

hosted by



HBaseConAsia2018

Gehua New Century Hotel Beijing, China

August 17, 2018

hosted by



HBase at DiDi

JingYi Yao

Content

hosted by



- 01 HBase Improvements & Applications
- 02 Phoenix Improvements & Applications
- 03 GeoMesa Applications
- 04 Monitor & Operations

hosted by



01 HBase Improvements & Applications

HBase Improvements & Applications

hosted by



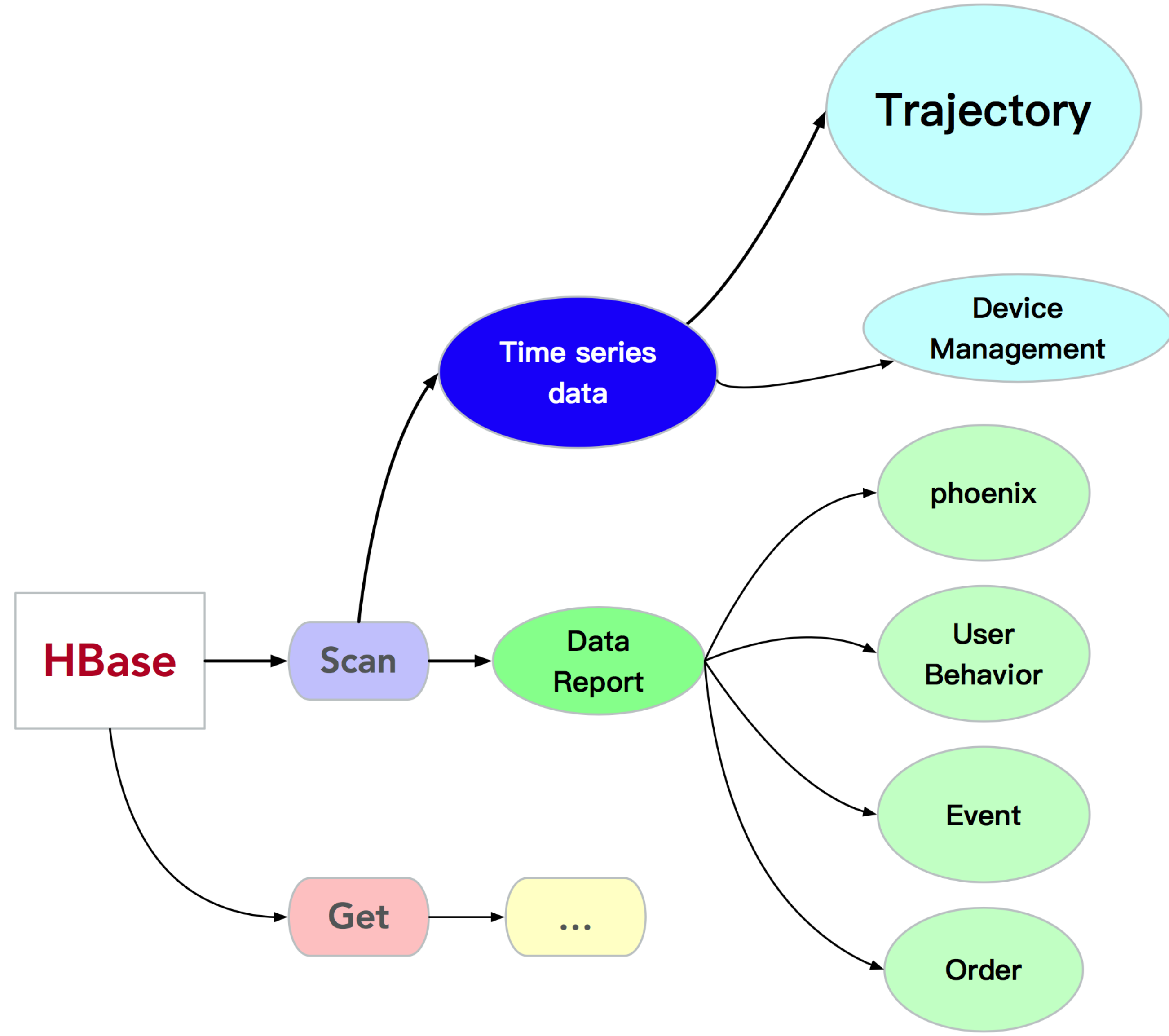
- ◇ 1. HBase Applications at DiDi
- ◇ 2. ACL Improvements
- ◇ 3. Replication Improvements
- ◇ 4. Reduce User Connection

HBase Improvements & Applications

hosted by

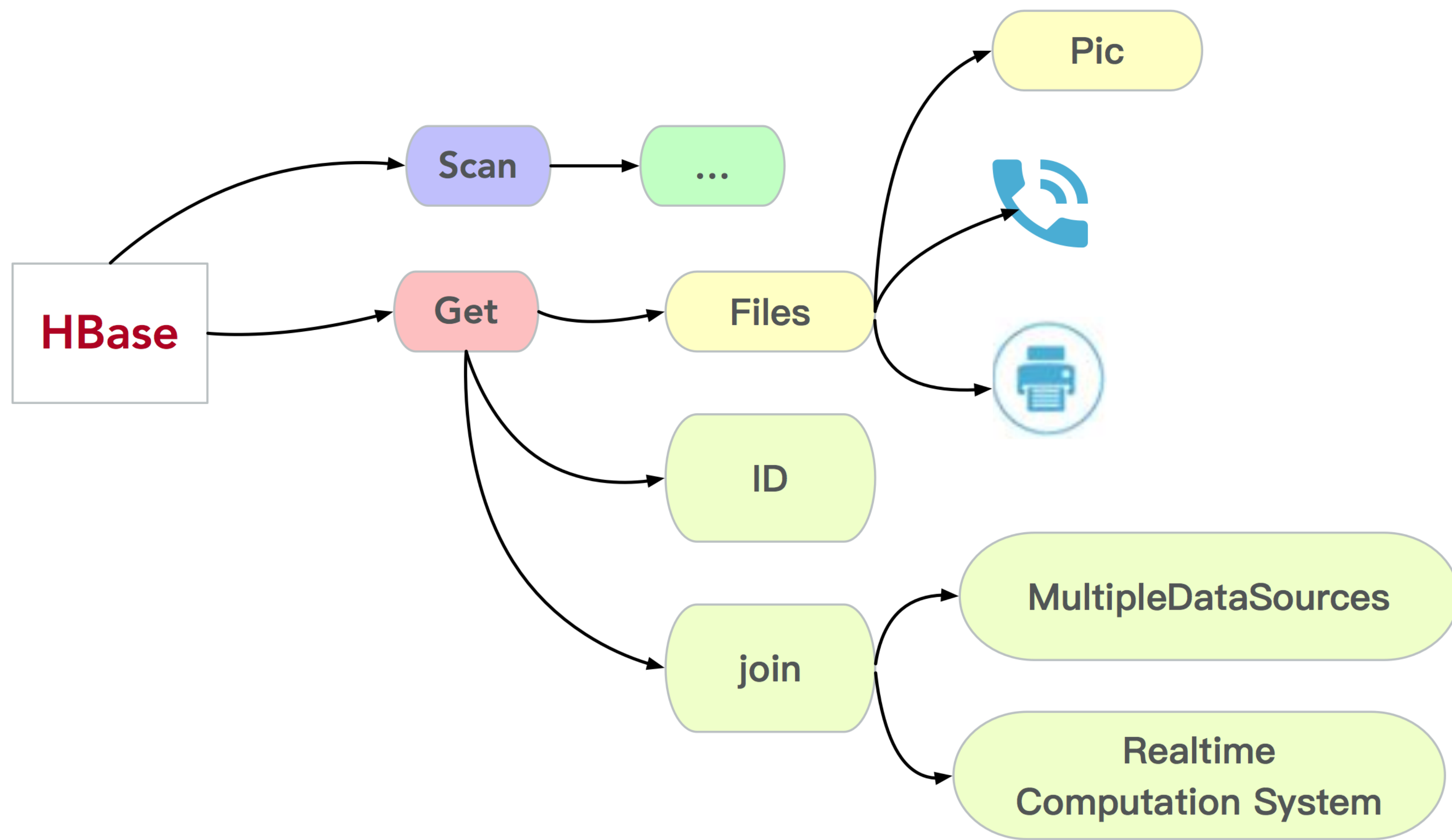


1. HBase Applications



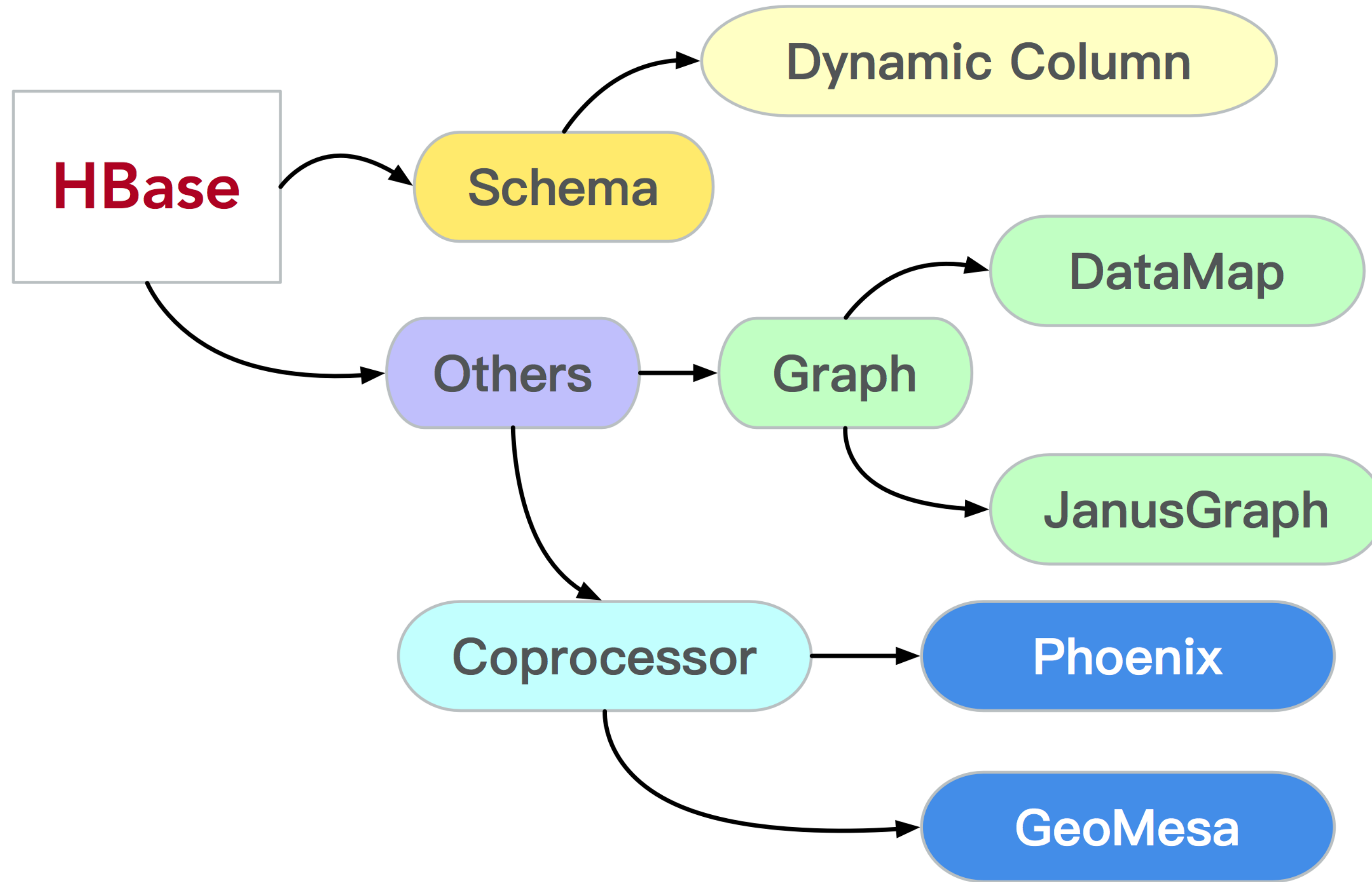
HBase Improvements & Applications

1. HBase Applications



HBase Improvements & Applications

1. HBase Applications



HBase Improvements & Applications

hosted by



1. HBase Applications

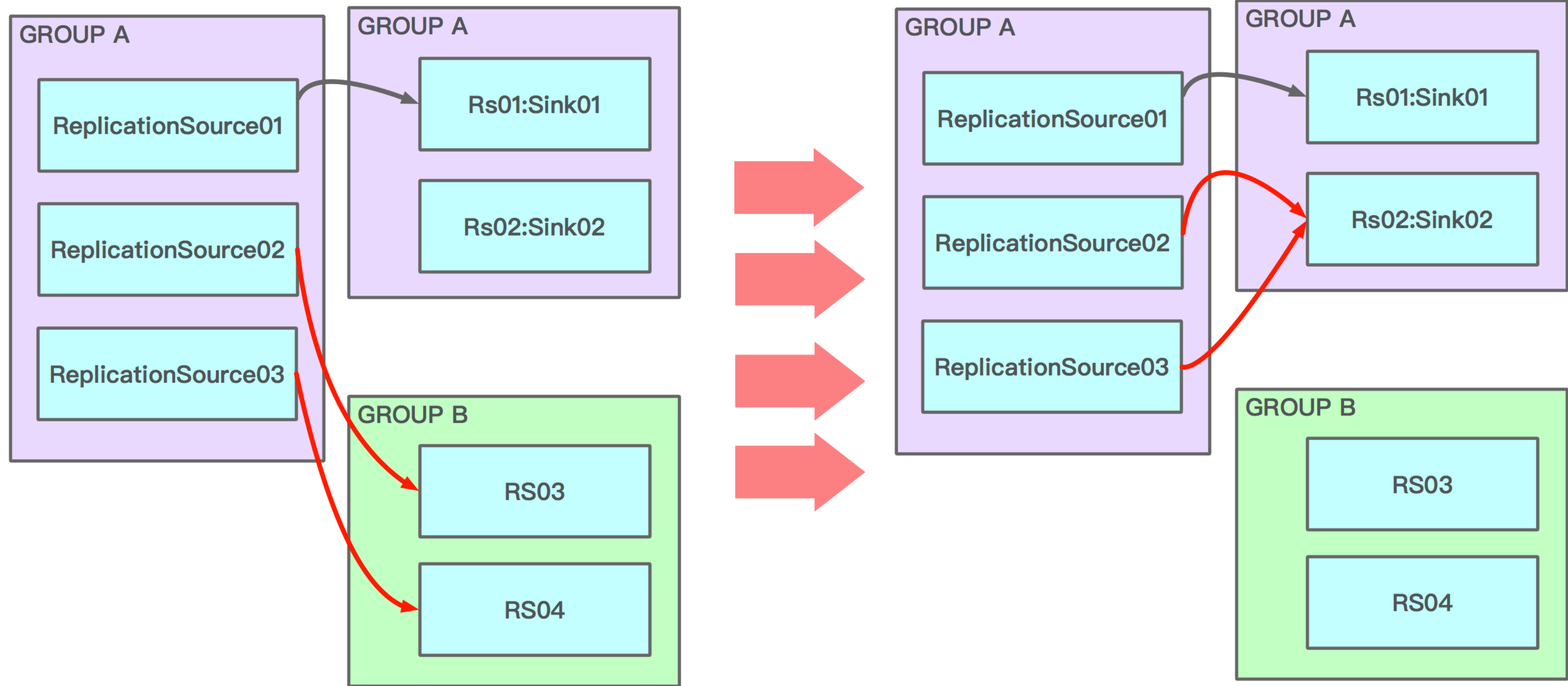
- ◇ Message Queue
- ◇ GeoGraphic Feature data
- ◇ Heat Map
- ◇ User Profile
- ◇ Log Analysis

HBase Improvements & Applications

hosted by



2. Replication Improvements



2. Replication Improvements

- ◇ (1). Replication in RS Group
- ◇ (2). Table Metric
- ◇ (3). RS Group ratio

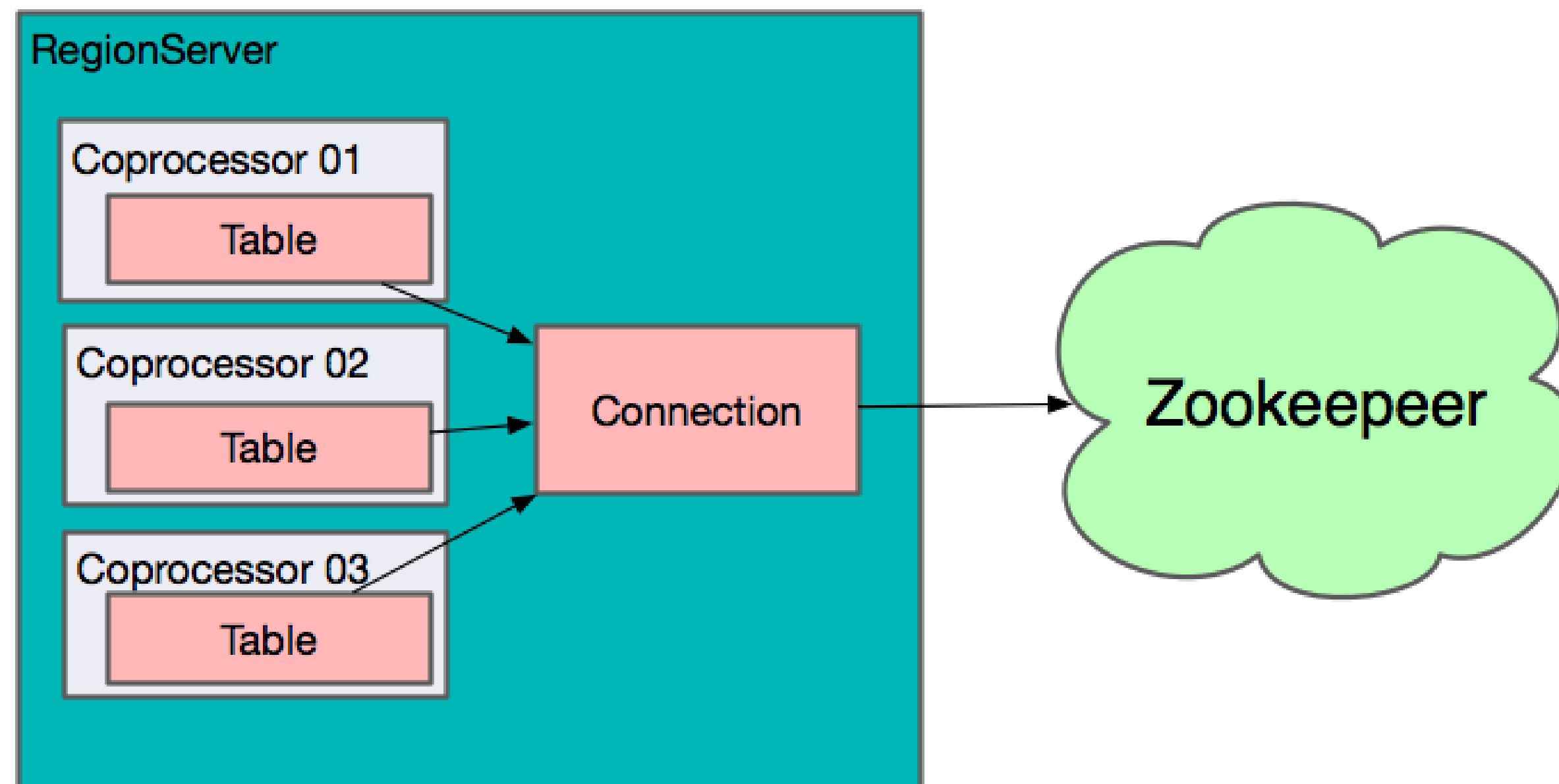
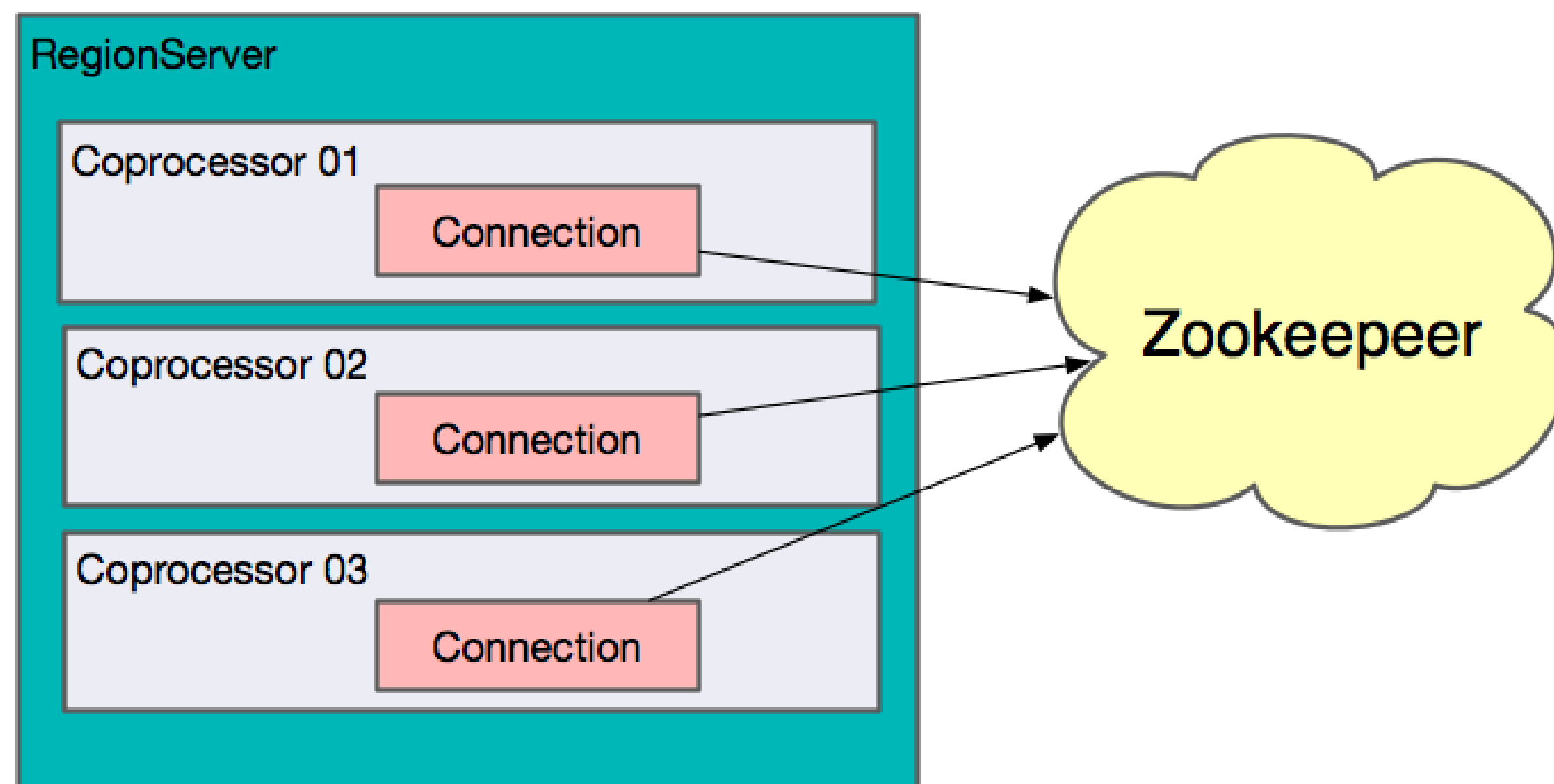
HBase Improvements & Applications

hosted by



3. Connection

- ◇ (1). users
- ◇ (2). Coprocessor:
- ◇ Regionserver clusterConnection



HBase Improvements & Applications

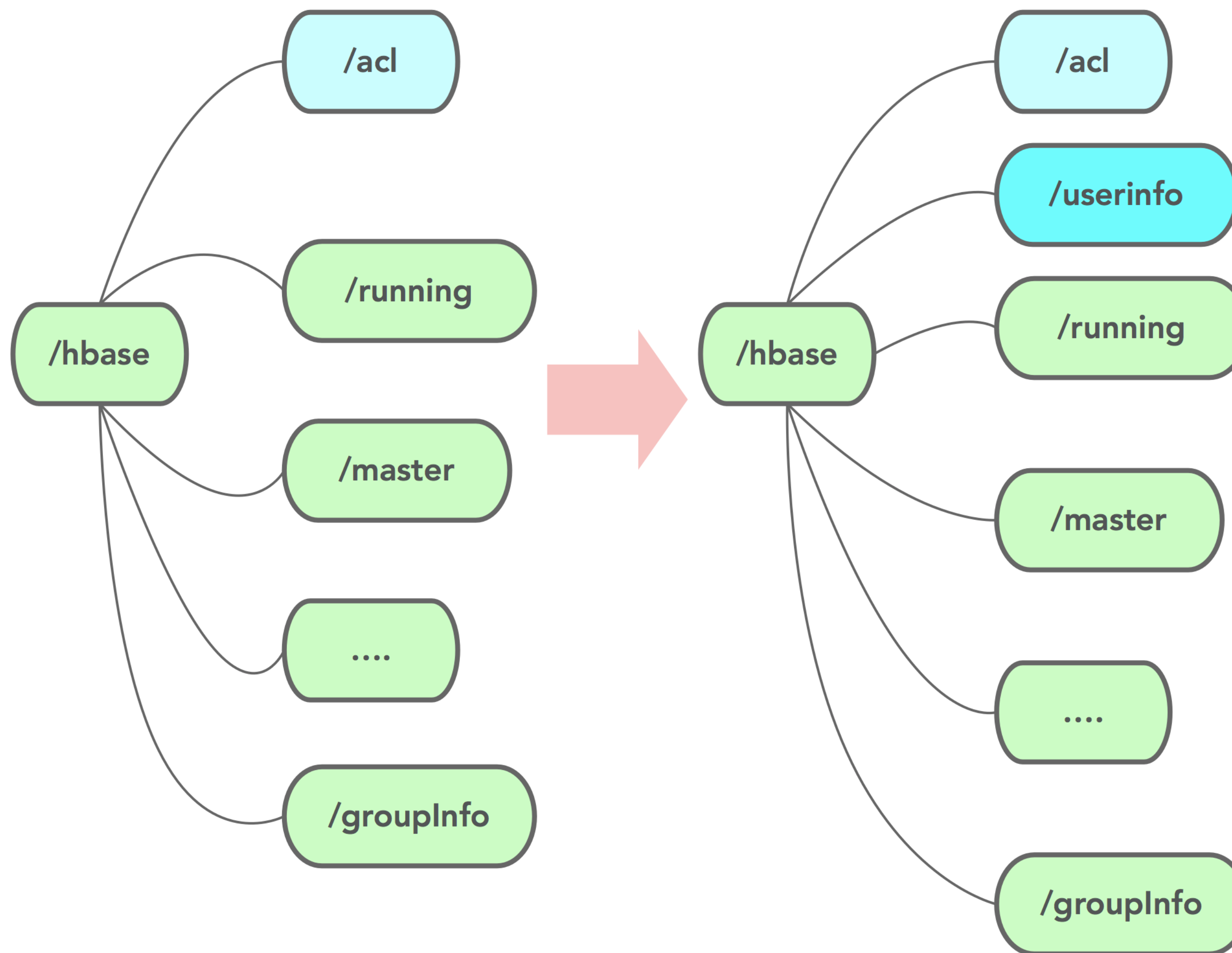
hosted by



4. ACL Improvements

hbase:userinfo

rowkey(username)	CF: l:p
hadoop	hadoop_password
username01	user01_password
username02	...
username03	...

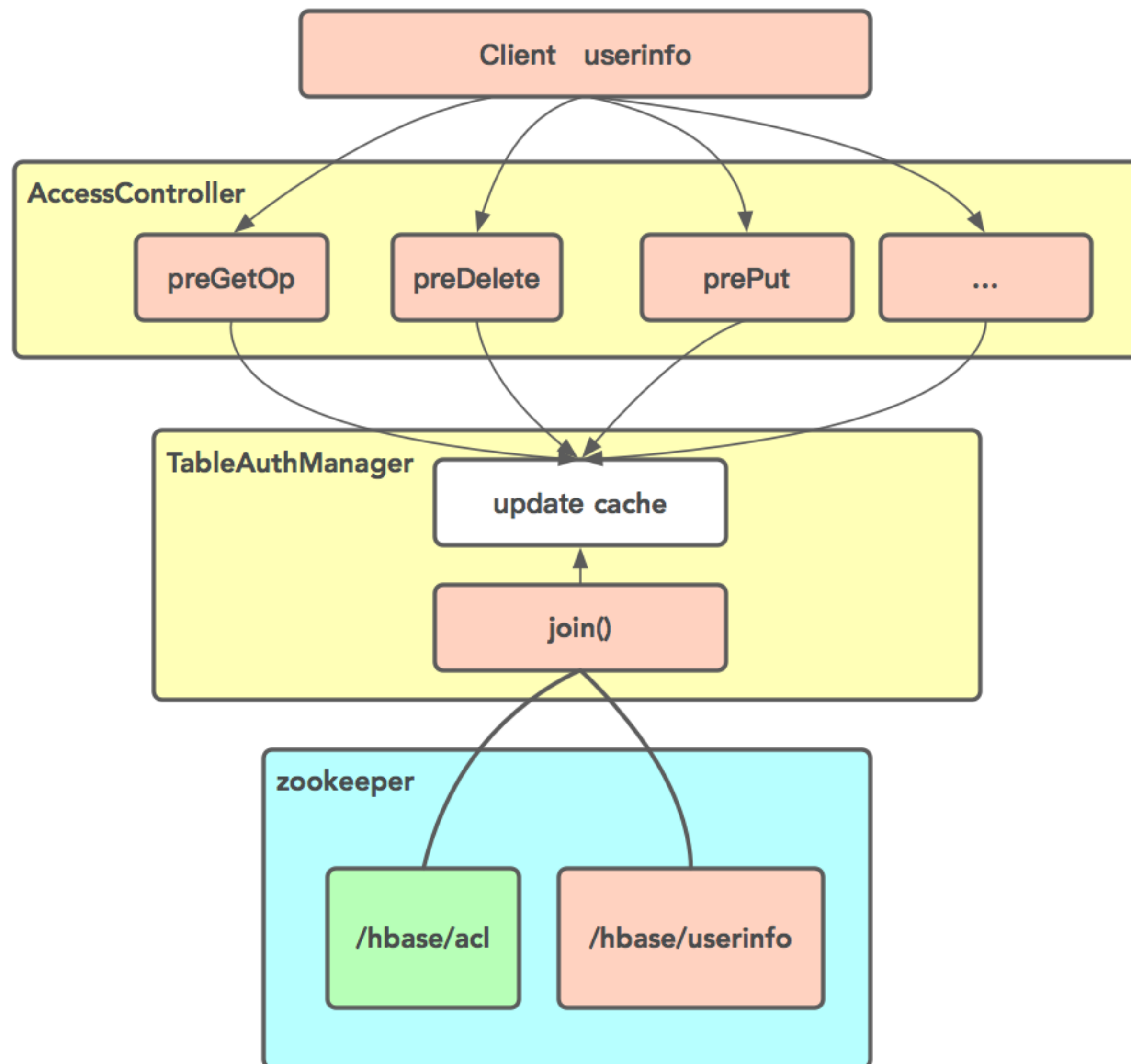


HBase Improvements & Applications

hosted by



4. ACL Improvements

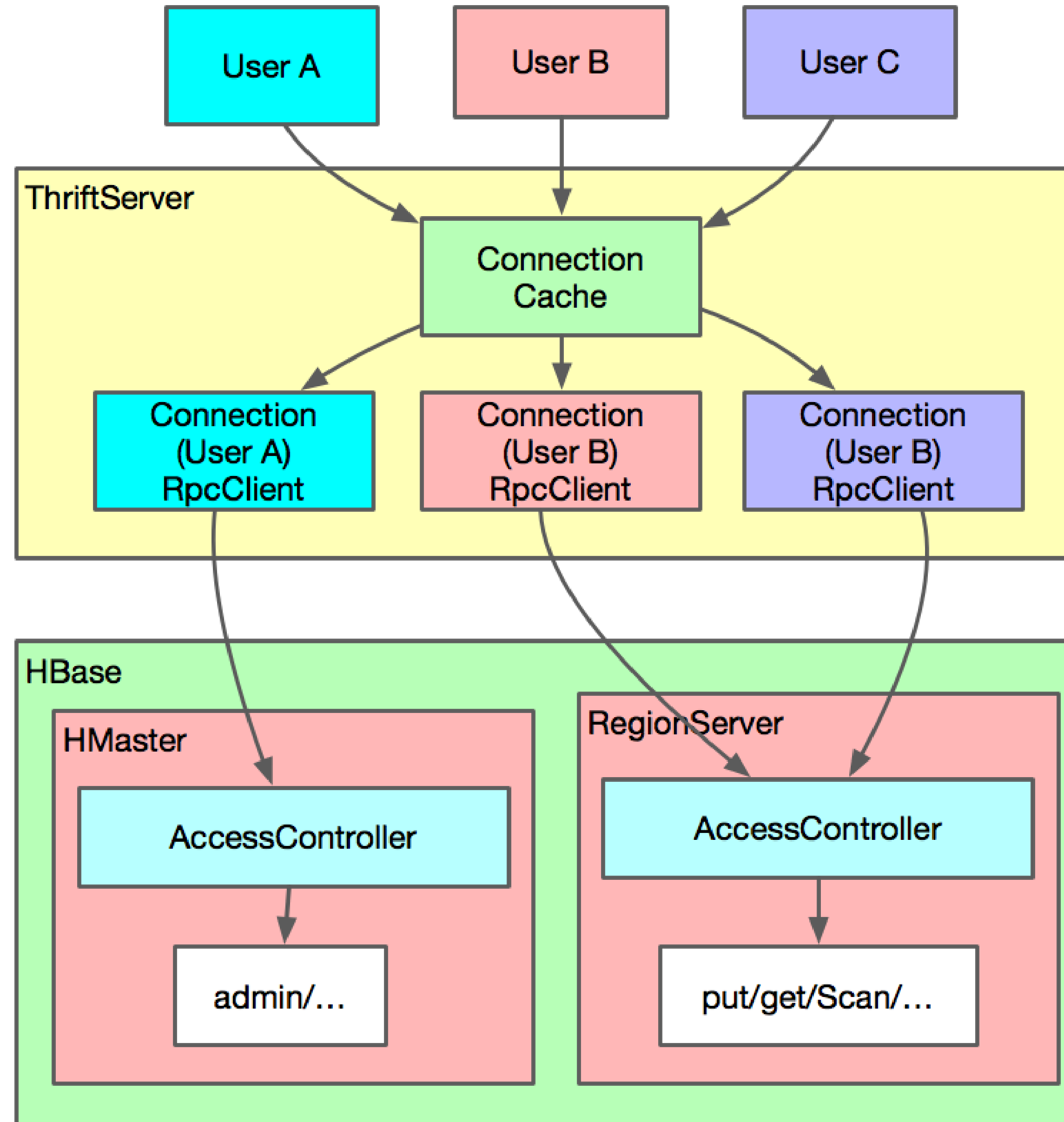


HBase Improvements & Applications

hosted by



4. ACL Improvements

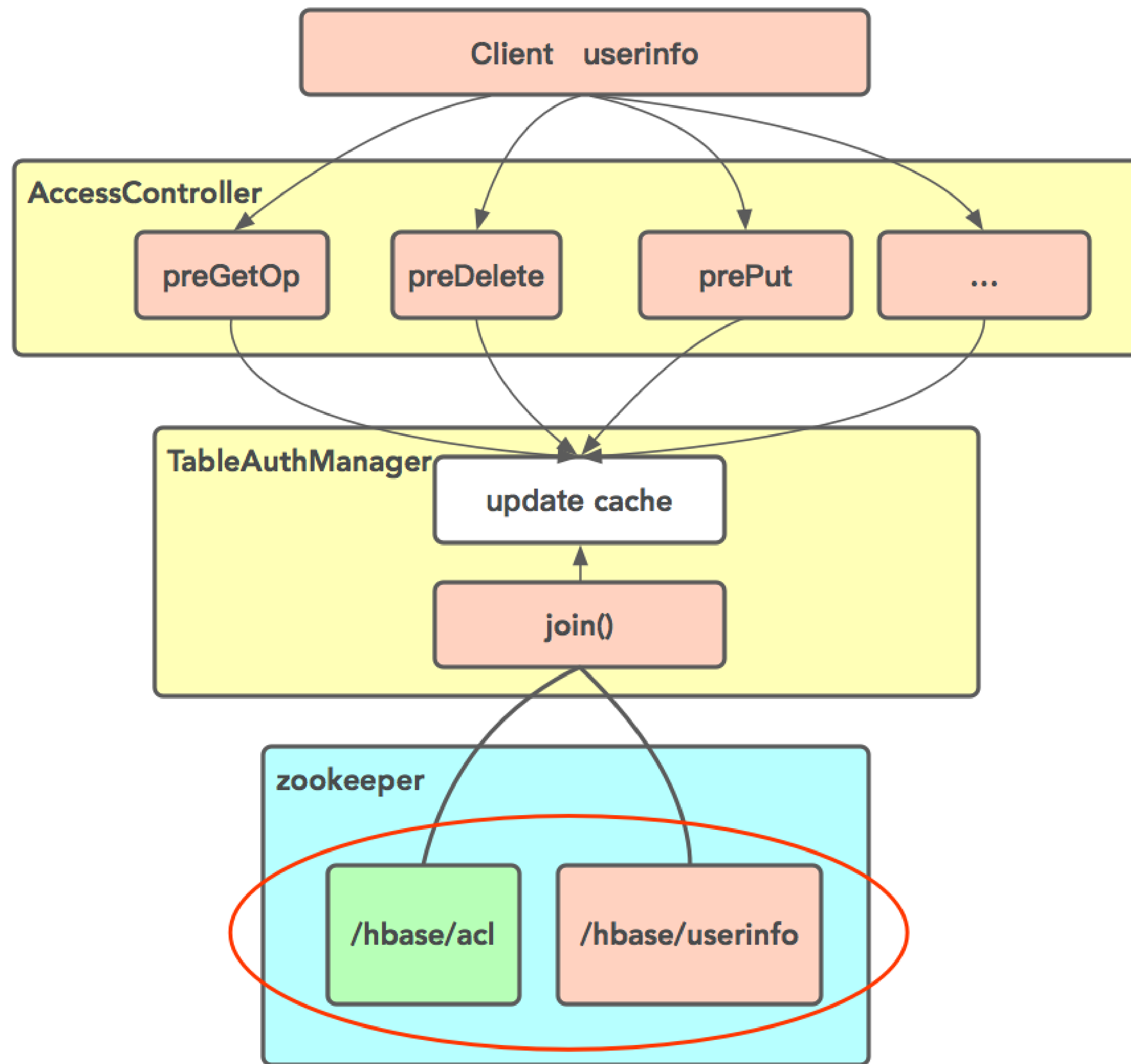


HBase Improvements & Applications

4. ACL Improvements

◇ reduce zookeeper znode

◇ and watcher



5. Others

- ◇ (1) RPC audit log
- ◇ (2) RSGroup bug fix
- ◇ (3) quota
- ◇ (4) safe drop table
- ◇ ...

hosted by



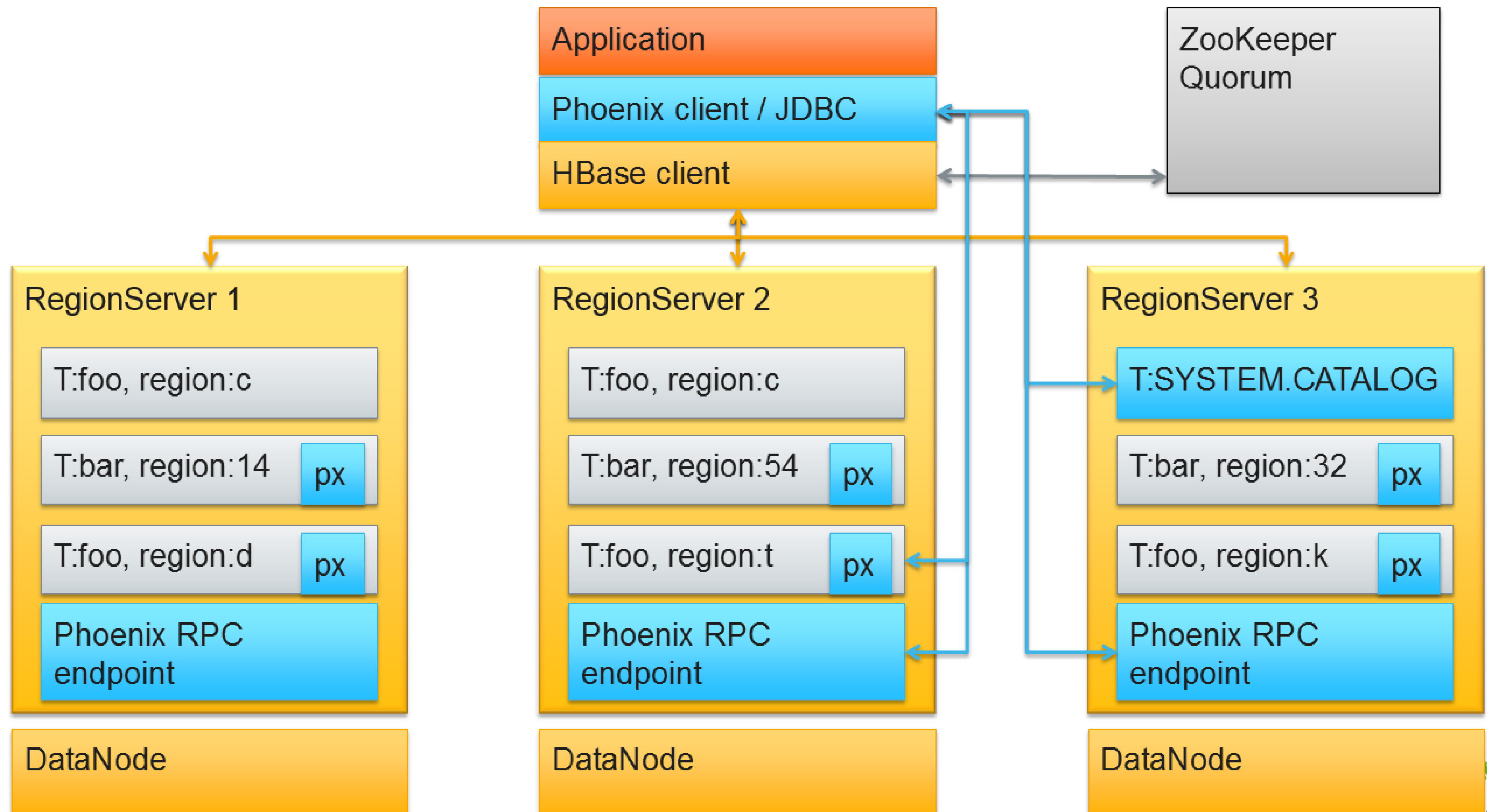
02 Phoenix Improvements & Applications

Phoenix Improvements & Applications

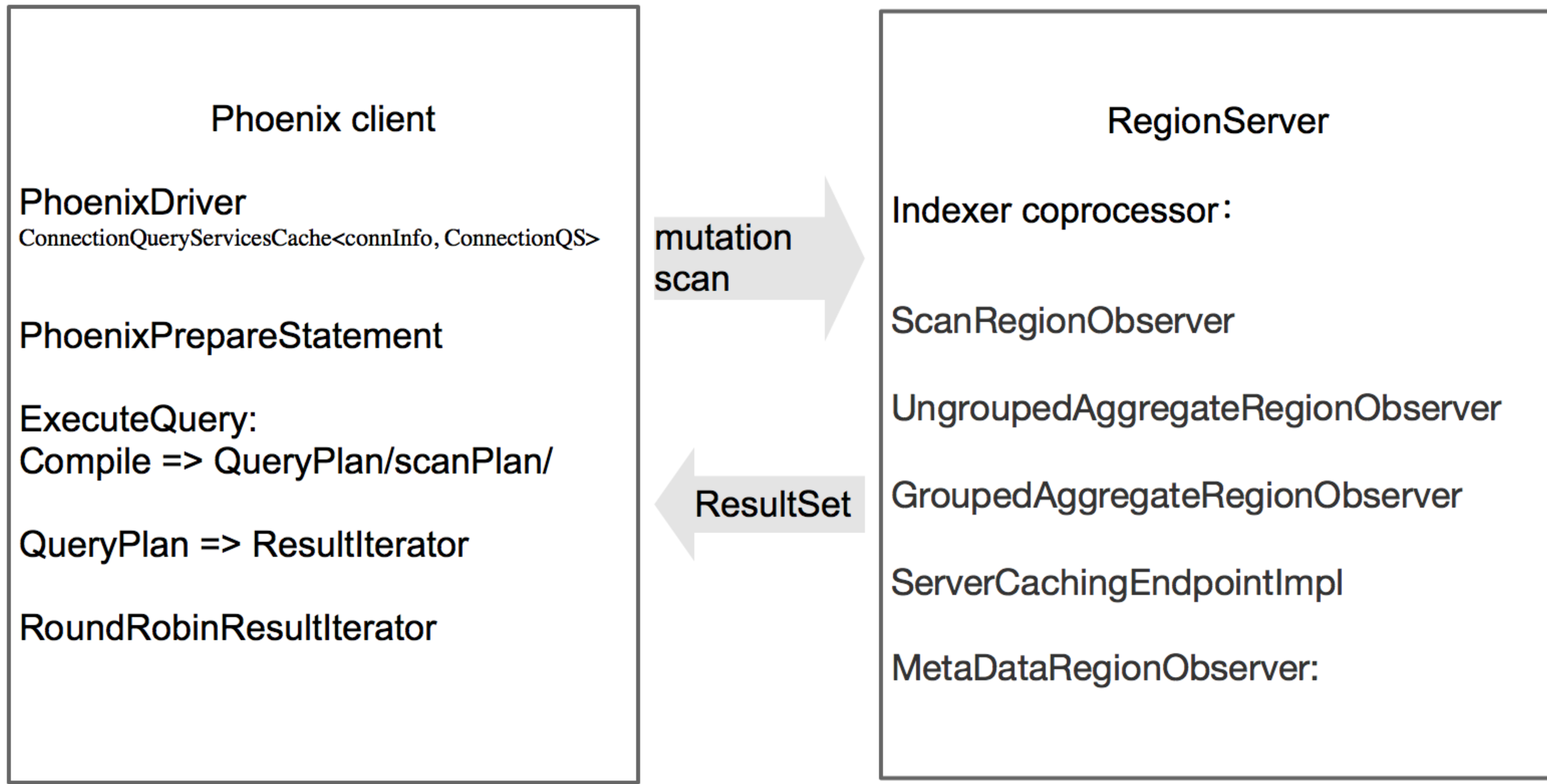
hosted by



Phoenix Architecture



Phoenix Improvements & Applications



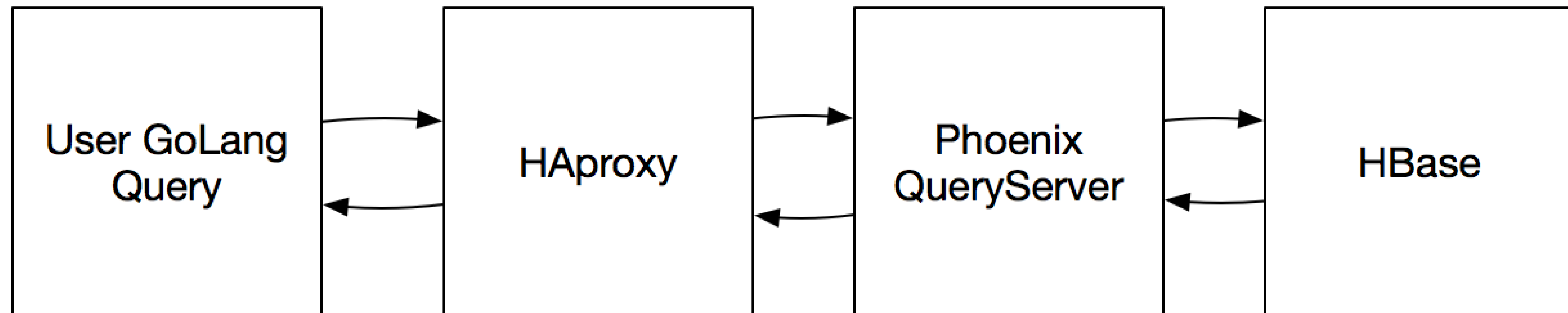
Phoenix Improvements & Applications

hosted by



5. Order History

- ◇ SLA 99.95%
- ◇ Secondary indexes
- ◇ Java-jdbc client P99 35 ms
- ◇ column default
- ◇ QueryServer



Phoenix Improvements & Applications

hosted by



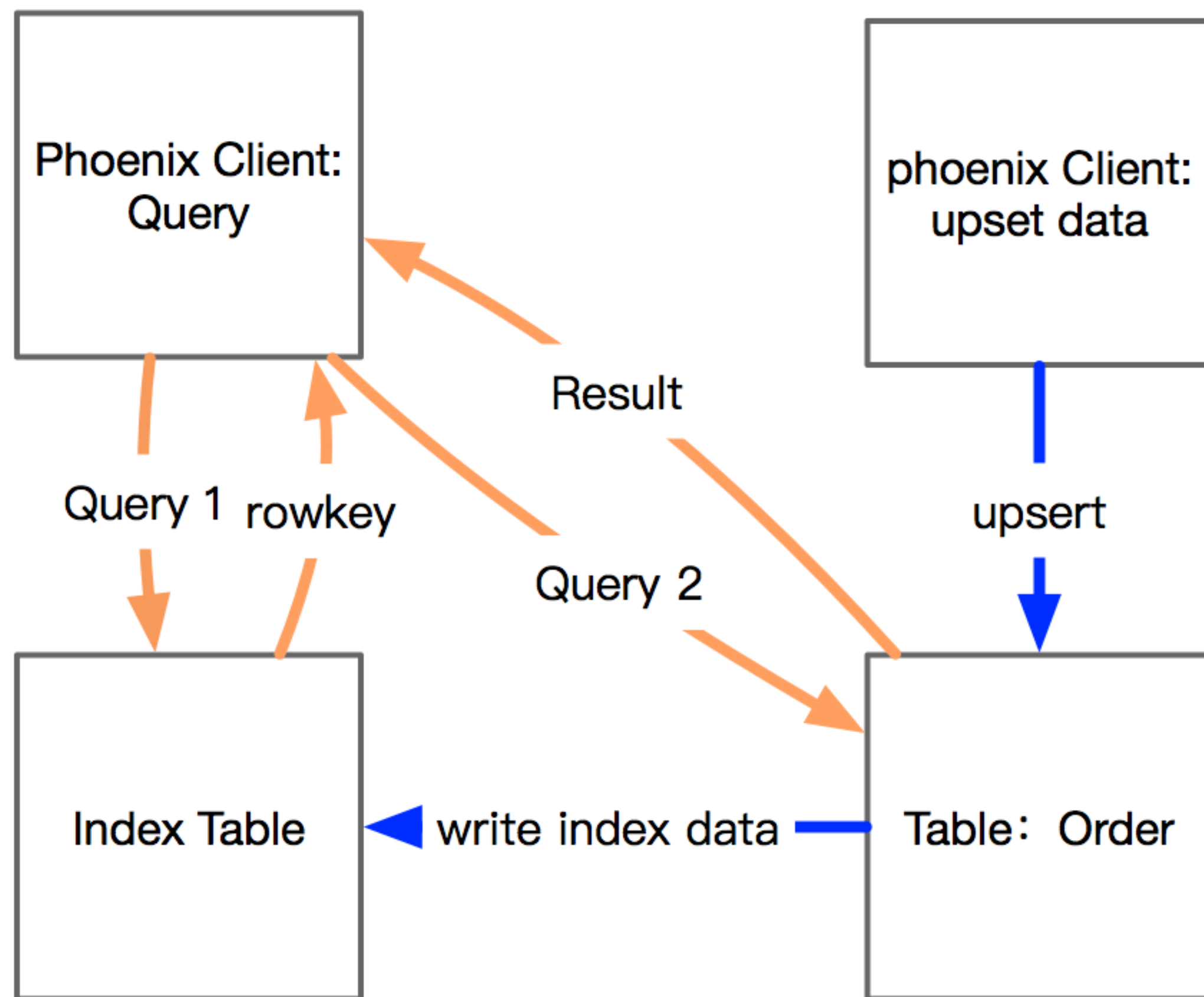
- ◇ Secondary indexes :
 - ◇ INDEX Type: Global Index => Covered Index
 - ◇ ASYNC INDEX
- ◇ SALT_BUCKETS => reverse(ID) & pre split
- ◇ coprocessor : reduce connection
- ◇ QueryServer metrics

Phoenix Improvements & Applications

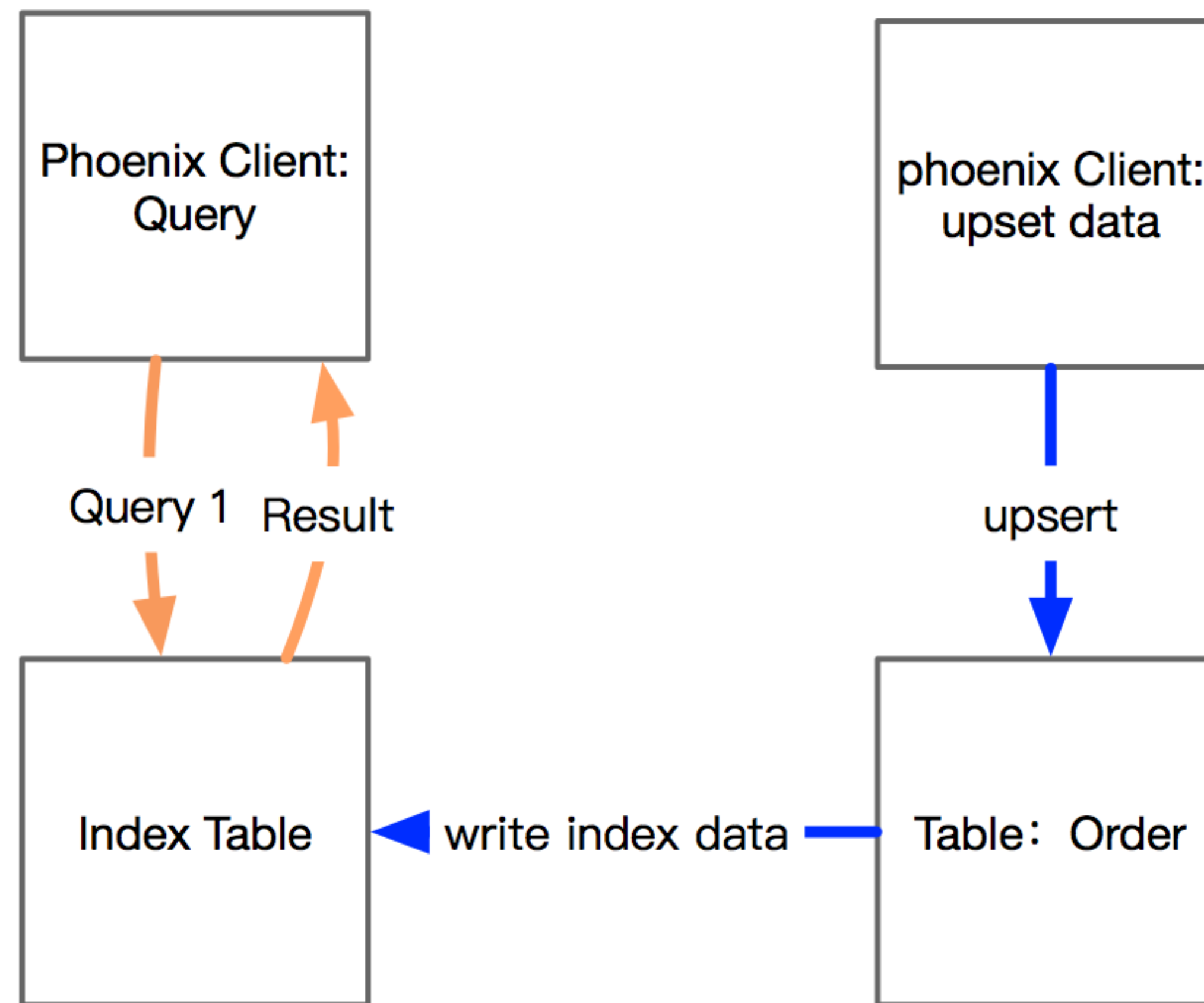
hosted by



Global Indexes



Covered Indexes



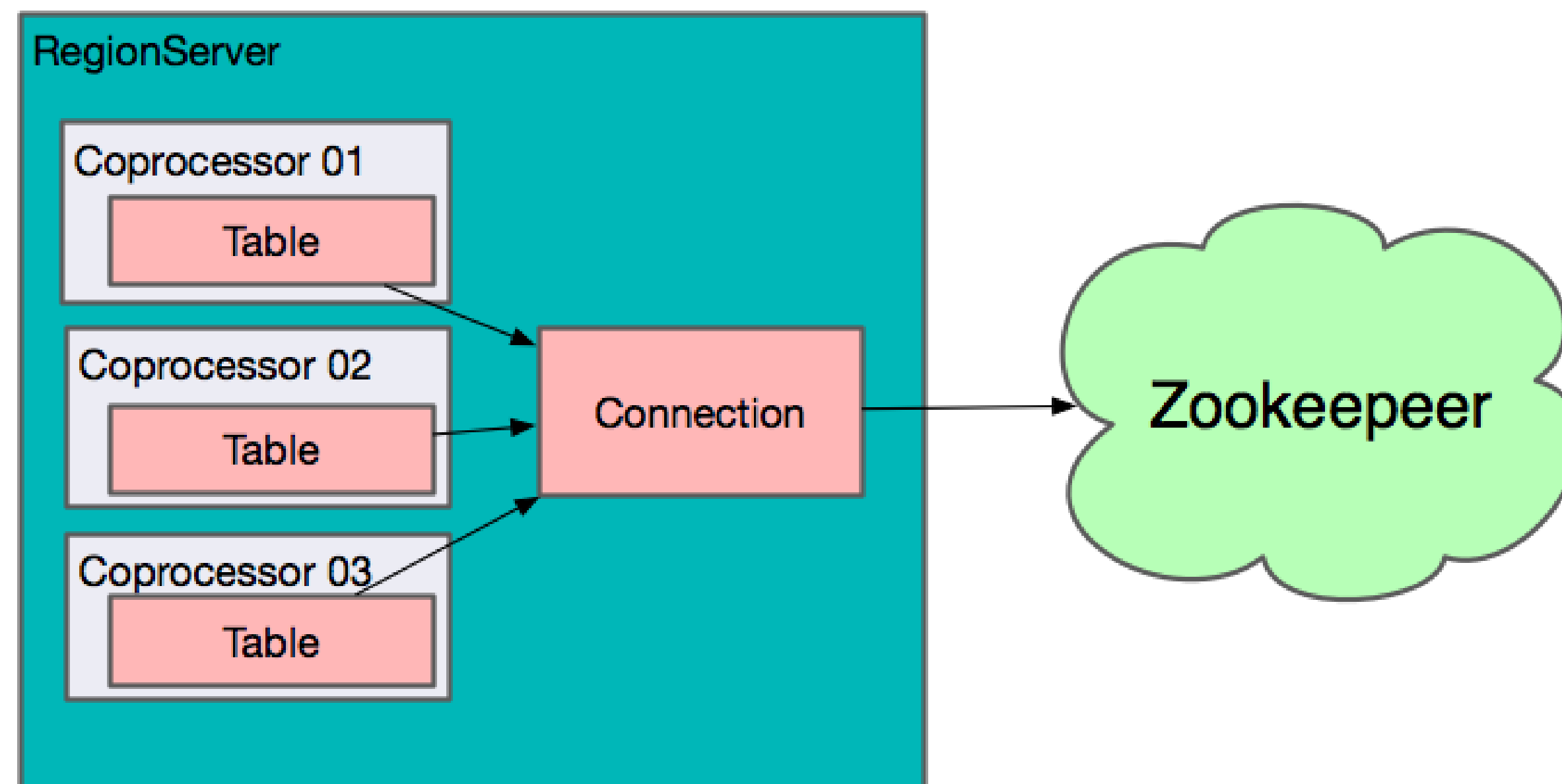
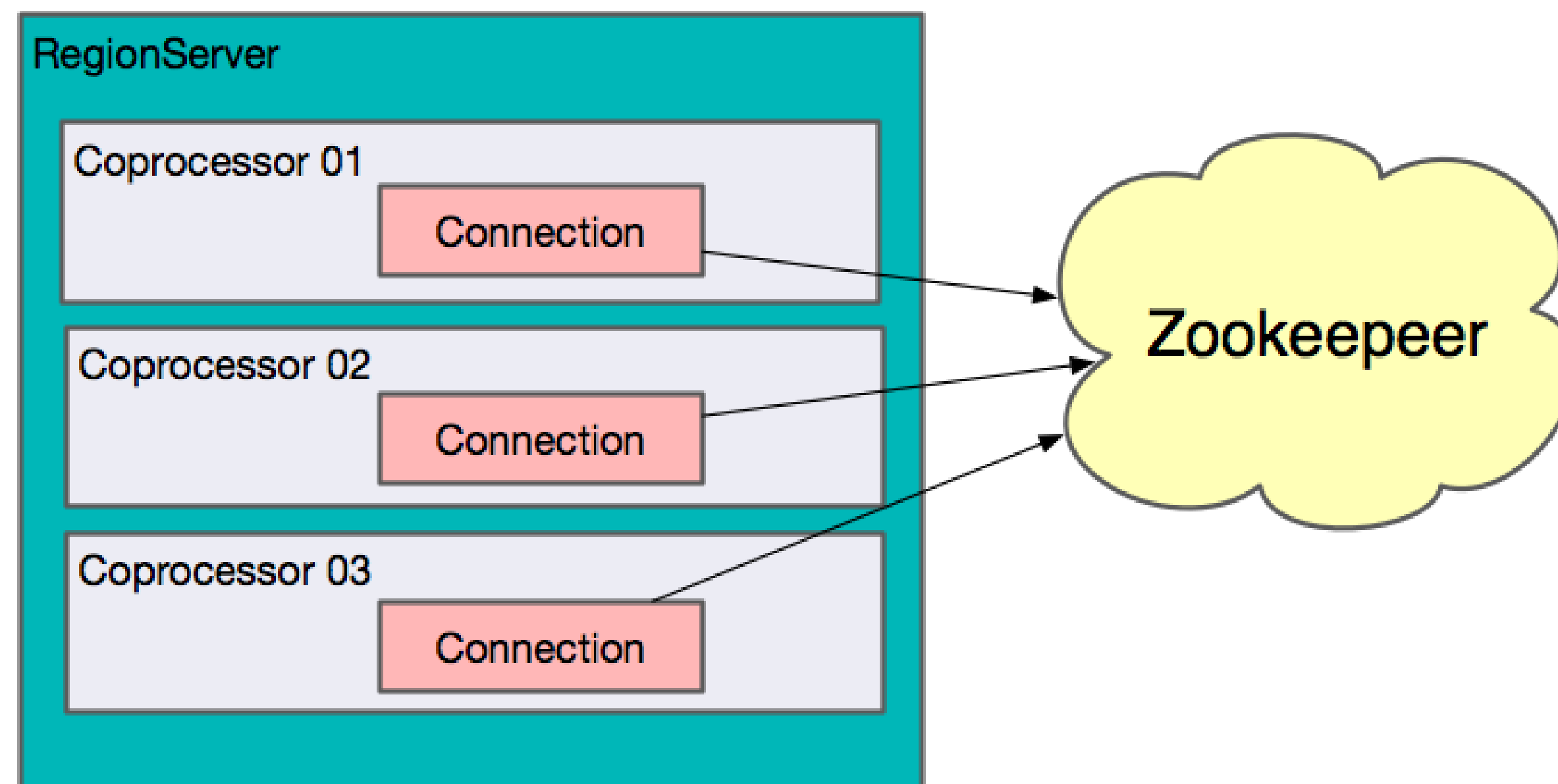
Phoenix Improvements & Applications

hosted by



◇ zk connection:

◇ region num * Index num => 1



hosted by



03 GeoMesa Applications

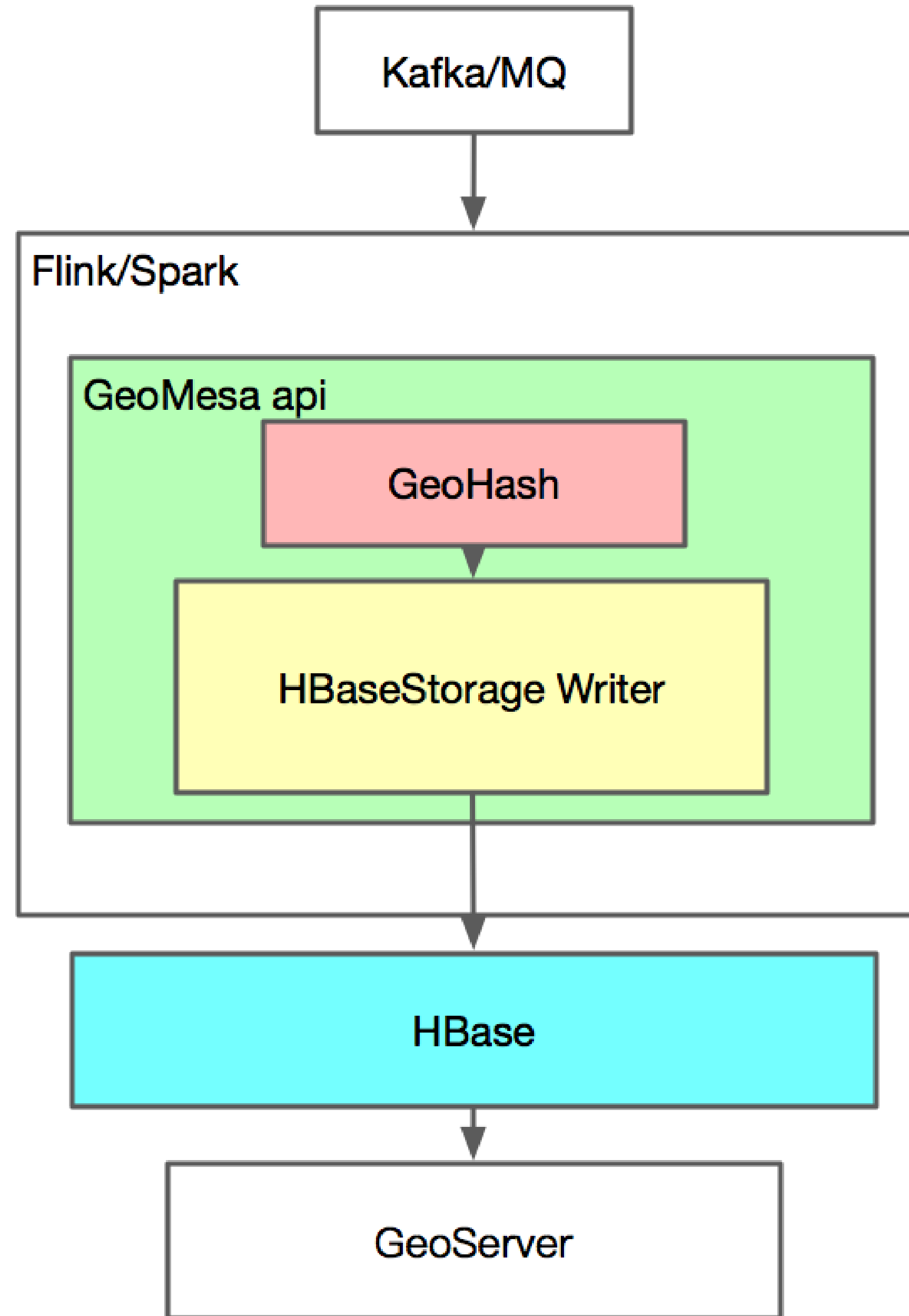
GeoMesa Applications

hosted by



GeoMesa

GeoMesa is an open source suite of tools that enables large-scale geospatial querying and analytics on distributed computing systems.

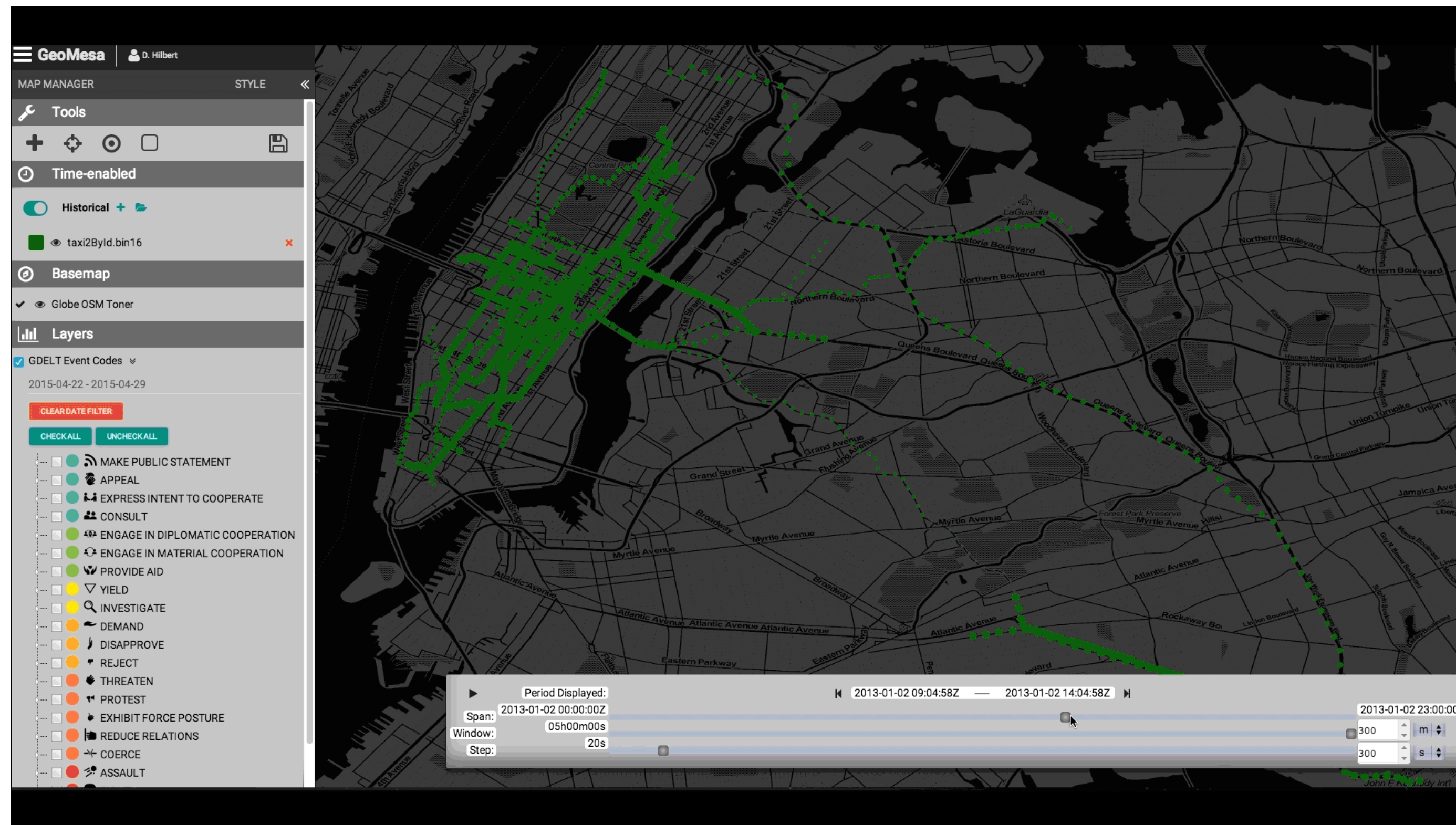


GeoMesa Applications

hosted by

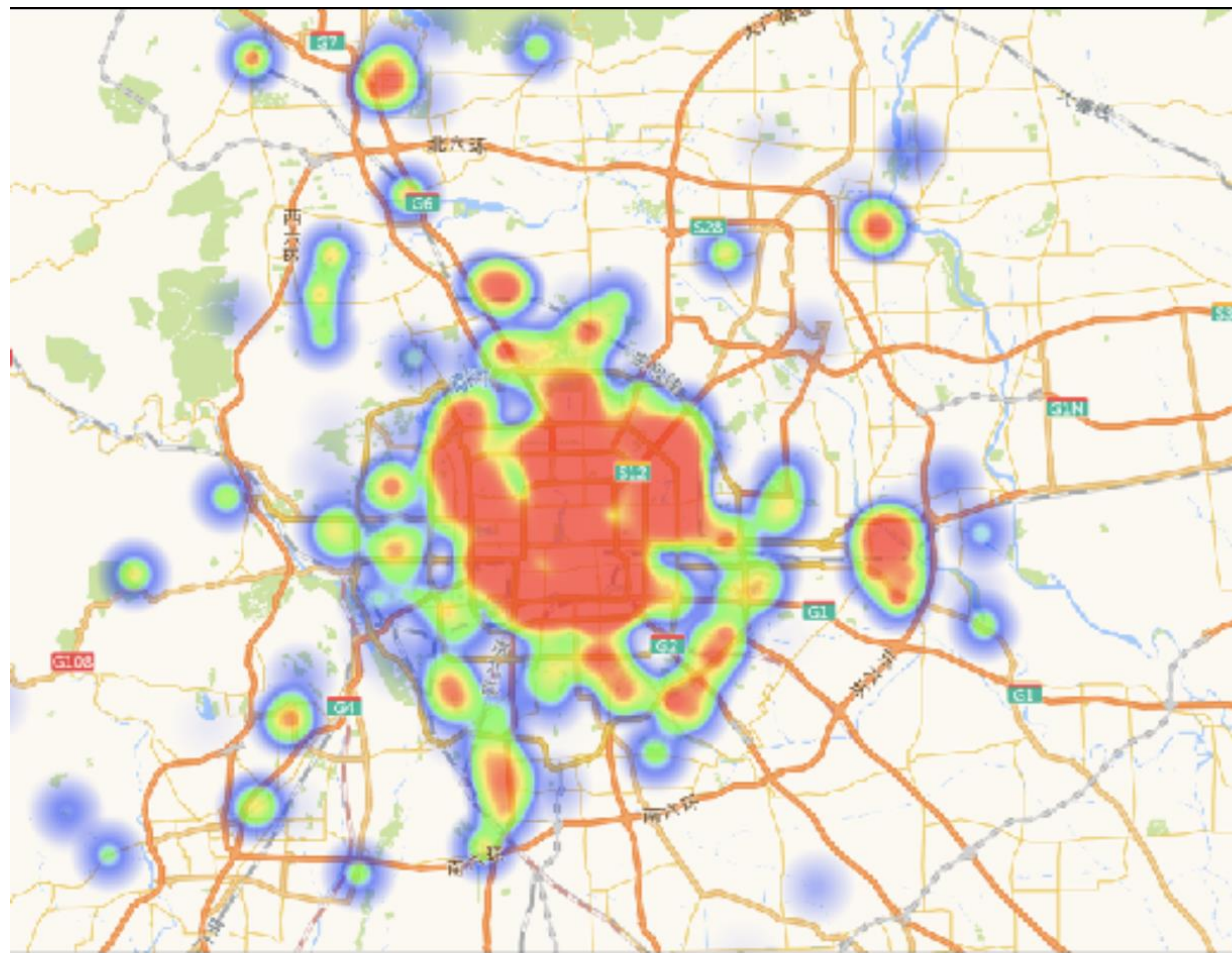


- ◇ 1. Road Detection
- ◇ 2. Device ID => Trajectory



GeoMesa Applications

hosted by



GeoMesa Applications

hosted by



Index

- ◇ Z2 : latitude and longitude
- ◇ Z3 : latitude, longitude, and time
- ◇ XZ2 : line strings or polygons
- ◇ XZ3 : non-Point geometry and has a time attribute

◇ Record/ID

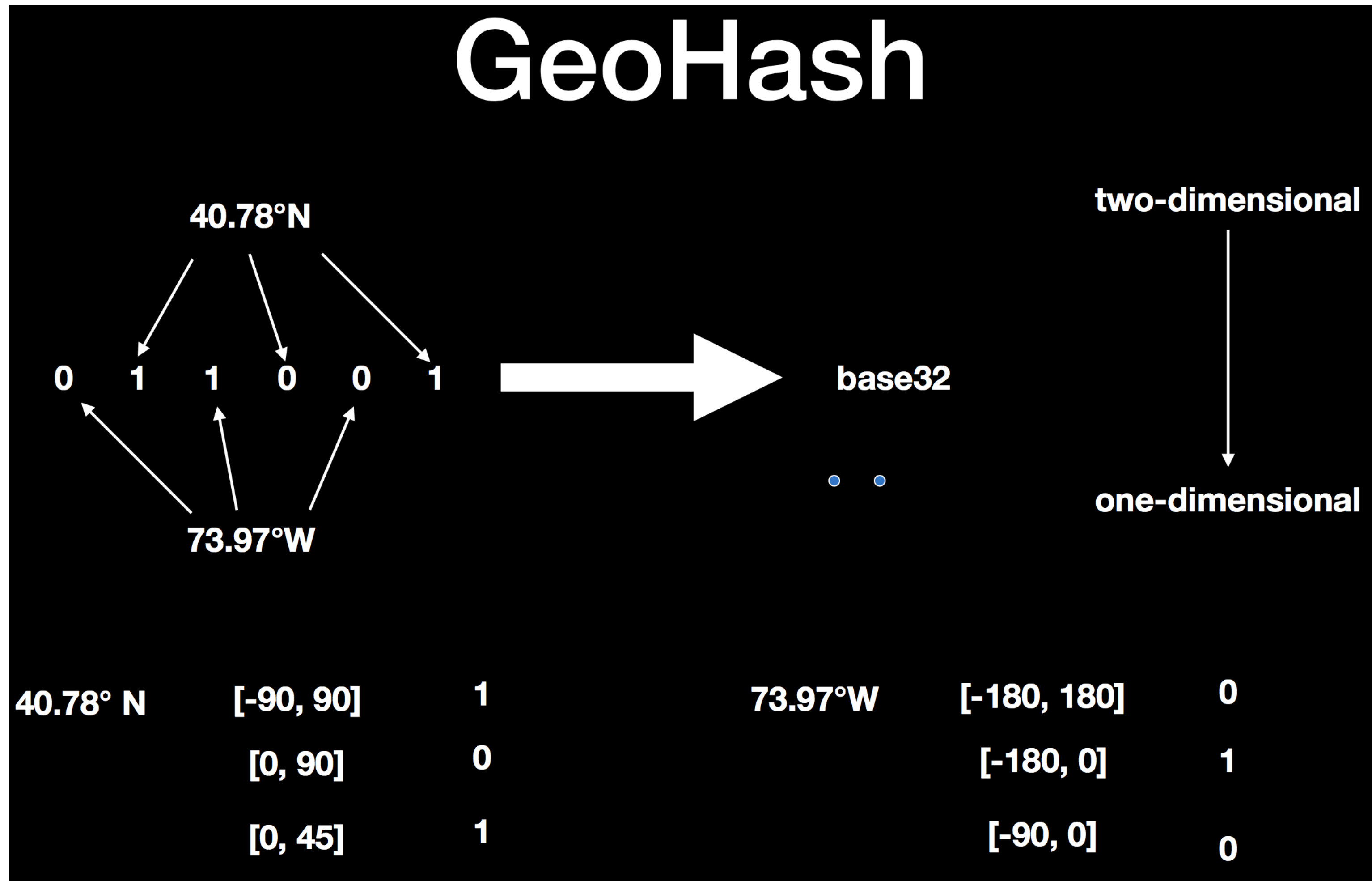
◇ Attribute

KEY						VALUE	
ROW			COLUMN		TIMESTAMP	VIZ	Byte-encoded SimpleFeature
			COLUMN FAMILY	COLUMN QUALIFIER			
Epoch Week 2 bytes	Z3(x,y,t) 8 bytes	Unique ID (such as UUID)	"F"	-	-	Security tags	

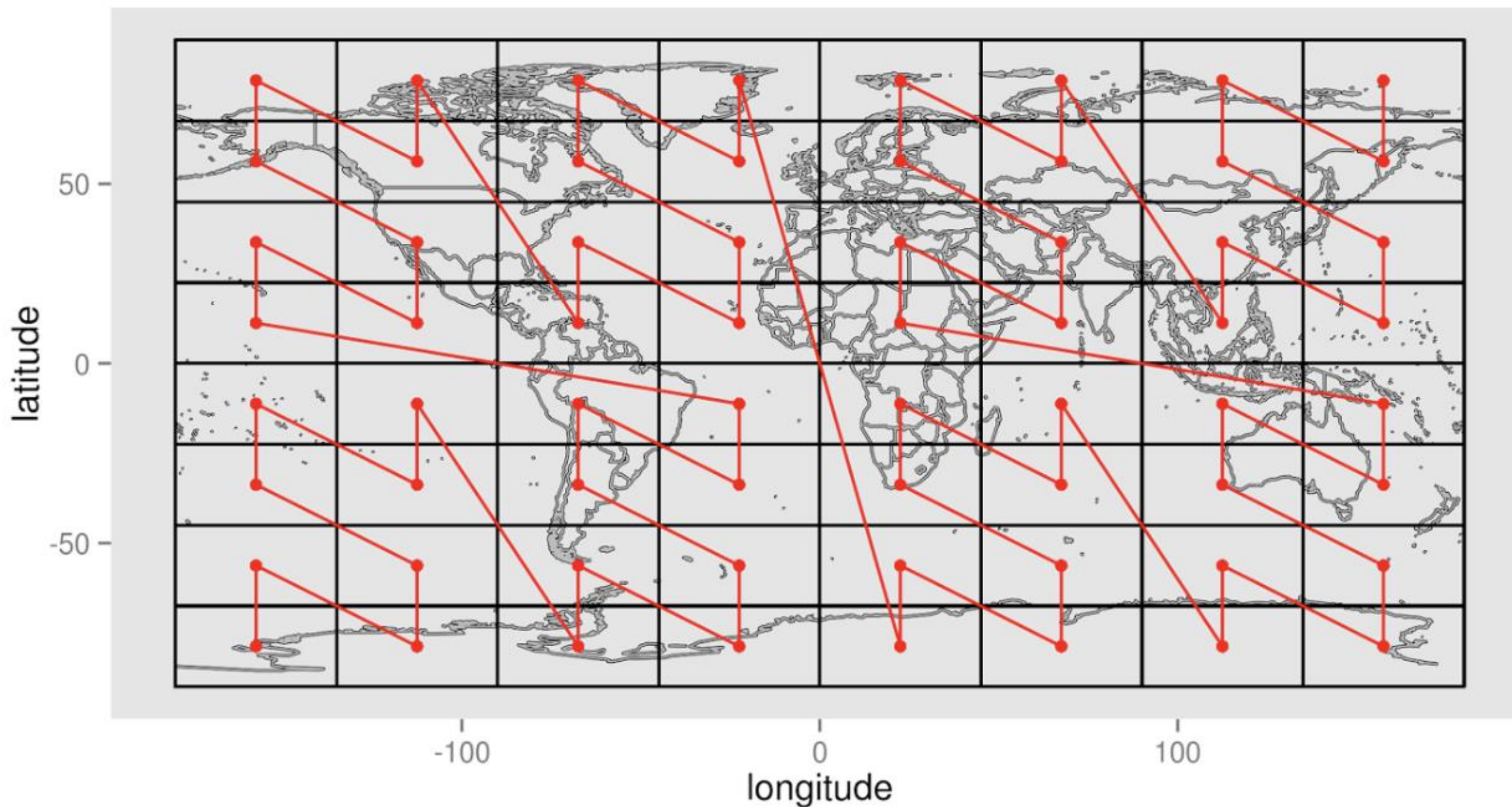
GeoMesa Applications

◇ GeoHash:

◇ geohashes have been used to first convert latitude and longitude into a single bit vector



Z-order curve

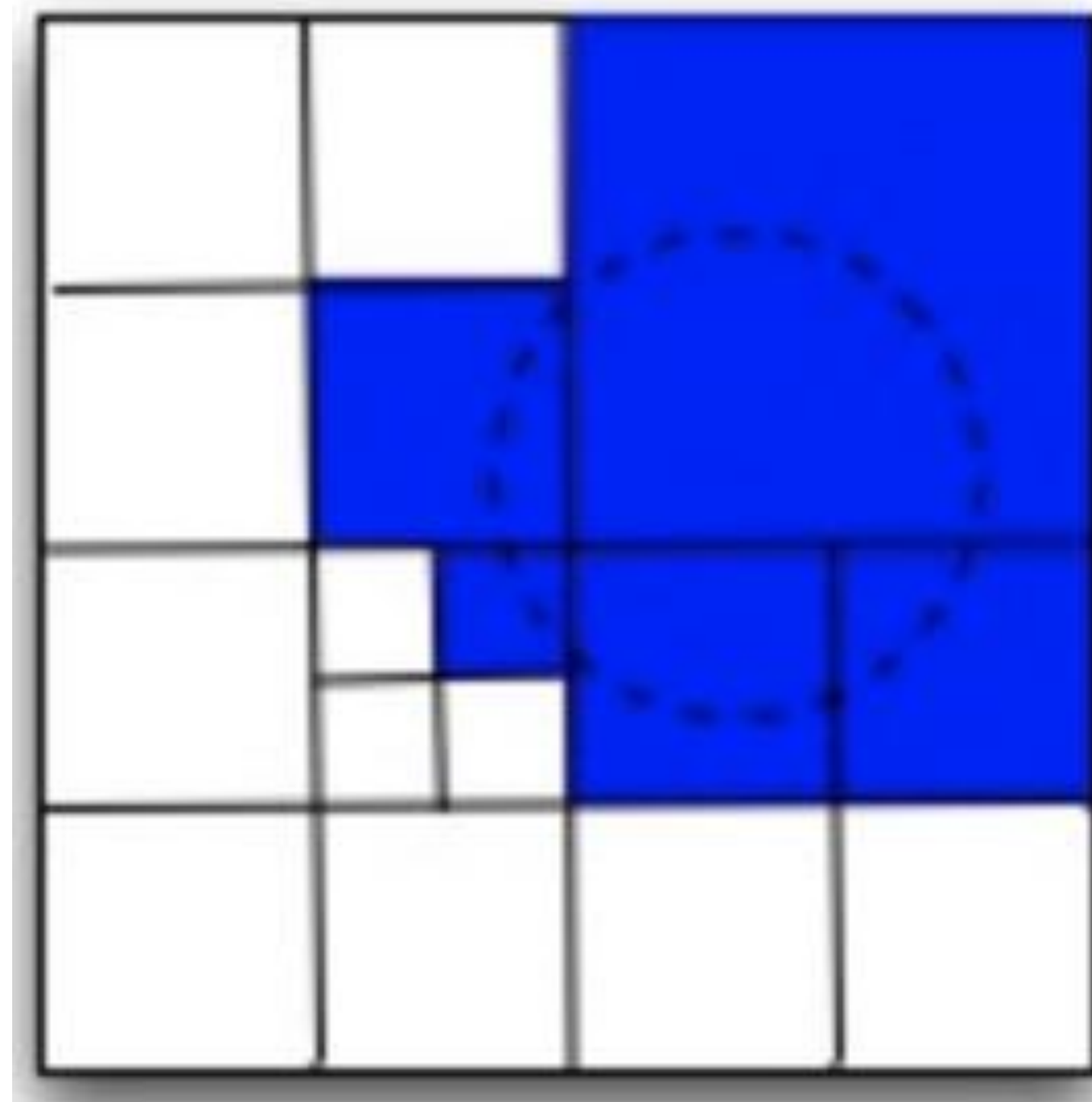


GeoMesa Applications

Z-order curve

◇ Weakness

◇ More scan



GeoMesa Applications

hosted by



Future

- ◇ Hilbert filling curve Index

Experience from Phoenix

- ◇ Geomesa plan
- ◇ Coprocessor
- ◇ GeoMesa Stats

hosted by



04 Monitor & Operations

Monitor & Operations

hosted by



1. Monitor

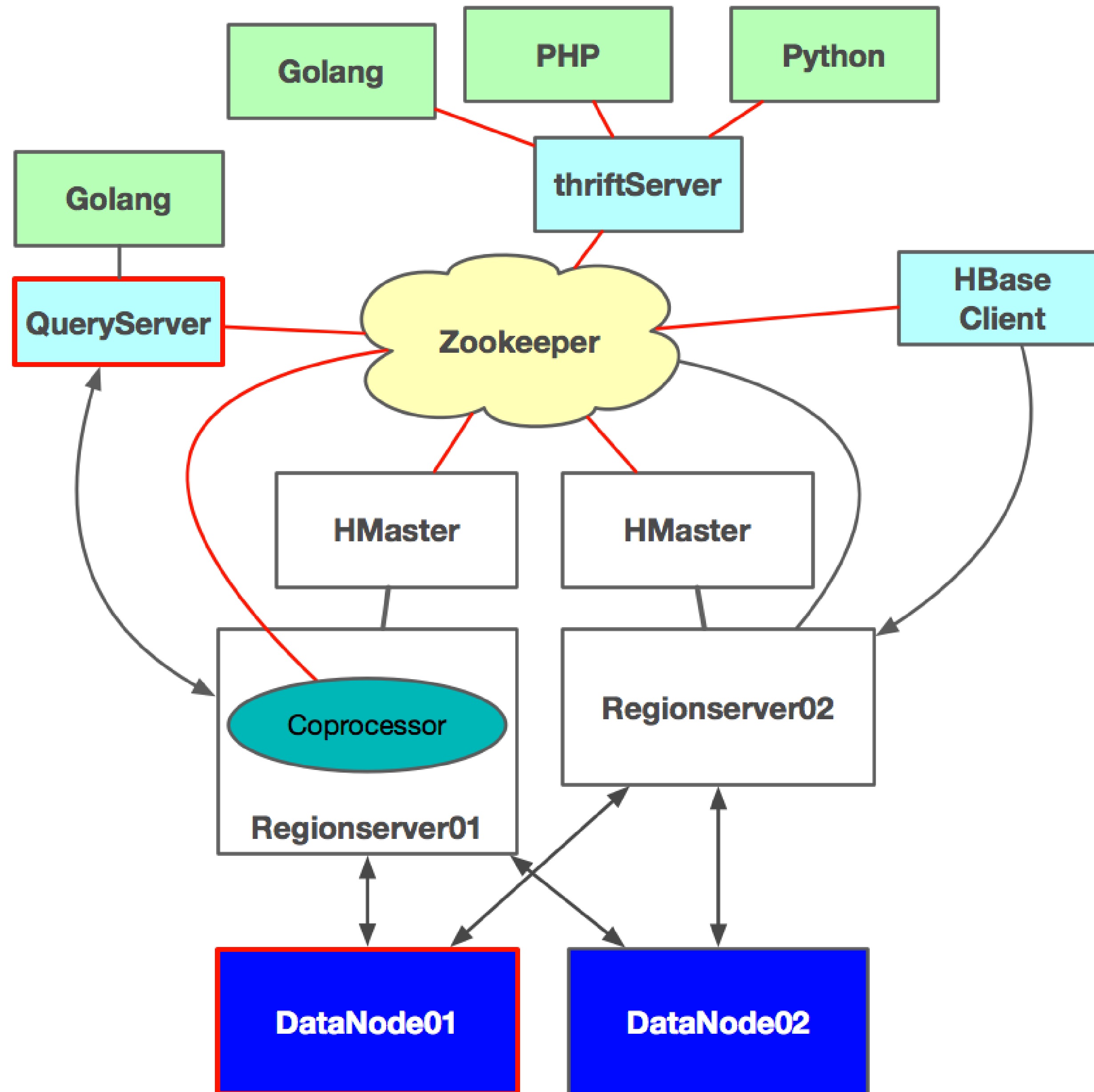
- ◇ (1). HBase
- ◇ (2). Zookeeper
- ◇ (3). HDFS
- ◇ (4). ThriftServer
- ◇ (5). Queryserver

Monitor & Operations

hosted by



Zookeeper



Monitor & Operations

hosted by



◇ echo mntr | nc \$zookeeper 2181

◇ zk_num_alive_connections

◇ zk_open_file_descriptor_count

◇ zk_outstanding_requests

◇ zk_min_latency

◇ zk_followers

Monitor & Operations

hosted by



ThriftServer

- ◇ [Put | Get | Scan ...]_99th_percentile
- ◇ Slow Read/Write ...

QueryServer

- ◇ print slow prepareStatement
- ◇ print slow query
- ◇ ExecuteQueryTime_99th_percentile

HDFS

◇ slow disk

◇ edits log => hdfs path

```
<RECORD>
  <OPCODE>OP_CLOSE</OPCODE>
  <DATA>
    <TXID>319911402</TXID>
    <LENGTH>0</LENGTH>
    <INODEID>0</INODEID>
    <PATH>/hbase/data/TEST_NAMESPACE/TABLENAME/a6d38aae20cad382935845ad6a088d75/.tmp/f2663bf56dec473790ab01a00c1f7b67</PATH>
    <REPLICATION>3</REPLICATION>
    <MTIME>1525836411898</MTIME>
    <ATIME>1525836411385</ATIME>
    <BLOCKSIZE>268435456</BLOCKSIZE>
    <CLIENT_NAME></CLIENT_NAME>
    <CLIENT_MACHINE></CLIENT_MACHINE>
    <BLOCK>
      <BLOCK_ID>1105579166</BLOCK_ID>
      <NUM_BYTES>4010867</NUM_BYTES>
      <GENSTAMP>31851229</GENSTAMP>
    </BLOCK>
    <PERMISSION_STATUS>
      <USERNAME>hadoop</USERNAME>
      <GROUPNAME>supergroup</GROUPNAME>
      <MODE>420</MODE>
    </PERMISSION_STATUS>
  </DATA>
</RECORD>
```

Monitor & Operations

hosted by



◇ DHS (Didi HBase Service):

◇ Visualizing Information: Region hot-spots from Region Metric

◇ Read, Write, StoreSize

Monitor & Operations

hosted by



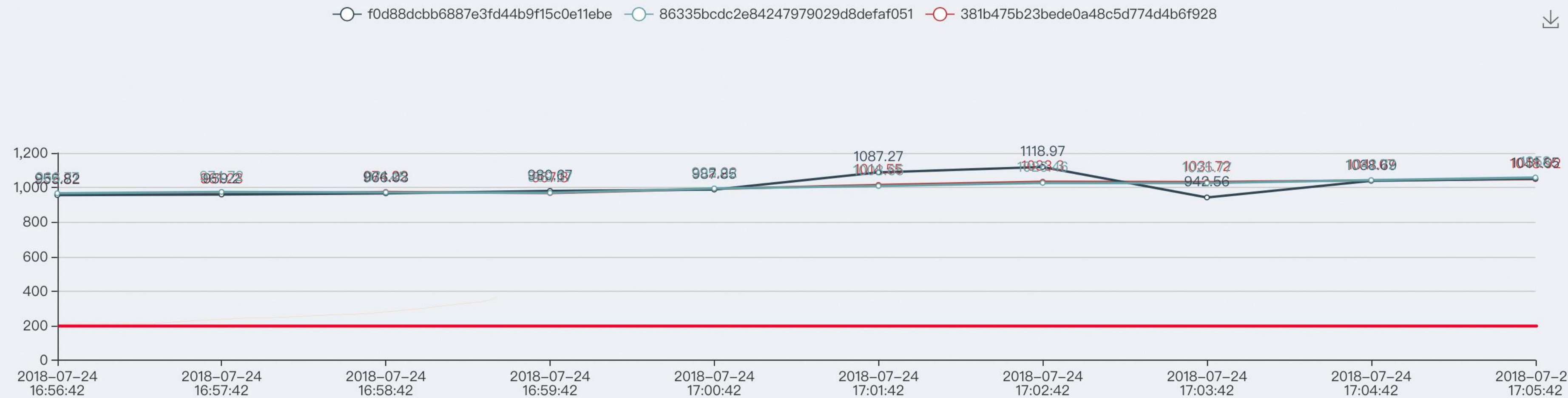
The highest 10
&
The lowest 10

GALAXY:OMEGA_ORALOG_CACHE region信息显示如下

regionName	startKey	endKey
GALAXY:OMEGA_ORALOG_CACHE,,1509349932814.86335bcd2e84247979029d8defaf051.		55555555
GALAXY:OMEGA_ORALOG_CACHE,55555555,1509349932814.381b475b23bede0a48c5d774d4b6f928.	55555555	aaaaaaaa
GALAXY:OMEGA_ORALOG_CACHE,aaaaaaaa,1509349932814.f0d88dccb6887e3fd44b9f15c0e11ebe.	aaaaaaaa	

Read Write StoreSize

ReadRequest/s



Monitor & Operations

hosted by



- ◇ Automatic computation: input/Output Size per sec.Store Size..
- ◇ Recommended values: Region Count

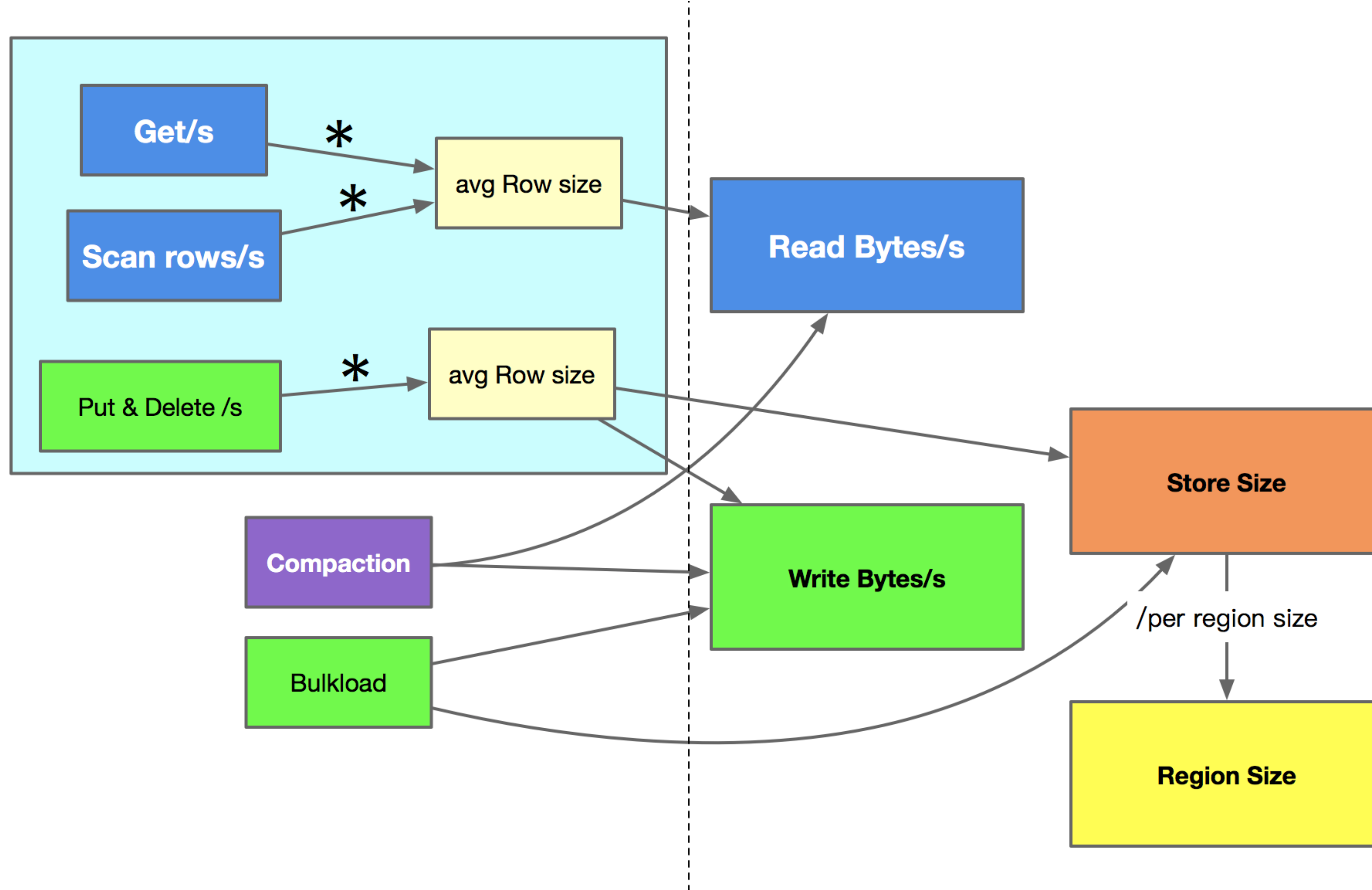
HBase Metric

- ◇ (1).read request、write request、Mutate_99th_percentile、Get_99th_percentile、ScanTime_99th_percentile、receivedBytes、sentBytes
- ◇ (2).compactionQueueLength、flushQueueLength、splitQueueLength
- ◇ (3).jvm metrics...

- ◇ slow request: Rpc p99、gc、compaction queue 、slow log

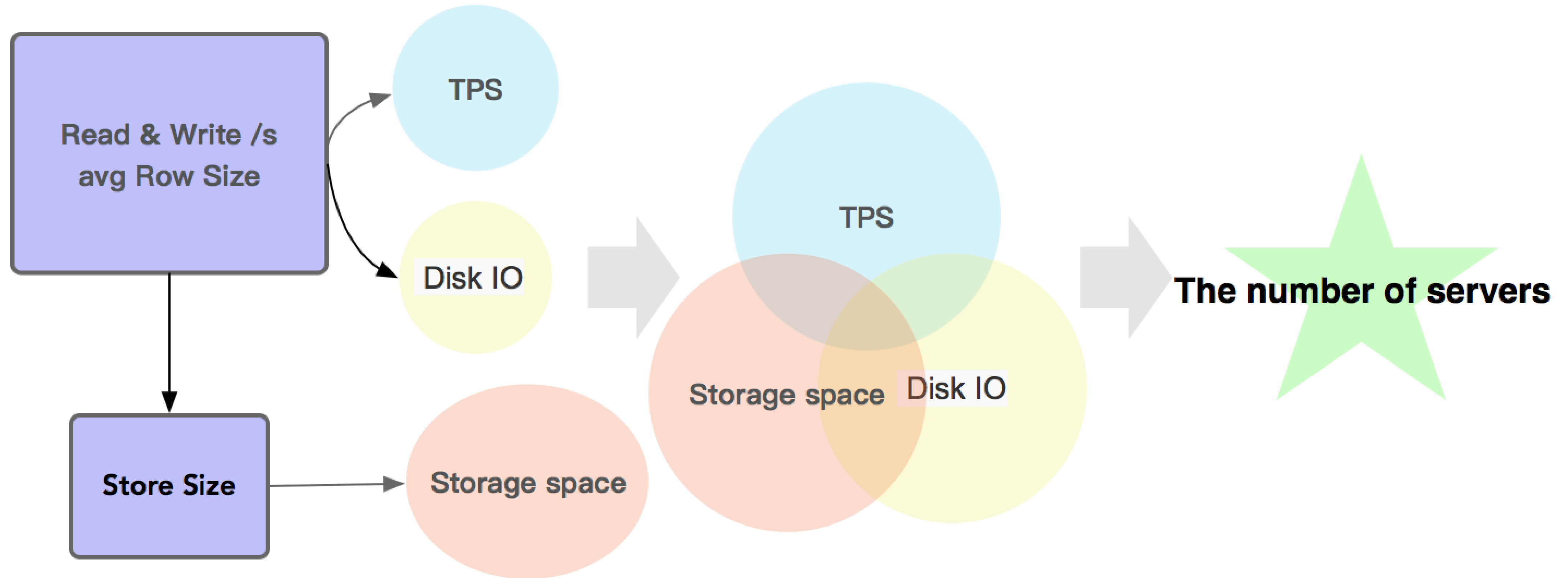
Monitor & Operations

hosted by



Monitor & Operations

hosted by



Monitor & Operations

hosted by



yaojingyi@didichuxing.com

hosted by



Thanks