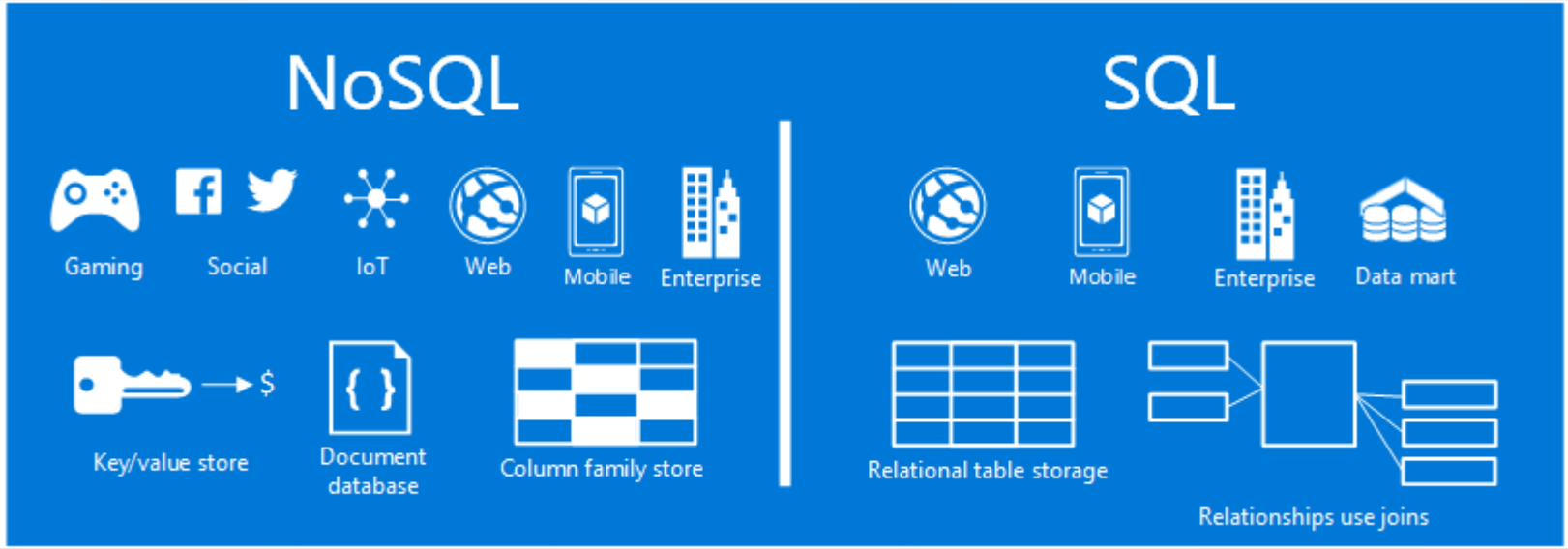


Hbase Shell

2017.2 XenRon

CONTENTS

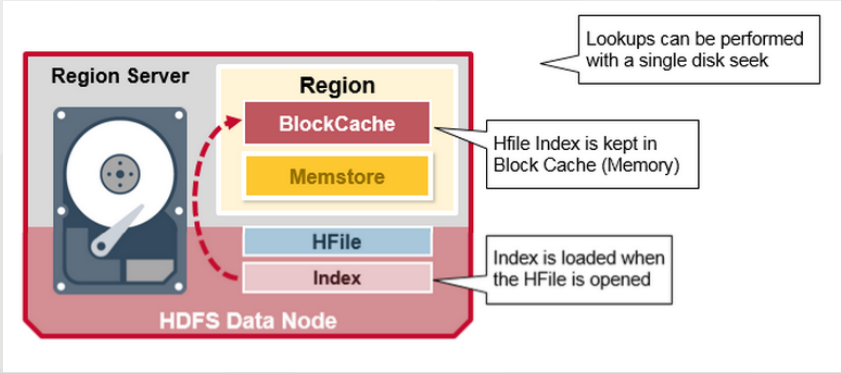
Structure 01

HDFS Storage 02

Shell 03

File Storage 04

05 Exercise



Review

Review

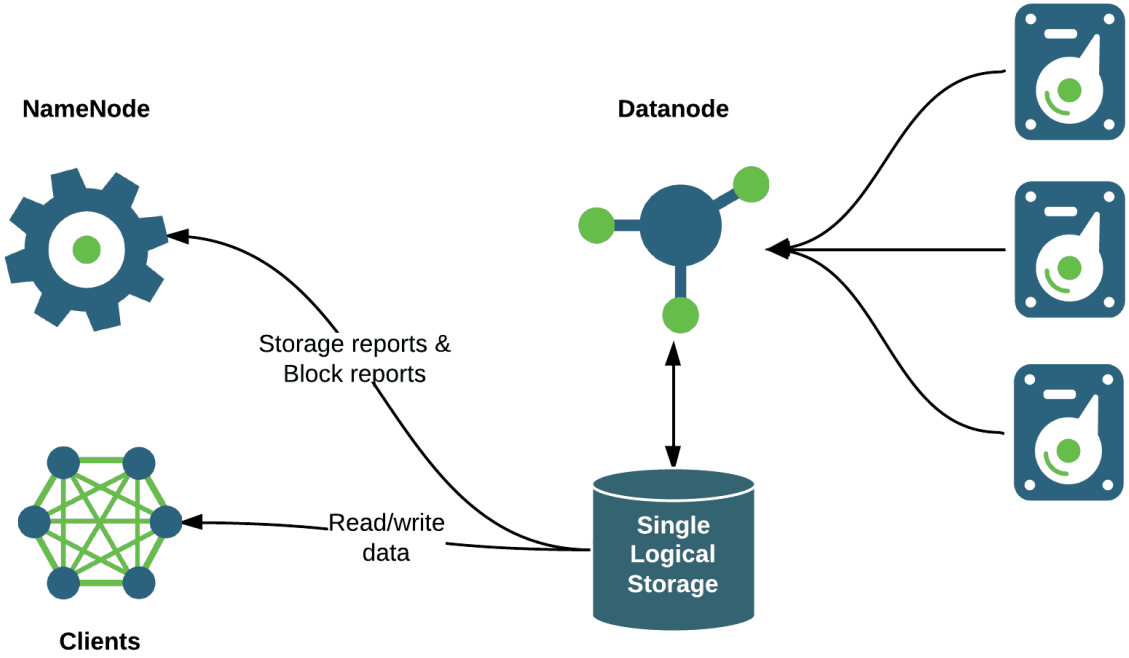
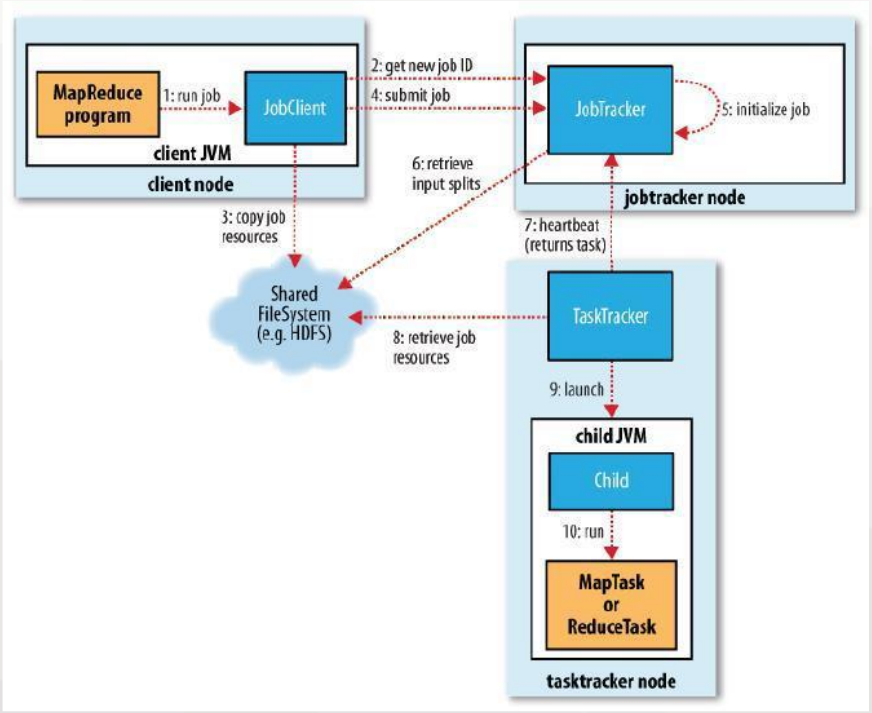


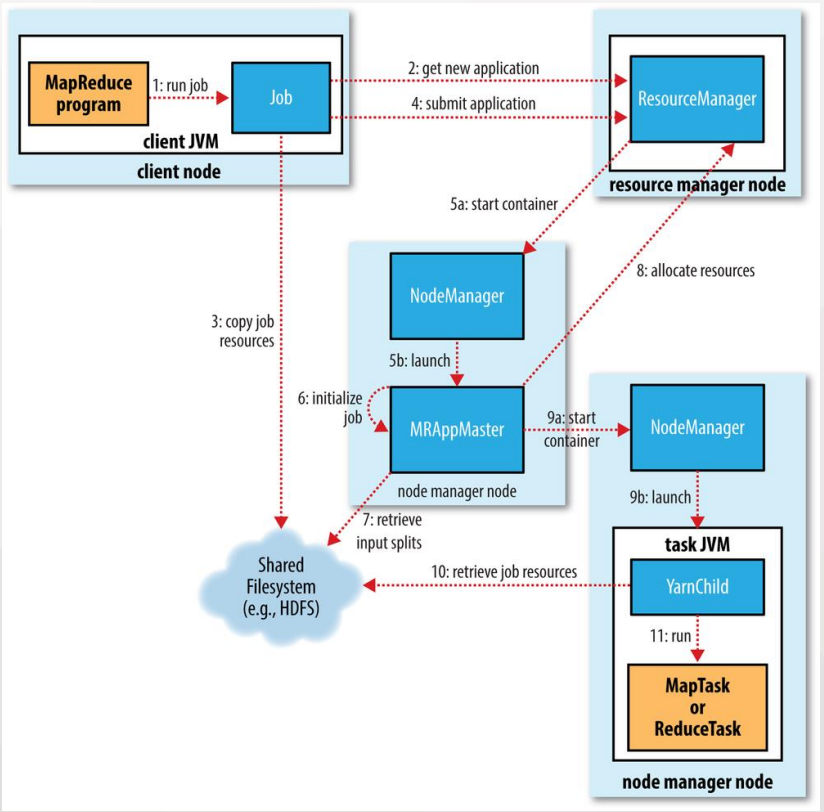
Figure 1: A DataNode presented itself as a single logical storage

Review

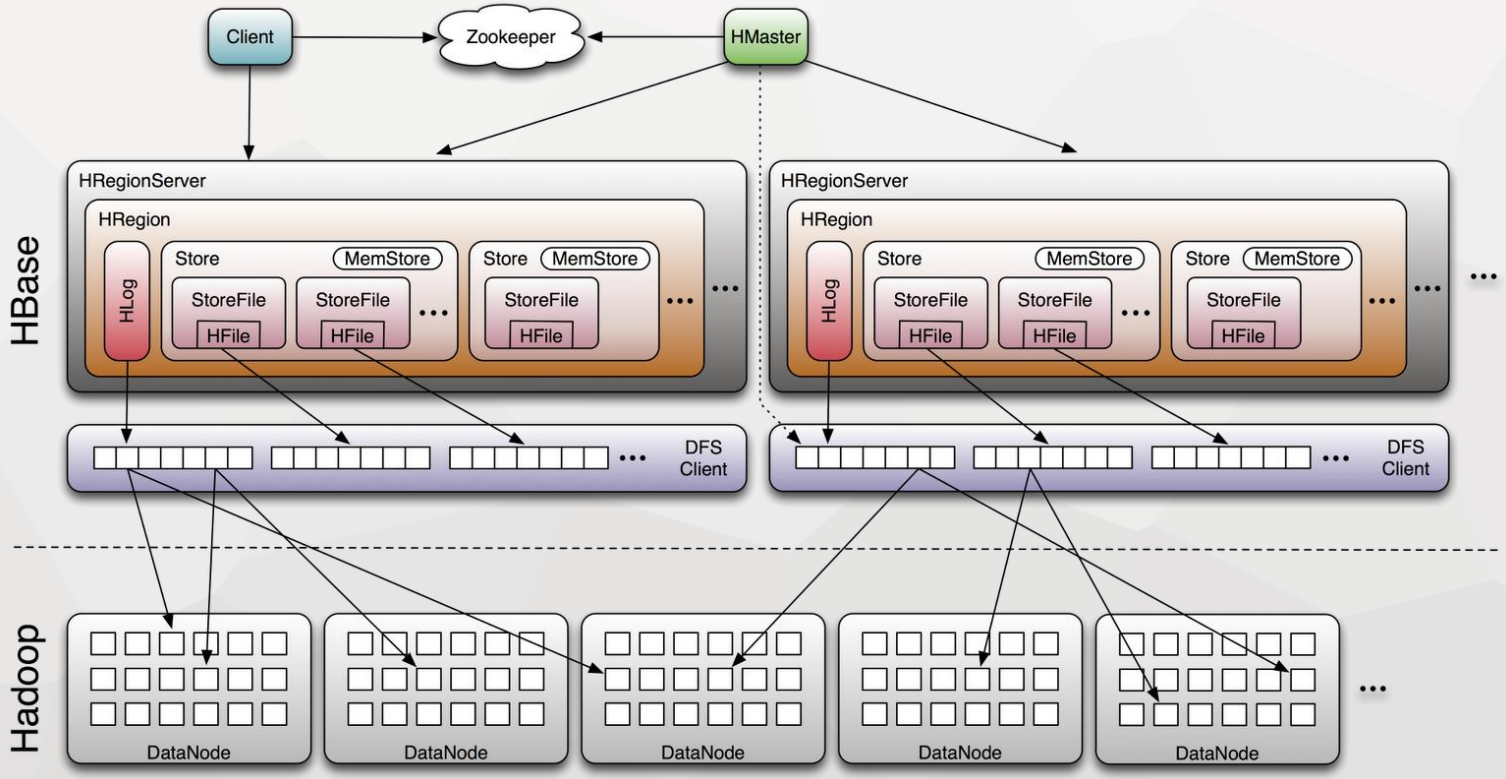
Hadoop 1.x



Hadoop 2.x

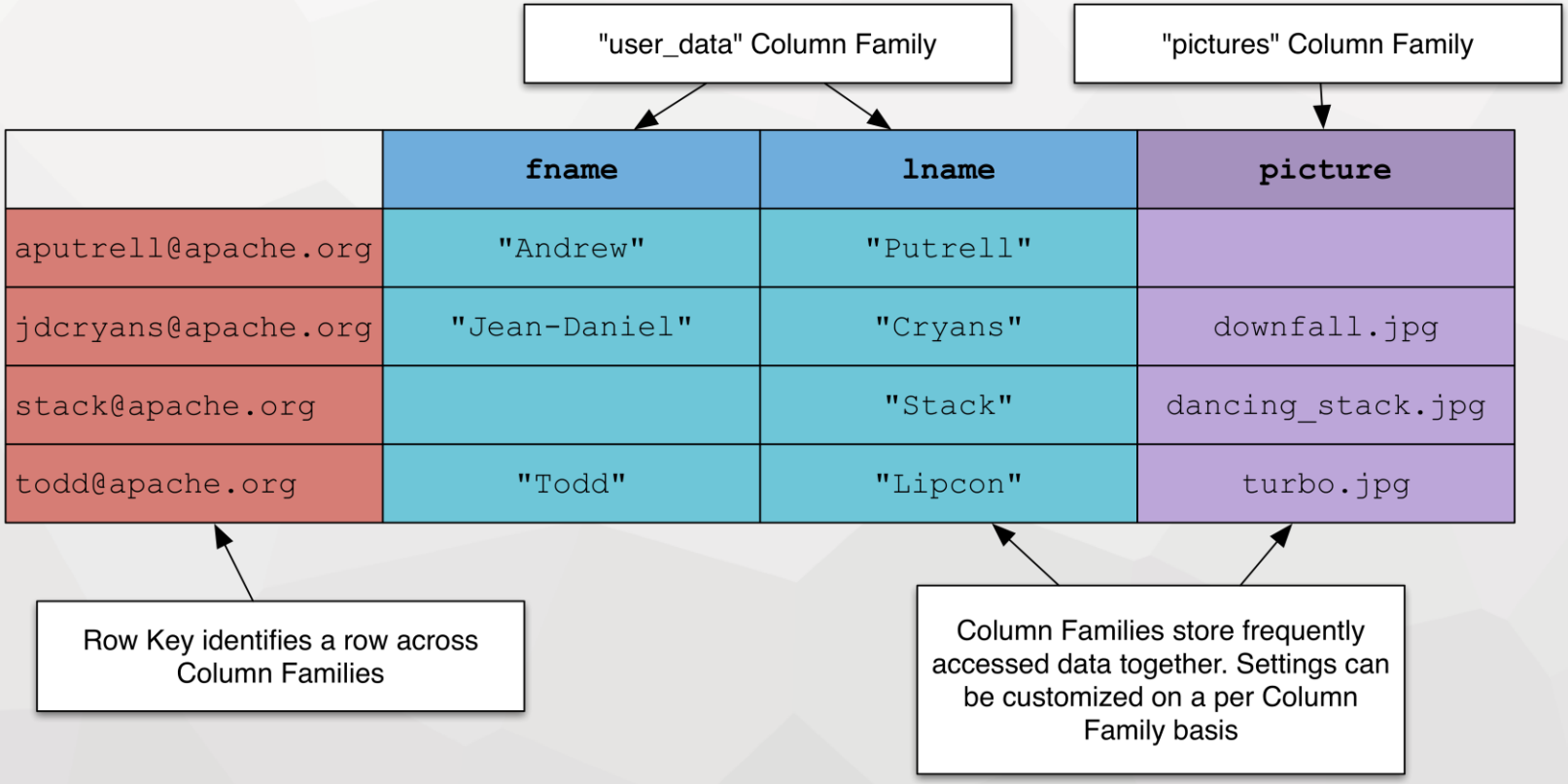


architecture

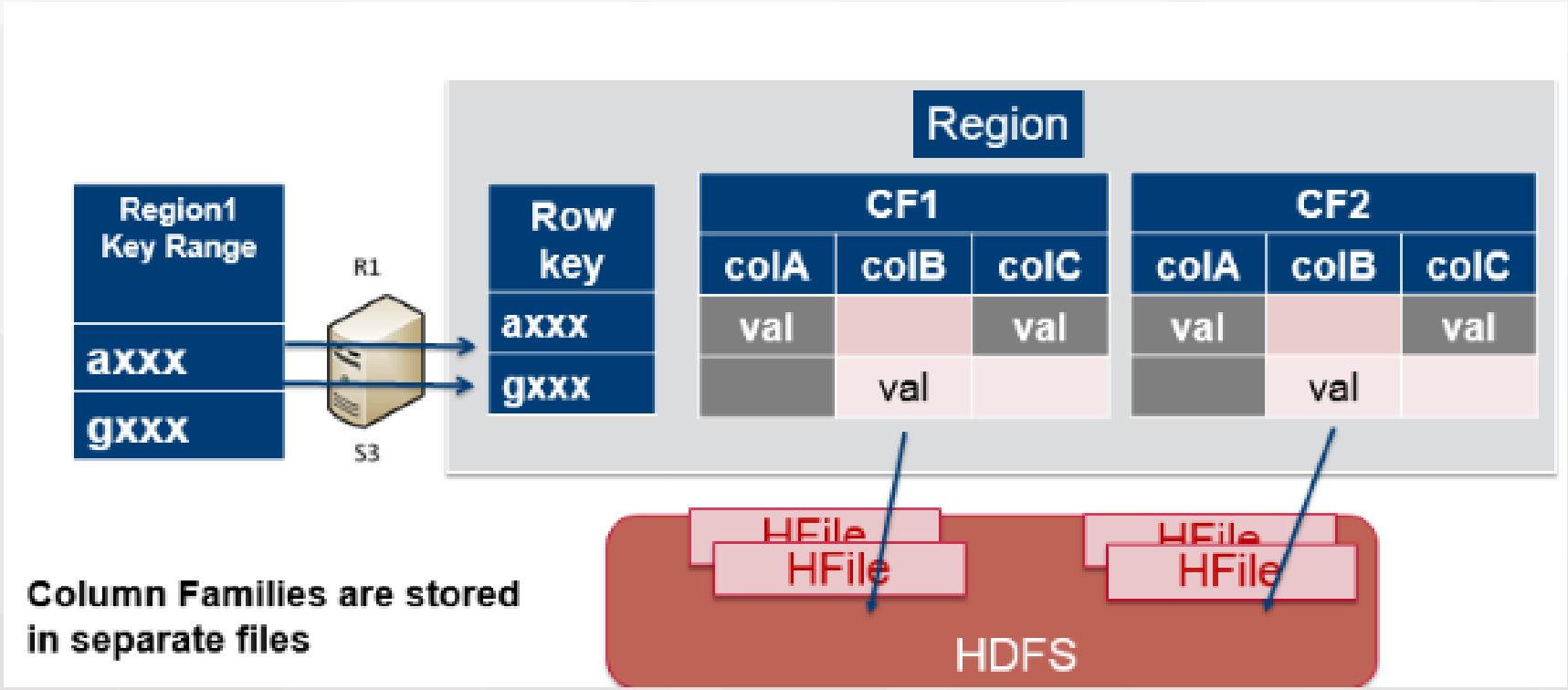


Structure

Column Family



Column Family



Column Family

Column families

Column names

Rows

	Contactinfo		Prifileinfo
Rowkey	Fname	Lname	Image
jdupont	Jean	Dupont	
jsmith	John	Smith	<smith.jpg>
mrossi	Mario	Rossi	<mario.jpg>

File

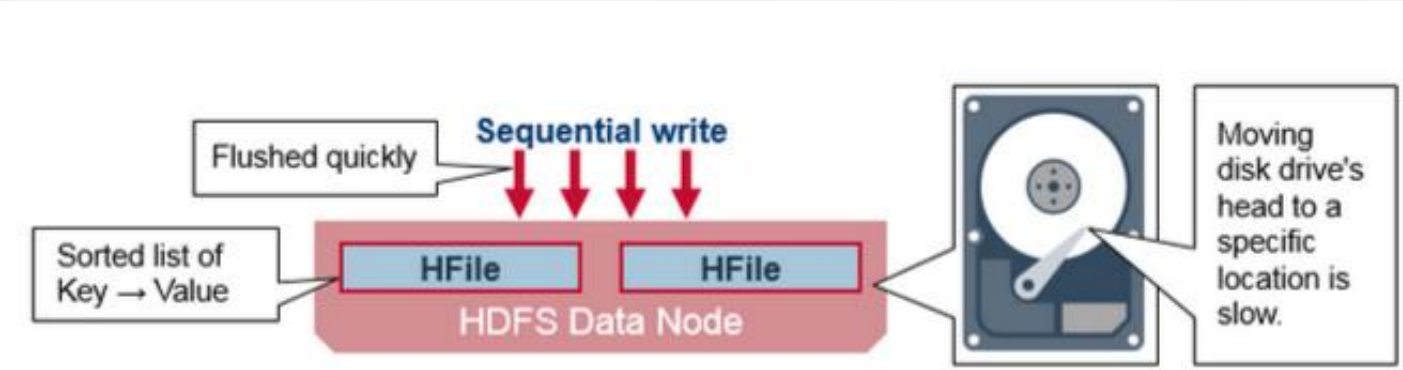
	Contactinfo	
Rowkey	Fname	Lname
jdupont	Jean	Dupont
jsmith	John	Smith
mrossi	Mario	Rossi

File

	Prifileinfo
Rowkey	Image
jdupont	
jsmith	<smith.jpg>
mrossi	<mario.jpg>

Table

Rowkey + column + timestamp -> value



Key Value

Key	CF1:Col	version	value
ra	cf1:ca	v1	1
rb	cf1:cb	v2	4
rb	cf1:cb	v1	3
rc	cf1:ca	v1	5

Key Value

Key	CF2:Col	version	value
ra	cf2:ca	v1	2
rc	cf2:ca	v2	7
rc	cf2:ca	v1	6
rc	cf2:cd	v1	8

Table

Rowkey + column + timestamp -> value

Rowkey	Column	Timestamp	Cell value
jdupont	Contactinfo:fname	1273746289103	Jean
jdupont	Contactinfo:lname	1273878447049	Dupont
jsmith	Contactinfo:fname	1273516197868	John
jsmith	Contactinfo:lname	1273871824184	Smith
mrossi	Contactinfo:fname	1273616297446	Mario
mrossi	Contactinfo:lname	1273971921442	Rossi

RDBMS VS HBase

	RDBMS	HBase
Data layout	Row- or column-oriented	Column family-oriented
Transactions	Yes	Single row only
Query language	SQL	get/put/scan
Security	Authentication/ Authorization	Access control at per-cell level, also at cluster, table, or row level
Indexes	Yes	Row key only
Max data size	TBs	PB+
Read/write throughput limits	1000s queries/second	Millions of queries/second

HDFS Storage

HBase

Browse Directory

/hbase Go!

Permission	Owner	Group	Size	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	0	0 B	.tmp
drwxr-xr-x	root	supergroup	0 B	0	0 B	MasterProcWALs
drwxr-xr-x	root	supergroup	0 B	0	0 B	WALs
drwxr-xr-x	root	supergroup	0 B	0	0 B	data
-rw-r--r--	root	supergroup	42 B	3	128 MB	hbase.id
-rw-r--r--	root	supergroup	7 B	3	128 MB	hbase.version
drwxr-xr-x	root	supergroup	0 B	0	0 B	oldWALs

/hbase/.tmp: 临时目录，当对表做创建和删除操作时，会将表move到该目录下，然后进行操作。

/hbase/WALs:RegionServer在处理数据插入和删除的过程中记录操作内容的一种日志，在0.94叫.logs

/hbase/data:核心目录，存储Hbase表的数据
默认情况下该目录下有两个目录
– Hbase/data/default:当在用户创建表的时候，没有指定namespace时，表就创建在此目录下
– Hbase/data/hbase： 系统内部创建的表，hbase:meta; namespace

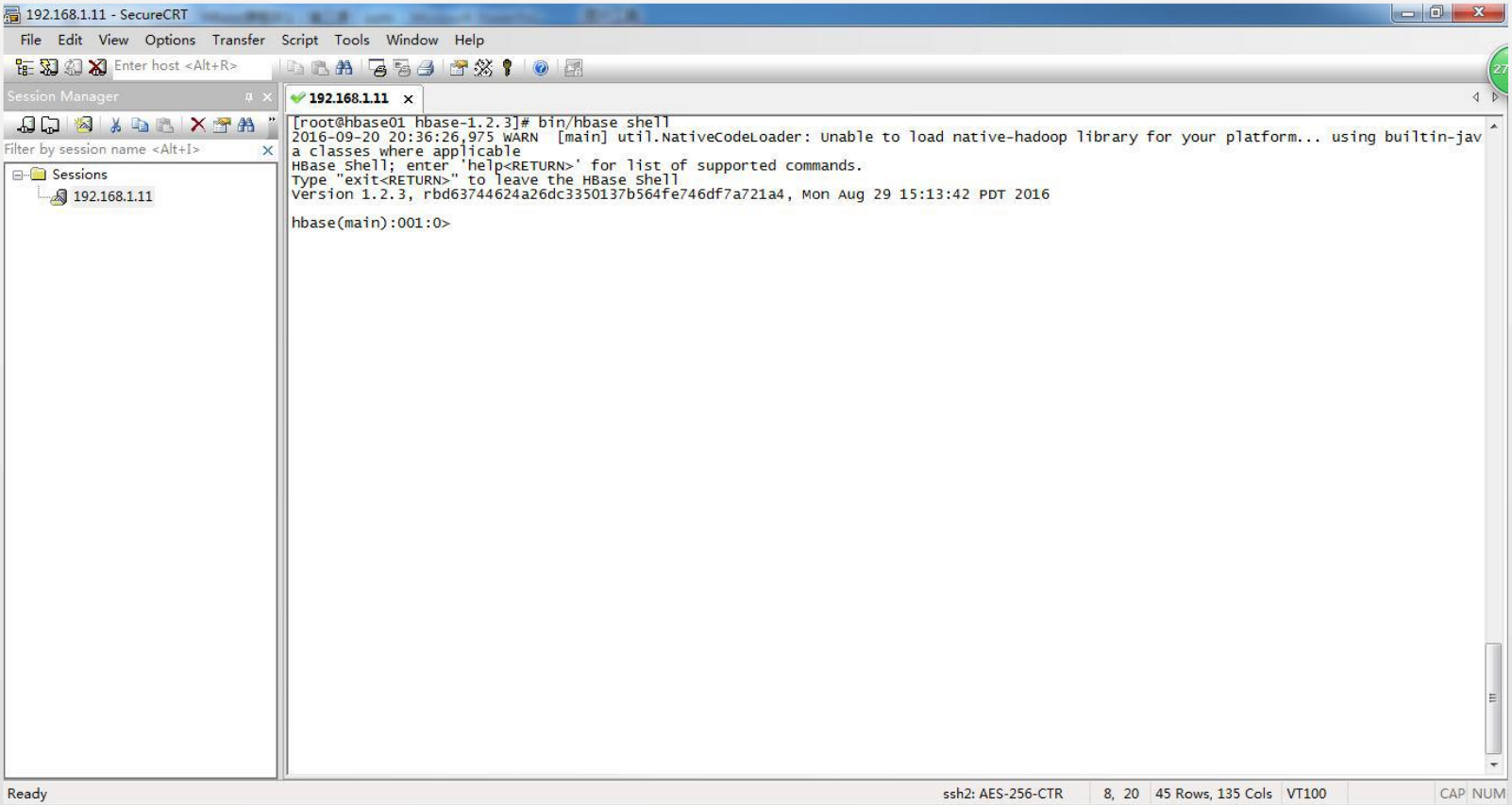
/hbase/hbase.id:存储集群唯一cluster id, (UUID)

/hbase/hbase.version:集群版本号

/hbase/oldWALs:
– 对应0.94版本中的.oldlogs目录，
– 当/hbase/WALs目录中的logs没有用之后，会将这些文件移到此目录下，Hmaster会定期进行清理

Shell

Shell


192.168.1.11 - SecureCRT
File Edit View Options Transfer Script Tools Window Help
Enter host <Alt+R>
Session Manager
Filter by session name <Alt+I>
Sessions
192.168.1.11
192.168.1.11
[root@hbase01 hbase-1.2.3]# bin/hbase shell
2016-09-20 20:36:26,975 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.3, rbd63744624a26dc3350137b564fe746df7a721a4, Mon Aug 29 15:13:42 PDT 2016
hbase(main):001:0>
Ready
ssh2: AES-256-CTR
8, 20
45 Rows, 135 Cols
VT100
CAP NUM

HBase

General HBase shell commands

status
Show cluster status. Can be ‘summary’, ‘simple’, or ‘detailed’. The default is ‘summary’.
hbase> status
hbase> status ‘simple’
hbase> status ‘summary’
hbase> status ‘detailed’

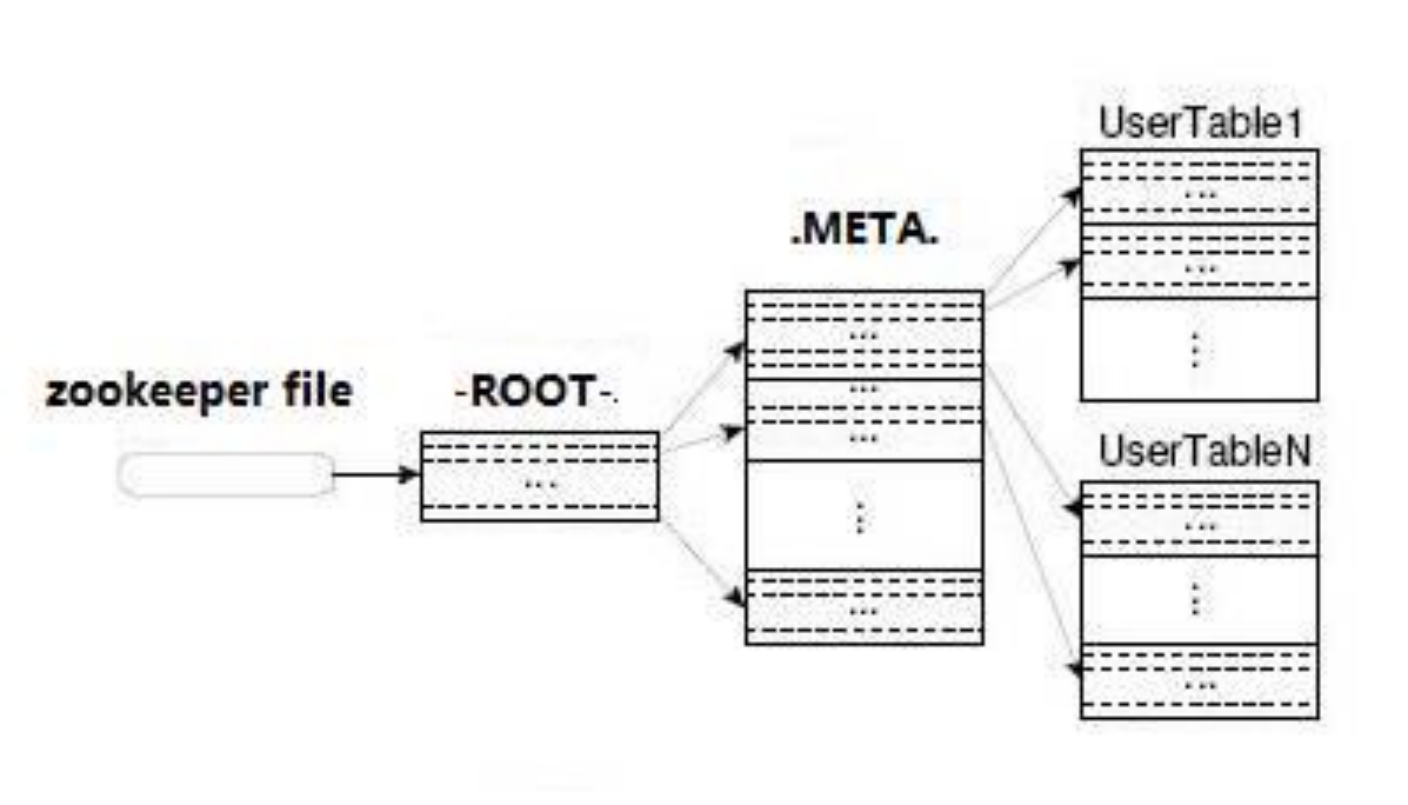
version
Output this HBase versionUsage:
hbase> version

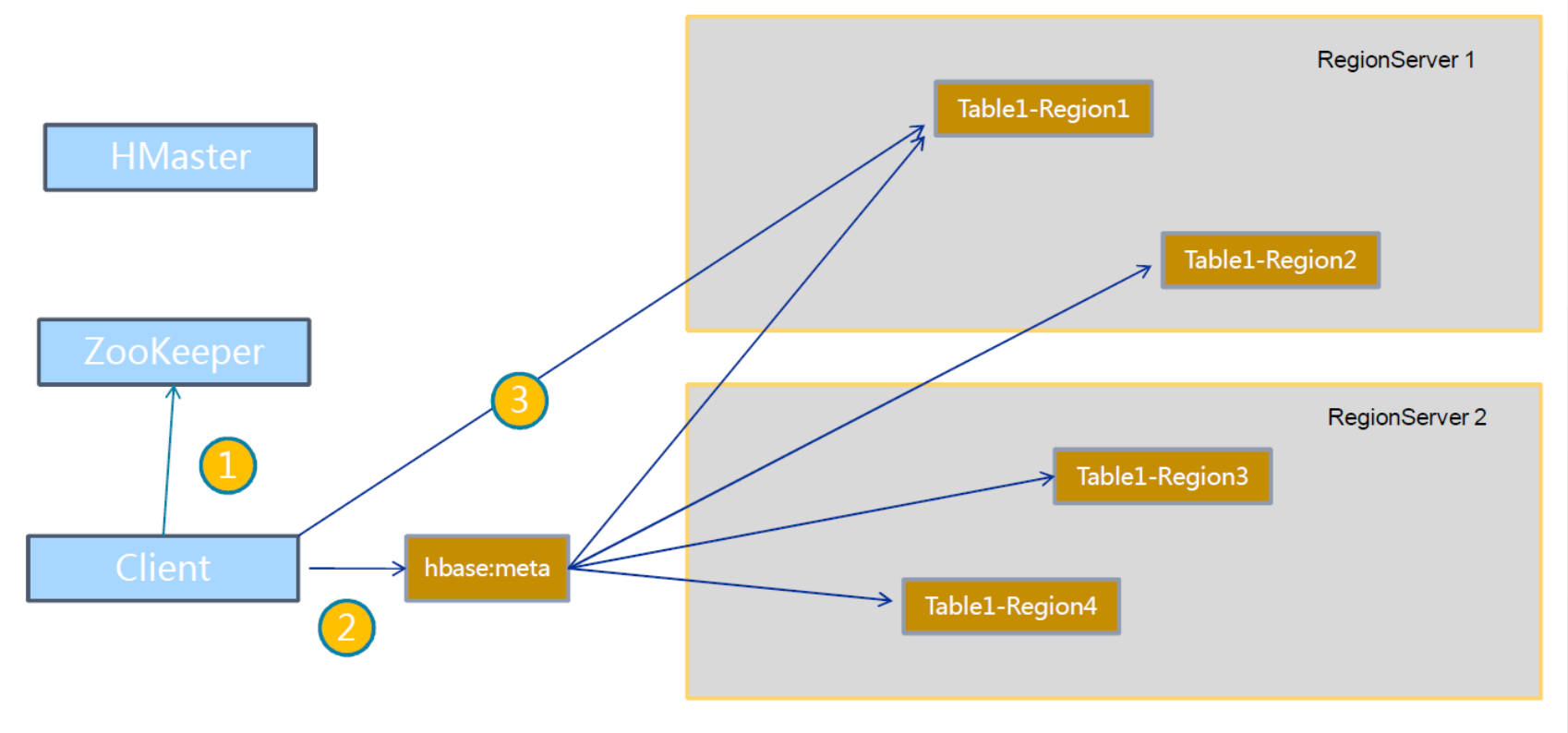
whoami
Show the current hbase user.Usage:
hbase> whoami

Data Versions

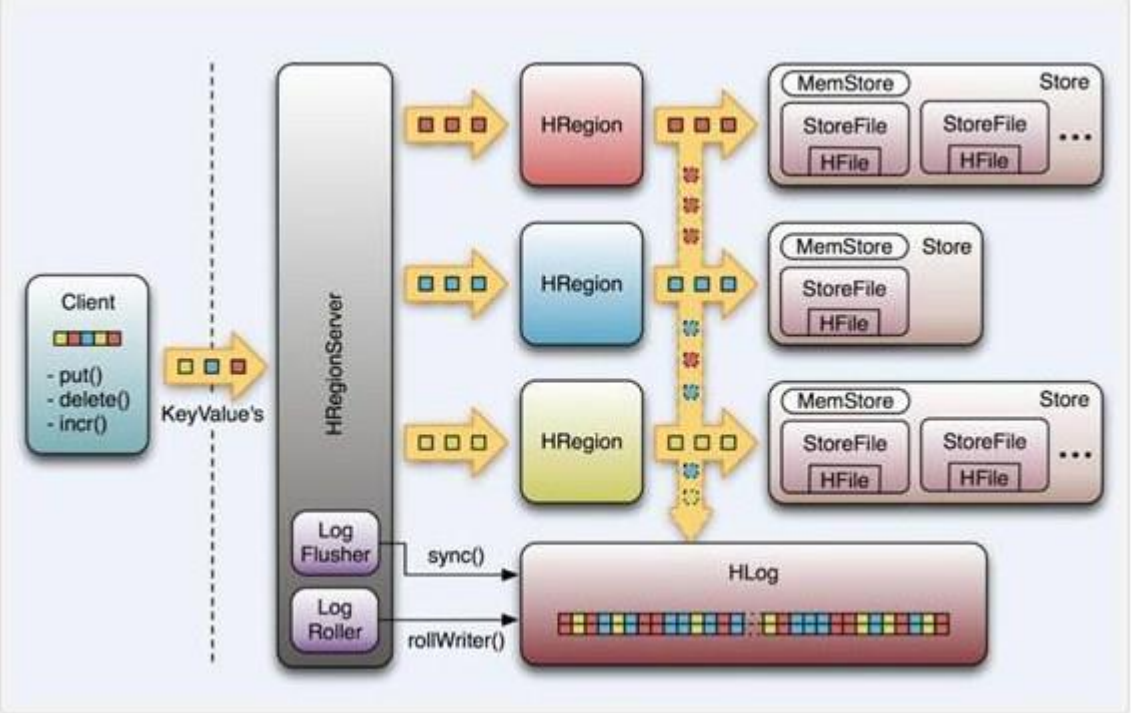
Row Key	Column	Timestamp	Cell Value
jsmith	profilephoto:image	1371851677671	<work.jpg>
jsmith	profilephoto:image	930001926438	<college.jpg>
jsmith	profilephoto:image	866929926351	<highschool.jpg>

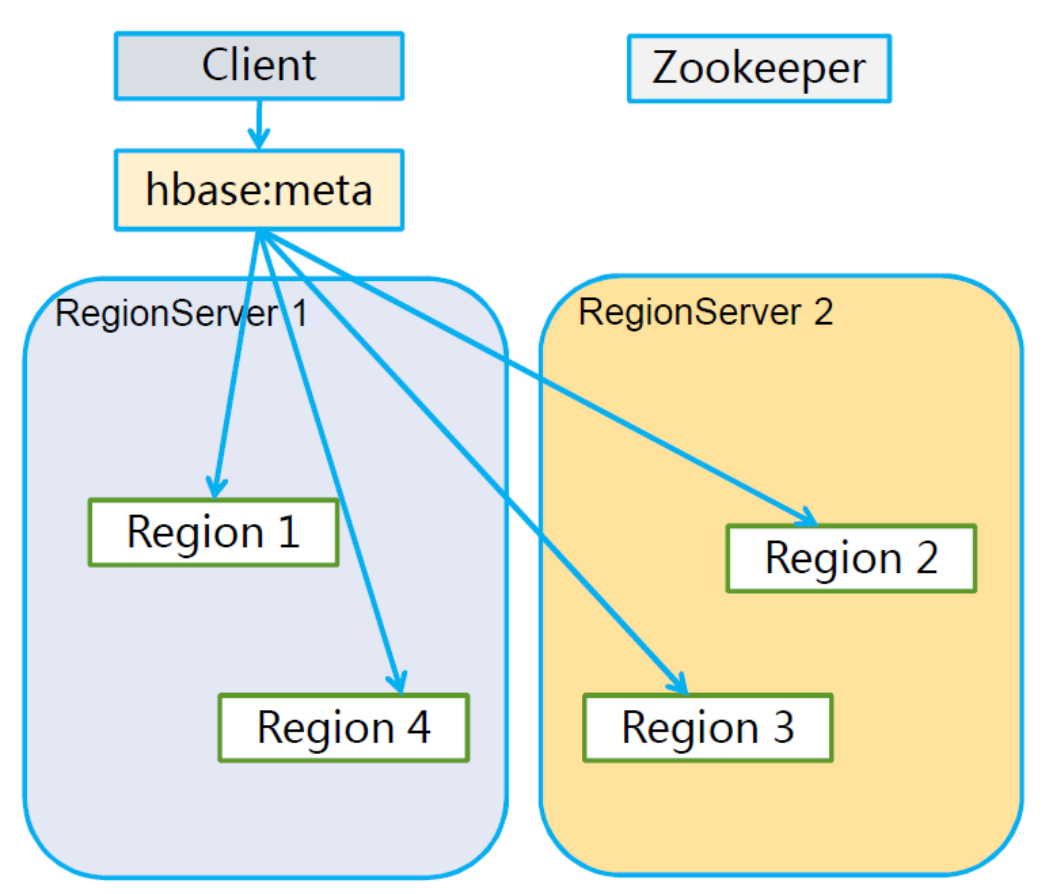
File Storage

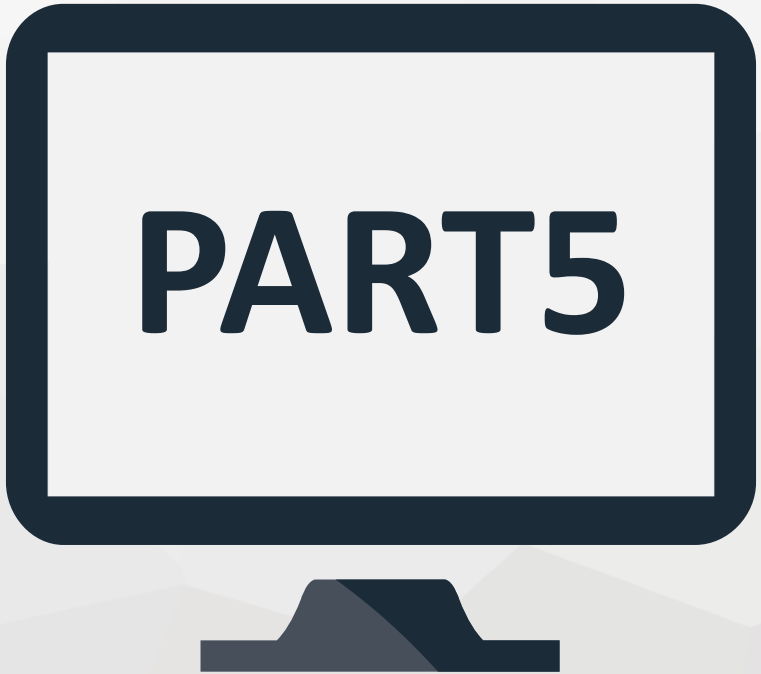

zookeeper file
-ROOT-
.META.
UserTable1
UserTableN


RegionServer 1
Table1-Region1
Table1-Region2
HMaster
ZooKeeper
3
1
RegionServer 2
Table1-Region3
Client
hbase:meta
Table1-Region4
2

HBase

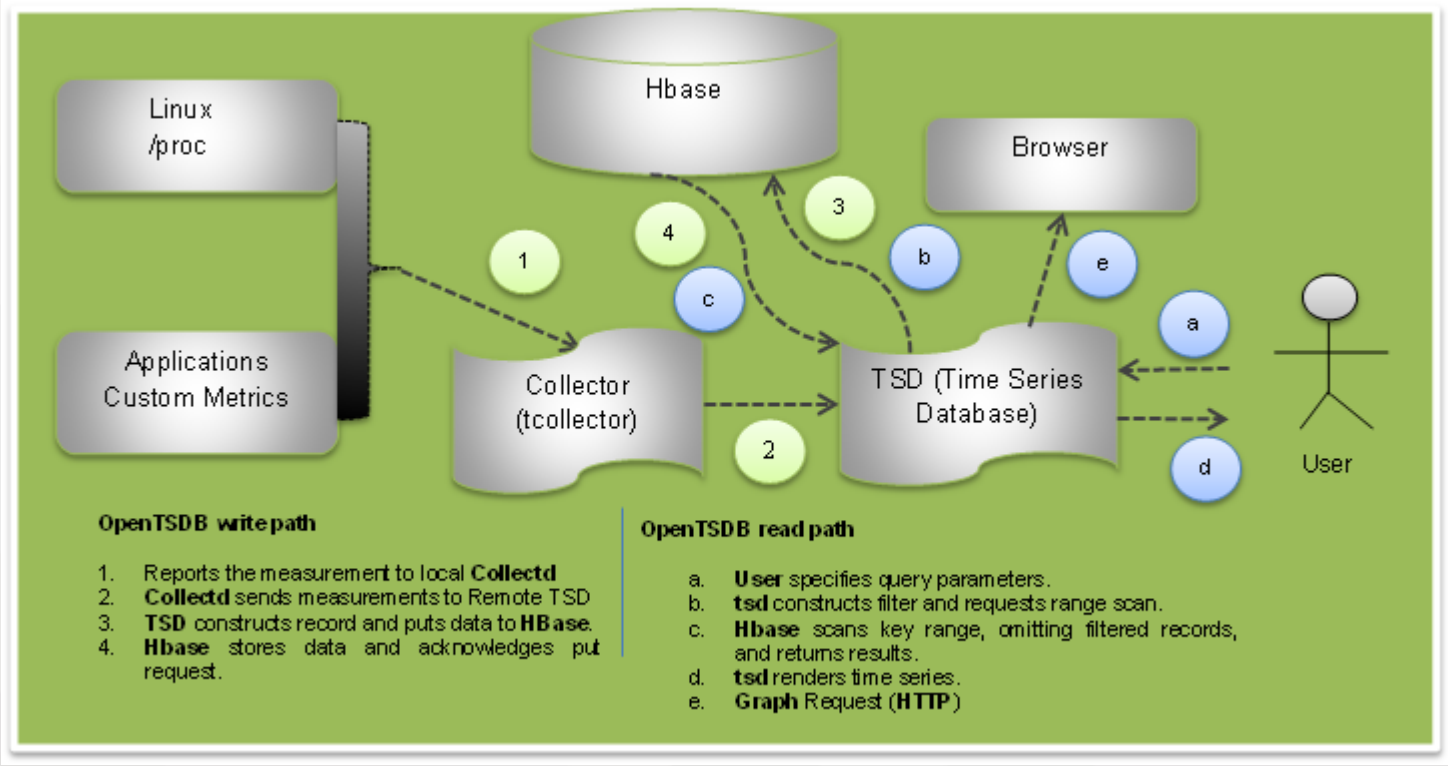

Client
- put()
- delete()
- incr()
KeyValue's
HRegionServer
Log Flusher
Log Roller
sync()
rollWriter()
HRegion
HRegion
HRegion
MemStore
Store
StoreFile
HFile
StoreFile
HFile
MemStore
Store
StoreFile
HFile
MemStore
Store
StoreFile
HFile
StoreFile
HFile
HLog


Client
Zookeeper
hbase:meta
RegionServer 1
RegionServer 2
Region 1
Region 2
Region 4
Region 3

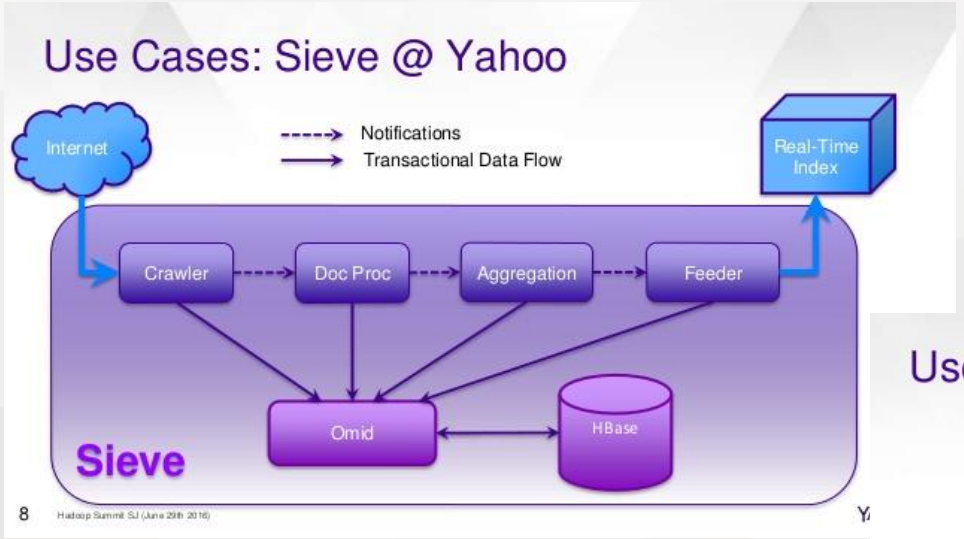


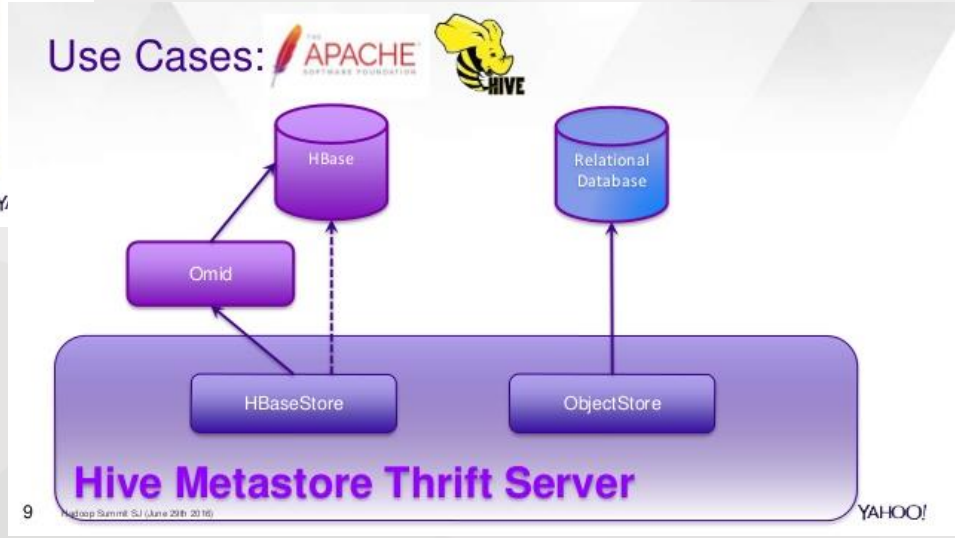
Use Case

Use Case


Hbase
Linux
/proc
Browser
Applications
Custom Metrics
Collector
(tcollector)
TSD (Time Series
Database)
User
1
2
3
4
a
b
c
d
e
OpenTSDB write path
1. Reports the measurement to local Collectd
2. Collectd sends measurements to Remote TSD
3. TSD constructs record and puts data to HBase.
4. Hbase stores data and acknowledges put request.
OpenTSDB read path
a. User specifies query parameters.
b. tsd constructs filter and requests range scan.
c. Hbase scans key range, omitting filtered records, and returns results.
d. tsd renders time series.
e. Graph Request (HTTP)

Use Case


Use Cases: Sieve @ Yahoo
Internet
Notifications
Transactional Data Flow
Real-Time Index
Crawler
Doc Proc
Aggregation
Feeder
Omid
HBase
Sieve
8


Use Cases:
APACHE
HIVE
HBase
Relational Database
Omid
HBaseStore
ObjectStore
Hive Metastore Thrift Server
9
YAHOO!


The End