

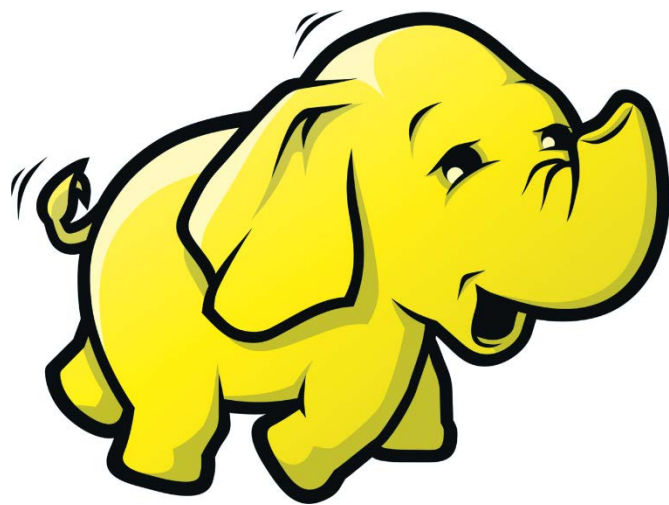
# Processing Sparse Data with HBase



Thomas M. Henson

@henson\_tm | [www.thomashenson.com](http://www.thomashenson.com)

# Overview



HBase What?

Building the Case for HBase

HBase from the Shell

Ingesting Data in HBase

Learn More About HBase

# HBase

NoSQL distributed and scalable database built on top of Hadoop. Real-time read/write in Hadoop.

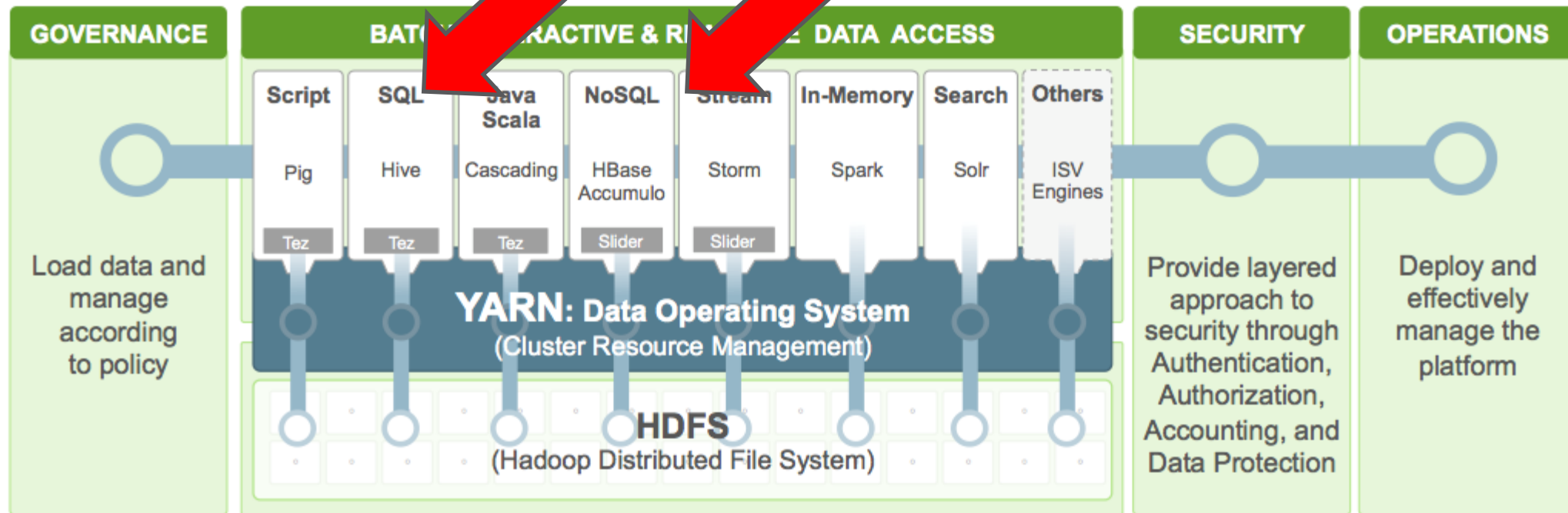


Host data with billions of rows and  
millions of columns.....

— Hbase Contributors



# HBase in Hadoop Stack



From Hortonworks: <http://hortonworks.com/hadoop/yarn/>

# How HBase Works

- Open Source
  - Apache
  - Java
- Run Modes
  - Distributed
  - Standalone
- Data Store
- Use Cases
  - Facebook
  - Netflix

APACHE  
HBASE

# Hbase Shell

\$Hbase Shell

Interactive

Create/Modify



## HBase Commands





```
$> hbase shell  
hbase>
```

## HBase Shell

Command used to enter command line interface (CLI)

```
hbase> create 'tableName', 'columnName'
```

## Create

Create table and columns

```
hbase> put 'tableName', 'rowNumber', 'columnName', 'value'
```

## Put

Inserts data in row and column

```
hbase> get 'tableName', 'rowNumber'
```

## Get

Query table and row

```
hbase> scan 'tableName'
```

## Scan

Display table columns and rows

```
hbase> disable 'tableName'
```

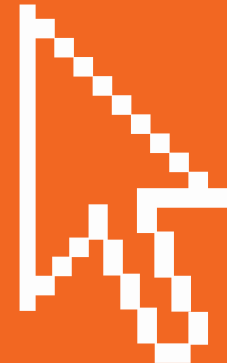
```
hbase> drop 'tablename'
```

## Drop

Delete entire table and records in table

# Hbase Shell Commands

Hbase CLI demo



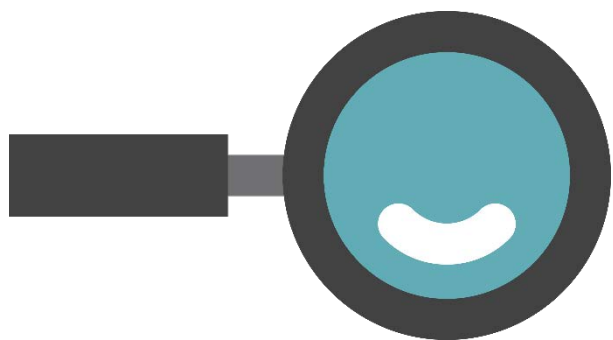


# Demo Moving Data

Moving data from csv to HBase



# Apple Daily Stock



1	Date, Open, High, Low, Close, Volume, Adj. Close
2	2015-07-09, 123.849998, 124.059998, 119.220001, 120.07, 77821600, 120.07
3	2015-07-08, 124.480003, 124.639999, 122.540001, 122.57, 60490200, 122.57
4	2015-07-07, 125.889999, 126.150002, 123.769997, 125.690002, 46716100, 125.690002
5	2015-07-06, 124.940002, 126.230003, 124.849998, 126.00, 27900200, 126.00
6	2015-07-02, 126.43, 126.690002, 125.769997, 126.440002, 27122500, 126.440002
7	2015-07-01, 126.900002, 126.940002, 125.989998, 126.599998, 30128600, 126.599998
8	2015-06-30, 125.57, 126.120003, 124.860001, 125.43, 43849800, 125.43
9	2015-06-29, 125.459999, 126.470001, 124.480003, 124.529999, 48911400, 124.529999
10	2015-06-26, 127.669998, 127.989998, 126.510002, 126.75, 42111000, 126.75
11	2015-06-25, 128.860001, 129.199997, 127.50, 127.50, 31816700, 127.50
12	2015-06-24, 127.209999, 129.800003, 127.120003, 128.110001, 54964900, 128.110001
13	2015-06-23, 127.480003, 127.610001, 126.879997, 127.029999, 30137100, 127.029999
14	2015-06-22, 127.489998, 128.059998, 127.080002, 127.610001, 33833500, 127.610001
15	2015-06-19, 127.709999, 127.82, 126.400002, 126.599998, 54181300, 126.599998
16	2015-06-18, 127.230003, 128.309998, 127.220001, 127.879997, 35241100, 127.879997
17	2015-06-17, 127.720001, 127.879997, 126.739998, 127.300003, 32768500, 127.300003
18	2015-06-16, 127.029999, 127.849998, 126.370003, 127.599998, 31404000, 127.599998
19	2015-06-15, 126.099998, 127.239998, 125.709999, 126.919998, 39842600, 126.919998

```
create 'stock', 'date', 'open', 'high', 'low', 'close',  
'volume', 'adj_close'
```

```
put 'stock', 'date', '07221993' 'open'.....
```



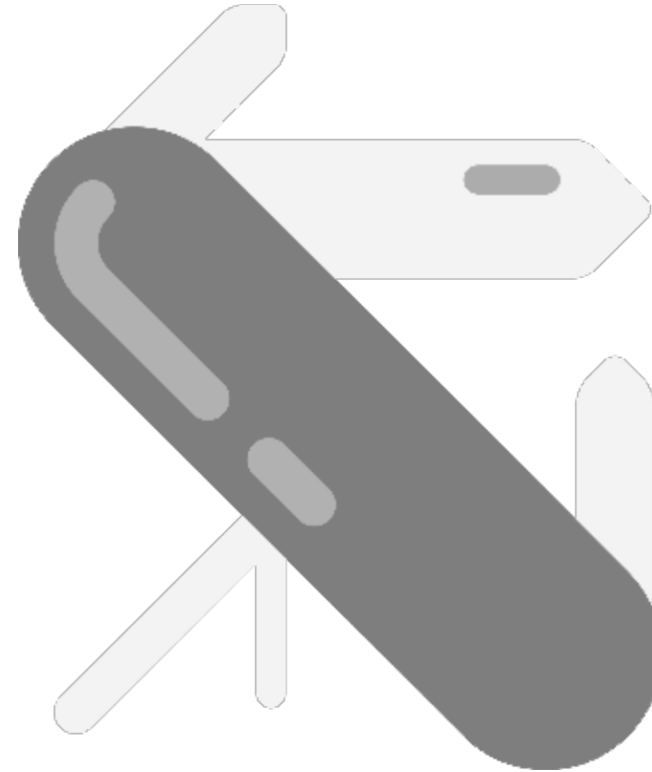
I'm happy  
to help.

# Insert .CSV into HBase

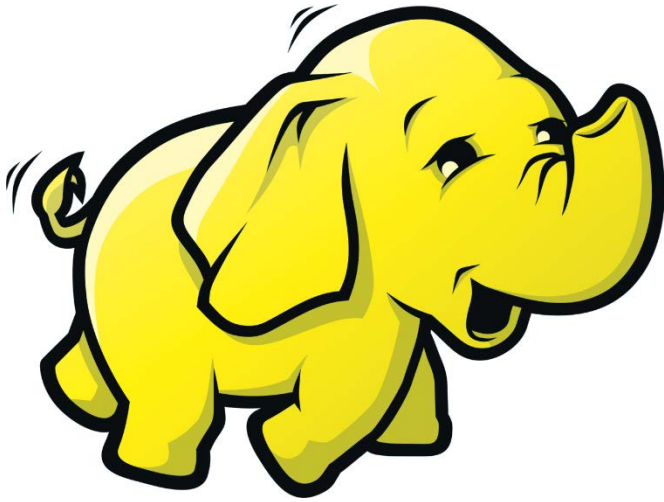


# Learn More

- Official Documentation
  - [hbase.apache.org](http://hbase.apache.org)
- Hadoop Distributors
  - Hortonworks
  - Cloudera
- Get involved
  - Join the mailing list
  - Source code



# Summary



Explained HBase NoSQL Approach

Described HBase Use Cases

Worked in the HBase Shell

Used HBase to Move Stock Data