NoSQL Systems

- Key-Value Store
  - Map (associative array)
  - Simple, non-trivial
  - Keys may be sorted to support range requests
  - Ex: Riak, Redis, etc.

- Document Store
  - Store data as a collection of documents
  - XML, JSON
  - Documents may have different fields
  - API to query the data in documents
  - Can support nested structures
  - Each document has a key associated with it
  - Ex: CouchDB

Column Store

- Good for OLAP workloads
- Store data in columns indexed by row

- Related to timeseries DB
  - Ex: C-Store
  - KDB

Large-Scale Distributed Data Processing

- 3Vs: Volume, Velocity, Variety

- Hadoop
  - Talk about MapReduce
  - 2 stages: map and reduce

- Storm
  - Spouts and bolts to define sources and processors
  - Batch processing of stream data
  - DAG topology
  - Streaming runs indefinitely

- Spark
  - Streaming model
  - RDBs - resilient disk sets
  - In memory, can be faster than Hadoop

- Streaming databases

- Graph Databases
  - See slides