

The Pilot Of HBase

2017.2

XenRon

CONTENTS

NoSQL

01

HBase Origin

02

Architecture

03

HBase Environment









04

05

Quick Start

06

Reference Books

Increasing Data Complexity ↓	Type	Examples
	Key-Value Store	 
	Wide Column Store	 
	Document Store	 
	Graph Store	 



PART1



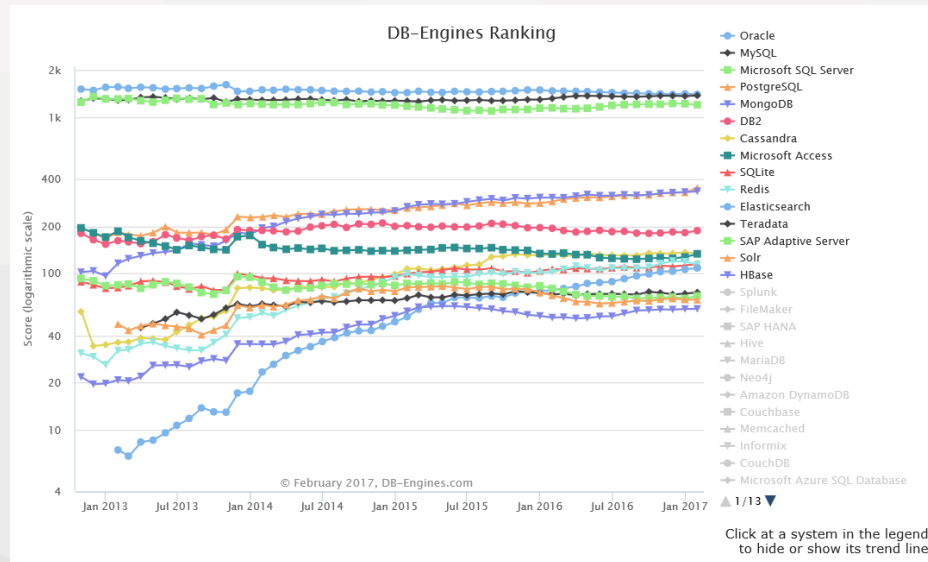
NoSQL

DB-Engines Ranking

4

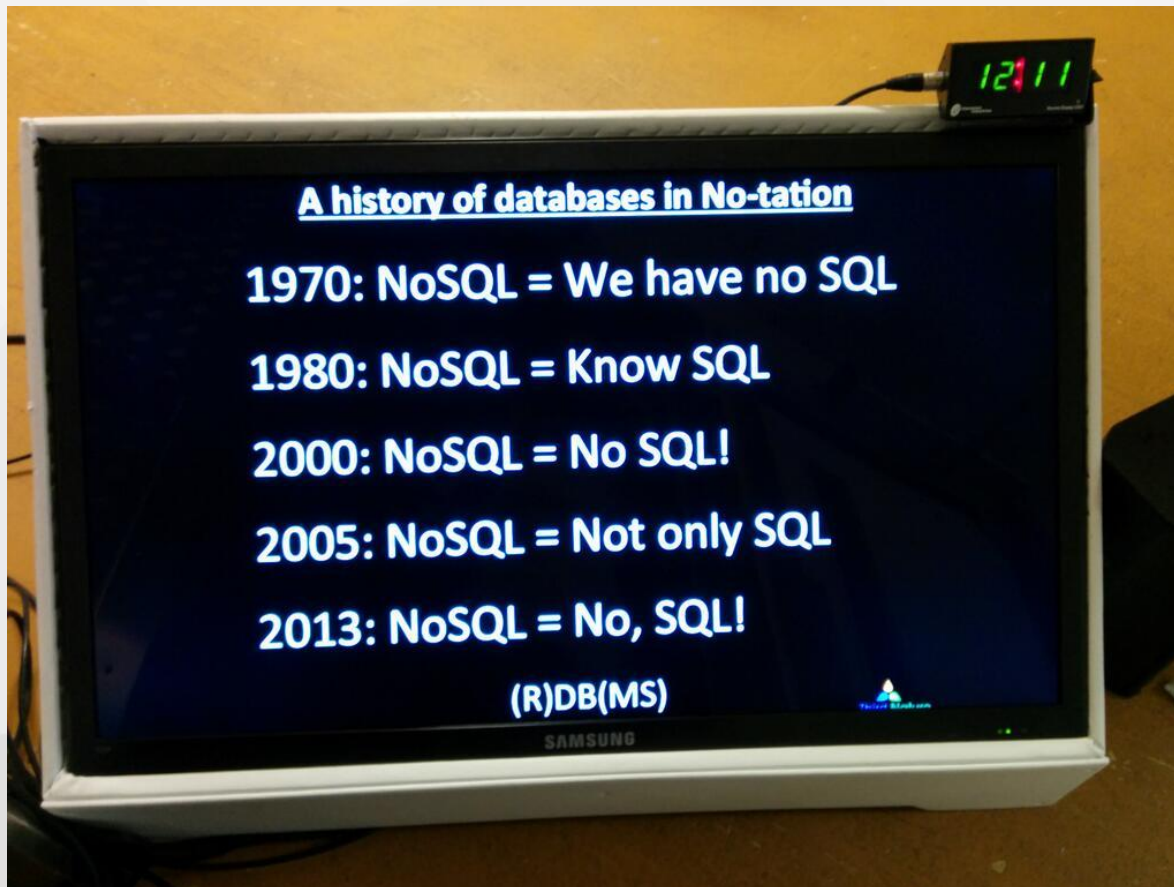
318 systems in ranking, February 2017

Rank	Feb 2017	Jan 2017	Feb 2016	DBMS	Database Model	Score	Feb 2017	Jan 2017	Feb 2016
1.	1.	1.		Oracle	Relational DBMS	1403.83	-12.89	-72.31	
2.	2.	2.		MySQL	Relational DBMS	1380.30	+14.02	+59.18	
3.	3.	3.		Microsoft SQL Server	Relational DBMS	1203.45	-17.50	+53.23	
4.	5.	5.	4.	PostgreSQL	Relational DBMS	353.68	+23.31	+65.02	
5.	4.	4.	5.	MongoDB	Document store	335.50	+3.60	+29.90	
6.	6.	6.	6.	DB2	Relational DBMS	187.90	+5.41	-6.58	
7.	7.	8.	7.	Cassandra	Wide column store	134.38	-2.06	+2.62	
8.	8.	7.	8.	Microsoft Access	Relational DBMS	133.39	+5.94	+0.31	
9.	10.	9.	9.	SQLite	Relational DBMS	115.31	+2.93	+8.53	
10.	9.	10.	10.	Redis	Key-value store	114.03	-4.66	+11.96	
11.	11.	12.	11.	Elasticsearch	Search engine	108.31	+2.14	+30.47	
12.	12.	13.	12.	Teradata	Relational DBMS	75.60	+1.43	+2.22	
13.	13.	11.	13.	SAP Adaptive Server	Relational DBMS	71.74	+2.63	-8.30	
14.	14.	14.	14.	Solr	Search engine	67.69	-0.39	-4.59	
15.	15.	16.	15.	HBase	Wide column store	59.24	+0.10	+7.22	
16.	16.	18.	16.	Splunk	Search engine	56.03	+0.54	+13.20	
17.	17.	17.	17.	FileMaker	Relational DBMS	55.19	+1.71	+8.16	
18.	18.	19.	18.	SAP HANA	Relational DBMS	52.45	+0.52	+14.37	
19.	19.	15.	19.	Hive	Relational DBMS	47.95	-3.19	-4.83	
20.	20.	23.	20.	MariaDB	Relational DBMS	45.35	+0.31	+16.57	
21.	21.	21.	21.	Neo4j	Graph DBMS	36.27	+0.00	+3.98	
22.	22.	26.	22.	Amazon DynamoDB	Document store	32.19	+1.16	+10.39	
23.	23.	24.	23.	Couchbase	Document store	31.18	+0.96	+5.79	
24.	24.	22.	24.	Memcached	Key-value store	30.53	+2.09	+1.60	
25.	25.	20.	25.	Informix	Relational DBMS	27.25	+0.82	-5.76	



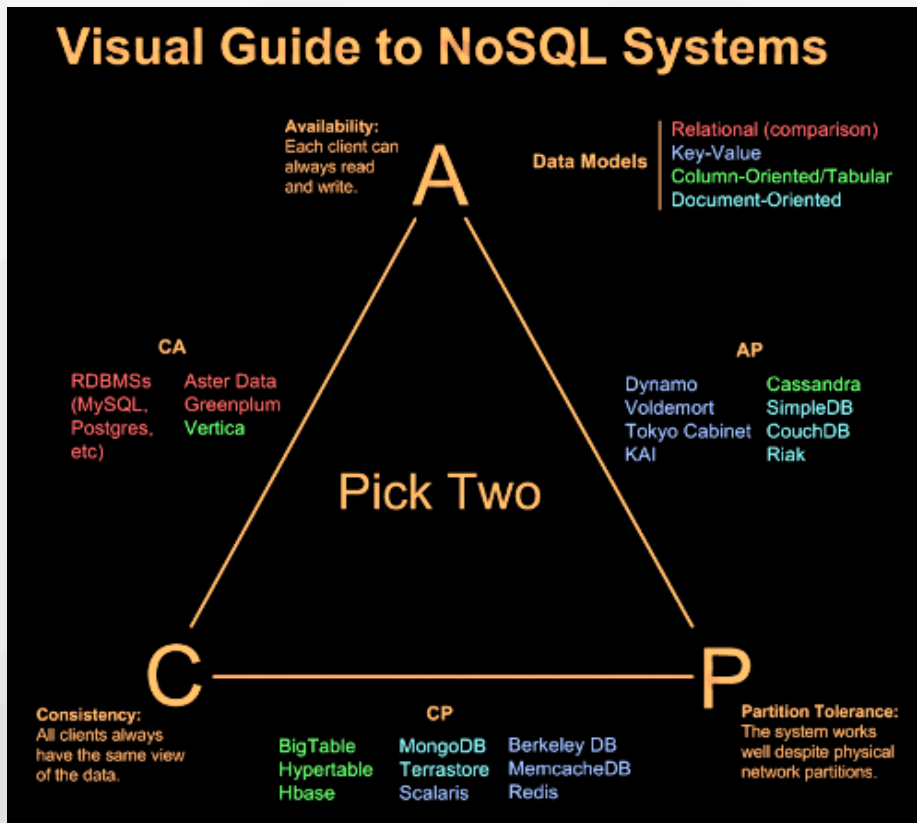


The History of SQL





CAP Theorem



Cassandra Write Data Flows

Single Region, Multiple Availability Zone

1. Client Writes to any Cassandra Node
2. Coordinator Node replicates to nodes and Zones
3. Nodes return ack to coordinator
4. Coordinator returns ack to client
5. Data written to internal commit log disk



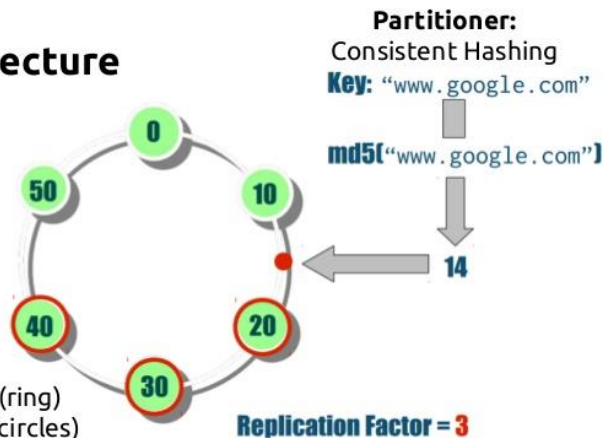
If a node goes offline, hinted handoff completes the write when the node comes back up.

Requests can choose to wait for one node, a quorum, or all nodes to ack the write

SSTable disk writes and compactions occur asynchronously



Architecture



- Cluster (ring)
- Nodes (circles)
- Peer-to-Peer Model
- Gossip Protocol



mongoDB Official Website : <https://www.mongodb.org/>
The latest stable Release : v3.0.4

```
{
  name: "sue",
  age: 26,
  status: "A",
  groups: [ "news", "sports" ]
}
```

← field: value
← field: value
← field: value
← field: value

```
db.users.insert (
  {
    name: "sue",
    age: 26,
    status: "A"
  }
)
```

← collection
← field: value
← field: value
← field: value } document

```
db.users.update(
  { age: { $gt: 18 } },
  { $set: { status: "A" } },
  { multi: true }
)
```

← collection
← update criteria
← update action
← update option

```
db.users.remove(
  { status: "D" }
)
```

← collection
← remove criteria

Collection Query Criteria Modifier

db.users.find({ age: { \$gt: 18 } }).sort({age: 1 })

{ age: 18, ... }
{ age: 28, ... }
{ age: 21, ... }
{ age: 38, ... }
{ age: 18, ... }
{ age: 38, ... }
{ age: 31, ... }

users

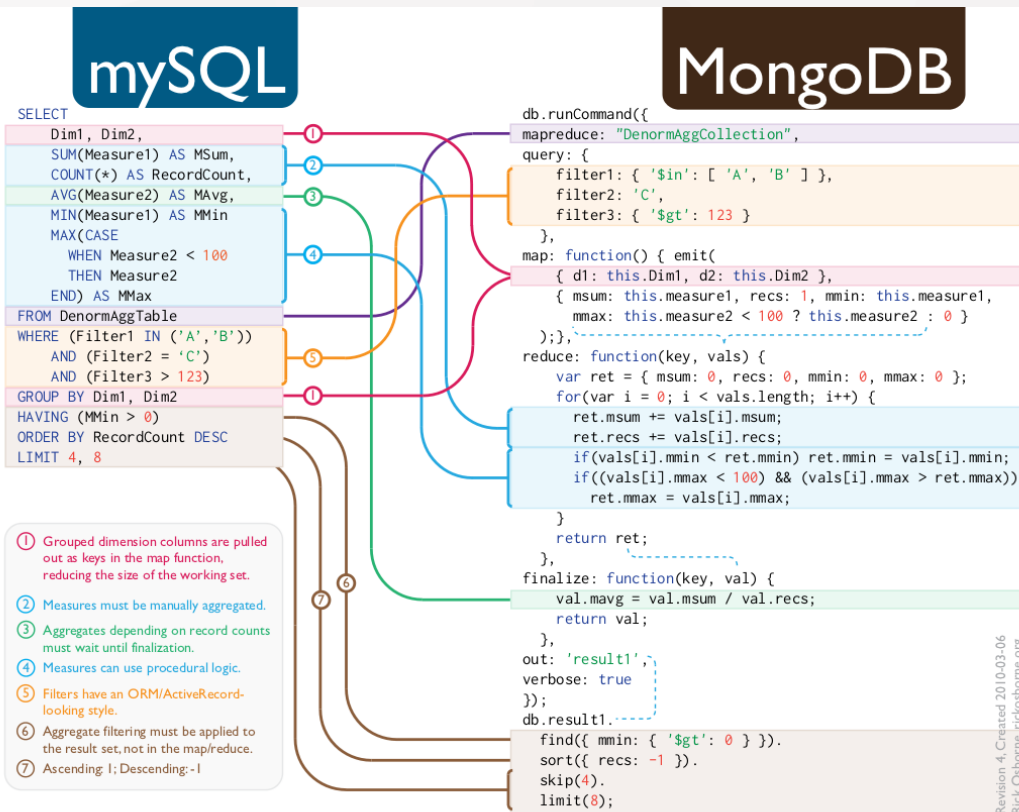
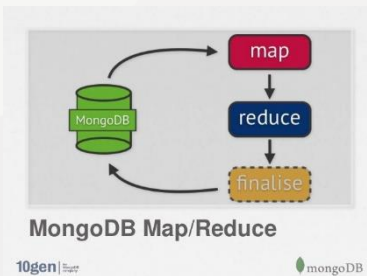
Query Criteria

{ age: 28, ... }
{ age: 21, ... }
{ age: 38, ... }
{ age: 38, ... }
{ age: 31, ... }

Modifier

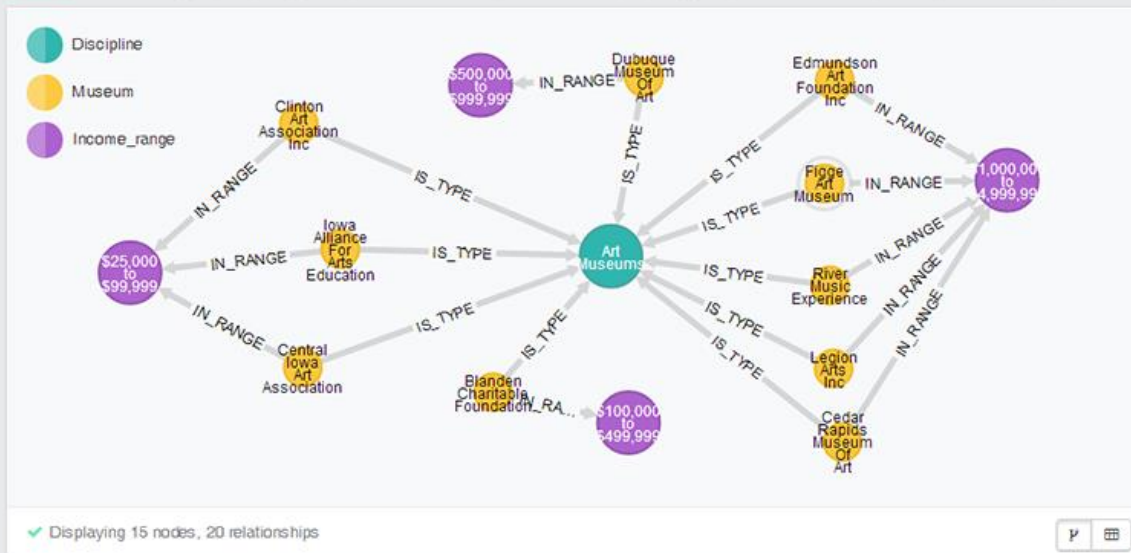
{ age: 21, ... }
{ age: 28, ... }
{ age: 31, ... }
{ age: 38, ... }
{ age: 38, ... }

Results



```
1 // Iowa art museums with income ranges greater than $10,000
2 MATCH (inc_rng)<-[:IN_RANGE]-(lam:Museum {state: "IA"})-[:IS_TYPE]->
  (disc:Discipline {code: "ART"})
3 WHERE toInt(inc_rng.code) > 1
4 RETURN lam, inc_rng, disc
```

CYPRER MATCH (inc_rng)<-[:IN_RANGE]-(lam:Museum {state: "IA"})-[:IS_TYPE]->(disc:Discipline {code: "ART"}) WHERE



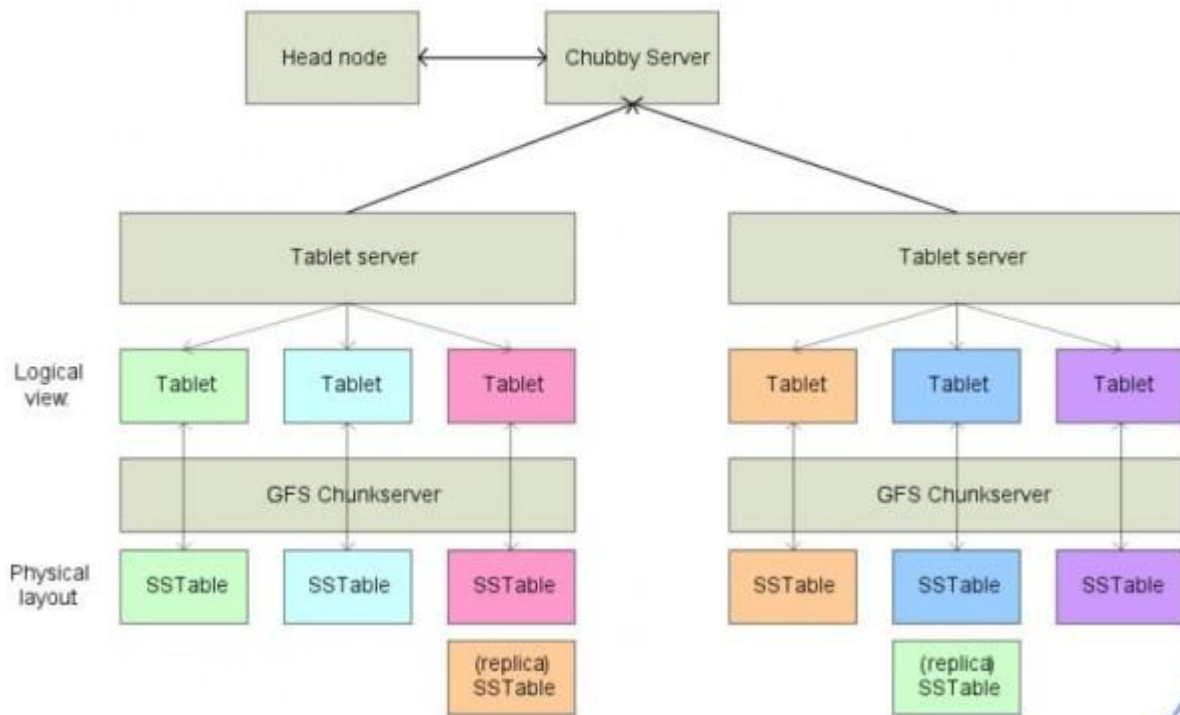


PART2



Hbase Origin

Bigtable Architecture





Java SE Public Updates			
Major Release	GA Date	End of Public Updates Notification	End of Public Updates
5.0	May 2004	Apr 2008	Oct 2009
6	Dec 2006	Feb 2011	Feb 2013
7	Jul 2011	Mar 2014	Apr 2015
8	Mar 2014	TBD	Sep 2017*

* or later, depending on factors described above.

<http://www.oracle.com/technetwork/java/eol-135779.html>

HBase Versions

Hadoop version support matrix

- "S" = supported
- "X" = not supported
- "NT" = Not tested

	HBase-0.94.x	HBase-0.98.x (Support for Hadoop 1.1+ is deprecated.)	HBase-1.0.x (Hadoop 1.x is NOT supported)	HBase-1.1.x	HBase-1.2.x	HBase-1.3.x	HBase-2.0.x
Hadoop-1.0.x	X	X	X	X	X	X	X
Hadoop-1.1.x	S	NT	X	X	X	X	X
Hadoop-0.23.x	S	X	X	X	X	X	X
Hadoop-2.0.x-alpha	NT	X	X	X	X	X	X
Hadoop-2.1.0-beta	NT	X	X	X	X	X	X
Hadoop-2.2.0	NT	S	NT	NT	X	X	X
Hadoop-2.3.x	NT	S	NT	NT	X	X	X
Hadoop-2.4.x	NT	S	S	S	S	S	X
Hadoop-2.5.x	NT	S	S	S	S	S	X
Hadoop-2.6.0	X	X	X	X	X	X	X
Hadoop-2.6.1+	NT	NT	NT	NT	S	S	S
Hadoop-2.7.0	X	X	X	X	X	X	X
Hadoop-2.7.1+	NT	NT	NT	NT	S	S	S

§ 4. Basic Prerequisites

This section lists required services and some required system configuration.

Table 2. Java

HBase Version	JDK 6	JDK 7	JDK 8
2.0	Not Supported	Not Supported	yes
1.3	Not Supported	yes	yes
1.2	Not Supported	yes	yes
1.1	Not Supported	yes	Running with JDK 8 will work but is not well tested.
1.0	Not Supported	yes	Running with JDK 8 will work but is not well tested.
0.98	yes	yes	Running with JDK 8 works but is not well tested. Building with JDK 8 would require removal of the deprecated <code>remove()</code> method of the <code>PoolMap</code> class and is under consideration. See HBASE-7608 for more information about JDK 8 support.
0.94	yes	yes	N/A



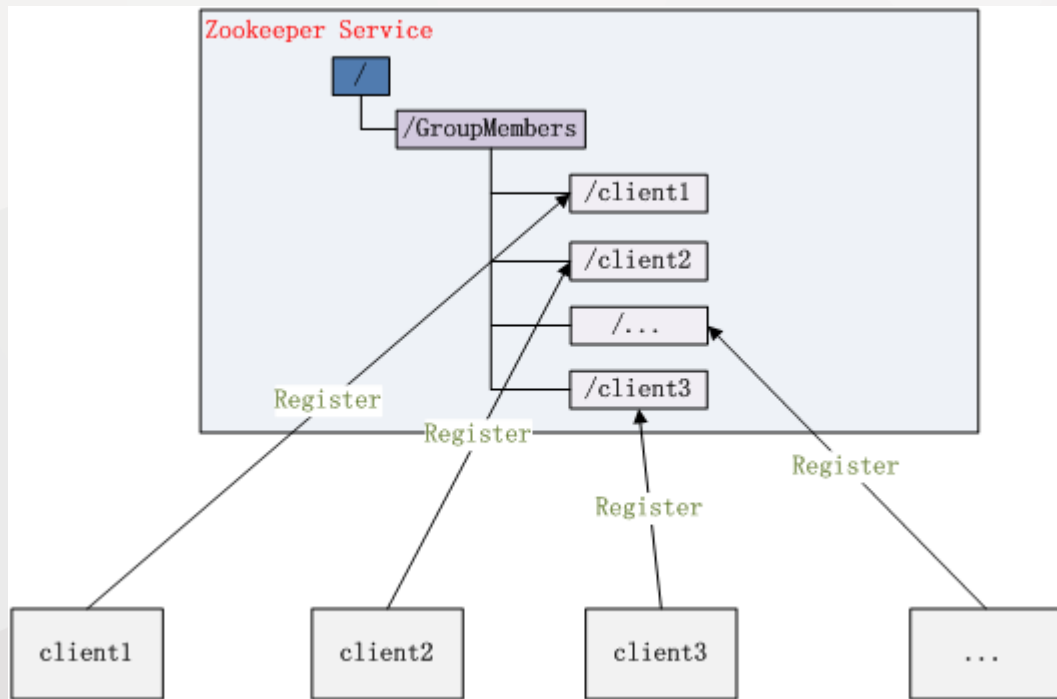
In HBase 0.98.5 and newer, you must set `JAVA_HOME` on each node of your cluster. `hbase-env.sh` provides a handy mechanism to do this.

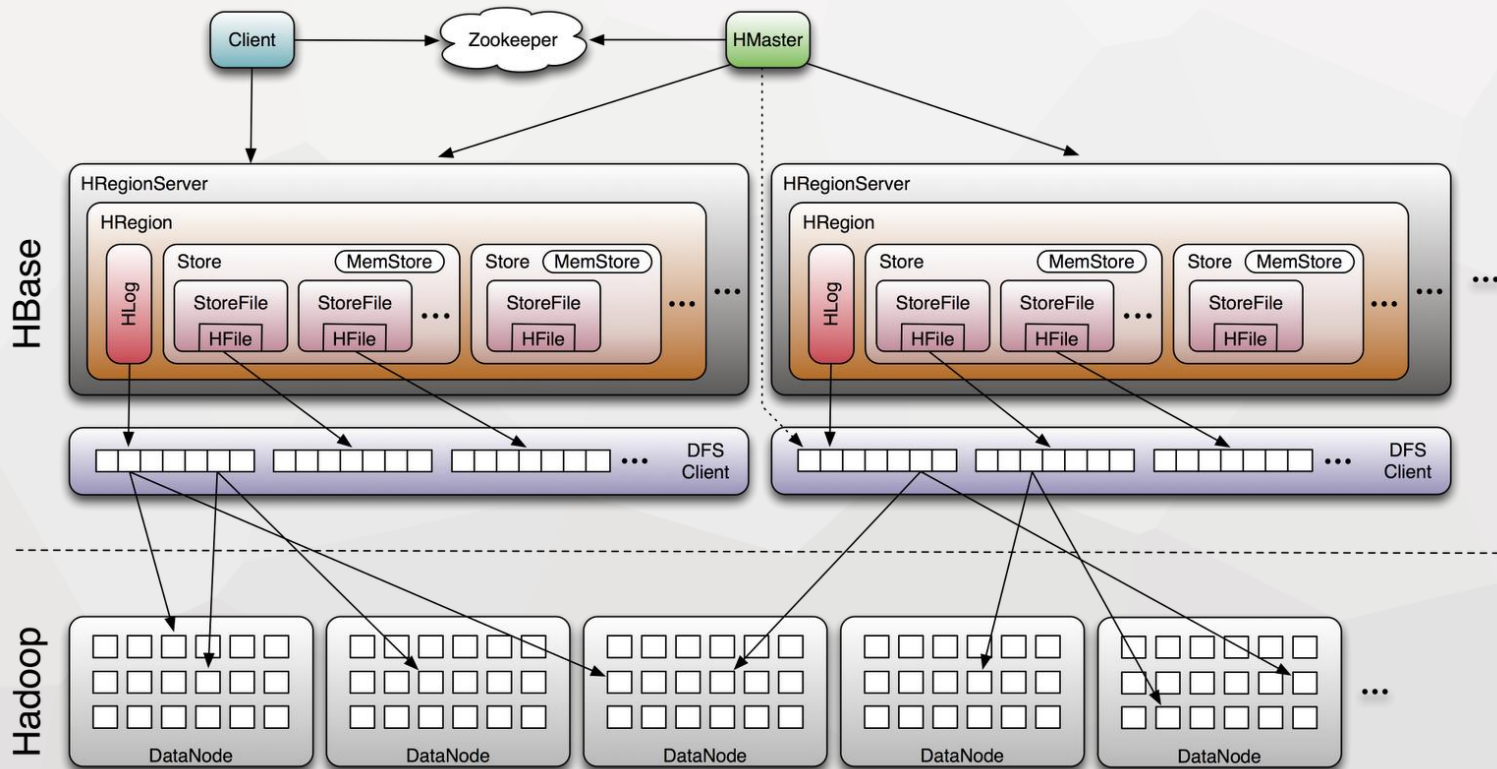


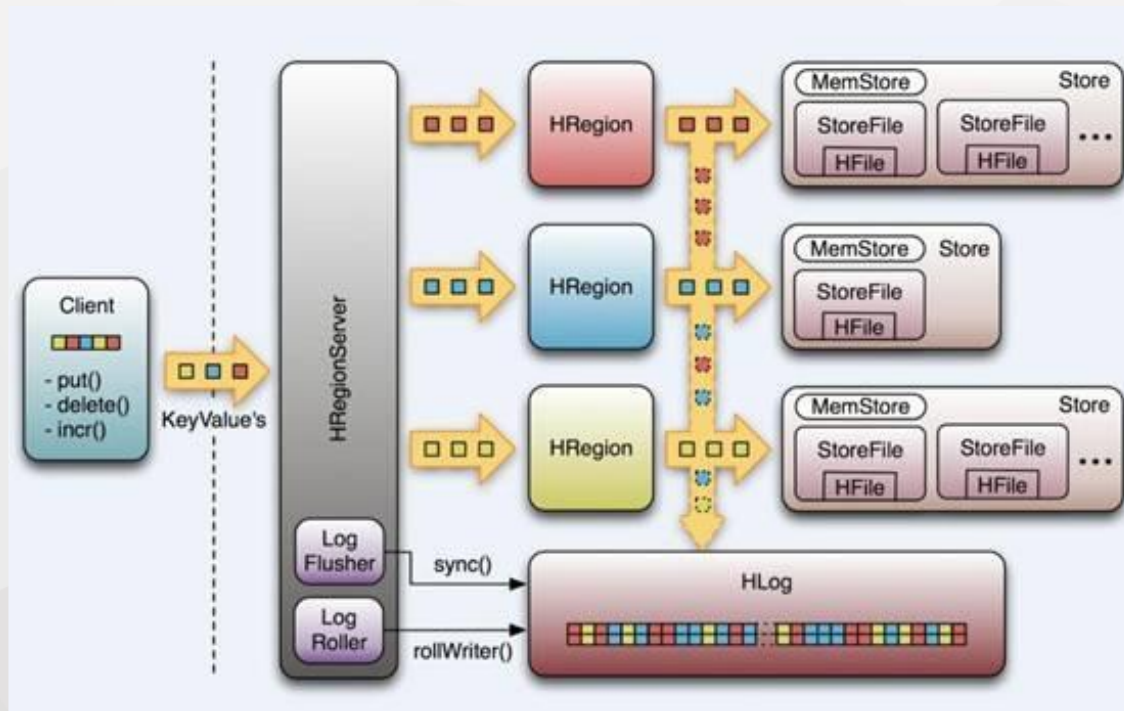
PART3

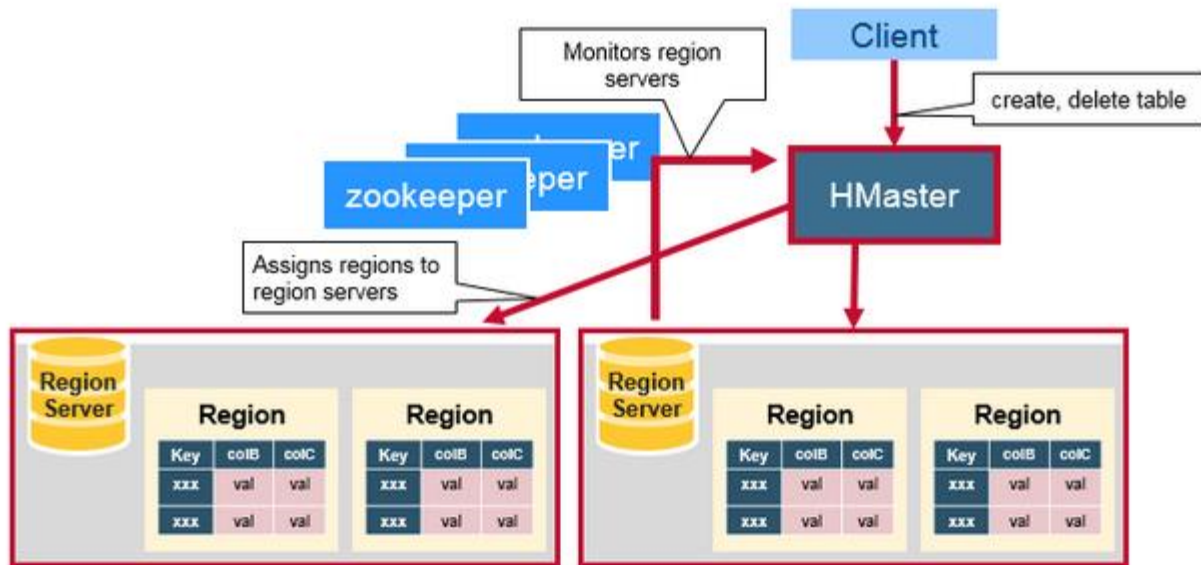


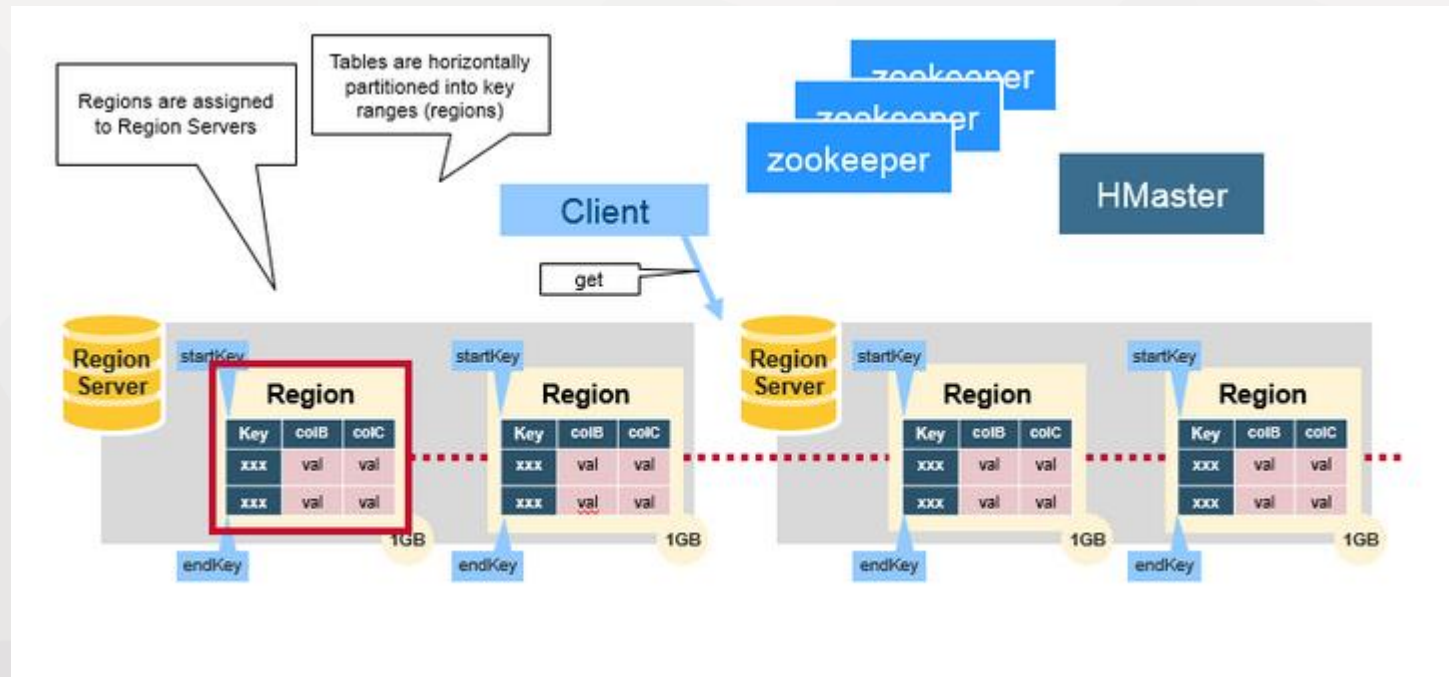
architecture













PART4

HBase Environment



PERFORMANCE

HBASE

[Home](#)[Local Logs](#)[Log Level](#)[Debug Dump](#)[Metrics Dump](#)[HBase Configuration](#)

Server Metrics

[Base Stats](#)[Memory](#)[Requests](#)[hlogs](#)[Storefiles](#)[Queues](#)[Block Cache](#)

Requests Per Second	Num. Regions	Block locality	Slow HLog Append Count
5404	4	100	0

Tasks

[Show All Monitored Tasks](#)[Show non-RPC Tasks](#)[Show All RPC Handler Tasks](#)[Show Active RPC Calls](#)[Show Client Operations](#)[View as JSON](#)

No tasks currently running on this node.

Block Cache

[Base Info](#)[Config](#)[Stats](#)[L1](#)[L2](#)

Attribute	Value	Description
Cache DATA on Read	true	True if DATA blocks are cached on read (INDEX & BLOOM blocks are always cached)
Cache DATA on Write	false	True if DATA blocks are cached on write.
Cache INDEX on Write	false	True if INDEX blocks are cached on write
Cache BLOOM on Write	false	True if BLOOM blocks are cached on write
Evict blocks on Close	false	True if blocks are evicted from cache when an HFile reader is closed
Compress blocks	false	True if blocks are compressed in cache
Prefetch on Open	false	True if blocks are prefetched into cache on open

Regions



PART5



Quick Start

```
hbase(main):001:0> help
HBase Shell, version 0.91.0-SNAPSHOT, r1130916, Sat Jul 23 12:44:34 CEST 2011
Type 'help "COMMAND"', (e.g. 'help "get"' -- the quotes are necessary) for
help on a specific command. Commands are grouped. Type 'help "COMMAND_GROUP"',
(e.g. 'help "general"') for help on a command group.
```

COMMAND GROUPS:

Group name: general

Commands: status, version

Group name: ddl

Commands: alter, create, describe, disable, drop, enable, exists,
is_disabled, is_enabled, list

```
hbase(main):0xx:0>status  
hbase(main):0xx:0>version  
hbase(main):0xx:0>create 'member','member_id','address','info'  
hbase(main):0xx:0>list  
hbase(main):0xx:0>describe 'member'
```

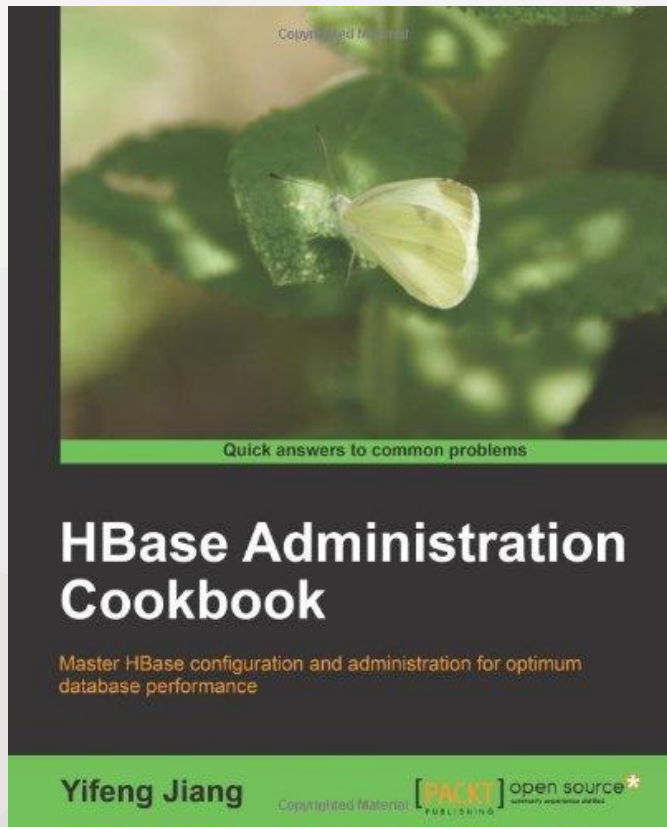
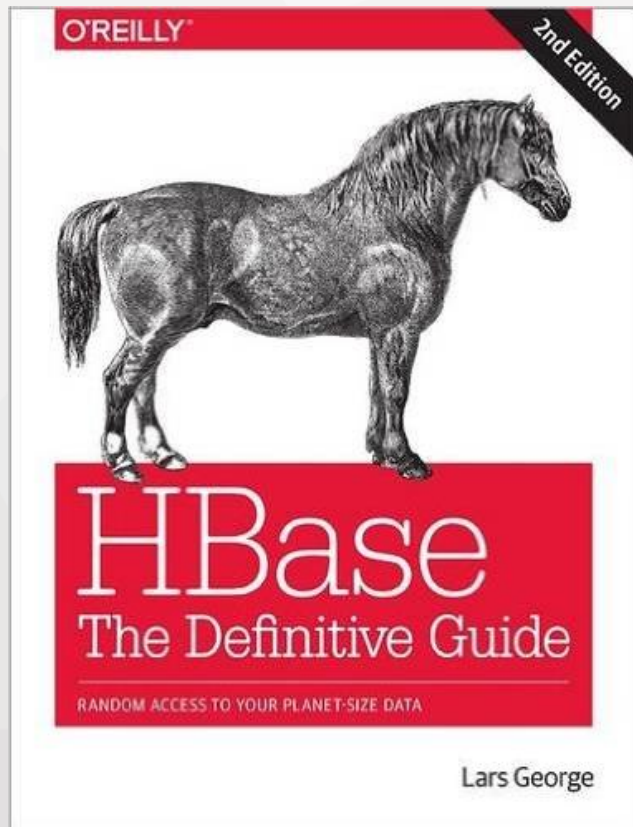
<https://hbase.apache.org/book.html>



PART6



Reference Books





The End