

Data Models

Tuesday, April 23, 2019 1:29 PM

- **NoSQL**
 - **Key-value stores** (the simplest)
 - e.g., Project Voldemort, Memcached
 - **Document stores** { JSON | SQL++
 - e.g., SimpleDB, CouchDB
 - **Extensible Record stores**
 - HBase, Cassandra
- **Key-Value Stores features**
 - Data Model: **(key, value)** pairs.
 - ↓
unique
(string/integer)
 - ↓
anything
 - Operations:
 - **get(key), put(key,value)** Use hash fun lib usually.
 - Operation on value not supported. e.g. just take sum of ASCII of string.
 - Distribution/ Partitioning - **hash function** generate an integer
 - No replication: **store on server hash** (Partitioning)
 - 3 way replication: **h1(k), h2(k), h3(k)**, **on three server** (Replication)
on update, propagate changes.
 - Usually use both
 - The return range of hash function is fixed.

It's difficult to find flights from Seattle, but easier to get information by keys.

- Document Stores

- Let value be structured data.
e.g., JSON, Protobuf, XML
- Called document but actually a data.

- Extensible Record Stores

- Base on Google's BigTable, just one table