Executing CRUD Operations Using the HBase Shell

Overview

Understand how data is stored in HBase

Create, Read, Update and Delete tables using the HBase shell

How Is Data Laid out in HBase?

Notification Data in a Traditional Database

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

This is a 2-dimensional data model

Notification Data in a Traditional Database

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

unique row id

column name

HBase has a **4-dimensional** data model

4-dimensional Data Model

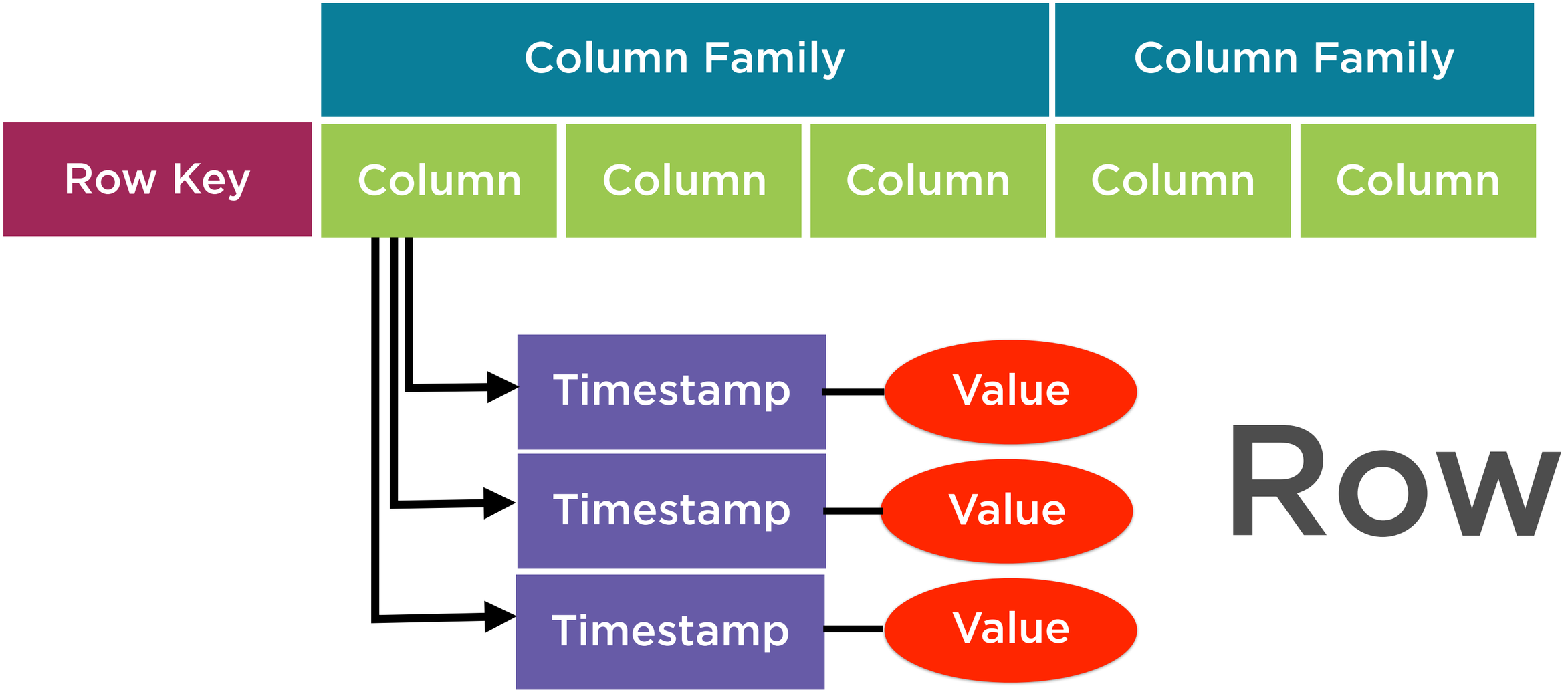
Row Key

Column Family

Column

Timestamp

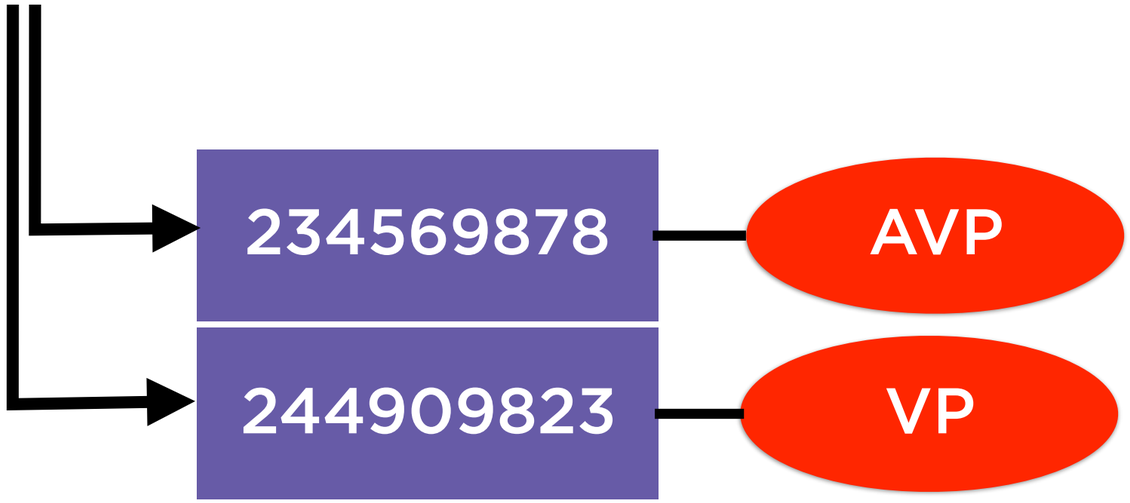
4-dimensional Data Model



A Table for Employee Data

	Work			Personal	
EmpId	Dept	Grade	Title	Name	SSN

A record for a single employee



Notification Data

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

Id	Column	Value
1	To	mike
1	Type	offer
1	Content	Offer on mobiles
2	To	john
2	Type	sale
2	Content	Redmi sale
3	To	jill
3	Type	order
3	Content	Order delivered
4	To	megan
4	Type	sale
4	Content	Clothes sale

Notification Data

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

Id	Column	Value
1	To	mike
1	Type	offer
1	Content	Offer on mobiles
2	To	john
2	Type	sale
2	Content	Redmi sale
3	To	jill
3	Type	order
3	Content	Order delivered
4	To	megan
4	Type	sale
4	Content	Clothes sale

Row Key

Notification Data

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

Id	Column	Value
1	To	mike
1	Type	offer
1	Content	Offer on mobiles
2	To	john
2	Type	sale
2	Content	Redmi sale
3	To	jill
3	Type	order
3	Content	Order delivered
4	To	megan
4	Type	sale
4	Content	Clothes sale

Column Family

Notification Data

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

Id	Column	Value
1	To	mike
1	Type	offer
1	Content	Offer on mobiles
2	To	john
2	Type	sale
2	Content	Redmi sale
3	To	jill
3	Type	order
3	Content	Order delivered
4	To	megan
4	Type	sale
4	Content	Clothes sale

Columns

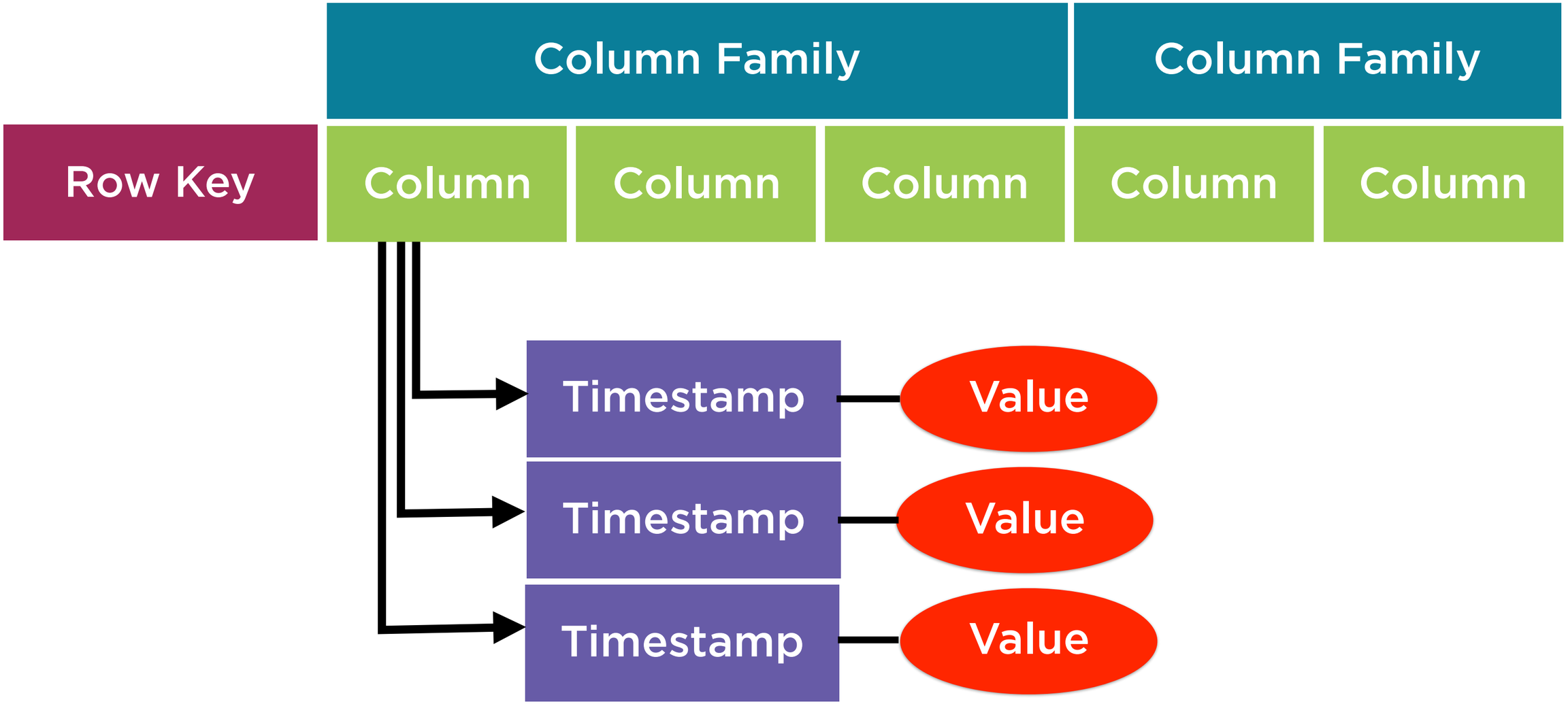
Notification Data

Id	To	Type	Content
1	mike	offer	Offer on mobiles
2	john	sale	Redmi sale
3	jill	order	Order delivered
4	megan	sale	Clothes sale

Id	Column	Value
1	To	mike
1	Type	offer
1	Content	Offer on mobiles
2	To	john
2	Type	sale
2	Content	Redmi sale
3	To	jill
3	Type	order
3	Content	Order delivered
4	To	megan
4	Type	sale
4	Content	Clothes sale

Value + Timestamp

Data Layout in HBase



Row Key

- Uniquely identifies a row
- Can be primitives, structures, arrays
- Represented internally as a byte array
- Sorted in ascending order

Column Family

All rows have the same set of column families

Each column family is stored in a separate data file

Set up at schema definition time

Can have different columns for each row

Columns are units within a column family

New columns can be added on the fly

ColumnFamily: ColumnName = Work:Department

Timestamp

Used as the version number for the values stored in a column

The value for any version can be accessed

Census Data Layout in HBase

	Personal			Professional	
Some ID	name	gender	marital_status	employed	field

Demo

Get started with simple commands on the HBase shell

- Create a table to store census data
- List and describe tables

Demo

Insert and update data using the HBase shell

- Add rows, each row represents data specific to one person
- Edit cells in a row

Demo

Access notification data using the HBase shell

- Get data in individual cells in a row
- Get data across multiple rows

Demo

Delete data using the HBase shell

- Delete data in individual cells

Demo

Delete tables using the HBase shell

- Disable tables
- Delete tables

Summary

Understood how HBase stores data internally

Used the CRUD operations on the HBase shell