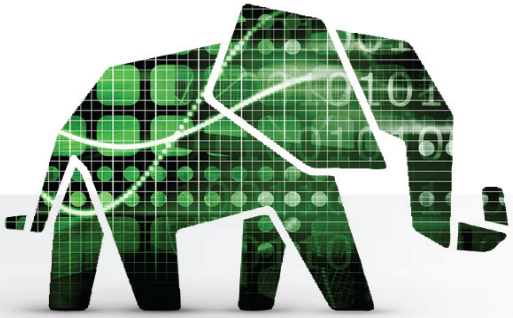
Apache HBase

For Architects

Nick Dimiduk, Hortonworks

Strata/Hadoop World Barcelona, 2014-11-21



Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

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?)

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

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

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

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

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

TL;DR

So what is HBase anyway?

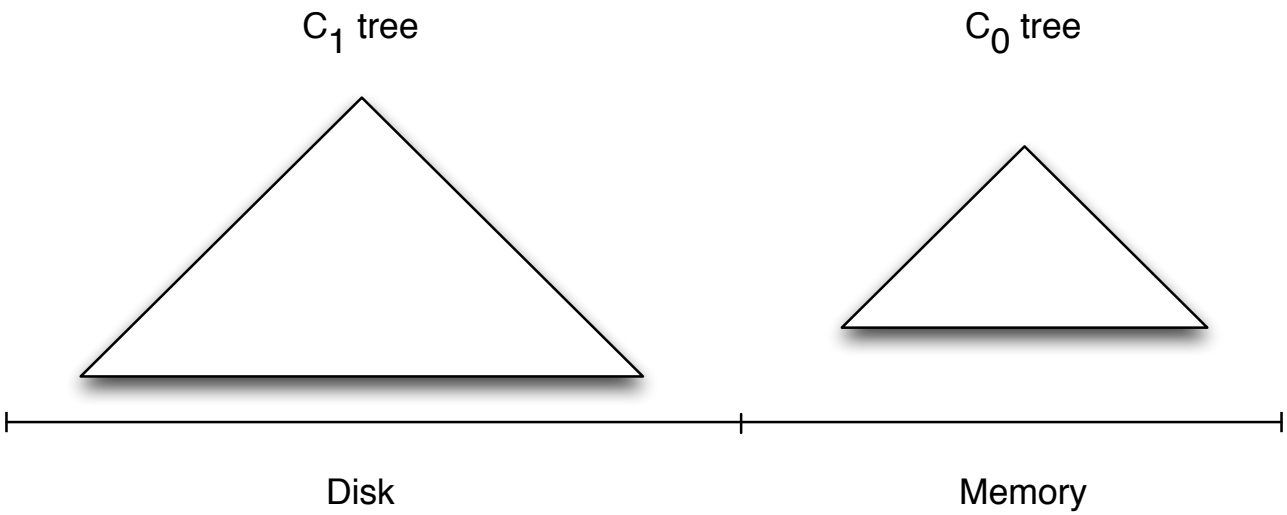
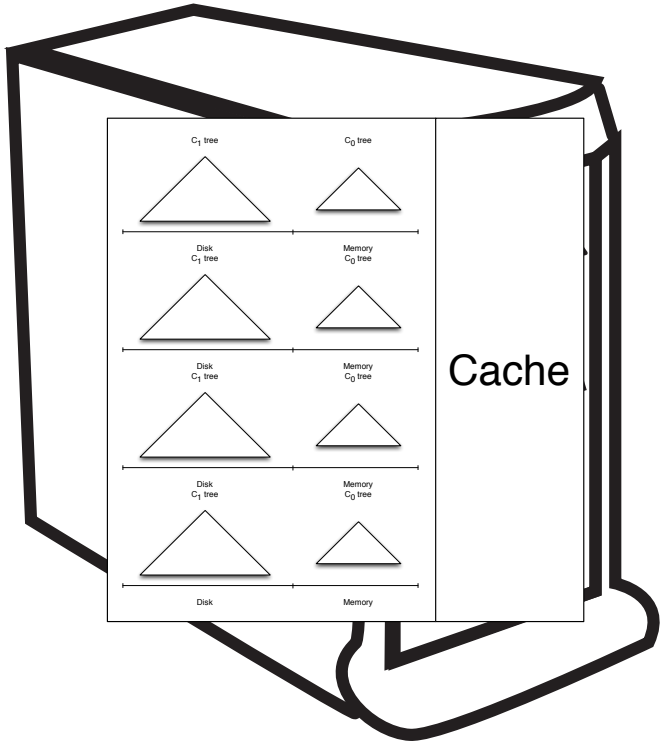


Figure 2.1. Schematic picture of an LSM-tree of two components

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

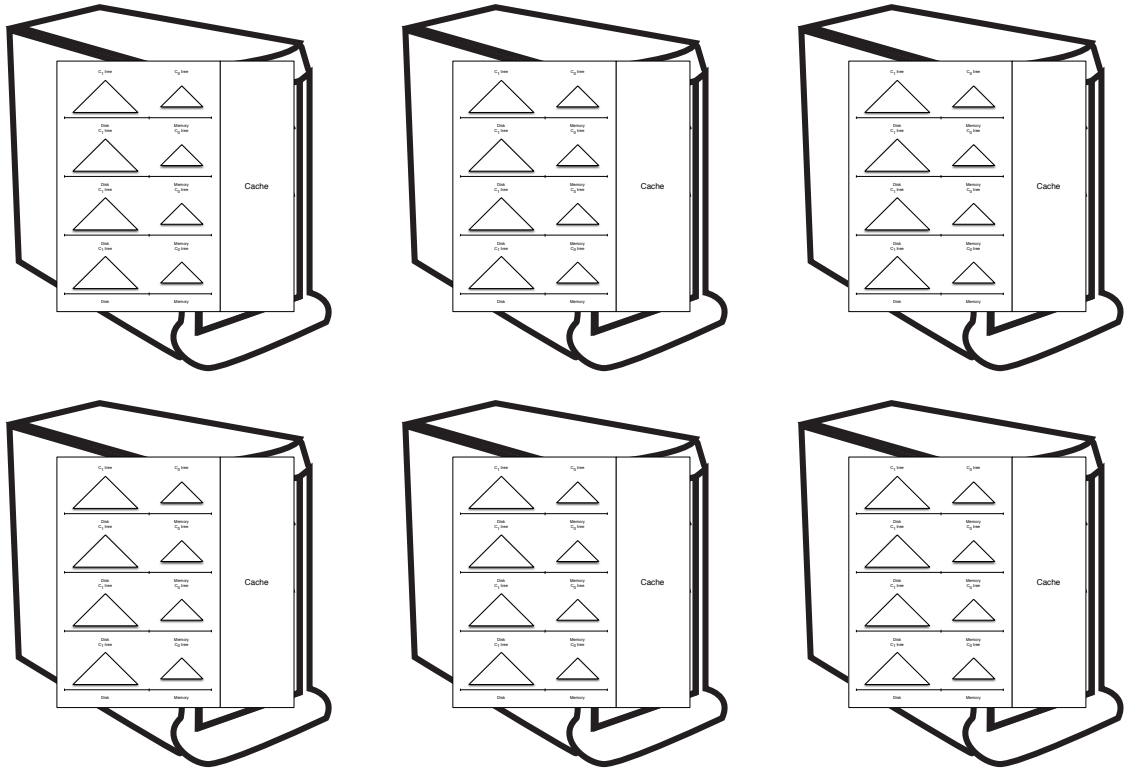
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

So what is HBase anyway?



Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

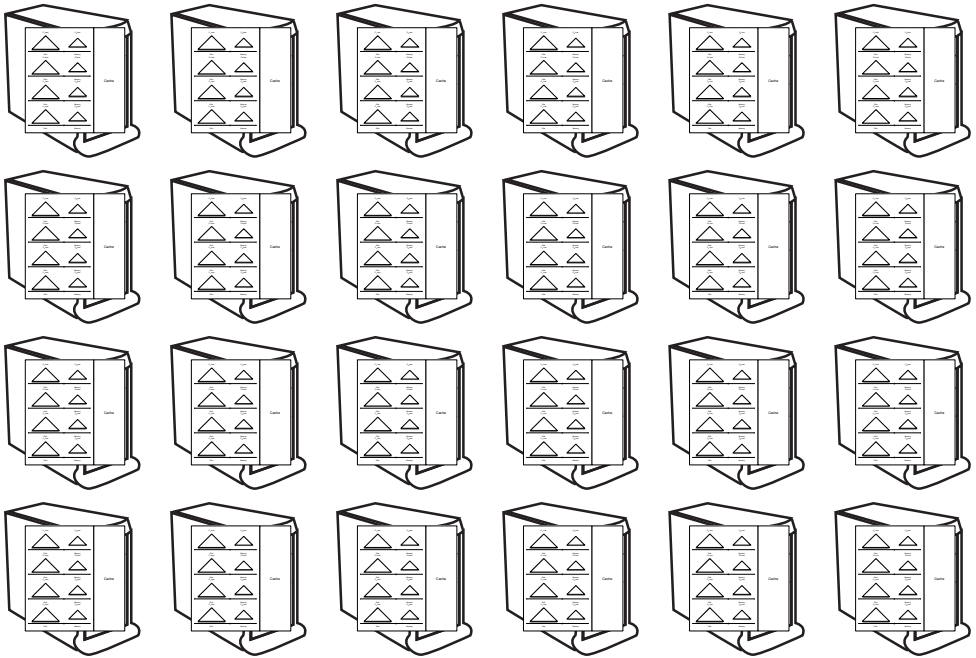
So what is HBase anyway?



Hortonworks

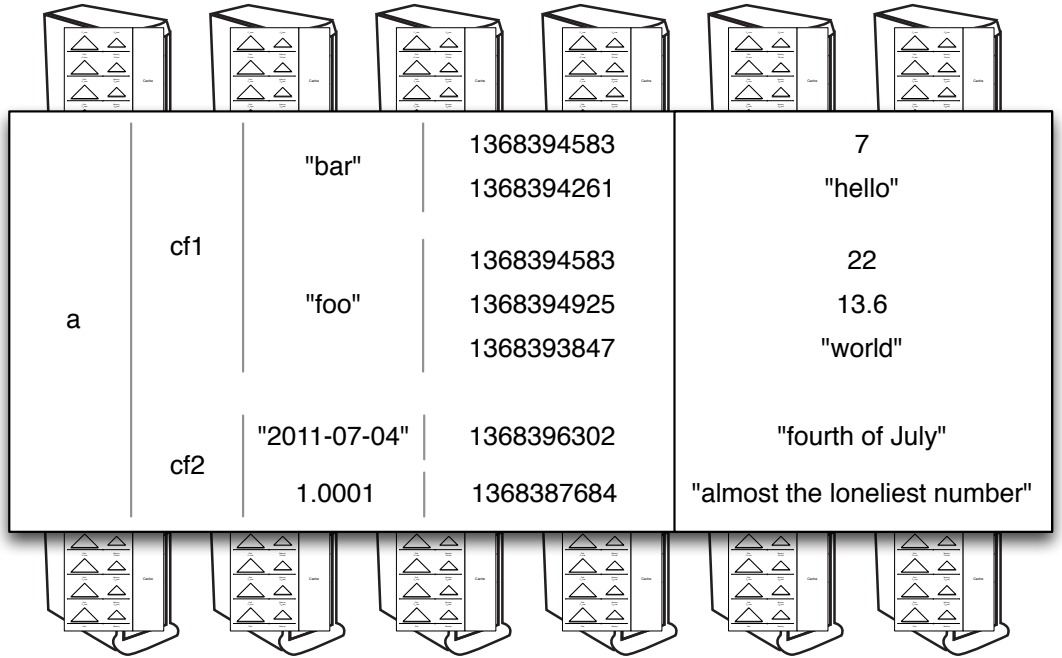
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

So what is HBase anyway?



Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

So what is HBase anyway?



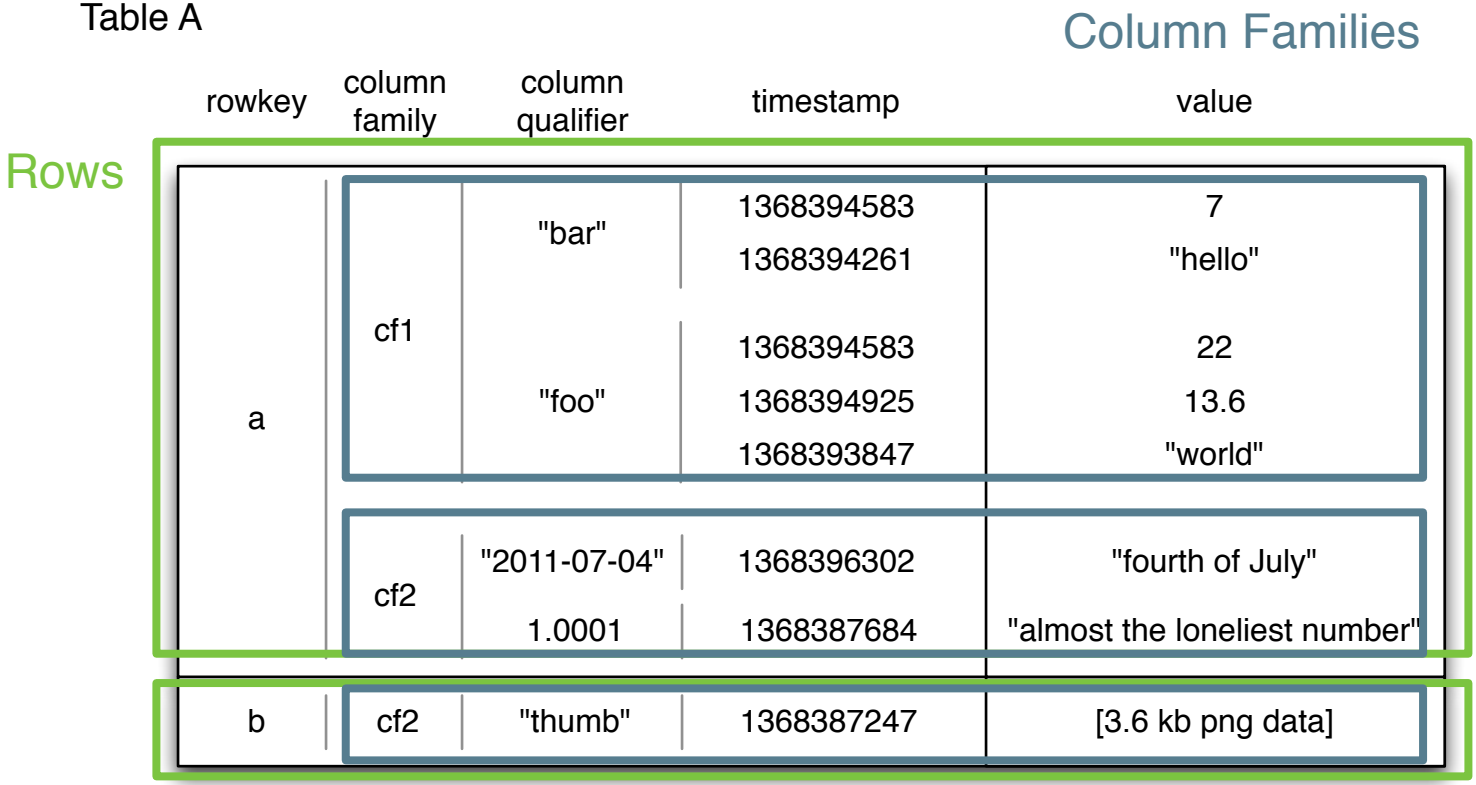
a	cf1	"bar"	1368394583	7
			1368394261	"hello"
		"foo"	1368394583	22
			1368394925	13.6
			1368393847	"world"
	cf2	"2011-07-04"	1368396302	"fourth of July"
		1.0001	1368387684	"almost the loneliest number"

Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

High-level Architecture

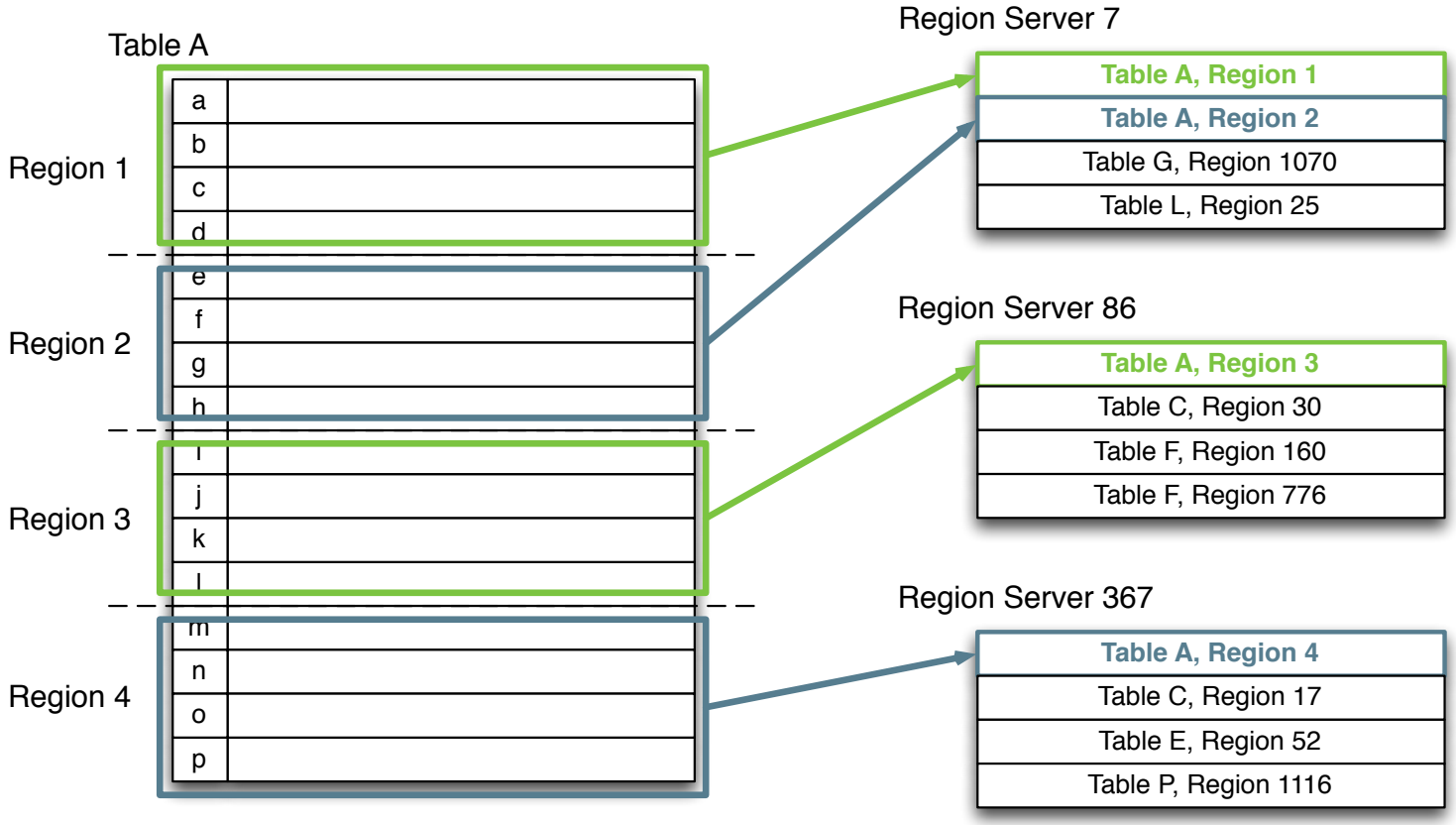
Logical Data Model



Hortonworks

Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

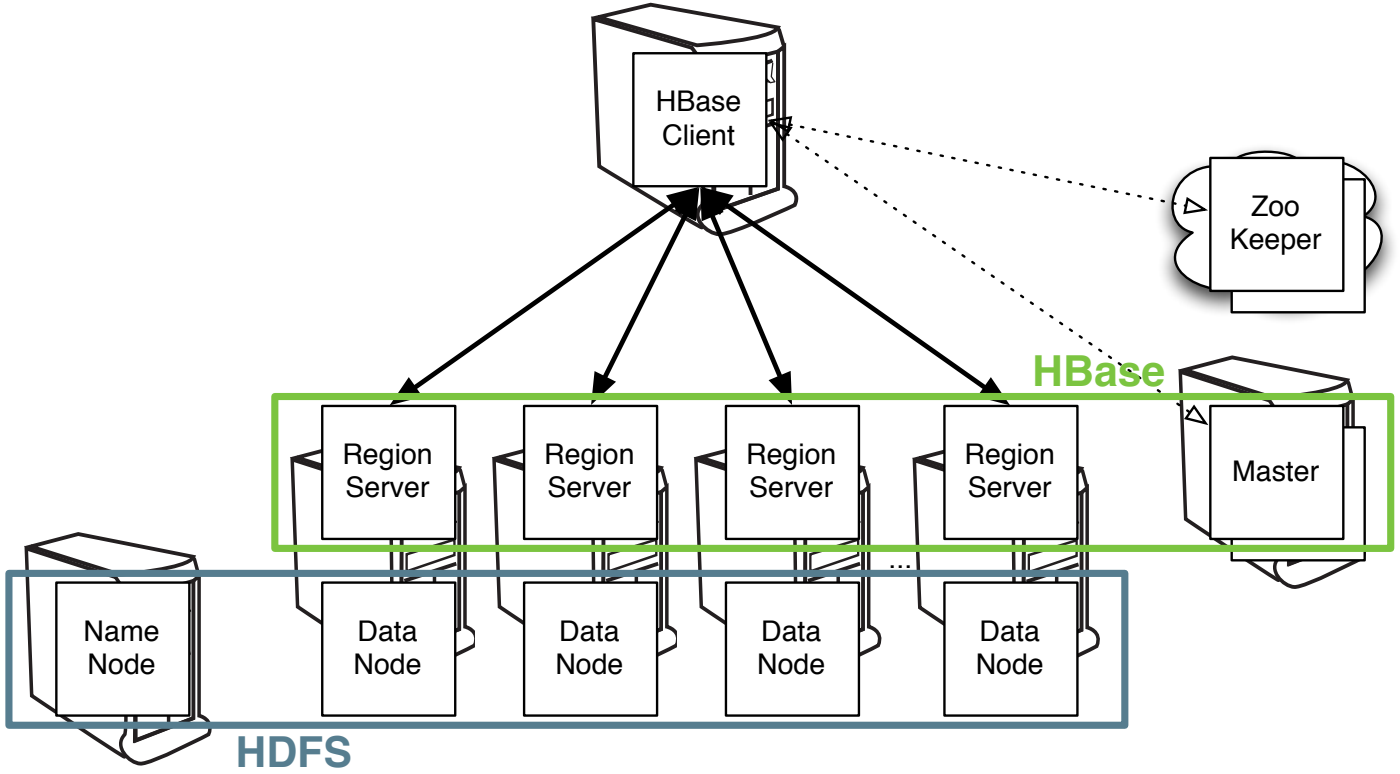
Logical Architecture



Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Physical Architecture



Hortonworks

Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

User API

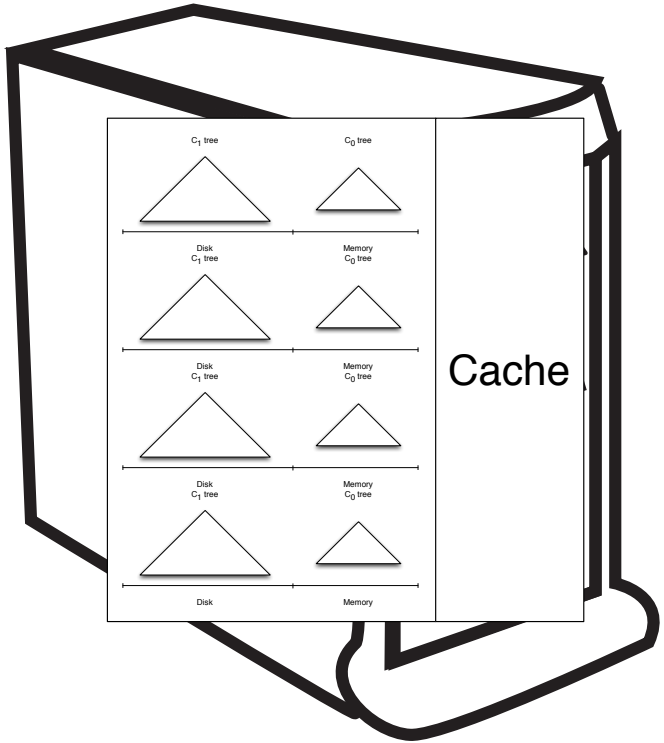
- **{rowkey => {family => {qualifier => {version => value}}}}**
  - Think: nested TreeMap (Java), OrderedDictionary (C#), OrderedDict (Python)
- **Basic data operations: GET, PUT, DELETE**
- **SCAN over range of key-values**
  - benefit of the sorted rowkey business
  - this is how you implement any 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.

Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

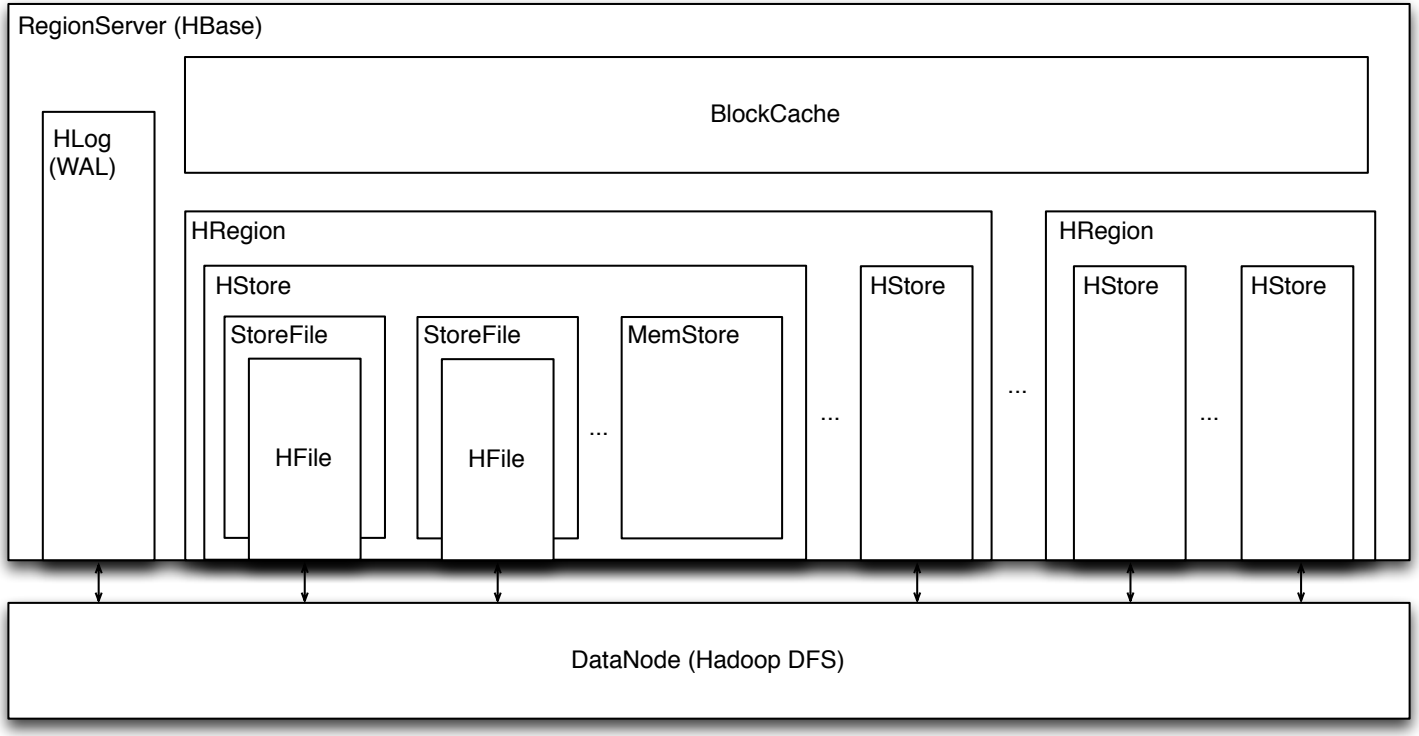
Anatomy of a RegionServer

So what is HBase anyway?



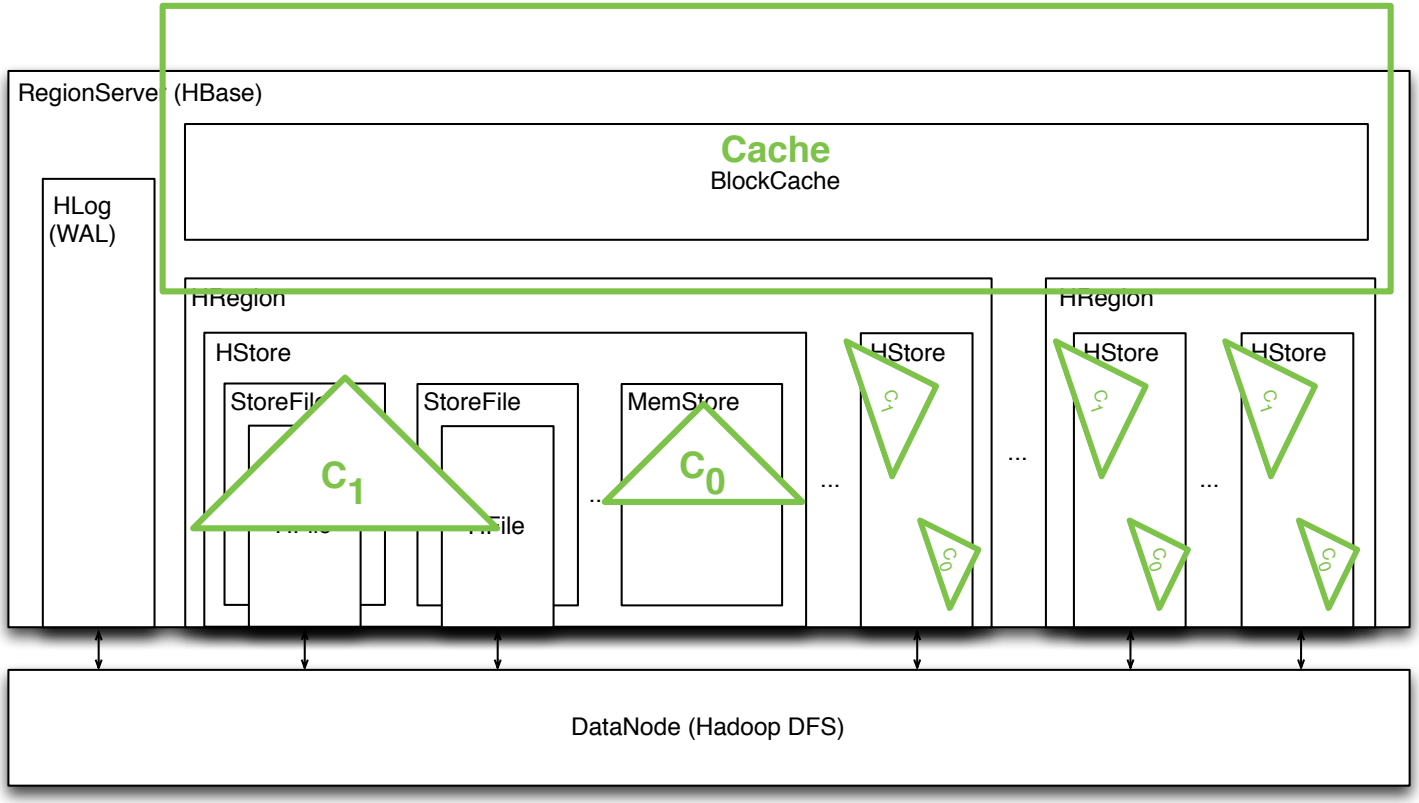
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Storage Machinery



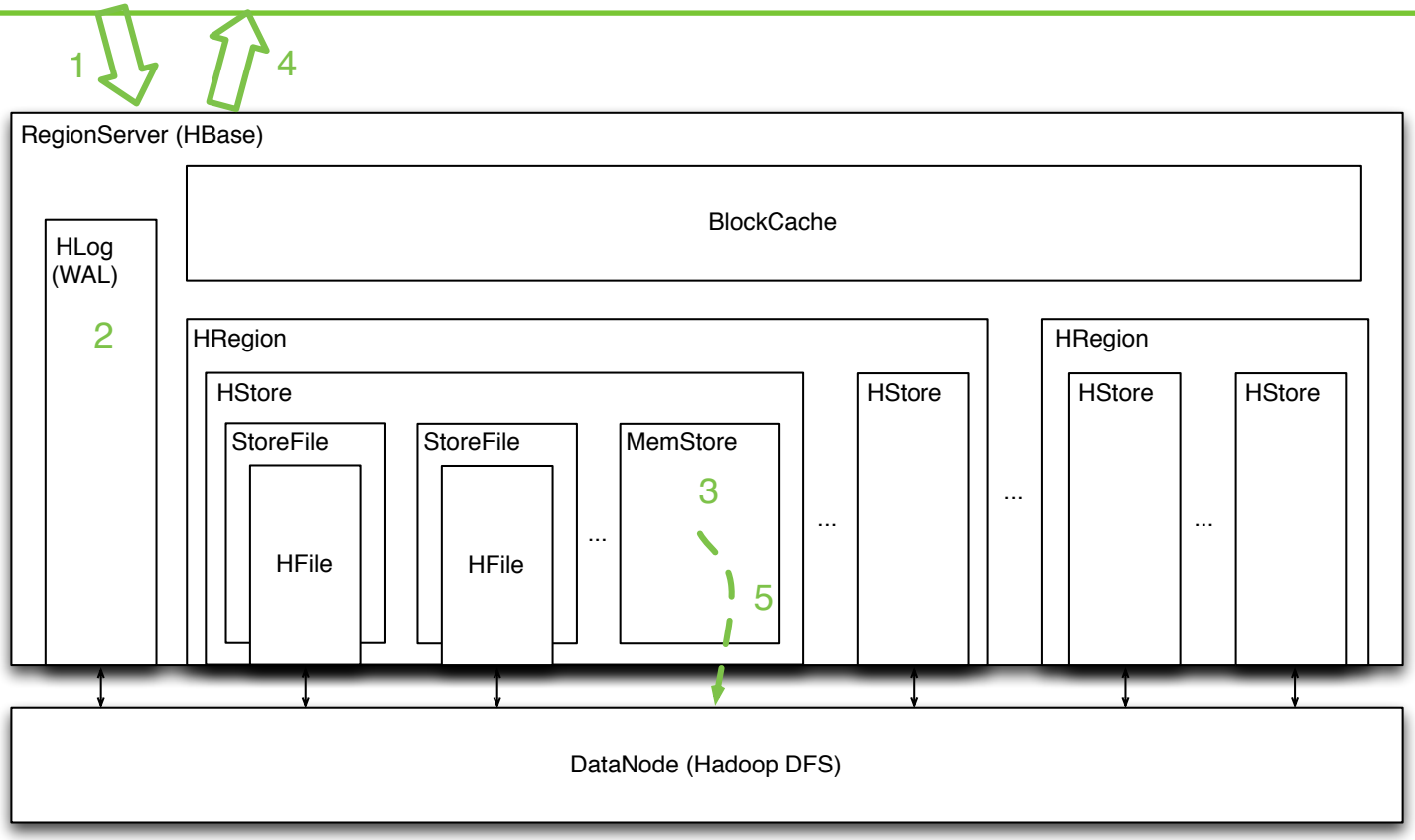
Hortonworks
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Storage Machinery



Hortonworks
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

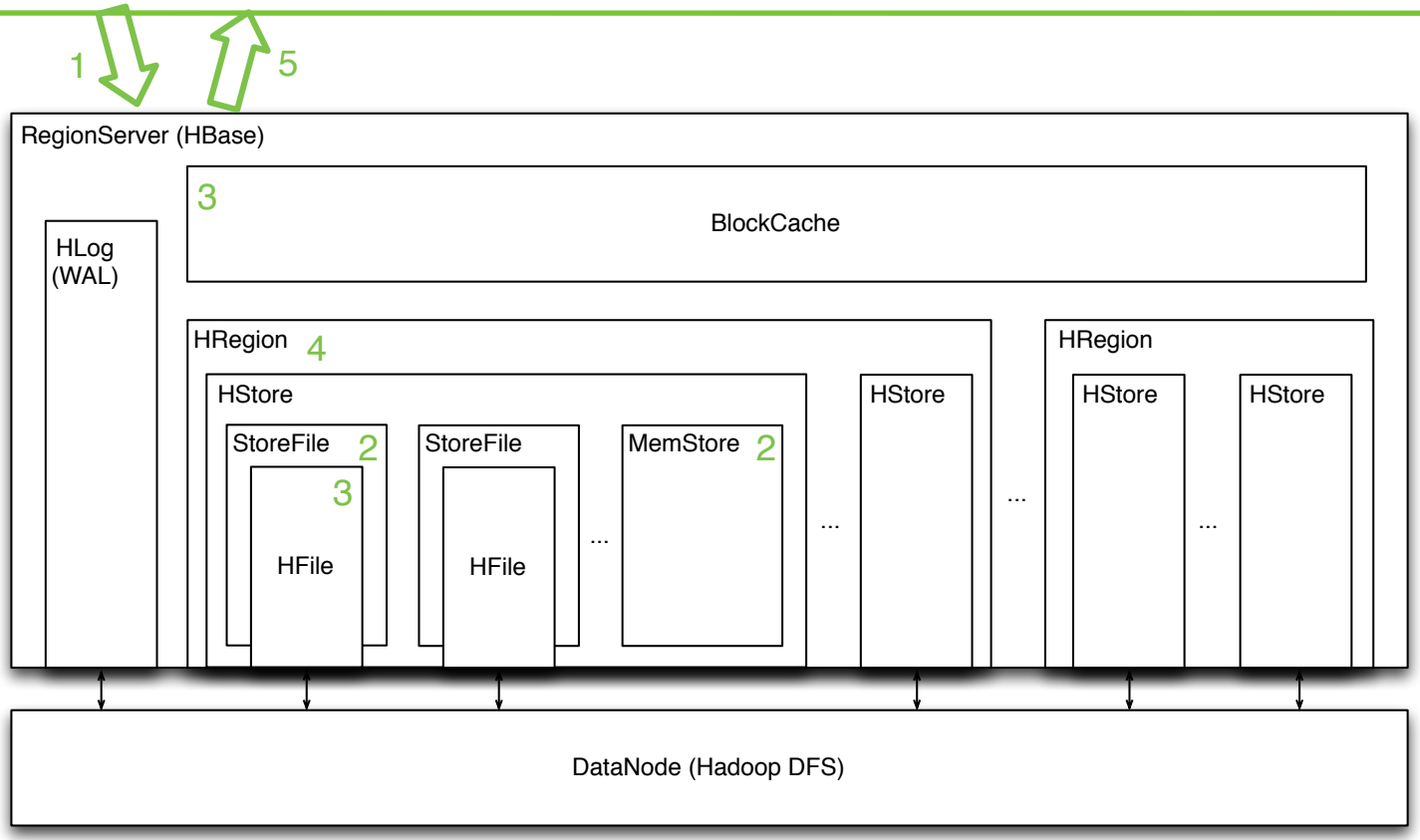
Storage Machinery: Write Path



Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

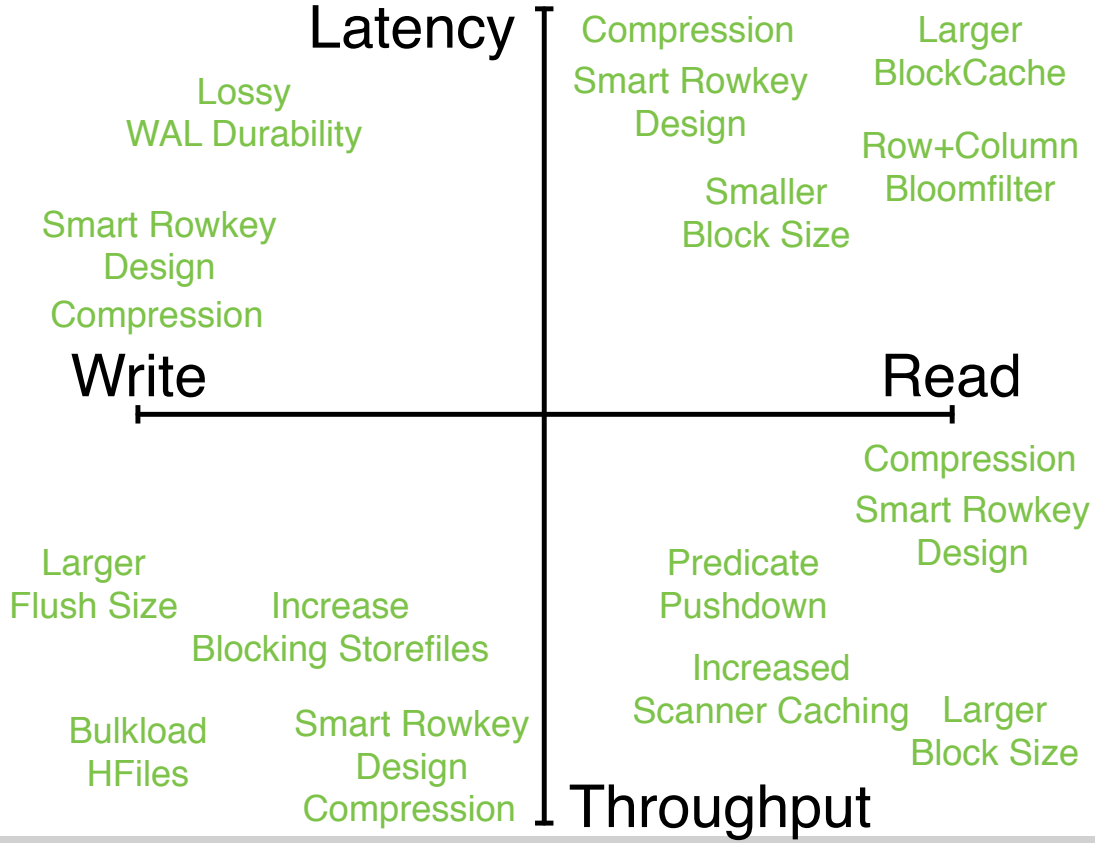
Storage Machinery: Read Path



Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

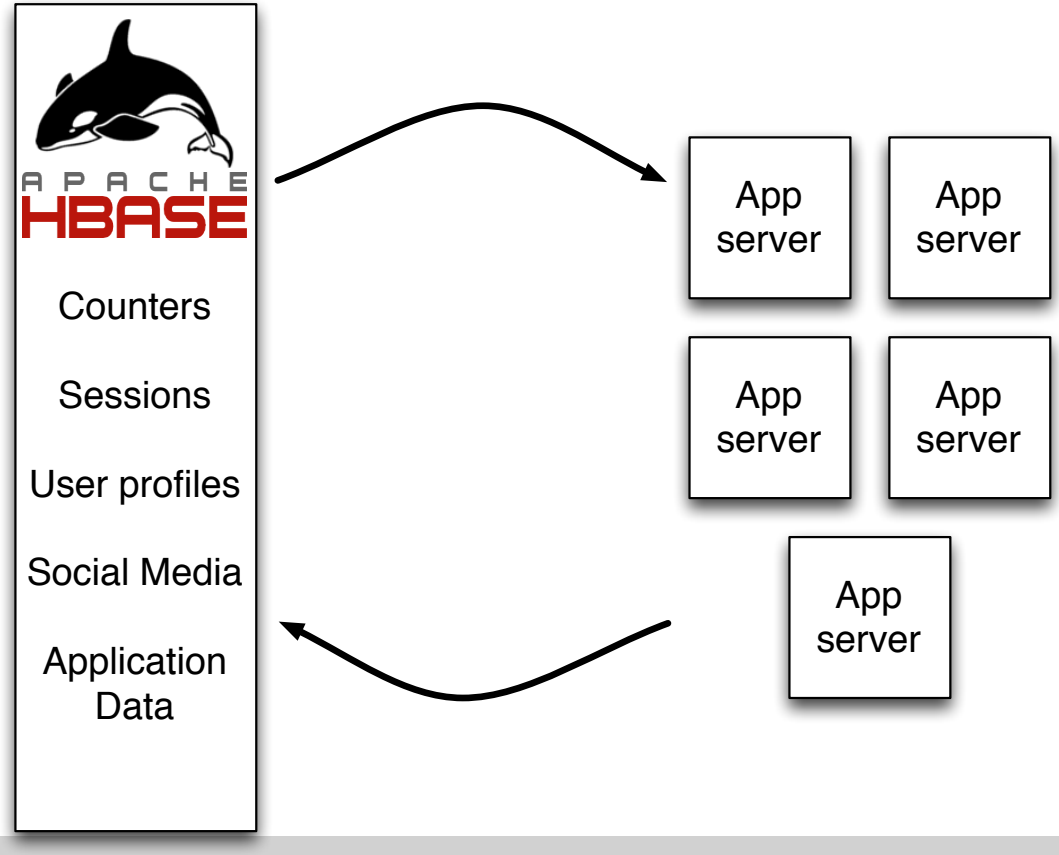
By Example

Database Dichotomy



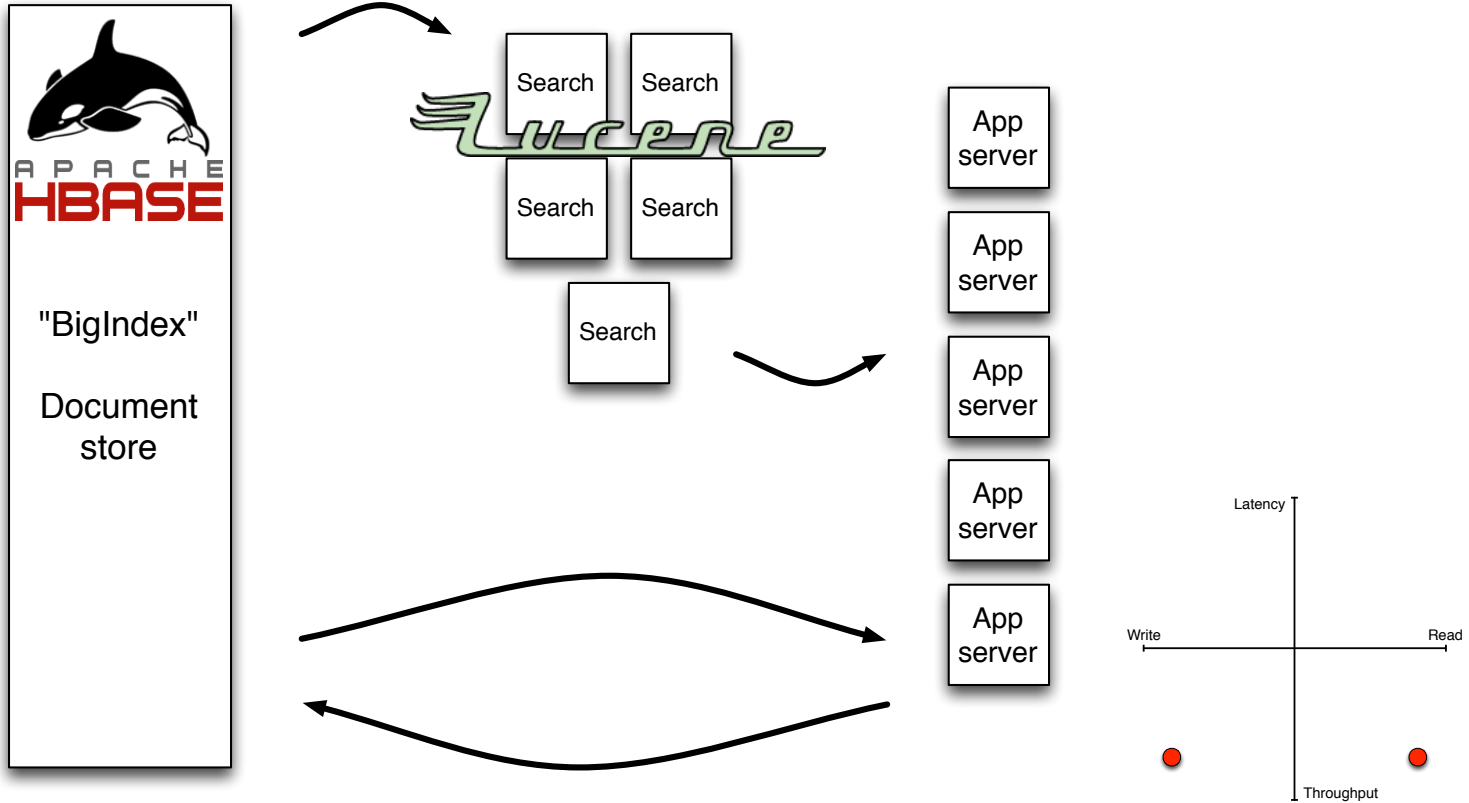
Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

Web-scale Database



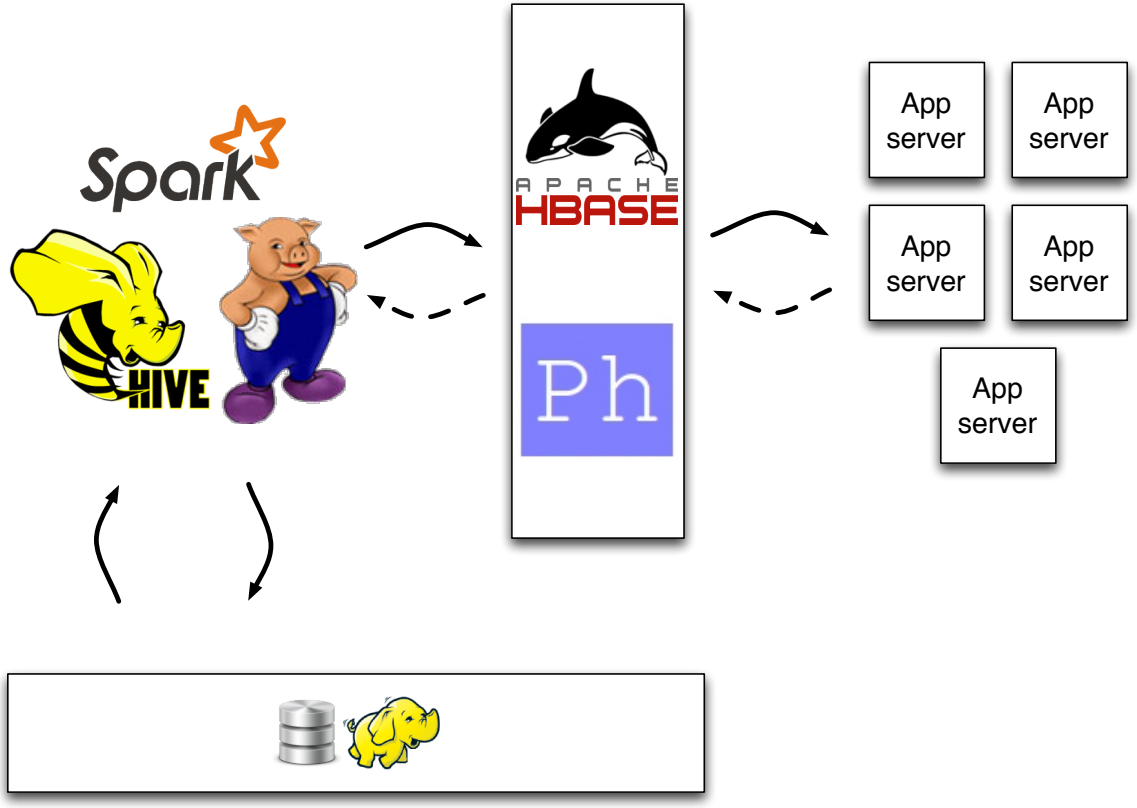
Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

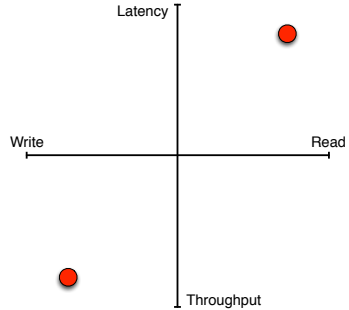
“BigIndex”



Hortonworks
Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

Materialized View

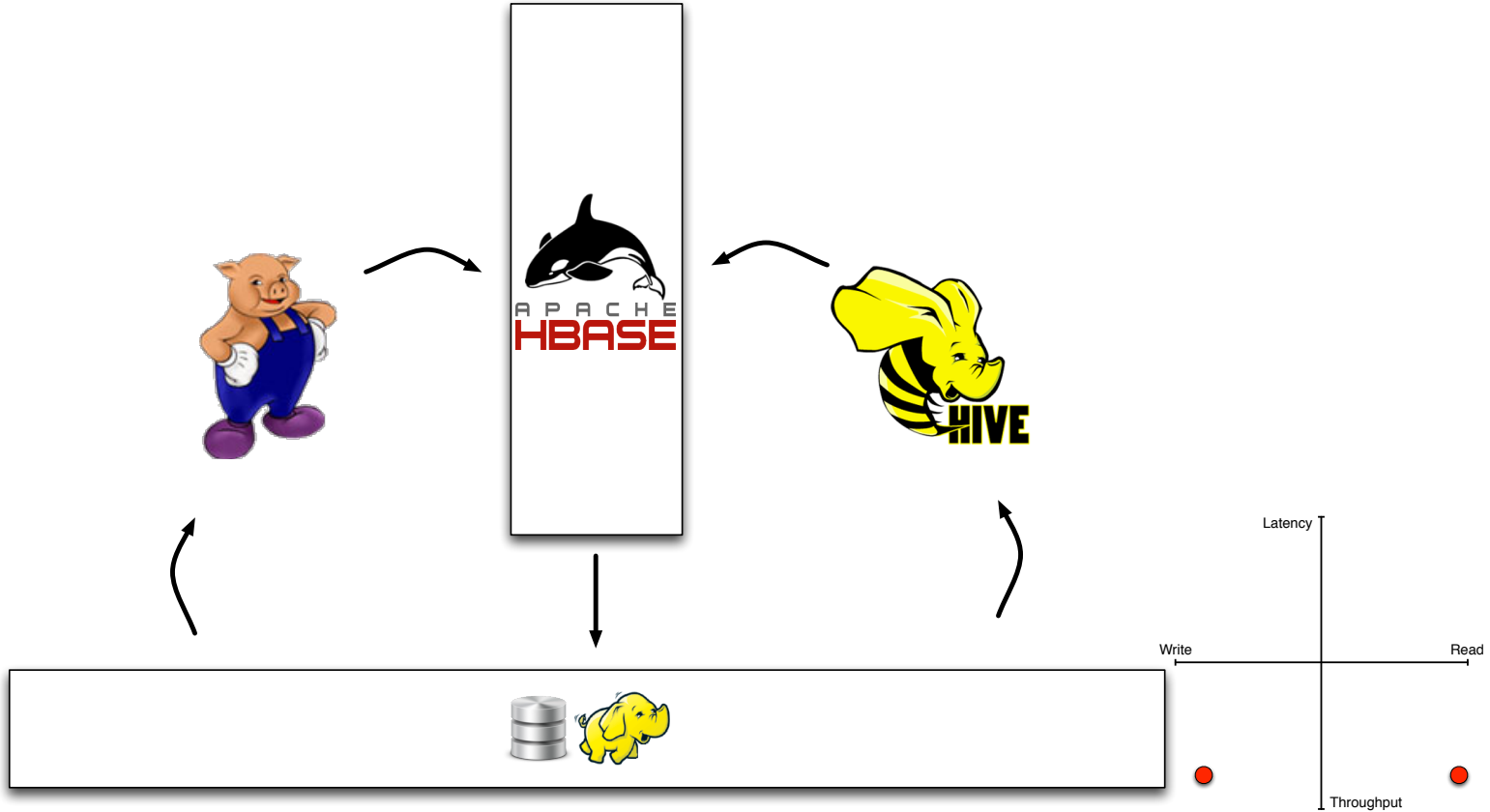




Hortonworks

Licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

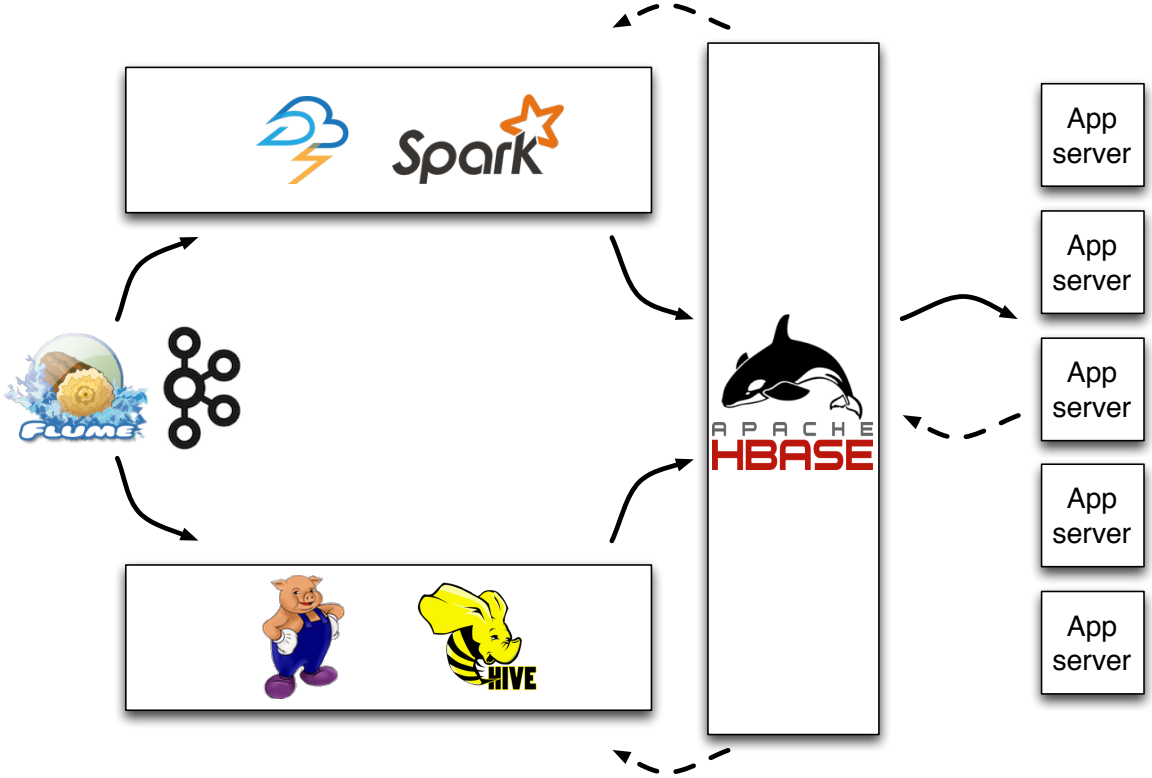
ETL Assist



Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Lambda Architecture



Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Resources

Join the Community!

- **hbase.apache.org**
  - hbase.apache.org/book.html
  - blogs.apache.org/hbase/
  - hbase.apache.org/mail-lists.html
- **IRC: irc.freenode.net #hbase**
- **JIRA: issues.apache.org/jira/browse/HBASE**
- **Source: git clone git://git.apache.org/hbase.git**
- **In person**
  - HBaseCon, hbasecon.com
  - Hadoop Summit, hadoopsummit.org
  - Strata / Hadoop World, strataconf.com
  - Local meetup near you!

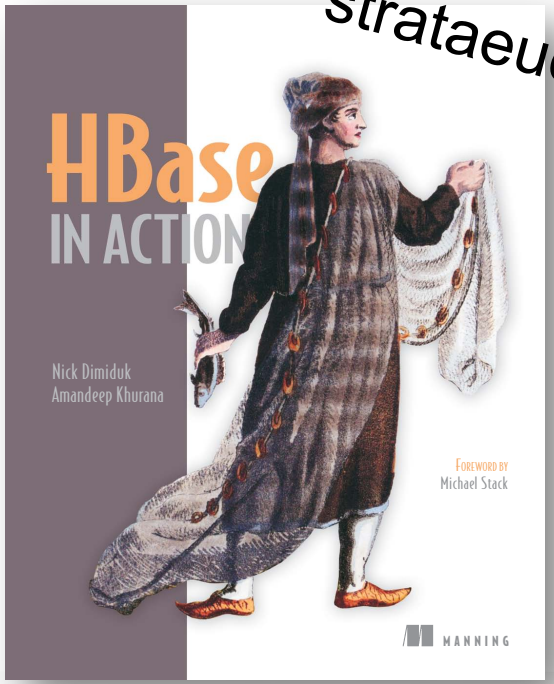
Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

HBase 0.98/1.0

- **Hardening**
  - Stability, Reliability, Availability, Performance
- **Horizontal Scalability**
  - 1000's of machines
- **Availability**
  - Speed of Recovery (MTTR), Region Replicas
- **Improved Multi-tenancy**
  - RPC priorities/QoS, namespace management
- **Cell-level security**
- **Semantic Versioning**
- **Client API cleanup**

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.

Thanks!



strataeucftw

hbaseinaction.com

Nick Dimiduk

github.com/ndimiduk

@xefyr

n10k.com

slideshare.net/xefyr/hbase-for-architects

Hortonworks

Licensed under a Creative Commons
Attribution-ShareAlike 3.0 Unported License.