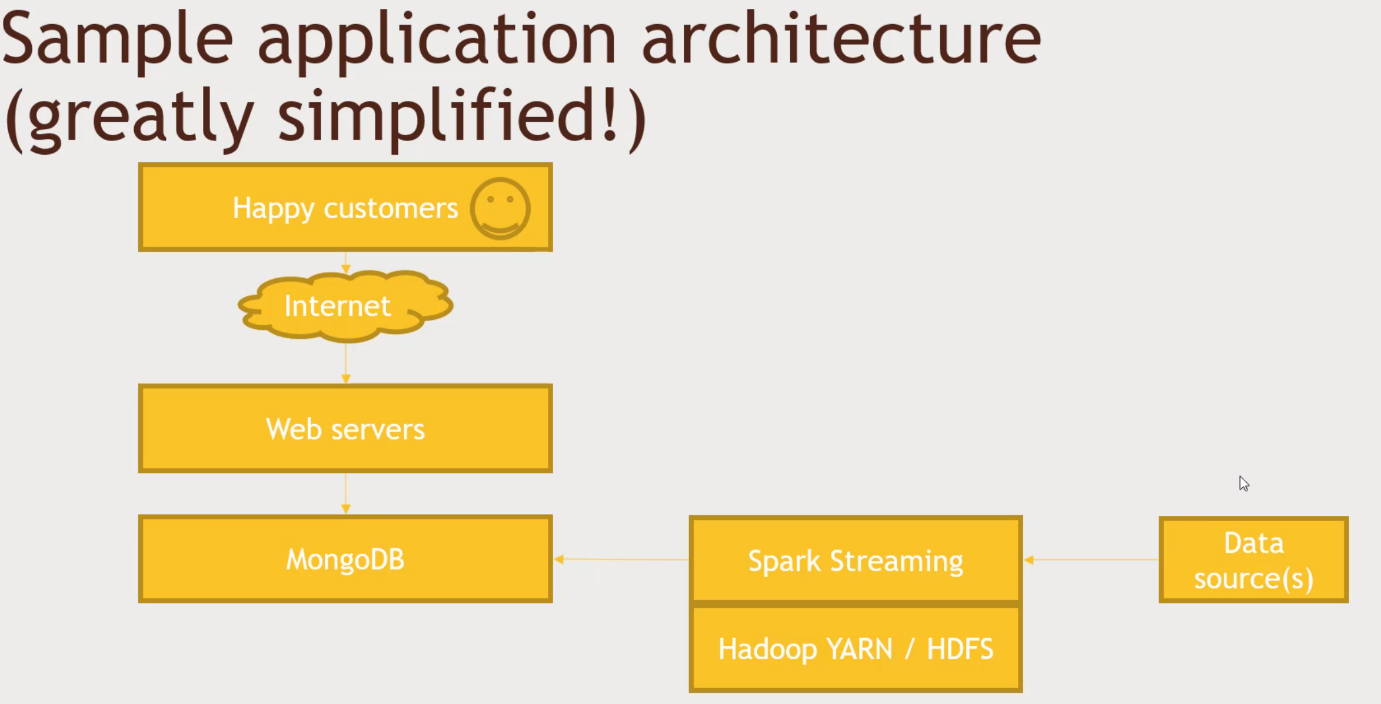
Materialize view – moving lots of data.

Removing Stored Procedures. Your database will melt down

Key value datastore at high level. You don’t need a fancy relational database.

Idea of NO SQL Database

* Client & Request router -> It knows where to get the data
* Shards that has a subset of customer’s data.
* Shard Backups if the node goes down, we are covered.

MongoDB or other NoSQL databases.

Every box can represent many machines. Thousands of webservers.