



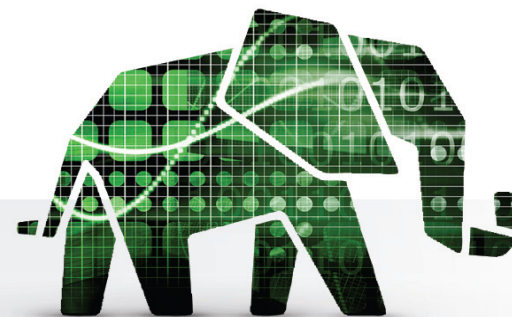
Apache HBase

For Architects

Nick Dimiduk

Member of Technical Staff, HBase

Strata/Hadoop World, 2013-10-29





Agenda

- **Background**
 - (how did we get here?)
- **TL;DR**
 - (don't waste my time!)
- **High-level Architecture**
 - (where are we?)
- **Anatomy of a RegionServer**
 - (how does this thing work?)
- **By Example**
 - (how do I use it?)
- **Resources**
 - (where do we go from here?)



Background

So what is HBase anyway?

- **BigTable paper from Google, 2006, Dean et al.**
 - “Bigtable is a sparse, distributed, persistent multi-dimensional sorted map.”
 - <http://research.google.com/archive/bigtable.html>
- **Key Features:**
 - Distributed storage across cluster of machines
 - Random, online read and write data access
 - Schemaless data model (“NoSQL”)
 - Self-managed data partitions

Apache Hadoop Dependencies

- **Apache Hadoop Distributed Filesystem (HDFS)**
 - Distributed, fault-tolerant, throughput-optimized data storage
 - The Google File System, 2003, Ghemawat et al.
 - <http://research.google.com/archive/gfs.html>
- **Apache Zookeeper (ZK)**
 - Distributed, available, reliable coordination system
 - The Chubby Lock Service ..., 2006, Burrows
 - <http://research.google.com/archive/chubby.html>
- **Apache Hadoop MapReduce (MR)**
 - Distributed, fault-tolerant, batch-oriented data processing
 - MapReduce: ..., 2004, Dean and Ghemawat
 - <http://research.google.com/archive/mapreduce.html>



TL;DR

So what is HBase anyway?

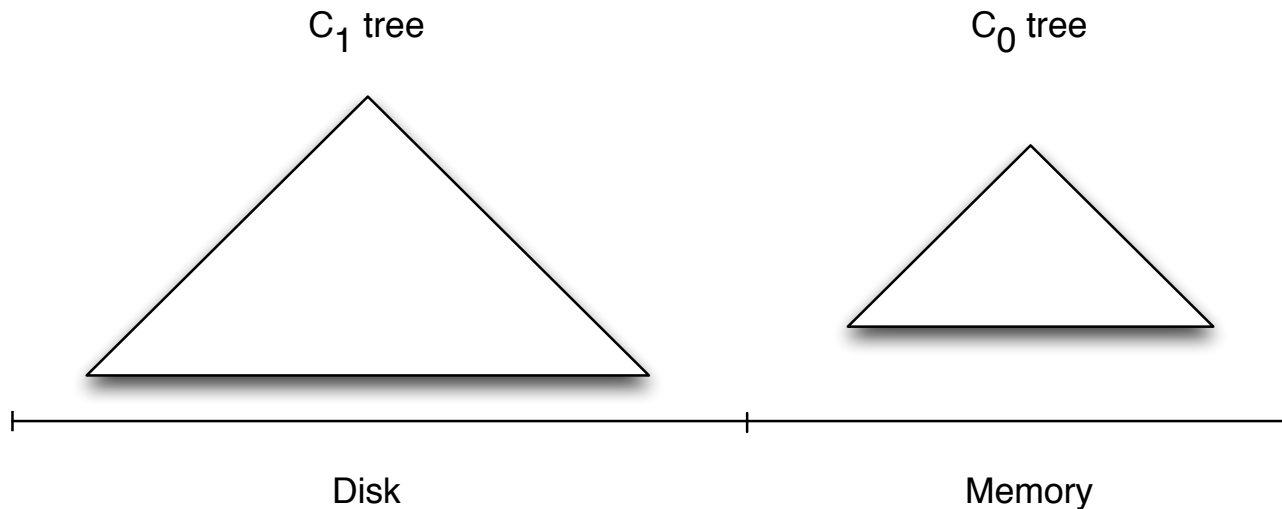
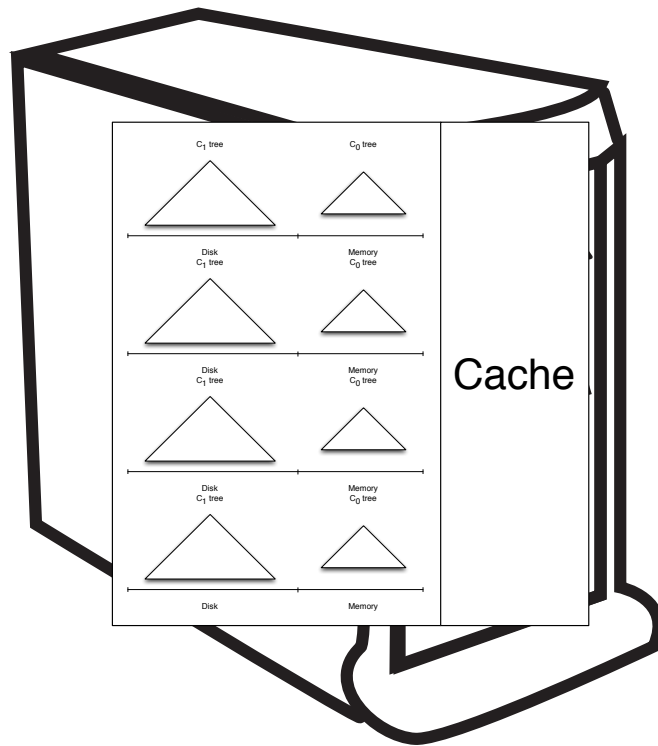


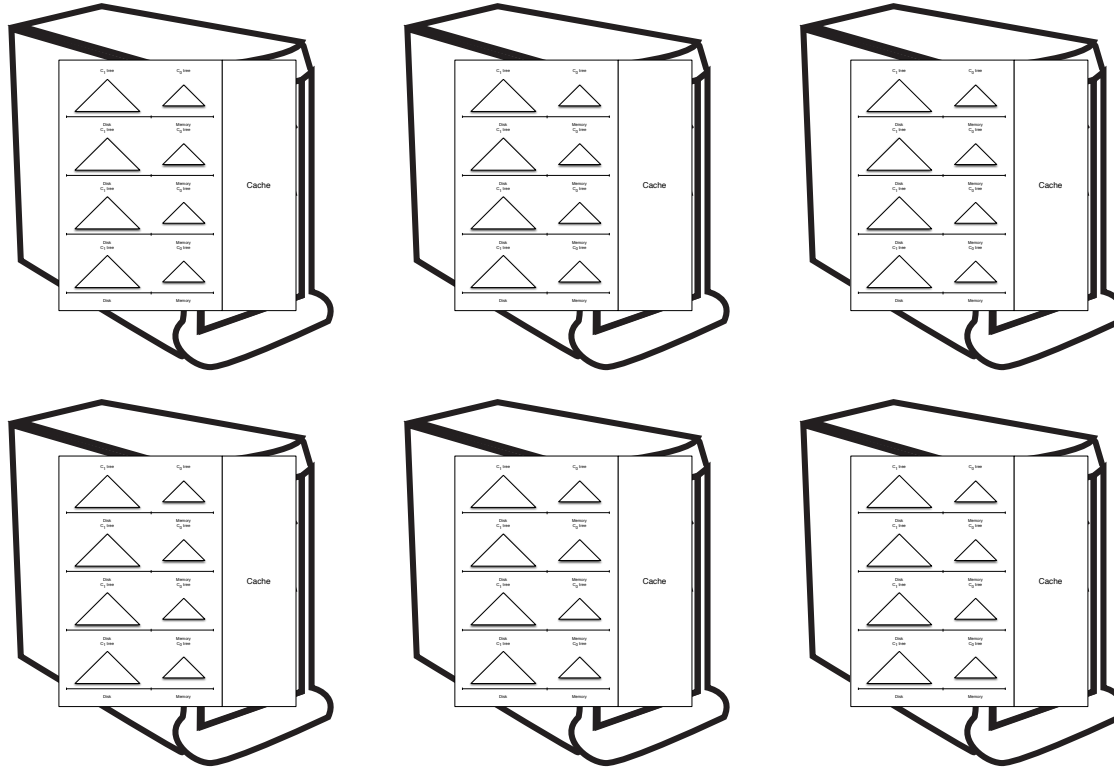
Figure 2.1. Schematic picture of an LSM-tree of two components

Figure 2.1 reproduced from O'Neil, Patrick, et al. "The log-structured merge-tree (LSM-tree)." Acta Informatica 33.4 (1996): 351-385.

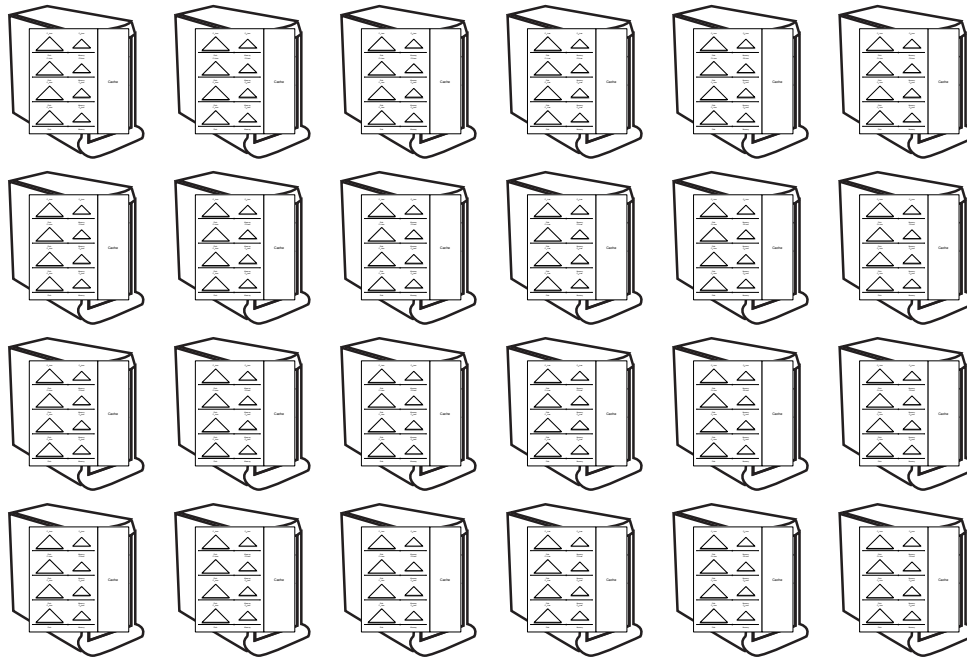
So what is HBase anyway?



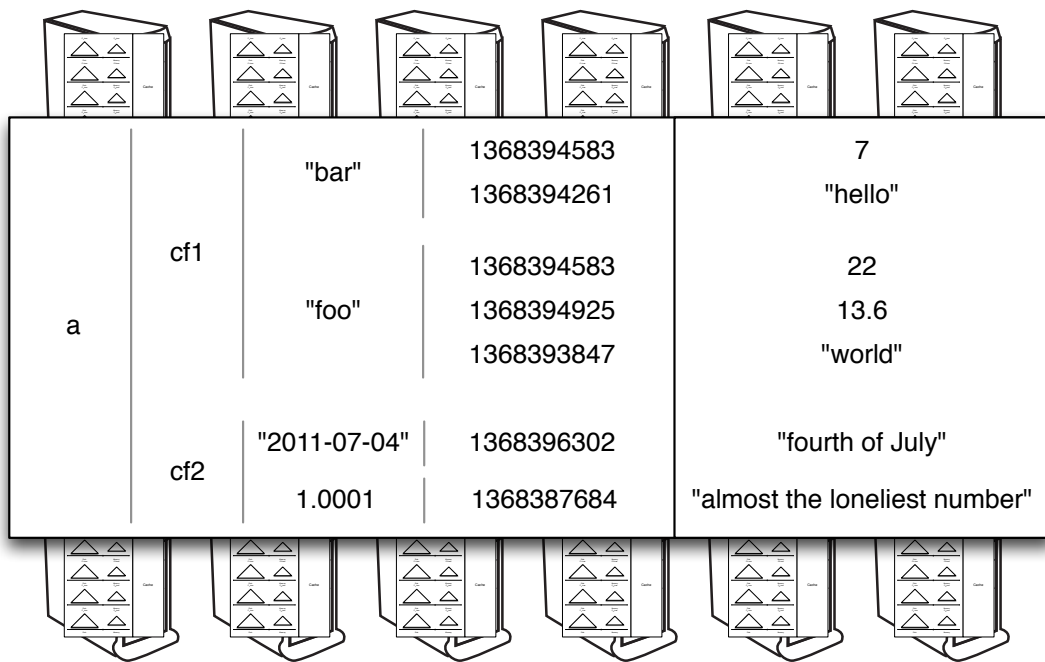
So what is HBase anyway?



So what is HBase anyway?



So what is HBase anyway?





High-level Architecture

Logical Data Model

Table A

rowkey	column family	column qualifier	timestamp	value
a	cf1	"bar"	1368394583	7
			1368394261	"hello"
		"foo"	1368394583	22
			1368394925	13.6
	cf2	"2011-07-04"	1368393847	"world"
	cf2	1.0001	1368396302	"fourth of July"
			1368387684	"almost the loneliest number"
b	cf2	"thumb"	1368387247	[3.6 kb png data]

Logical Data Model

Table A

Column Families

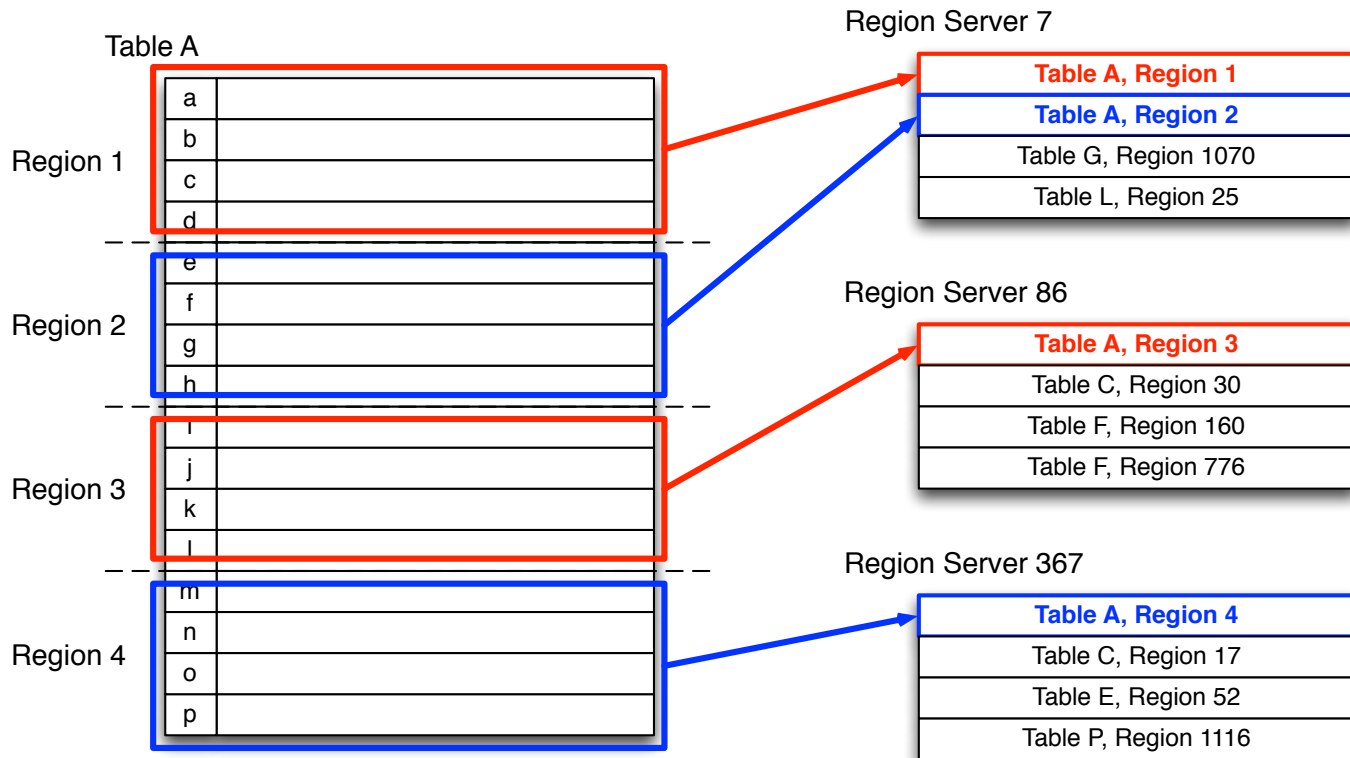
rowkey	column family	column qualifier	timestamp	value
a	cf1	"bar"	1368394583	7
			1368394261	"hello"
	cf1	"foo"	1368394583	22
			1368394925	13.6
			1368393847	"world"
	cf2	"2011-07-04"	1368396302	"fourth of July"
		1.0001	1368387684	"almost the loneliest number"
b	cf2	"thumb"	1368387247	[3.6 kb png data]

User API

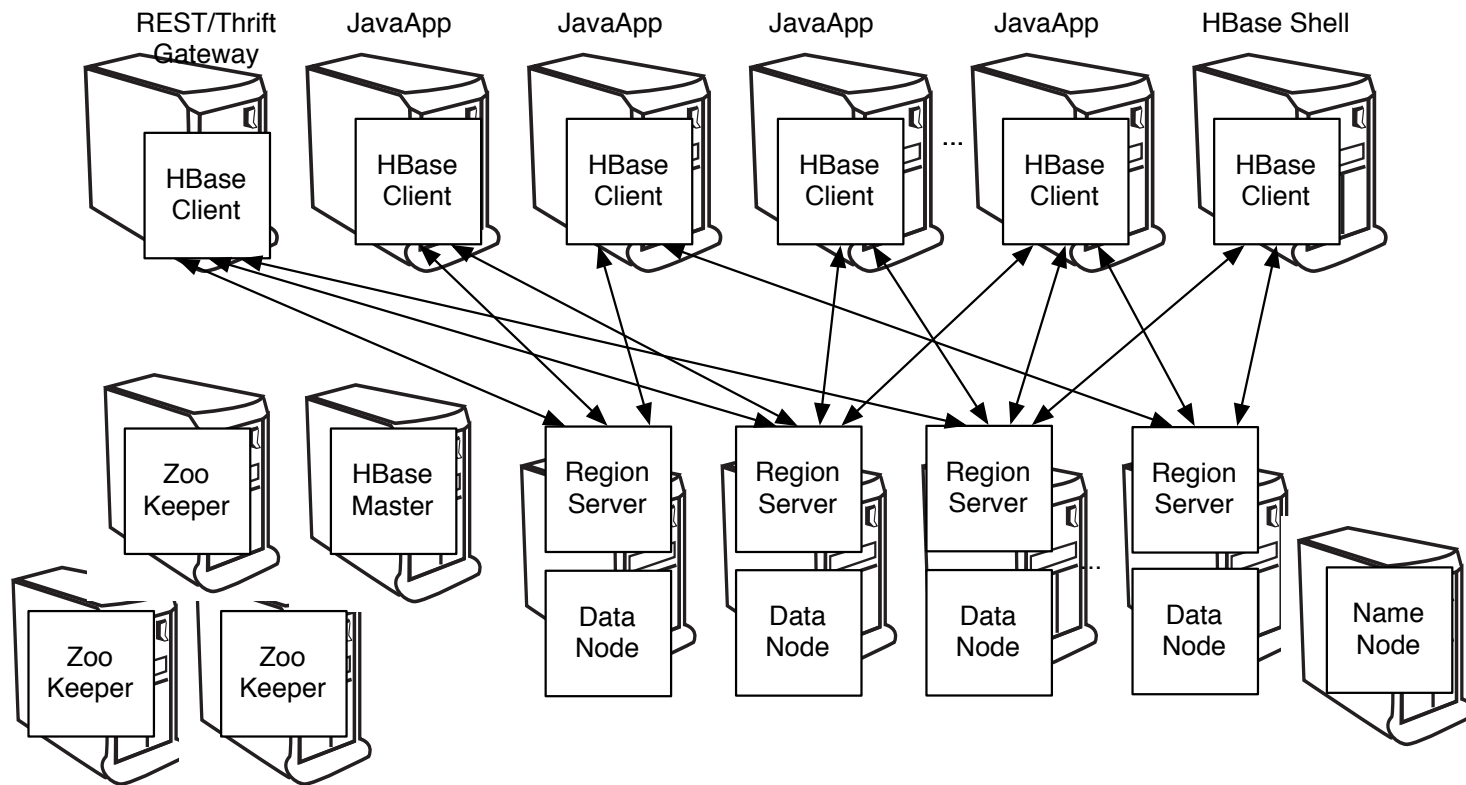
- **{rowkey => {family => {qualifier => {version => value}}}}**
 - Think: nested OrderedDictionary (C#), TreeMap (Java)
- **Basic data operations: GET, PUT, DELETE**
- **SCAN over range of key-values**
 - benefit of the sorted rowkey business
 - this is how you implement any kind of "complex query" *
- **GET, SCAN support Filters**
 - Push application logic to RegionServers
- **INCREMENT, APPEND, CheckAnd{Put,Delete}**
 - Server-side, atomic data operations, can be contentious!

* This is also a foundational component in what we refer to as “schema design” in this “schemaless” database.

Logical Architecture



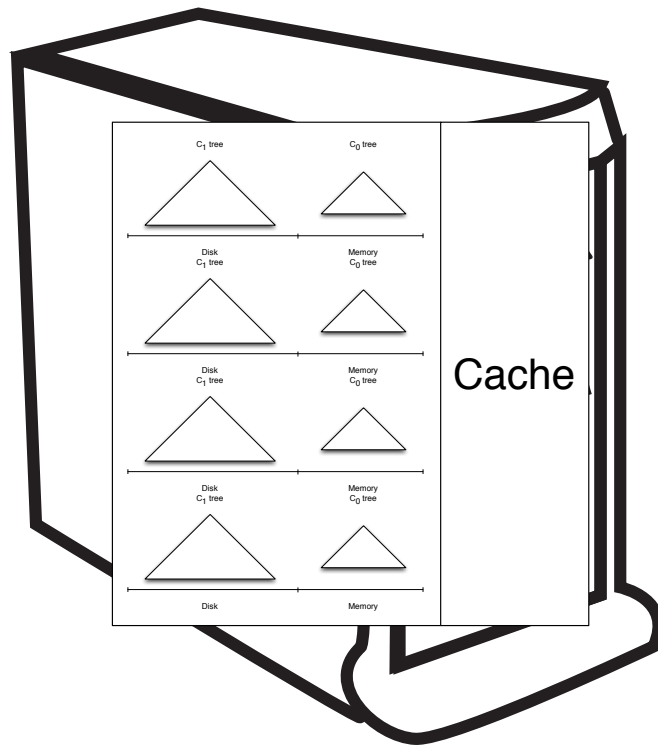
Physical Architecture



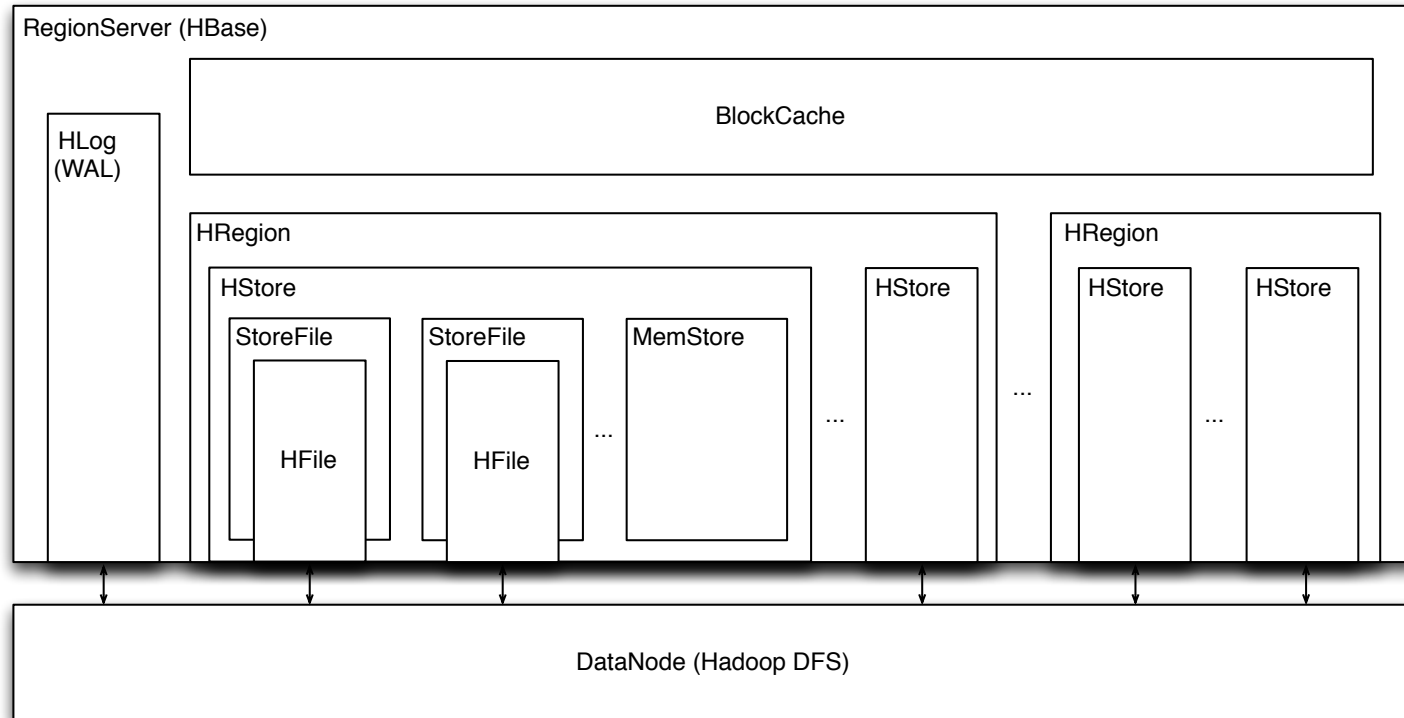


Anatomy of a RegionServer

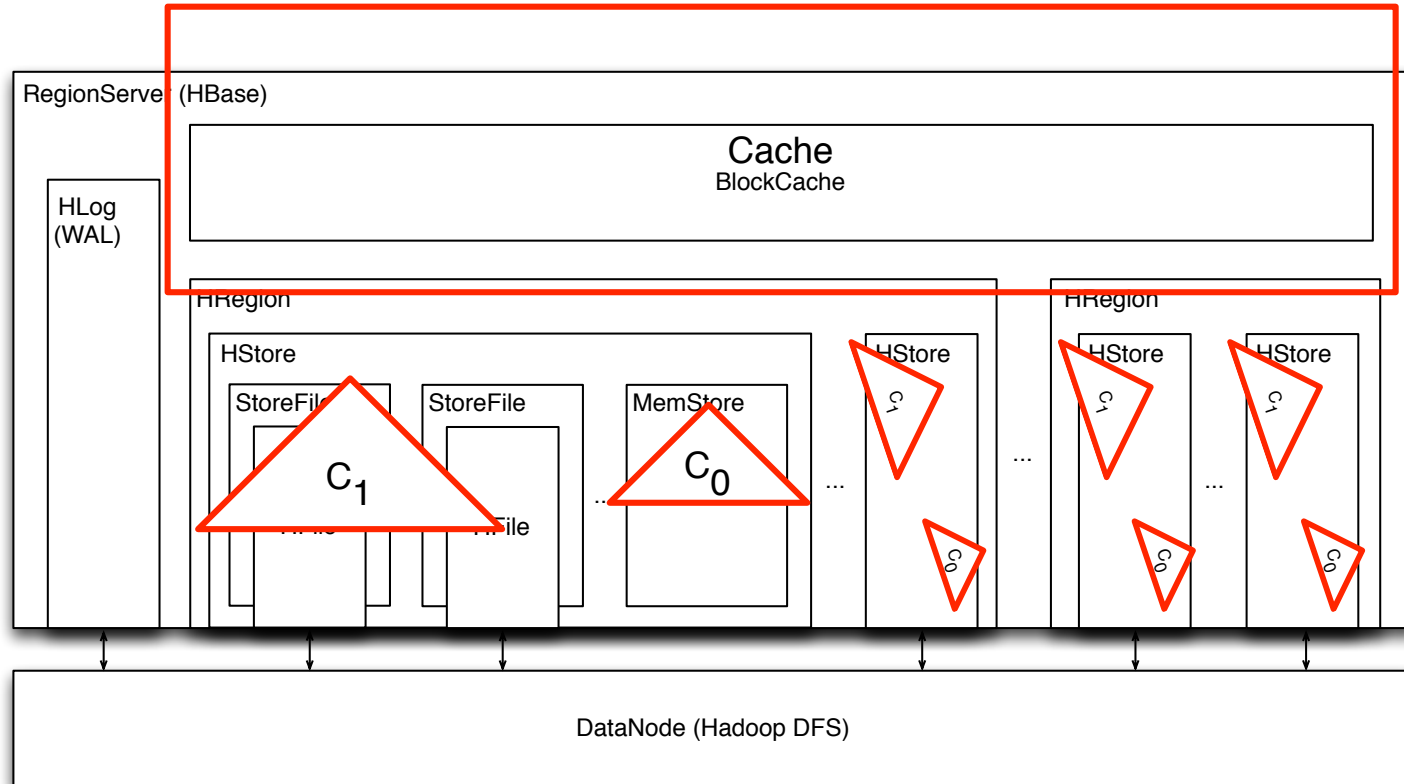
So what is HBase anyway?



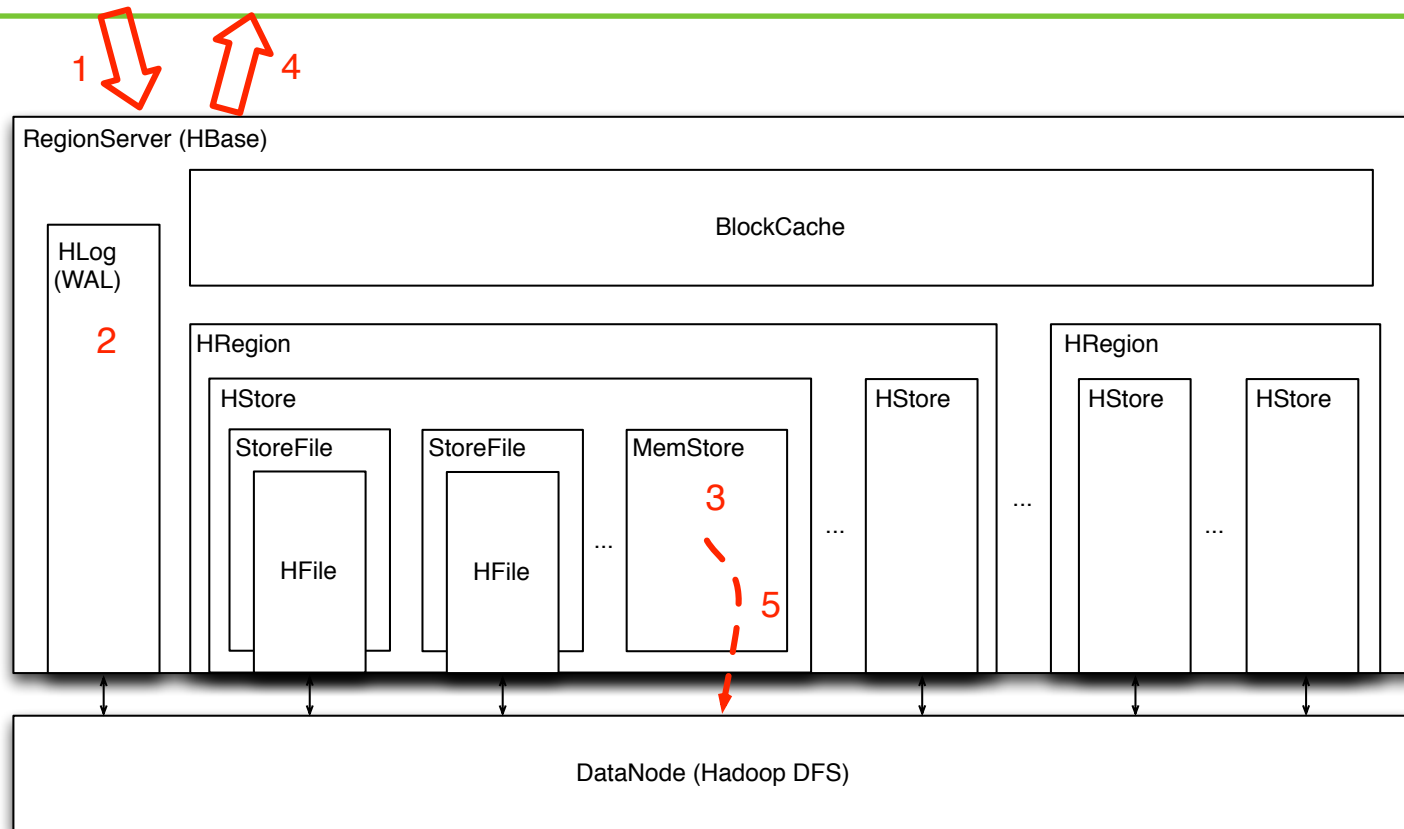
Storage Machinery



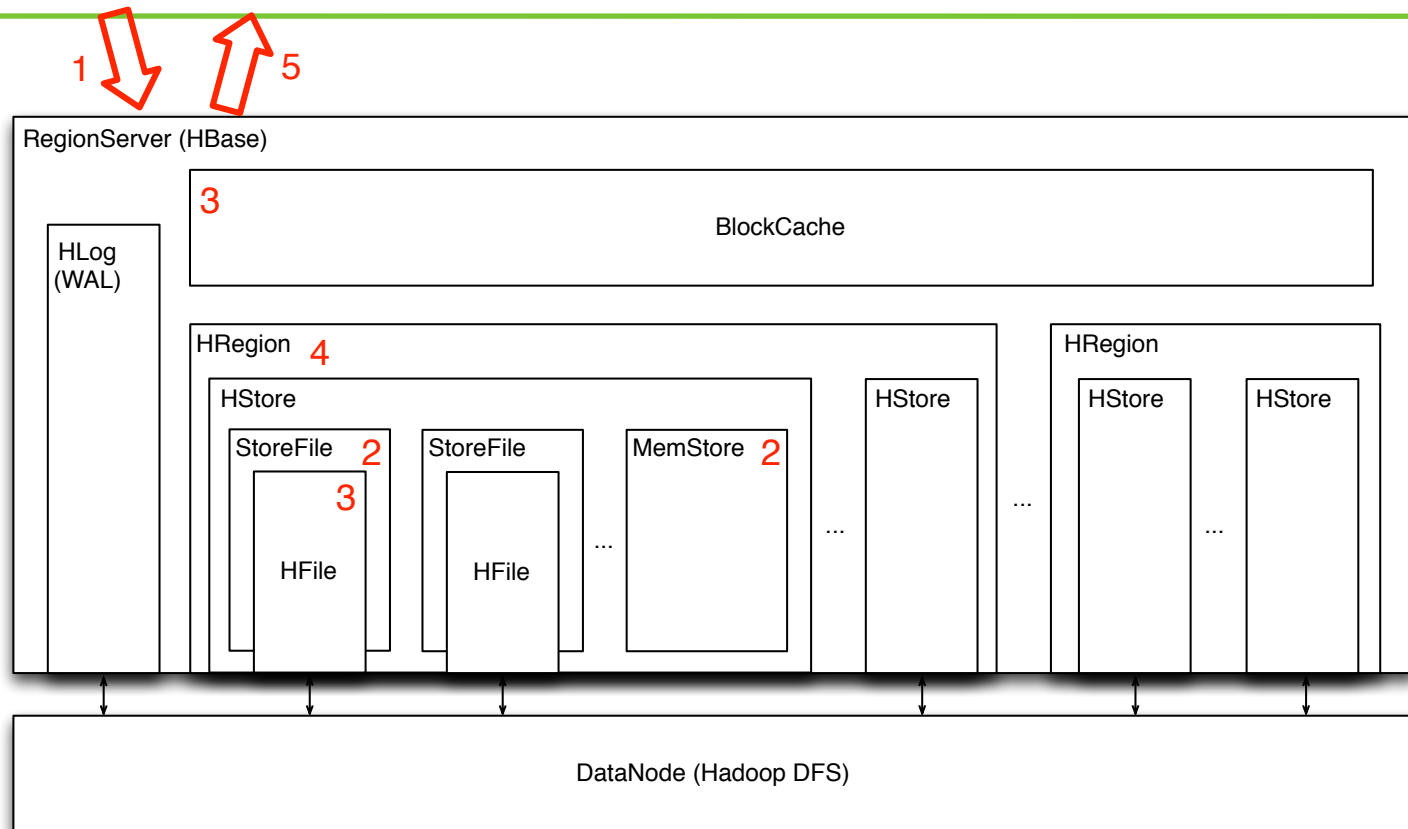
Storage Machinery



Storage Machinery: Write Path



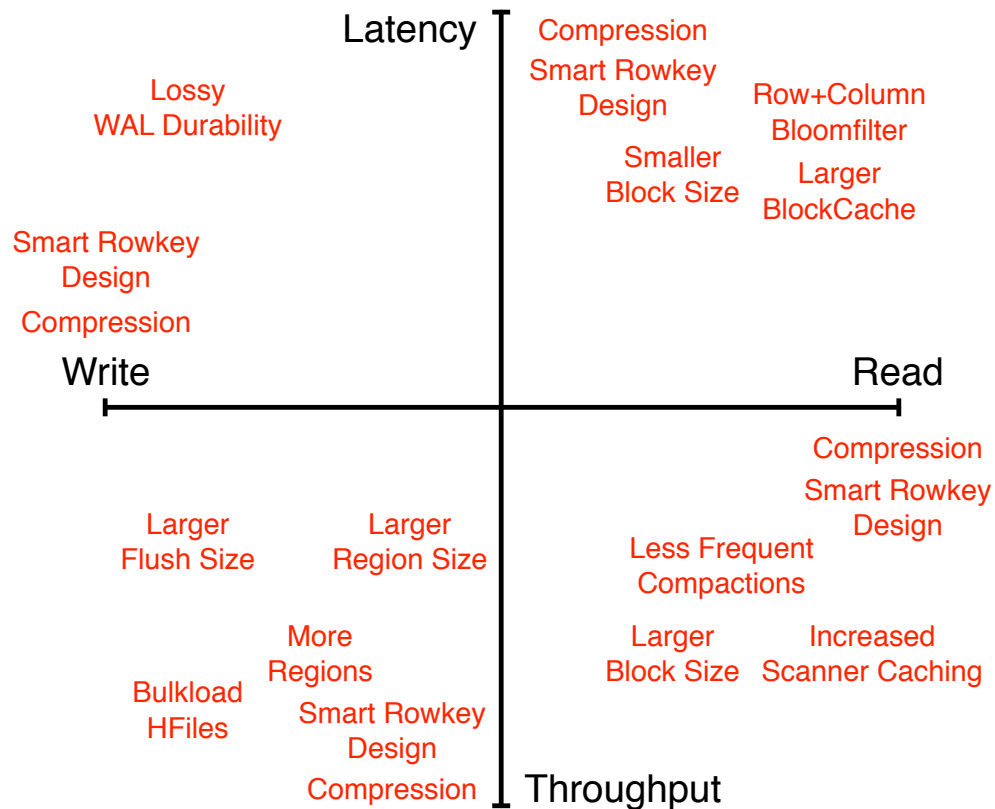
Storage Machinery: Read Path



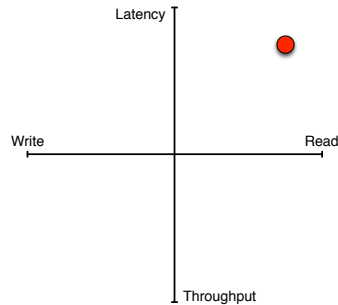
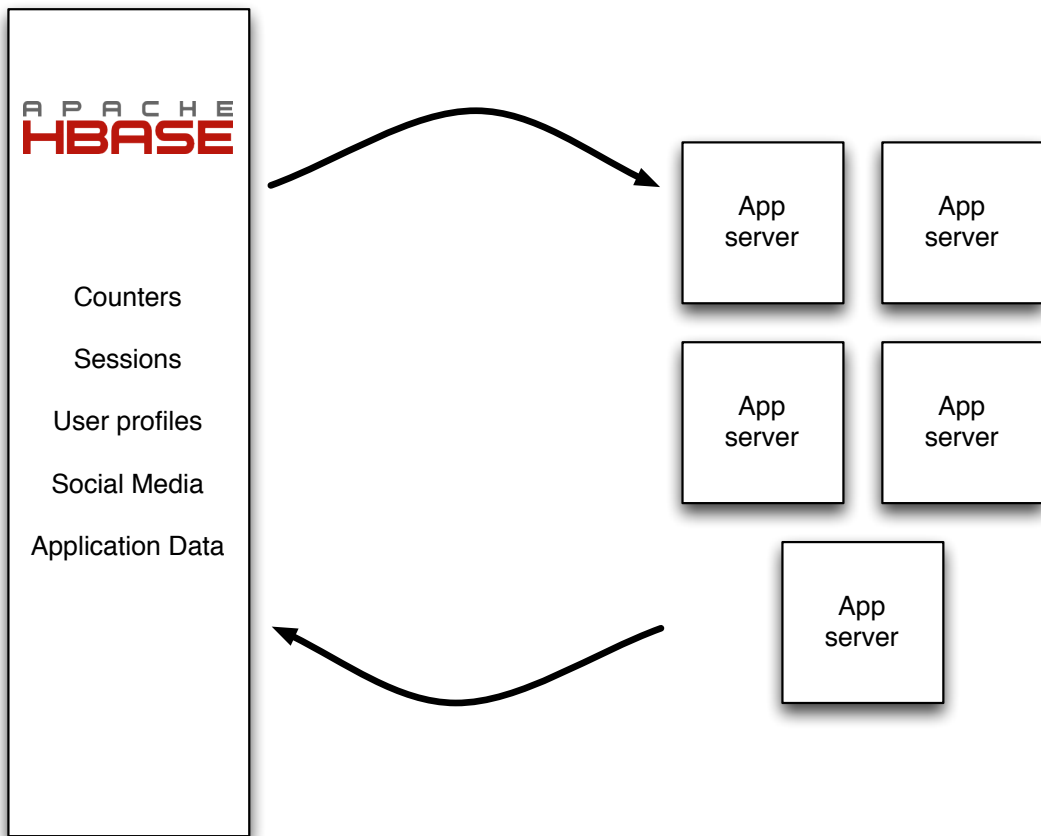


By Example

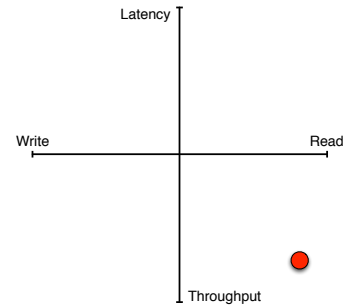
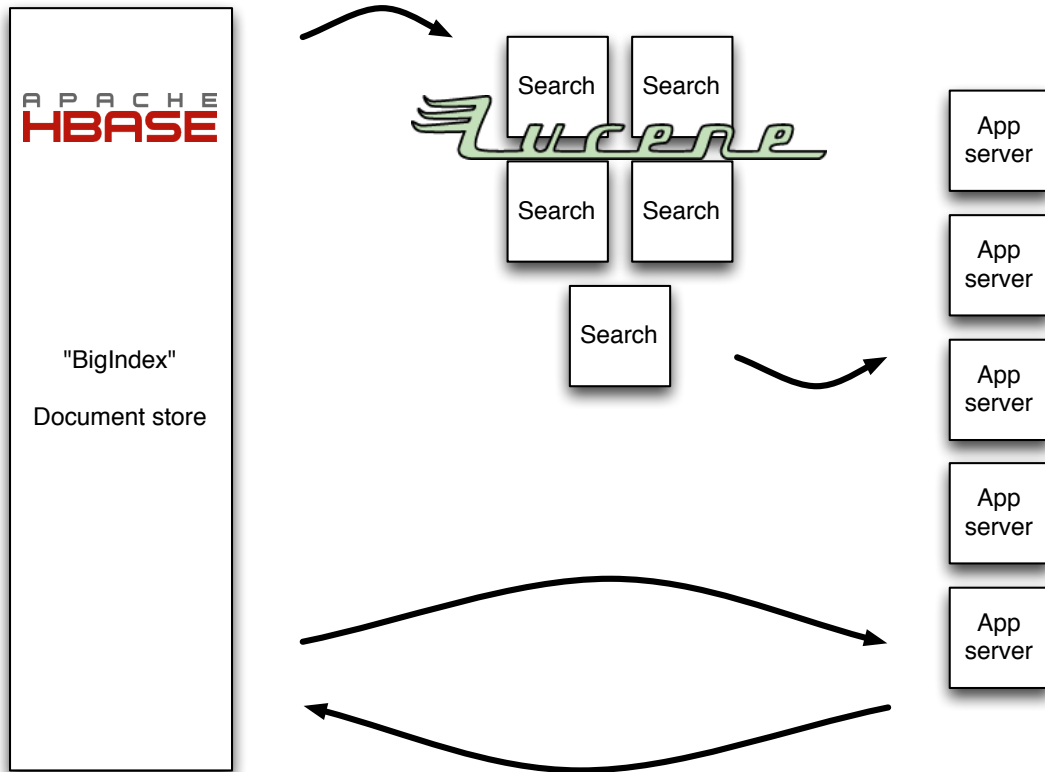
Database Dichotomy



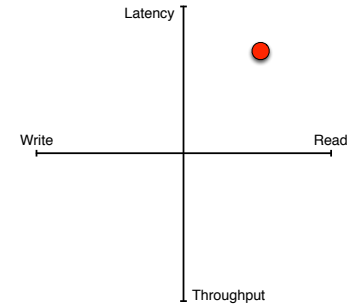
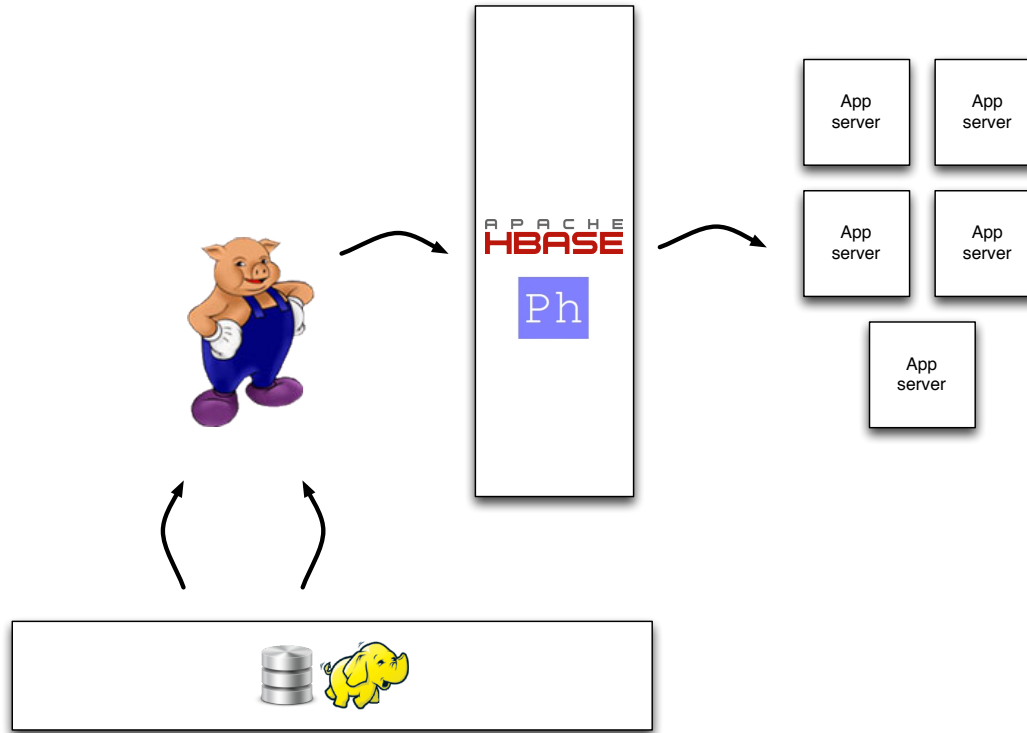
Web-scale Database



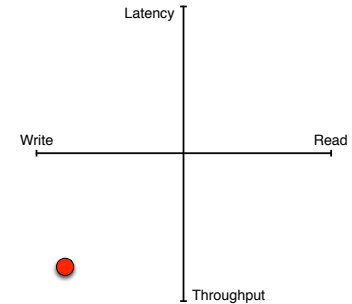
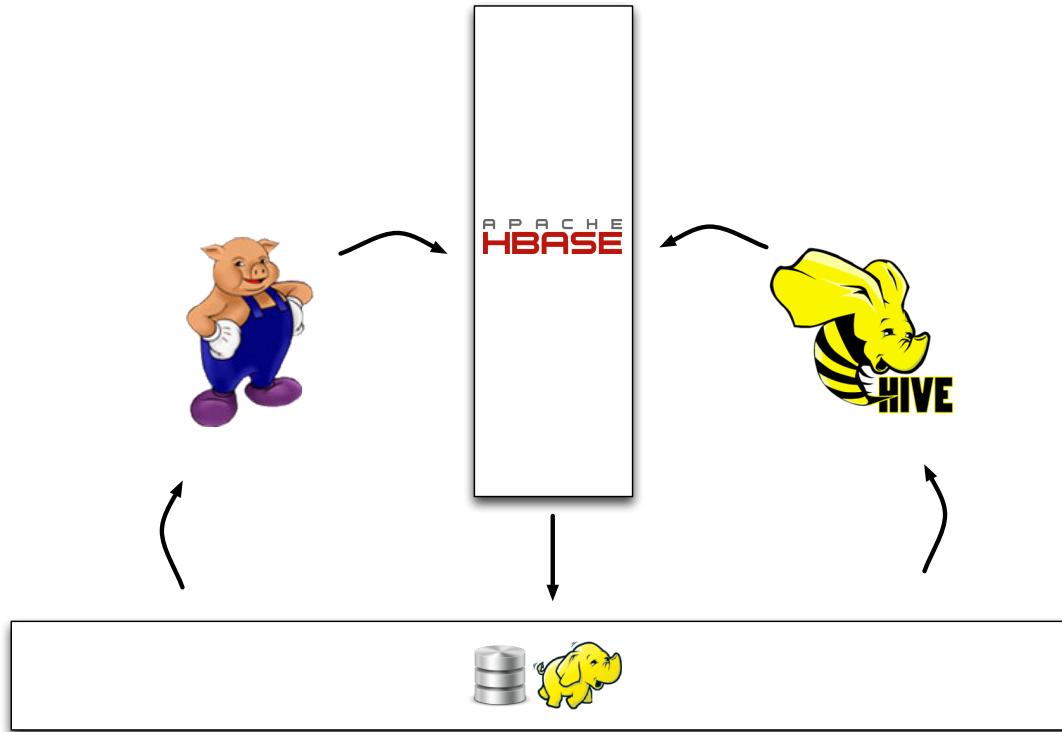
“BigIndex”



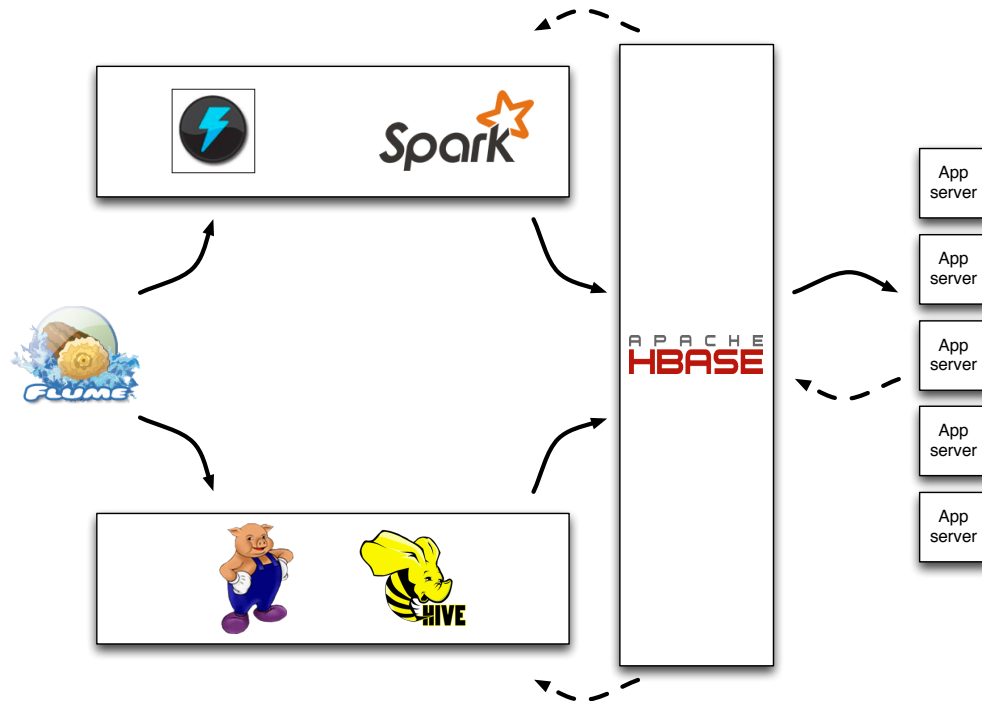
Materialized View



ETL Assist



“Lambda Architecture”





Resources

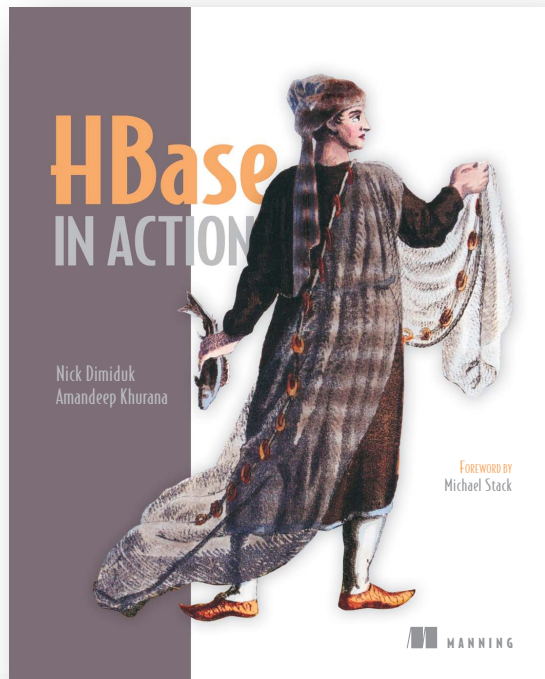
Join the Community!

- **hbase.apache.org**
 - blogs.apache.org/hbase/
 - hbase.apache.org/mail-lists.html
- **IRC: [irc.freenode.net#hbase](https://irc.freenode.net/#hbase)**
- **JIRA: issues.apache.org/jira/browse/HBASE**
- **Source**
 - `git clone git://git.apache.org/hbase.git`
 - `svn checkout http://svn.apache.org/repos/asf/hbase/trunk hbase`
- **In person**
 - HBaseCon, hbasecon.com
 - Hadoop Summit, hadoopsummit.org
 - Local meetup near you

HBase 0.96.x

- **Stability**
- **Horizontal Scalability (1000's of machines)**
- **Speed of Recovery (MTTR)**
- **Operations**
- **Future-proofing RPC (protobuf)**
- **Improved Multi-tenancy**
- **Hadoop1, Hadoop2**
- **Plenty more enhancements, new features:**
 - https://blogs.apache.org/hbase/entry/hbase_0_96_0_released
- ***Backward incompatible***

Thanks!



hbaseinaction.com

Nick Dimiduk

 github.com/ndimiduk

 [@xefyr](https://twitter.com/xefyr)

 n10k.com

