

Security Level:

Faster HBase queries

Introducing hindex – Secondary indexes for HBase

ApacheCon North America 2014

www.huawei.com

Rajeshbabu Chintaguntla
rajeshbabu@apache.org

HUAWEI TECHNOLOGIES CO., LTD.



\$ whoami

- Apache HBase Committer
- Senior Software Engineer, Huawei India R&D
- Developed **hindex** – Secondary indexes for HBase

Agenda

- **HBase – A brief introduction**
- **Introduction to hindex**
- **Usage**
- **Test Results**

Agenda

- **HBase – A brief introduction**
 - Introduction to hindex
 - Usage
 - Test Results

Apache HBase

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

- an open-source, distributed, versioned, non-relational database
- modeled after Google's BigTable
- leverages distributed data storage provided by HDFS
- allows random, read/write access to data in HDFS

- GOAL: hosting very large tables - billions of rows X millions of columns

HBase – Introduction

Sorted

lexicographically

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

HBase – Introduction

Sorted

Sparse

any number of columns

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M		SE	534	1230	...
135	Ram	M		SSE			...
141	Anu	M			123		...
142	Pia	F			326		...
145	Jay		SOA	SA	521	4300	...
148	Suma		SOA	SSE		1550	...
152	Som	M	OIH	SE		1270	...
...

HBase – Introduction

Sorted

Sparse

Multi-dimensional

data is versioned

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SSE	534	1230	...
135	Ram	M		TL			...
141	Anu	M			123		...
142	Pia	F			326		...
145	Jay		SOA	SA	521	4300	...
148	Suma		SOA	SSE		1550	...
152	Som	M	OIH	SSE		1270	...
...

Key=123		cf1	cf2
name	T1	Raj	
gender	T1	M	
dept	T1	OIH	
title	T1	SE	
title	T2	SSE	
mobile	T1	534	
salary	T3	1230	
dob	T1	19880830	

(key, column family, column, timestamp) -> value

HBase – Introduction

Sorted

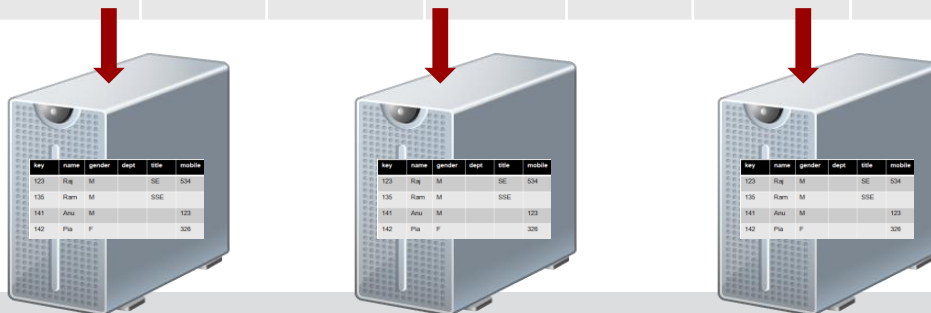
Sparse

Multi-dimensional

Distributed

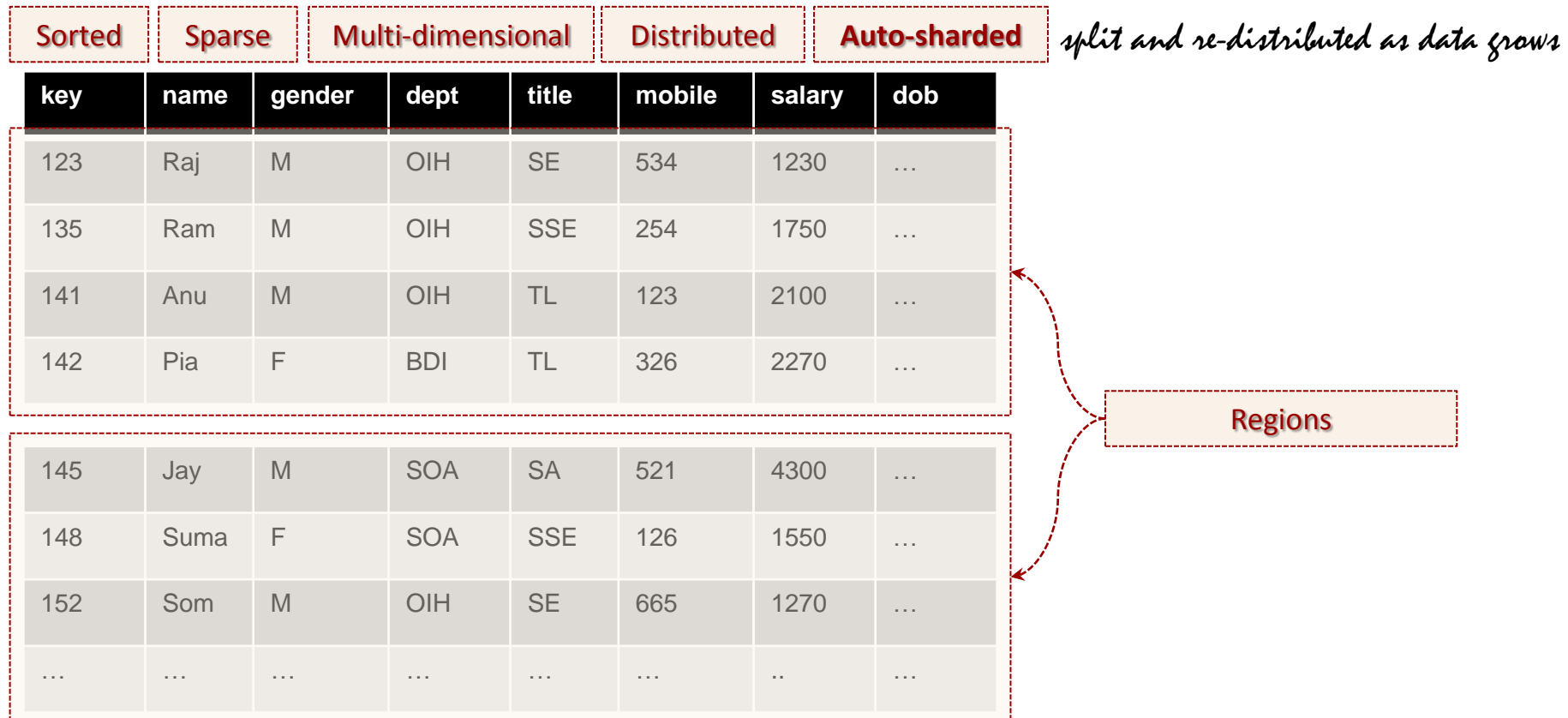
leverages HDFS, data replicated across nodes

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M		SE	534	1230	...
135	Ram	M		SSE			...
141	Anu	M			123		...
142	Pia	F			326		...
145	Jay		SOA	SA	521	4300	...
148	Suma		SOA	SSE		1550	...
152	Som	M	OIH	SE		1270	...
...



Protects against failing node

HBase – Introduction



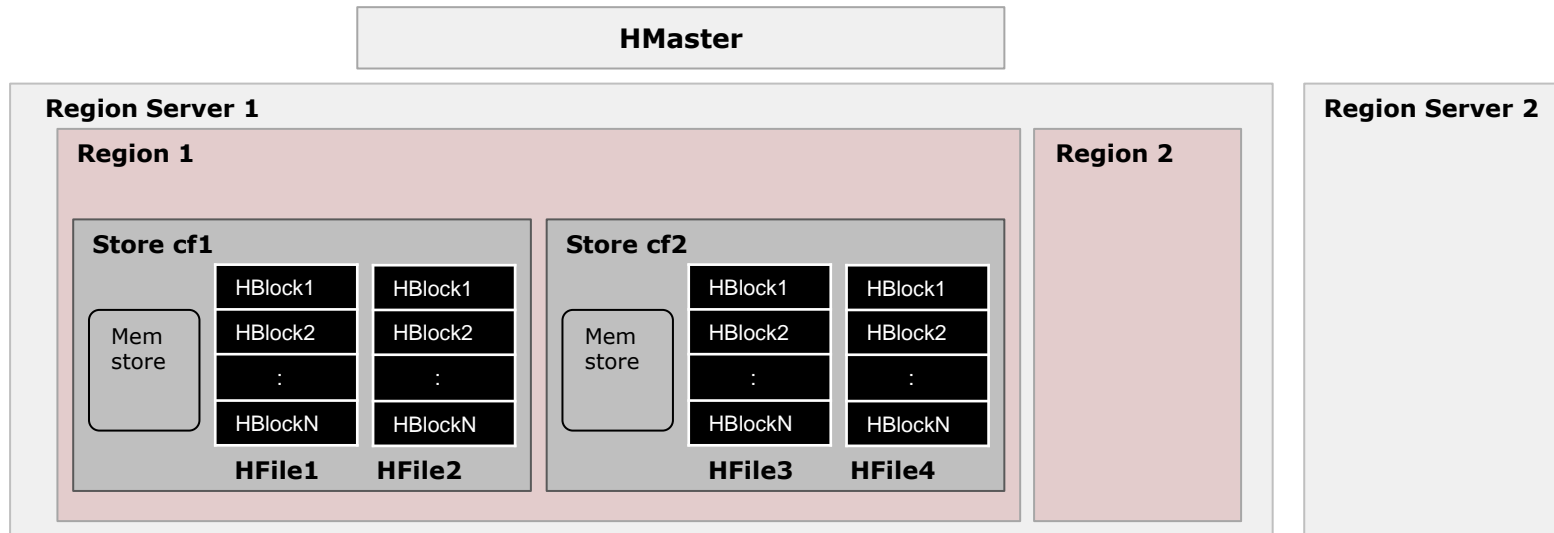
HBase – Introduction

Sorted	Sparse	Multi-dimensional	Distributed	Auto-sharded			
key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

metadata	
Region	Range
R1	120-145
R2	145-170
R3	...
...	...

Regions

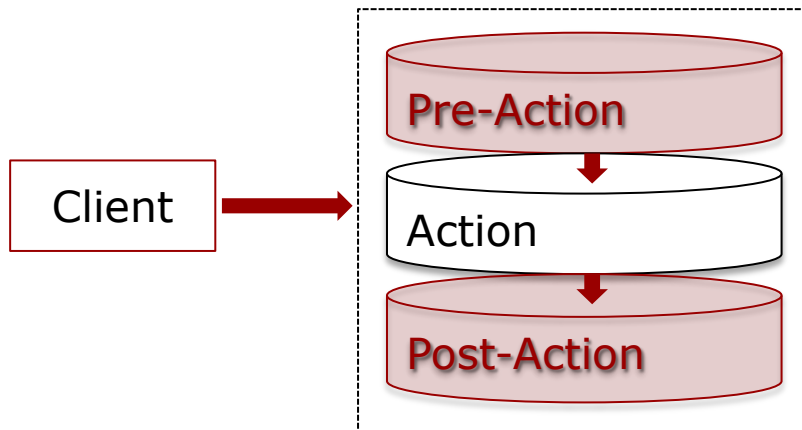
HBase – Introduction



- Master, region servers and zookeeper
- Table horizontally divided into regions
- Columns grouped into Column families – Vertical partition of tables
- Memstore, HFiles in DFS. HFiles logically split into smaller blocks. Data read write happen as blocks
- MapReduce integration

Coprocessors

- Allow to run client-supplied code on server-side,
- Can extend the functionality of HBase without changing the kernel



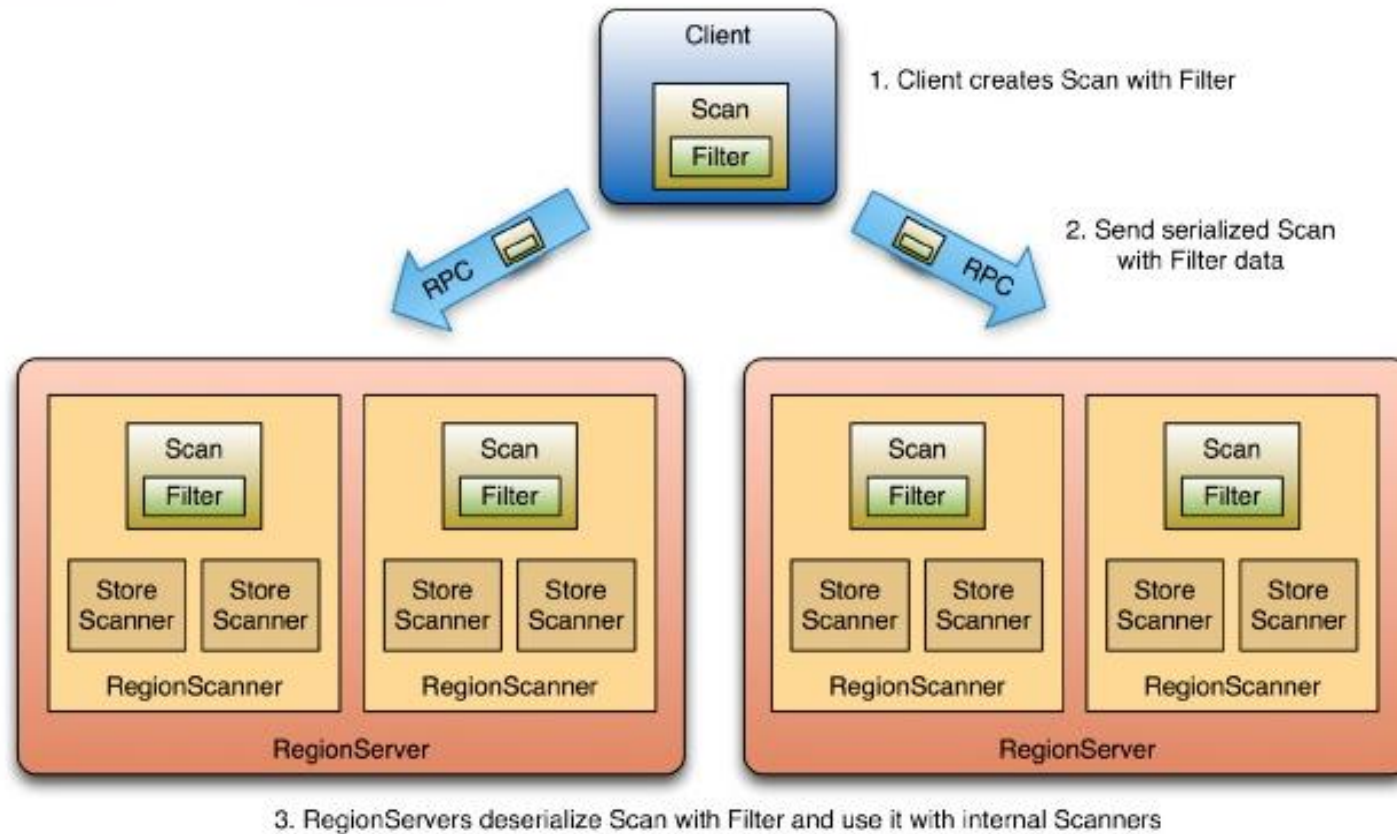
- **Observers – Like triggers**

- Runs extended functionality before or after an action through hooks provided by coprocessor framework.

- **EndPoints – Like Stored procedures**

- Can run any time from client
- The endpoint implementation will then be executed remotely at the target region(s)
- results from those executions will be returned to the client.

Filters



Source: Lars, George, HBase The Definitive Guide, O'Reilly Media. 2011

HBase – Query

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE KEY=141
```

NOTE: HBase does not have native support for SQL like query interface. It is used here for the sake of easy understanding. Similar query can be done using scanners and filters.

HBase – Query by Rowkey

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE KEY=141
```


HBase – Query by Column (w/o Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

HBase – Query by Column (w/o Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

HBase – Query by Column (w/o Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

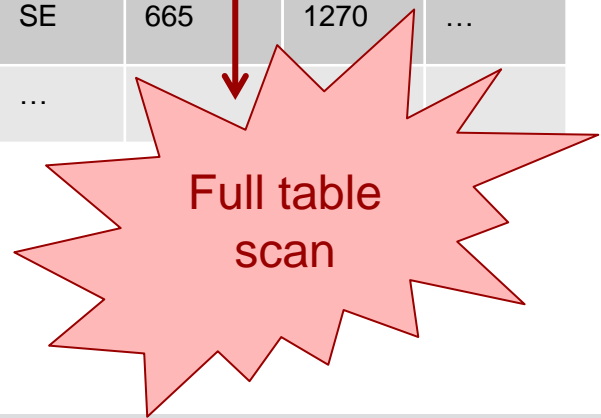
```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```



HBase – Query by Column (w/o Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

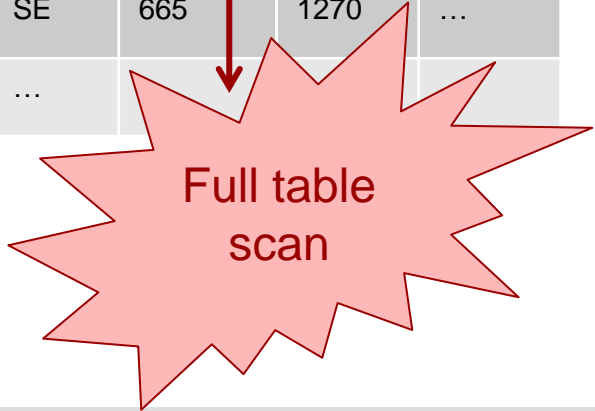


- Billions of records → full table scan evil to performance

HBase – Query by Column (w/o Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```



- Billions of records → full table scan evil to performance
- Side effects – Client timeouts, lease expiring

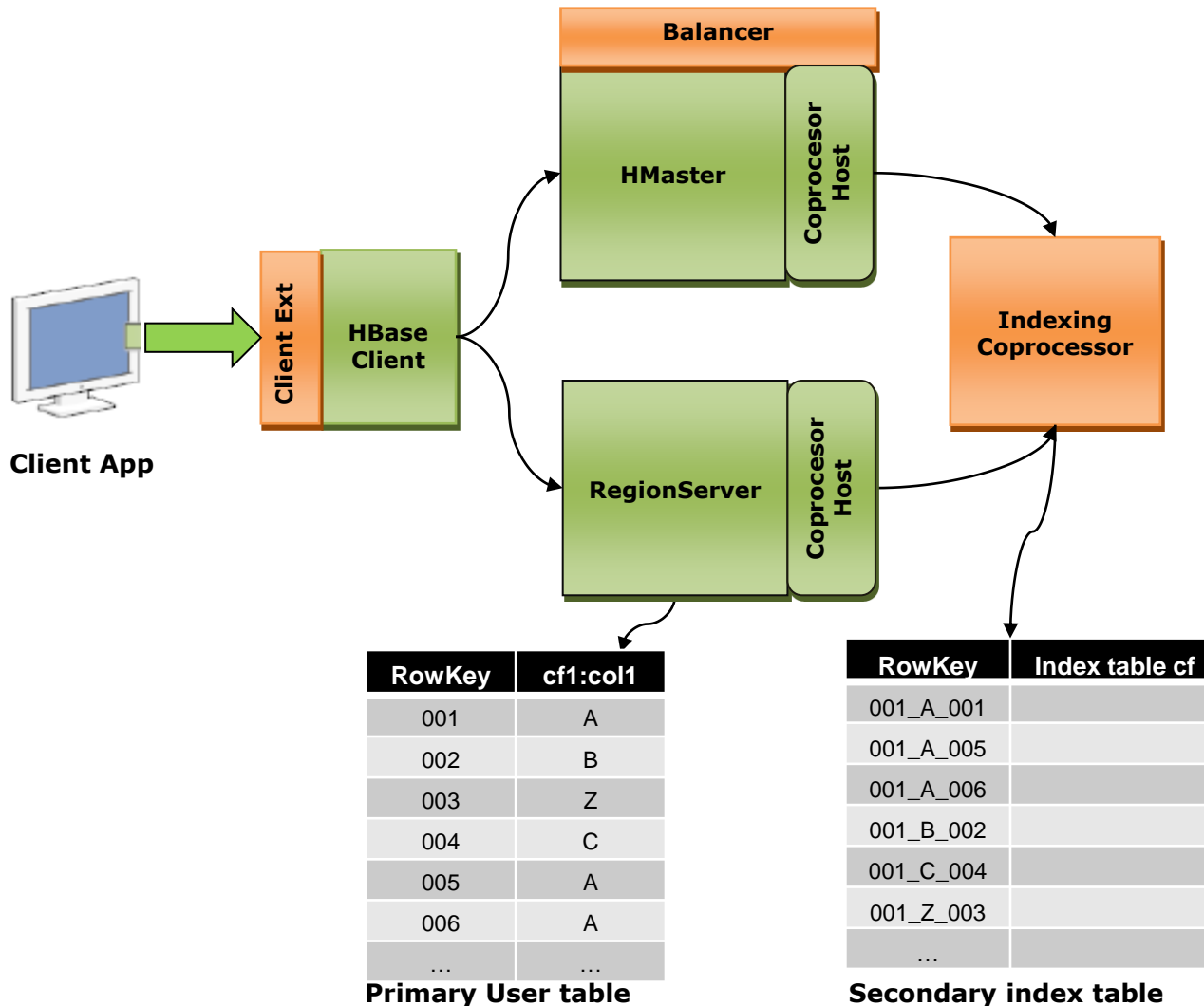
Agenda

- HBase – A brief introduction
- **Introduction to hindex**
- Usage
- Test Results

Introducing hindex

- Coprocessor based server side implementation of secondary indexing solution
- Separate index table, used by all indexes of a table.
- Region wise indexing (aka local indexing)
- Custom load balancer co-locates the index table regions with actual table regions.
- Index table rowkey construction is:
region start key + index name+ indexed column(s) value + user table rowkey

hindex: Architecture



- Coprocessor handles the index data
- A custom `LoadBalancer` does collocation
- Client Extn allows specifying index details while creating table, not needed for read/write

HBase – Query by Column (w/ Index)

key	name	gender	dept	title	mobile	salary	dob
123	Raj	M	OIH	SE	534	1230	...
135	Ram	M	OIH	SSE	254	1750	...
141	Anu	M	OIH	TL	123	2100	...
142	Pia	F	BDI	TL	326	2270	...
145	Jay	M	SOA	SA	521	4300	...
148	Suma	F	SOA	SSE	126	1550	...
152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

Index column: MOBILE

HBase – Query by Column (w/ Index)

idx1		key	name	gender	dept	title	mobile	salary	dob
123	141	123	Raj	M	OIH	SE	534	1230	...
254	135	135	Ram	M	OIH	SSE	254	1750	...
326	142	141	Anu	M	OIH	TL	123	2100	...
534	123	142	Pia	F	BDI	TL	326	2270	...
126	148	145	Jay	M	SOA	SA	521	4300	...
521	145	148	Suma	F	SOA	SSE	126	1550	...
665	152	152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

HBase – Query by Column (w/ Index)

idx1	key	name	gender	dept	title	mobile	salary	dob
123_141	123	Raj	M	OIH	SE	534	1230	...
254_135	135	Ram	M	OIH	SSE	254	1750	...
326_142	141	Anu	M	OIH	TL	123	2100	...
534_123	142	Pia	F	BDI	TL	326	2270	...
126_148	145	Jay	M	SOA	SA	521	4300	...
521_145	148	Suma	F	SOA	SSE	126	1550	...
665_152	152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

To ensure user table row key order when a column value repeated in multiple rows

HBase – Query by Column (w/ Index)

idx1	key	name	gender	dept	title	mobile	salary	dob
123_141	123	Raj	M	OIH	SE	534	1230	...
254_135	135	Ram	M	OIH	SSE	254	1750	...
326_142	141	Anu	M	OIH	TL	123	2100	...
534_123	142	Pia	F	BDI	TL	326	2270	...
126_148	145	Jay	M	SOA	SA	521	4300	...
521_145	148	Suma	F	SOA	SSE	126	1550	...
665_152	152	Som	M	OIH	SE	665	1270	...
...

```
SELECT NAME  
FROM PERSON  
WHERE MOBILE=123
```

HBase – Query by Column (w/ Index)

idx1	key	name	gender	dept	title	mobile	salary	dob
123_141	123	Raj	M	OIH	SE	534	1230	...
254_135	135	Ram	M	OIH	SSE	254	1750	...
326_142	141	Anu	M	OIH	TL	123	2100	...
534_123	142	Pia	F	BDI	TL	326	2270	...
126_148	145	Jay	M	SOA	SA	521	4300	...
521_145	148	Suma	F	SOA	SSE	126	1550	...
665_152	152	Som	M	OIH	SE	665	1270	...
...

Regions

Index maintained per region, not globally

HBase – Query by Column (w/ Index)

idx1	key	name	gender	dept	title	mobile	salary	dob
123_141	123	Ra			SE	534	1230	...
254_135	135	Ram	M	OIH	SSE	254	1750	...
326_142	141	Anu	M	OIH	TL	123	2100	...
534_123	142	Pia	F	BDI	TL	326	2270	...
126_148	145	Jay	M	SOA	SA	521	4300	...
521_145	148	Suma	F	SOA	SSE	126	1550	...
665_152	152	Som	M	OIH	SE	665	1270	...
...

Network calls avoided

Regions

Index maintained per region, not globally

HBase – Query by Column (w/ Index)

idx1	key	name	gender	dept	title	mobile	salary	dob
123_141	123	Raj	M	OIH	SE	534	1230	...
254_135	135	Ram	M	OIH	SSE	254	1750	...
326_142	141	Anu	M	OIH	TL	123	2100	...
534_123	142	Pia	F	BDI	TL	326	2270	...
126_148	145	Jay	M	SOA	SA	521	4300	...
521_145	148	Suma	F	SOA	SSE	126	1550	...
665_152	152	Som	M	OIH	SE	665	1270	...
...

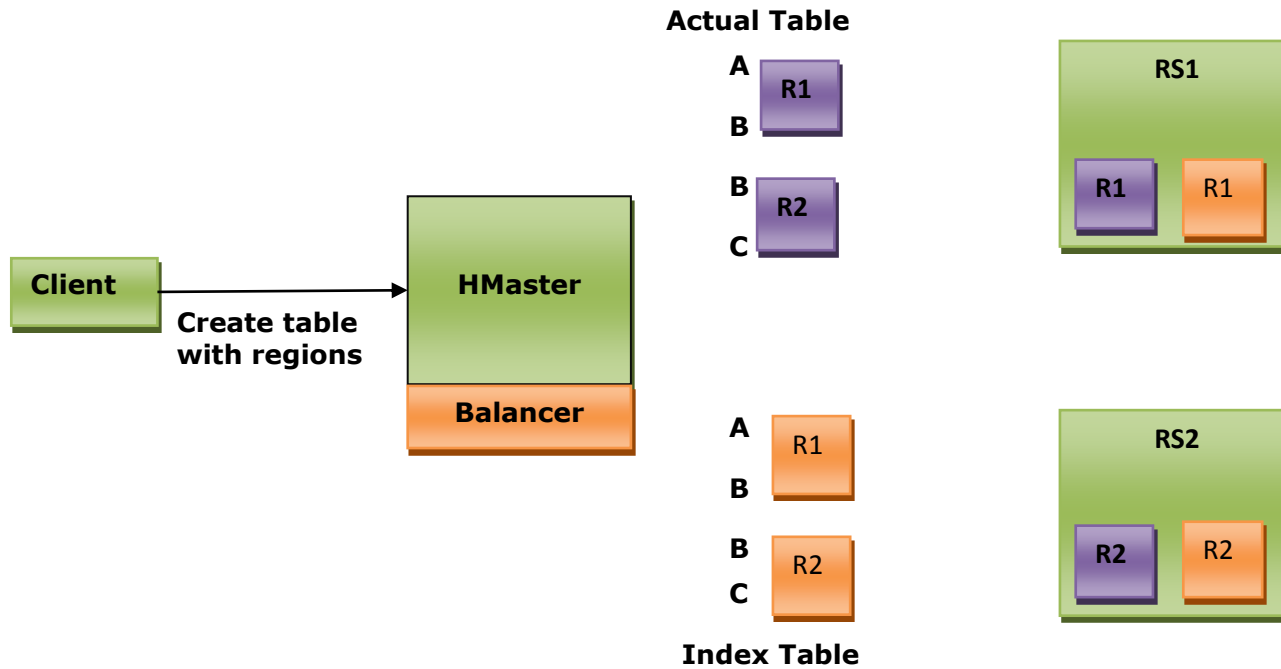
Regions

Index maintained per region, not globally

Handle

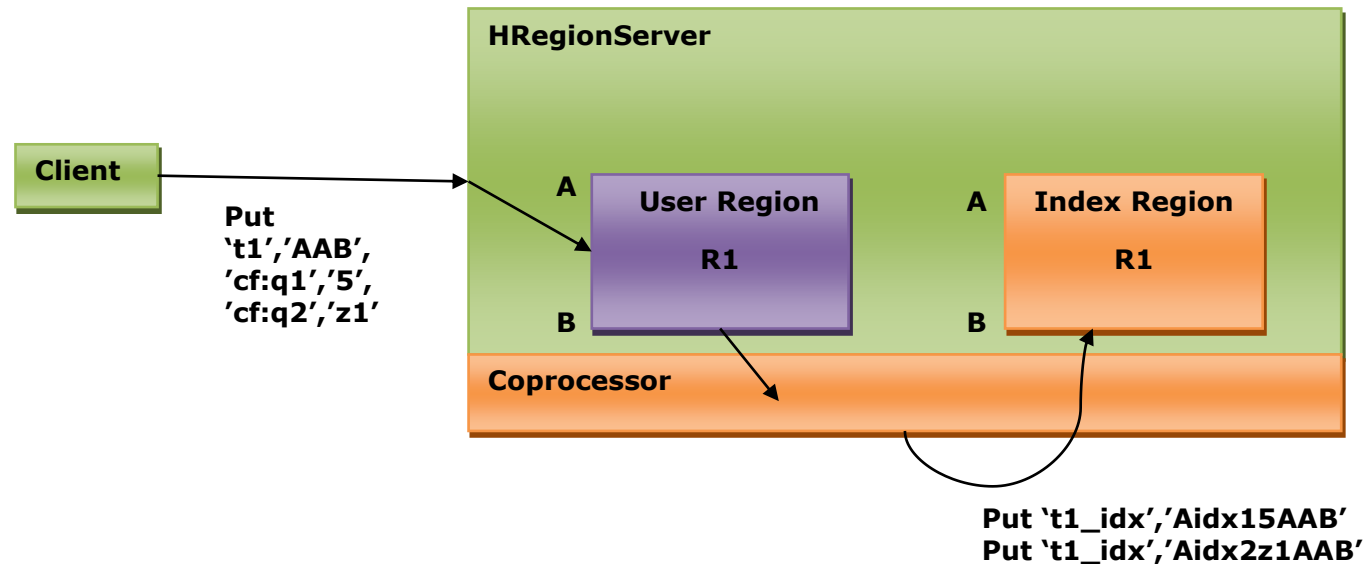
- Region Movement
- Region Splits

Regions Co-location



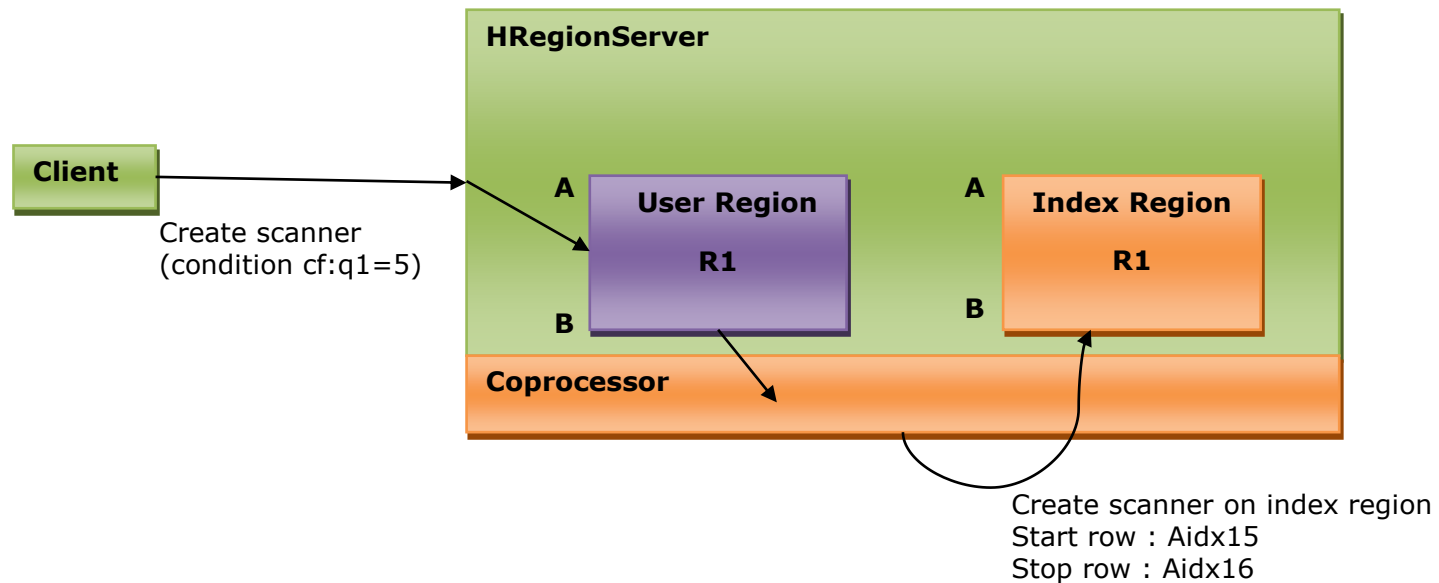
Put operation

- Table-> t1 & Column family -> cf
- Index-> **idx1(cf:q1) & idx2(cf:q2)**
- Index table -> t1_idx



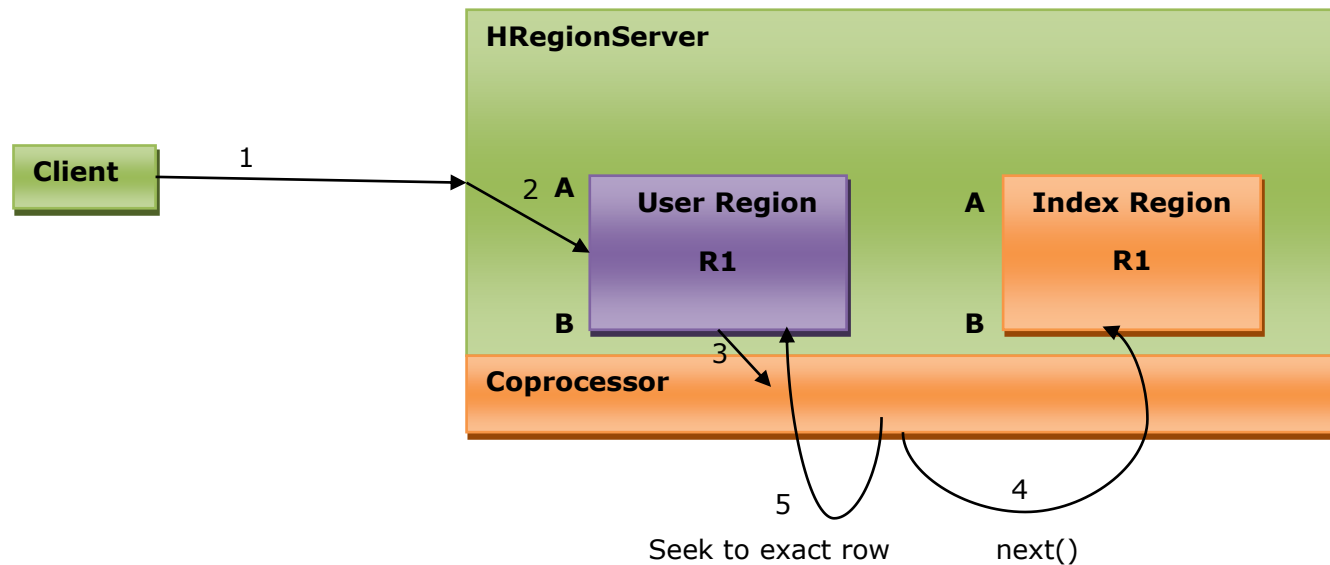
Scan Operation

1) Create scanner for index region at server side



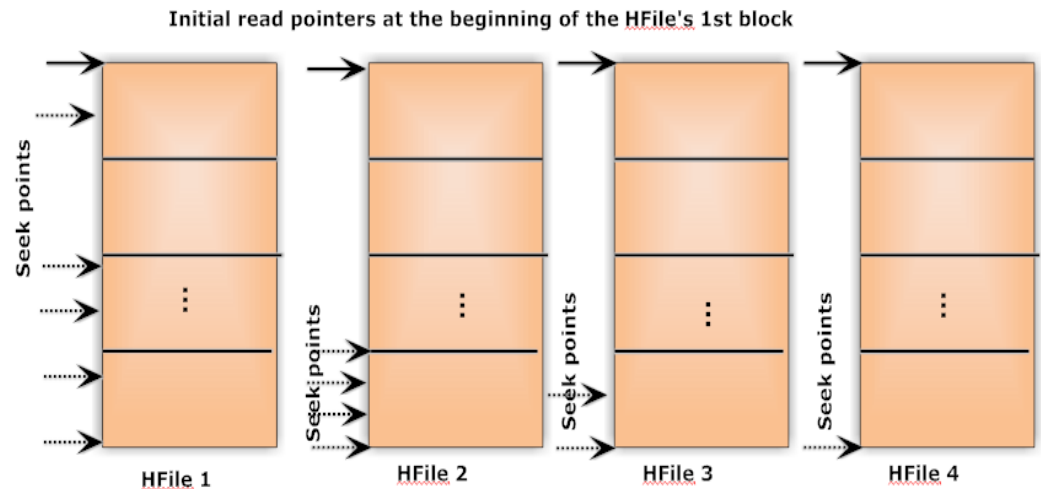
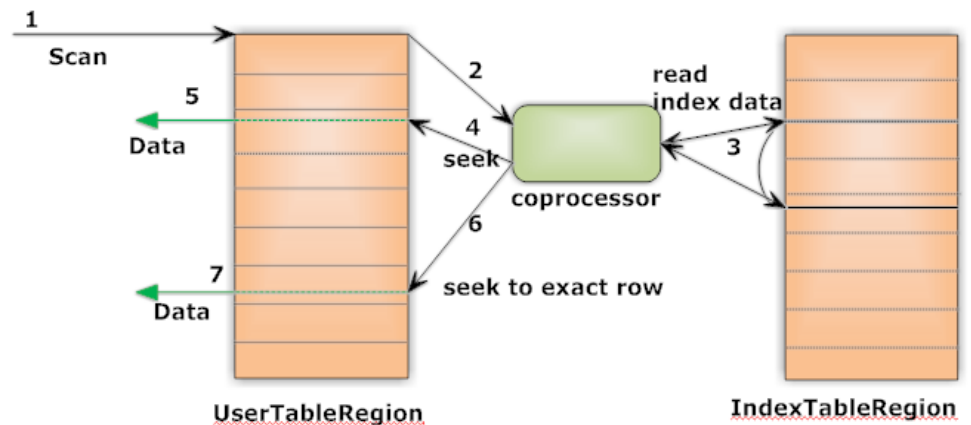
Scan Operation

2) Scan index table at server side and seek to exact rows in the user table



Scan Operation

- Coprocessors read index and seek to exact row in the user table
- Doing seeks on HFiles based on rowkey obtained from index data
- HFiles reads as block by block
- Default block size is 64kb
- Skipping block reads from HDFS where data not at all present
- Some times skipping full HFile
- No need to read index details back to client avoiding network extra usage.



hindex: Usage

```
SELECT NAME  
FROM PERSON  
WHERE (DEPT='OIH' OR TITLE='TL')  
AND (400 > MOBILE AND MOBILE > 500)
```

Filters (with equal or range conditions)

AND → Filters list with MUST_PASS_ALL
OR → Filters list with MUST_PASS_ONE

HBase – Query with AND

idx1	idx2
BDI_142	TL_141
OIH_123	TL_142
OIH_141	TL_152
OIH_142	SA_145
OIH_152	SE_123
SOA_145	SSE_135
SOA_148	SSE_148
...	...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	OIH	SSE	254
141	Anu	M	OIH	TL	123
142	Pia	F	BDI	TL	326
145	Jay	M	SOA	SA	521
148	Suma	F	SOA	SSE	126
152	Som	M	OIH	TL	665
...

```
SELECT NAME  
FROM PERSON  
WHERE DEPT=OIH  
AND TITLE=TL
```

HBase – Query with AND

idx1_BDI_142
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_145
idx1_SOA_148
...
idx2_TL_141
idx2_TL_142
idx2_TL_152
idx2_SA_145
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	OIH	SSE	254
141	Anu	M	OIH	TL	123
142	Pia	F	BDI	TL	326
145	Jay	M	SOA	SA	521
148	Suma	F	SOA	SSE	126
152	Som	M	OIH	TL	665
...

```
SELECT NAME  
FROM PERSON  
WHERE DEPT=OIH  
AND TITLE=TL
```

- 1) Single index table per user table, easier to collocate
- 2) Index table row keys have index name to store each index data together

HBase – Query with AND

idx1_BDI_142

idx1_OIH_123

idx1_OIH_141

idx1_OIH_142

idx1_OIH_152

idx1_SOA_145

idx1_SOA_148

...

idx2_TL_141

idx2_TL_142

idx2_TL_152

idx2_SA_145

...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	OIH	SSE	254
141	Anu	M	OIH	TL	123
142	Pia	F	BDI	TL	326
145	Jay	M	SOA	SA	521
148	Suma	F	SOA	SSE	126
152	Som	M	OIH	TL	665
...

```
SELECT NAME  
FROM PERSON  
WHERE DEPT=OIH  
AND TITLE=TL
```

Create two scanners

- 1) start: **idx1_OIH**, end: **idx1_OII**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND



key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND

- idx1_BDI_160
- idx1_OIH_123**
- idx1_OIH_141
- idx1_OIH_142
- idx1_OIH_152
- idx1_SOA_135
- idx1_SOA_148
- Idx1_SOA_150
- ...

- **idx2_TL_142**
- idx2_TL_145
- Idx2_TL_150
- idx2_TL_152
- Idx2_TL_160
- idx2_SA_141
- ...

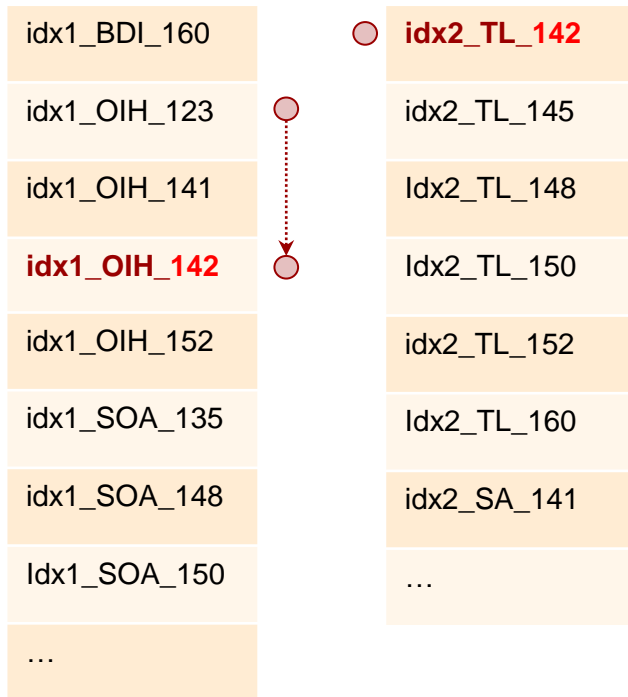
123 < 142
scanner-1 can jump to **idx1_OIH_142**

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
			SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

SELECT NAME
FROM PERSON
WHERE **DEPT=OIH**
AND TITLE=TL

two scanners
1) start: **idx1_OIH**, end: **idx1_OIH**
2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND



key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

idx2_TL_142
idx2_TL_145
Idx2_TL_148
Idx2_TL_150
idx2_TL_152
Idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay				
148	Suma				
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

Pia

142 = 142 (meets our condition)
 - Can fetch required data
 - Move both scanners to next

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

idx2_TL_142
idx2_TL_145
Idx2_TL_148
Idx2_TL_150
idx2_TL_152
...

152 > 145
scanner-2 can jump to **idx2_TL_152**

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
			SOA	TL	521
			SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

Pia

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
idx1_SOA_150
...

idx2_TL_142
Idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
Idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay				
148	Suma				
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

Pia

Som

152 = 152 (meets our condition)
 - Can fetch required data
 - Move both scanners to next

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with AND

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

idx2_TL_142
Idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141



Scanner-1 reaches end
Close both scanners

two scanners
1) start: **idx1_OIH**, end: **idx1_OIH**
2) Start: **idx2_TL**, end: **idx2_TM**

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
155	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

Pia
Som

HBase – Query with AND (w/ 2-column index)

idx1_BDI_TL_160
idx1_OIH_SA_141
idx1_OIH_TL_142
idx1_OIH_TL_152
idx1_SOA_SSE_135
idx1_SOA_TL_145
Idx1_SOA_TL_148
Idx1_SOA_TL_150
Idx1_SOA_TL_150
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME  
FROM PERSON  
WHERE DEPT=OIH  
AND TITLE=TL
```

1 scanner:
start: **idx1_OIH_TL**, end: **idx1_OIH_TM**

HBase – Query with AND (w/ 2-column index)

idx1_BDI_TL_160
idx1_OIH_SA_141
idx1_OIH_TL_142
idx1_OIH_TL_152
idx1_SOA_SSE_135
idx1_SOA_TL_145
idx1_SOA_TL_148
idx1_SOA_TL_150
idx1_SOA_TL_150
...



key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME  
FROM PERSON  
WHERE DEPT=OIH  
AND TITLE=TL
```

1 scanner:

start: **idx1_OIH_TL**, end: **idx1_OIH_TM**

HBase – Query with AND (w/ 2-column index)

- idx1_BDI_TL_160
- idx1_OIH_SA_141
- idx1_OIH_TL_142
- idx1_OIH_TL_152
- idx1_SOA_SSE_135
- idx1_SOA_TL_145
- idx1_SOA_TL_148
- idx1_SOA_TL_150
- idx1_SOA_TL_150
- ...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
141	Anu	M	SOA	SSE	254
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

Meets condition
- Can fetch required data

SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL

Pia

1 scanner:
start: **idx1_OIH_TL**, end: **idx1_OIH_TM**

HBase – Query with AND (w/ 2-column index)

- idx1_BDI_TL_160
- idx1_OIH_SA_141
- idx1_OIH_TL_142
- idx1_OIH_TL_152
- idx1_SOA_SSE_135
- idx1_SOA_TL_145
- idx1_SOA_TL_148
- Idx1_SOA_TL_150
- Idx1_SOA_TL_150
- ...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
			OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL

- Pia
- Som

Meets condition
- Can fetch required data

1 scanner:
start: idx1_OIH_TL, end: idx1_OIH_TM

HBase – Query with AND (w/ 2-column index)

- idx1_BDI_TL_160
- idx1_OIH_SA_141
- idx1_OIH_TL_142
- idx1_OIH_TL_152
- idx1_SOA_SSE_135
- idx1_SOA_TL_145
- Idx1_SOA_TL_148
- Idx1_SOA_TL_150
- Idx1_SOA_TL_150
- ...



Scanner reaches end
Close the scanner

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
AND TITLE=TL
```

- Pia
- Som

1 scanner:
start: idx1_OIH_TL, end: idx1_OIH_TM

HBase – Query with OR

- idx1_BDI_160
- idx1_OIH_123**
- idx1_OIH_141
- idx1_OIH_142
- idx1_OIH_152
- idx1_SOA_135
- idx1_SOA_148
- Idx1_SOA_150
- ...

- idx2_TL_142**
- Idx2_TL_145
- idx2_TL_148
- ...
- idx2_TL_160
- idx2_SA_141
- ...

123 <= 142
 - Get required results
 - Move scanner-1 to next till exceeds **142**

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
...
...
...
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160	idx2_TL_142	key	name	gender	dept	title	mobile
idx1_OIH_123	idx2_TL_145	123	Raj	M	OIH	SE	534
idx1_OIH_141	idx2_TL_148	135	Ram	M	SOA	SSE	254
idx1_OIH_142					OIH	SA	123
idx1_OIH_152					OIH	TL	326
idx1_SOA_135	idx2_TL_160	145	Jay	M	SOA	TL	521
idx1_SOA_148	idx2_SA_141	148	Suma	F	SOA	TL	126
idx1_SOA_150	...	150		M	SOA	TL	325
...		152	Som	M	OIH	TL	665
		160	Su	F	BDI	TL	928
	

SELECT NAME FROM PERSON WHERE DEPT=OIH OR TITLE=TL

123 <= 142
 - Get required results
 - Move scanner-1 to next till exceeds 142

Raj

two scanners
 1) start: idx1_OIH, end: idx1_OIH
 2) Start: idx2_TL, end: idx2_TM

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...



idx2_TL_142
Idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
...

141 <= 142
 - Get required results
 - Move scanner-1 to next till exceeds **142**

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
...	SOA	TL	521
...	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

Raj

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

- idx1_BDI_160
- idx1_OIH_123
- idx1_OIH_141**
- idx1_OIH_142
- idx1_OIH_152
- idx1_SOA_135
- idx1_SOA_148
- Idx1_SOA_150
- ...

- **idx2_TL_142**
- Idx2_TL_145
- idx2_TL_148
- idx2_TL_150
- idx2_TL_152
- ...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
...	SOA	TL	521
...	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

SELECT NAME
FROM PERSON
WHERE **DEPT=OIH**
OR TITLE=TL

- Raj
- Anu

141 <= 142
- Get required results
- Move scanner-1 to next till exceeds **142**

two scanners
1) start: **idx1_OIH**, end: **idx1_OIH**
2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

- idx1_BDI_160
- idx1_OIH_123
- idx1_OIH_141
- idx1_OIH_142**
- idx1_OIH_152
- idx1_SOA_135
- idx1_SOA_148
- Idx1_SOA_150
- ...



- idx2_TL_142**
- Idx2_TL_145
- idx2_TL_148
- idx2_TL_150
- idx2_TL_152
- idx2_TL_160
- idx2_SA_141

142 == 142
 - Get required results
 - Move both scanners in this case

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
...	OIH	TL	665
...	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

- Raj
- Anu

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

idx2_TL_142
Idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
			OIH	TL	665
			BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

Raj
Anu
Pia

142 == 142
 - Get required results
 - Move both scanners in this case

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...



idx2_TL_142
idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
...	BDI			TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

Raj
Anu
Pia

152 >= 145 (scanner-2 is behind)
 - Get required results
 - Move scanner-2 to next till exceeds **152**

two scanners
 1) start: **idx1_OIH**, end: **idx1_OIH**
 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
idx1_SOA_150
...

idx2_TL_142
idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
...	BDI			TL	928
...

SELECT NAME
FROM PERSON
WHERE **DEPT=OIH**
OR TITLE=TL

Raj
Anu
Pia
Jay

152 >= 145 (scanner-2 is behind)
- Get required results
- Move scanner-2 to next till exceeds **152**

two scanners
1) start: **idx1_OIH**, end: **idx1_OIH**
2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

○ idx2_TL_142
● idx2_TL_145
● idx2_TL_148
○ idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

```
SELECT NAME
FROM PERSON
WHERE DEPT=OIH
OR TITLE=TL
```

Raj
Anu
Pia
Jay
Suma

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

○ idx2_TL_142
● idx2_TL_145
● idx2_TL_148
● idx2_TL_150
○ idx2_TL_152
idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

SELECT NAME
FROM PERSON
WHERE **DEPT=OIH**
OR TITLE=TL

Raj
Anu
Pia
Jay
Suma

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

HBase – Query with OR

idx1_BDI_160
idx1_OIH_123
idx1_OIH_141
idx1_OIH_142
idx1_OIH_152
idx1_SOA_135
idx1_SOA_148
Idx1_SOA_150
...

idx2_TL_142
idx2_TL_145
idx2_TL_148
idx2_TL_150
idx2_TL_152
idx2_TL_160
idx2_SA_141
...

key	name	gender	dept	title	mobile
123	Raj	M	OIH	SE	534
135	Ram	M	SOA	SSE	254
141	Anu	M	OIH	SA	123
142	Pia	F	OIH	TL	326
145	Jay	M	SOA	TL	521
148	Suma	F	SOA	TL	126
150		M	SOA	TL	325
152	Som	M	OIH	TL	665
160	Su	F	BDI	TL	928
...

SELECT NAME
FROM PERSON
WHERE **DEPT=OIH**
OR TITLE=TL

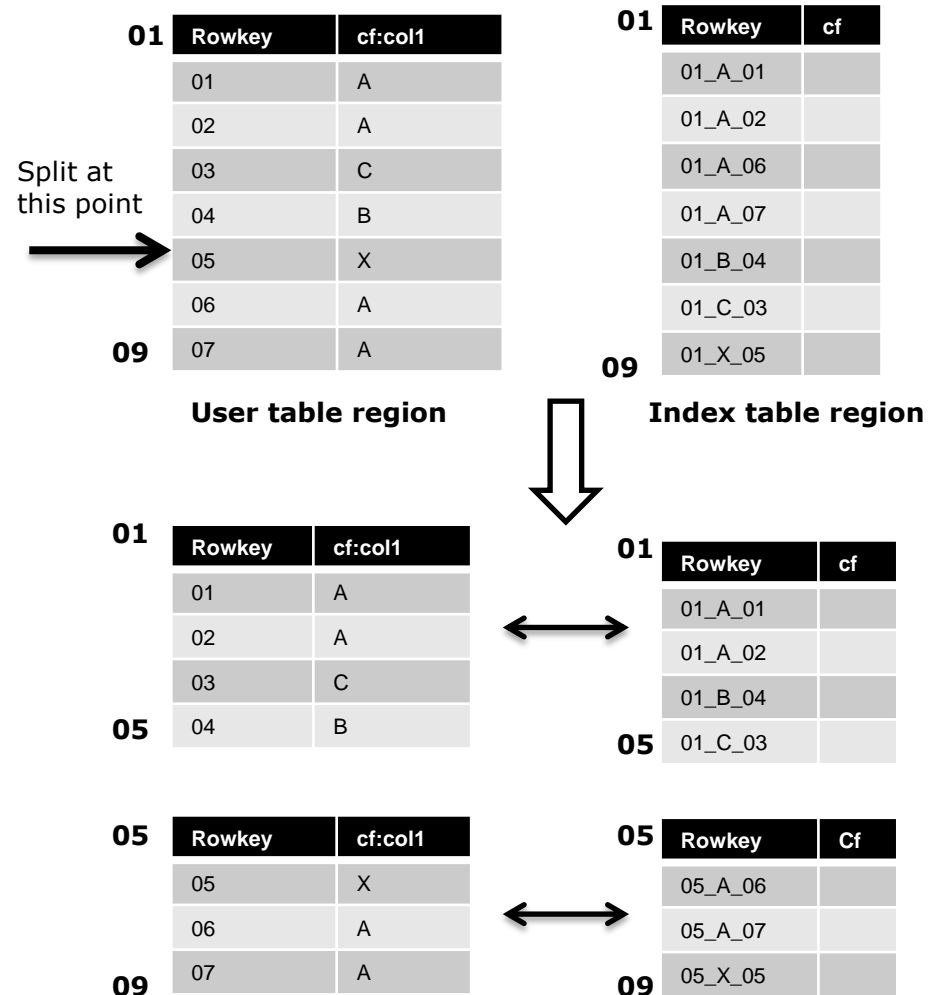
Raj
Anu
Pia
Jay
Suma
Som

two scanners

- 1) start: **idx1_OIH**, end: **idx1_OIH**
- 2) Start: **idx2_TL**, end: **idx2_TM**

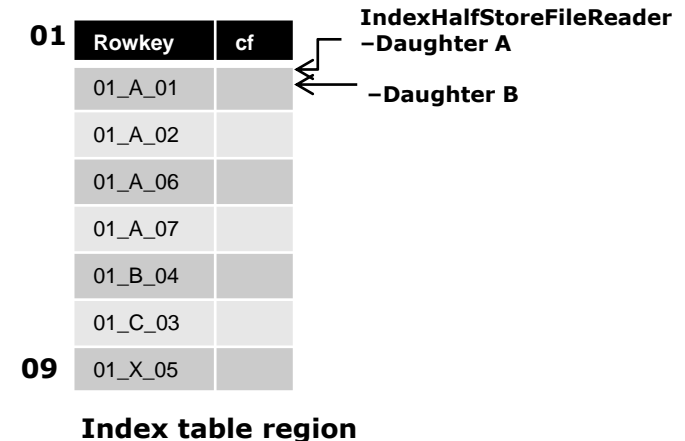
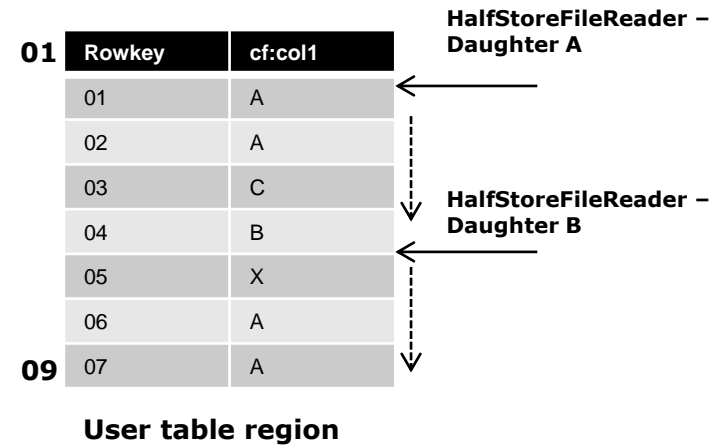
Region Split

- Explicit split on index region is avoided (using custom split policy for index tables)
- When user table region splits, corresponding index region also splits
- Split key for index region same as that of user region



Region Split

- Custom `HalfStoreFileReader` for reading index daughter regions.
- `IndexHalfStoreFileReader` – Both half store file readers start at same point i.e. beginning of file
- Checks actual table rowkey and decide KV corresponding to it or not
- `IndexHalfStoreFileReader` for daughter B - Changes the key as per the daughter start key.



Agenda

- HBase – A brief introduction
- Introduction to hindex
- **Usage**
- Test Results

Usage: Getting Started

- **For HBase 0.94.x**
 - <https://github.com/Huawei-Hadoop/hindex/>
- **For HBase 0.98 or trunk**
 - <https://issues.apache.org/jira/browse/HBASE-10222>

Usage: Configurations

Name	Value
<code>hbase.coprocessor.master.classes</code>	<code>org.apache.hadoop.hbase.index.coprocessor.master.IndexMasterObserver</code>
<code>hbase.coprocessor.region.classes</code>	<code>org.apache.hadoop.hbase.index.coprocessor.regionserver.IndexRegionObserver</code>
<code>hbase.coprocessor.wal.classes</code>	<code>org.apache.hadoop.hbase.index.coprocessor.wal.IndexWALObserver</code>
<code>hbase.master.loadbalancer.class</code>	<code>org.apache.hadoop.hbase.index.SecIndexLoadBalancer</code>
<code>hbase.use.secondary.index</code>	<code>true</code>

Usage: Creating Index

```
HBaseAdmin admin = new IndexAdmin(conf);
HTableDescriptor htd = new HTableDescriptor(TableName.valueOf(tableName));
HColumnDescriptor hcd = new HColumnDescriptor(columnFamily);
htd.addFamily(hcd);
IndexSpecification iSpec = new IndexSpecification(indexName);
iSpec.addIndexColumn(hcd, indexColumnQualifier, ValueType.String, 10);
TableIndices tableIndices = new TableIndices();
tableIndices.addIndex(iSpec);
htd.setValue(Constants.INDEX_SPEC_KEY, tableIndices.toByteArray());
admin.createTable(htd);
```

Usage: Tools

- TableIndexer tool to create index(es) for existing data

```
$HBASE_HOME/bin/hbase org.apache.hadoop.hbase.index.mapreduce.TableIndexer  
-Dtablename.to.index=table -Dtable.columns.index= 'IDX1=>cf1:[q1->datatype& length]'
```

- Bulk load tool to load user data to user table and index it at same time

```
$HBASE_HOME/bin/hbase org.apache.hadoop.hbase.index.mapreduce.IndexImportTsv  
-Dimporttsv.columns=a,b,c -Dimporttsv.bulk.output=hdfs://storefile-outputdir <tablename>  
<hdfs-data-inputdir>
```

```
$HBASE_HOME/bin/hbase org.apache.hadoop.hbase.index.mapreduce.IndexLoadIncrementalHFiles  
<hdfs://storefileoutput> <tablename>
```

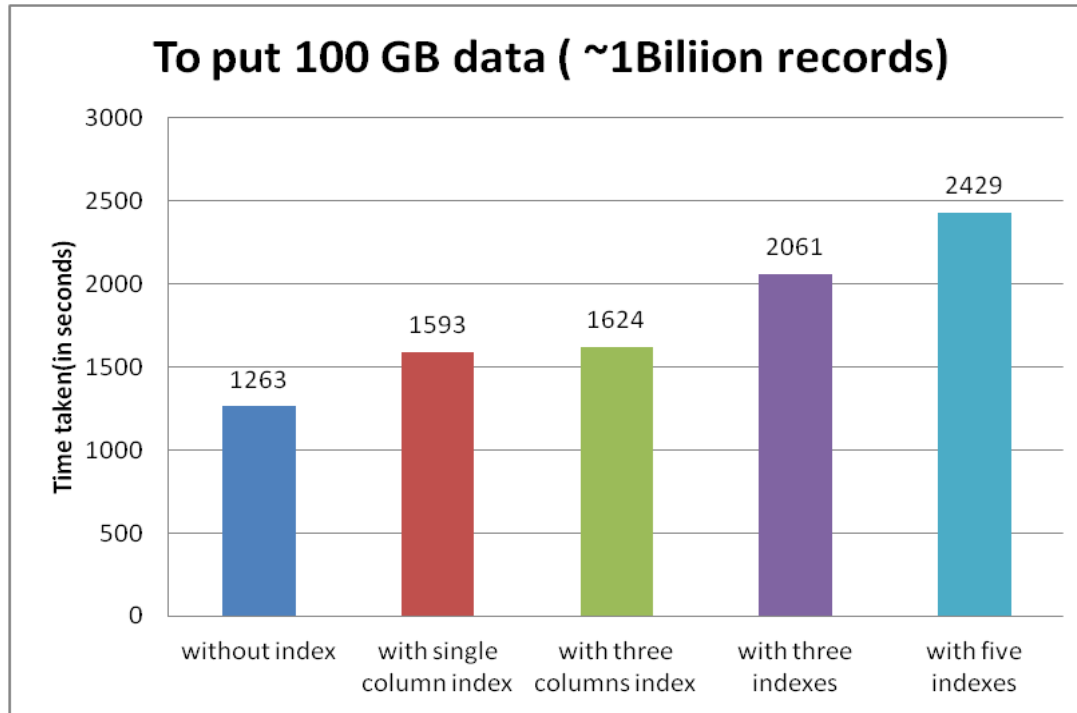
- Tool to check regions co-location and repair if any co-location mismatches

```
$HBASE_HOME/bin/hbase org.apache.hadoop.hbase.index.util.SecondaryIndexColocator
```

Agenda

- HBase – A brief introduction
- Introduction to hindex
- Usage
- **Test Results**

Test Results: Put Performance

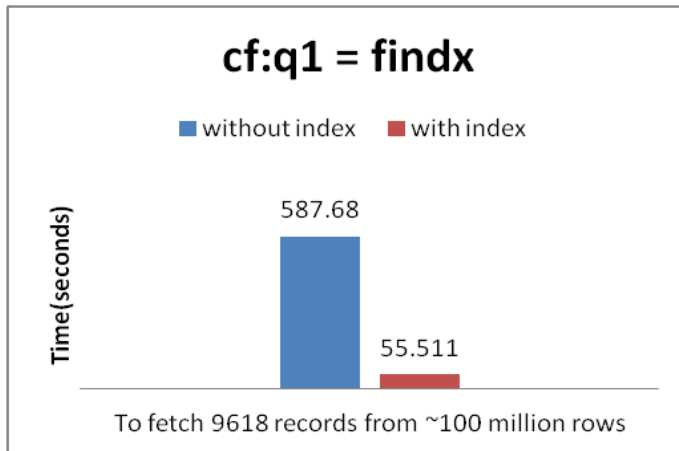


Hardware	Architecture : x86_64 CPU(s) : 24 (2 threads per core) RS Heap size: 8GB
Topology	5 Region Servers 100 Regions (user table)
Data	100 GB data 500 bytes per record

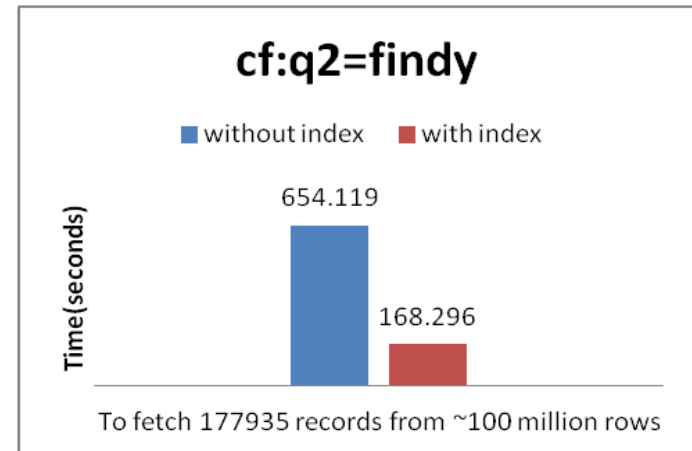
Test Results: Scan Performance

Search for a column value

idx1 → cf:q1



idx2 → cf:q2

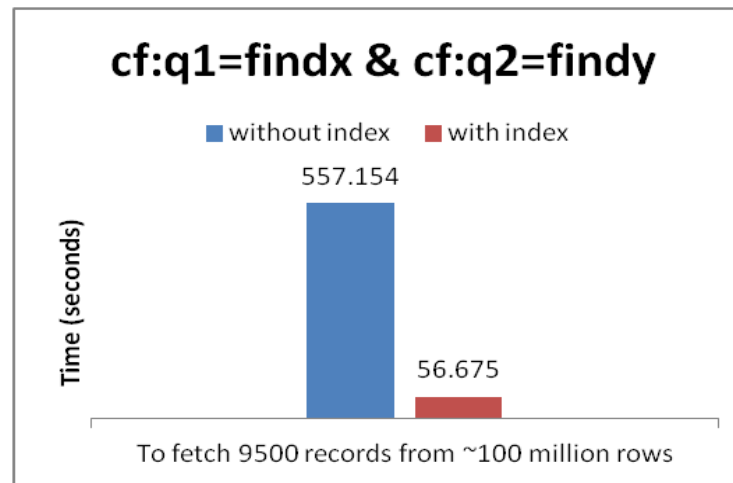


Hardware	Architecture : x86_64 CPU(s) : 24 (2 threads per core) RS Heap size: 8GB
Topology	5 Region Servers 100 Regions (user table)
Data	50 GB data 500 bytes per record

Test Results: Query with AND

idx1 → cf:q1

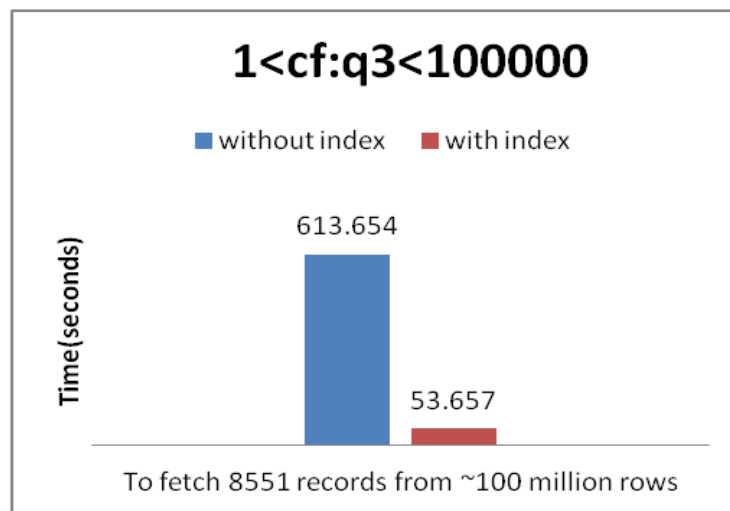
idx2 → cf:q2



Hardware	Architecture : x86_64 CPU(s) : 24 (2 threads per core) RS Heap size: 8GB
Topology	5 Region Servers 100 Regions (user table)
Data	50 GB data 500 bytes per record

Test Results: Scan w/ Range Query

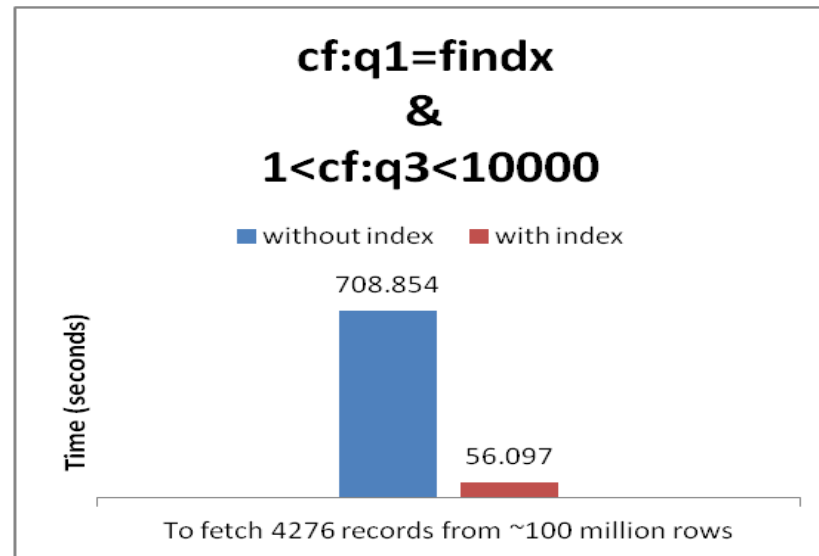
idx3 → cf:q3



Hardware	Architecture : x86_64 CPU(s) : 24 (2 threads per core) RS Heap size: 8GB
Topology	5 Region Servers 100 Regions (user table)
Data	50 GB data 500 bytes per record

Test Results: Scan w/ Multi Column Index

idx4 -> cf:q1,cf:q3



Hardware	Architecture : x86_64 CPU(s) : 24 (2 threads per core) RS Heap size: 8GB
Topology	5 Region Servers 100 Regions (user table)
Data	50 GB data 500 bytes per record

Summary

■ Design

- Supports multiple indexes and multi-column indexes on a table
- Supports indexing on part of a column value
- Supports equal and range condition scans using index
- Supports dynamic add/drop index
- Supports hints to skip index scan or specific indexes to use in the scan.
- Intelligent Filter evaluation

■ Application usage

- No changes required to perform read and write operations.
- Use `IndexAdmin` (client extension) to perform admin operations like create, enable, disable and drop on indexed table.
- Need not perform admin operations separately on index table.

■ Upgrade/Integration

- Minimal code changes in HBase kernel. HBase version upgrade is very easy.

Roadmap

- Contribute to HBase community (In progress – refer [HBASE-9203](#))
- Integration with Apache Phoenix
- HBACK tool support for Secondary index tables
- Pluggable Scan-Evaluation

<https://github.com/Huawei-Hadoop/hindex/>

Q & A

Thank you

www.huawei.com

Copyright©2011 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.