CS 488P/588P

Yufei Tao, Mi Yon Kim

- **System Investigation Table**

	MongoDB	Cassandra	HBase
Data Model	Document store: The documents in a collection can have different fields and data types.	Wide column store: each column is stored separately on disk.	Wide column store: each column is stored separately on disk.
Indices	MongoDB provides default _id index. In addition, the user can make their own index to traverse documents efficiently. E.g., indexes on foreign key references and primary keys	The Cassandra provides a partition key. But we can make a secondary index as well.	The HBase provides row key similar data retrieval benefits as the primary index. So, when you create a secondary index, it is good to choose elements that are different from the row key.
Consistency	Immediate Consistency	Eventual Consistency	Immediate Consistency
Sharding / Partitioning	Data will be partitioned by the shard key.	Data will be partitioned by the partition key.	HBase has auto sharding. The tables are dynamically distributed by the system to different region servers when the data become too large.
Replication	There is a primary and secondary node. The primary: Writes and Reads from the Application.	The data gets replicated on each node according to your replication strategy.	There is a master cluster and the slave cluster. The master cluster pushes the data to the slave cluster.

	The Secondary: The primary replicated into the secondary node.		
Joins	MongoDB has the Lookup equivalent to Joins.	No Joins are supported by Cassandra.	No Joins are supported.