

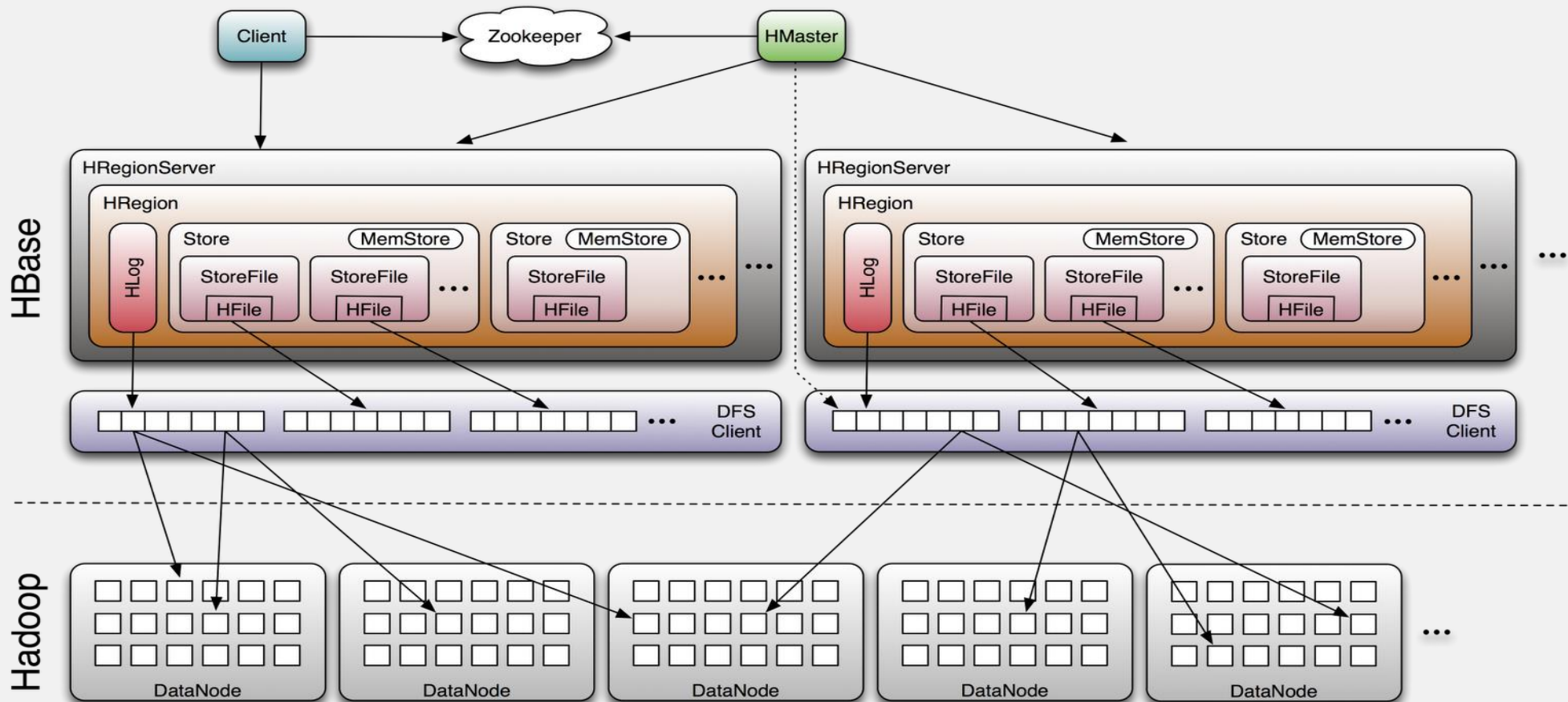


CLOUD COMPUTING APPLICATIONS

HBase Internals - Part 1

Roy Campbell & Reza Farivar

HBase



HBase Building Blocks

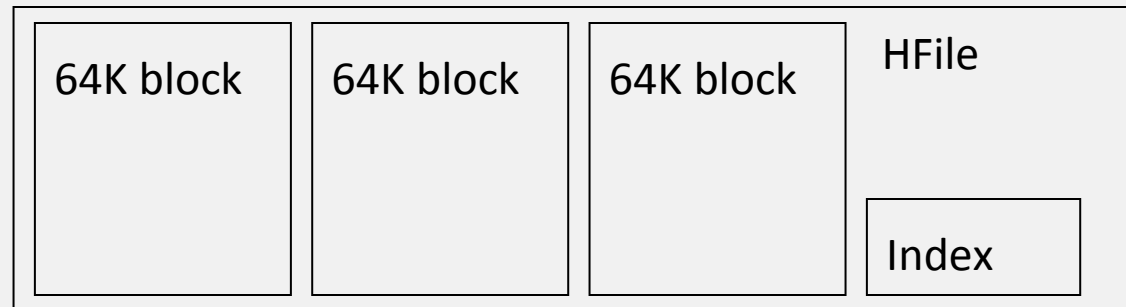
- HDFS
- Apache ZooKeeper
 - ZooKeeper uses ZAB (ZooKeeper's Atomic Broadcast)
- HFile

HFile

- Basic building block of HBase
- On-disk file format representing a map from string to string
- Persistent, ordered immutable map from keys to values
 - Stored in HDFS

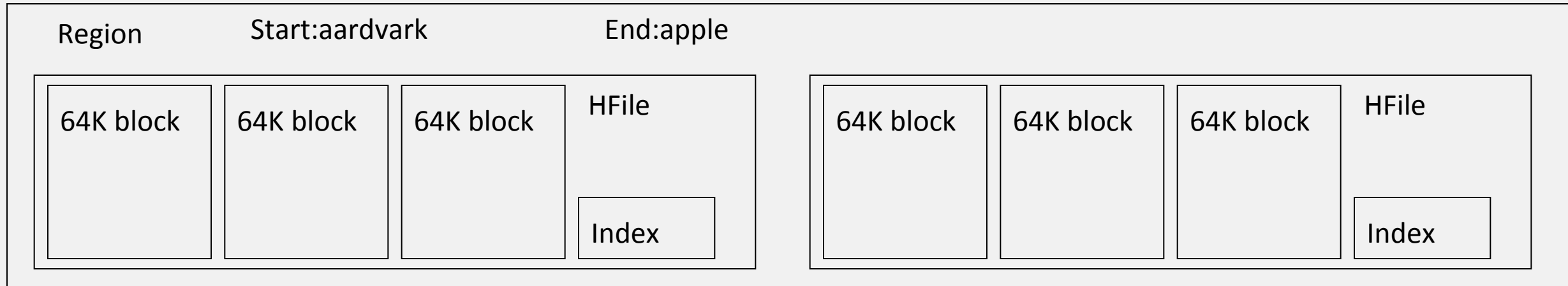
HFile

- Sequence of blocks on disk plus an index for block lookup
 - Can be completely mapped into memory
 - MemStore
- Supported operations:
 - Lookup value associated with key
 - Iterate key/value pairs within a key range



HRegion

- Dynamically partitioned range of rows
- Built from multiple HFiles



HBase Master

- Assigns HFiles to HRegion servers
- Detects addition and expiration of HRegion servers
- Balances HRegion server load. HRegions are distributed randomly on nodes of the cluster for load balancing.
- Handles garbage collection
- Handles schema changes

HRegion Servers

- Each HRegion server manages a set of regions
 - Typically between 10 to 1,000 regions
 - Each 100-200 MB by default
- Handles read and write requests to the regions
- Splits regions that have grown too large

HRegion Location

