Filtering Rows Based on Conditions

Janani Ravi
CO-FOUNDER, LOONYCORN
www.loonycorn.com

Overview

Filter records using row keys

Filter records using values in columns

Apply multiple filters to filter records

SQL vs. HBase Shell Commands

SQL	HBase Shell
select * from census	scan 'census'

SQL vs. HBase Shell Commands

SQL	HBase Shell
select name from census	scan 'census', {COLUMNS => ['personal:name']}

SQL vs. HBase Shell Commands

SQL	HBase Shell
select * from census limit 1	scan 'census', {LIMIT => 1}

SQL vs. HBase Shell Commands

SQL | **HBase Shell**

select * from census where rowkey = 1 | **get 'census', 1**

SQL vs. HBase Shell Commands

SQL | **HBase Shell**

select * from census
where rowkey > 1

SQL vs. HBase Shell Commands

SQL	HBase Shell

select * from census where name = "Mike Jones"

SQL vs. HBase Shell Commands

SQL	HBase Shell

select * from census where name = "Mike Jones" and field = "politics"

Filters allow you to control what data is returned from a scan operation

Built-in Filters

Conditions on row keys

Conditions on columns

Multiple conditions on columns

Timestamp range

Demo

Filter records based on conditions on row keys

Demo

Filter records based on conditions on column values

Demo

Filter records using multiple conditions

Summary

Understood how advanced filtering techniques work on HBase records

Learnt how to write code using the HBase Java libraries to implement these filters