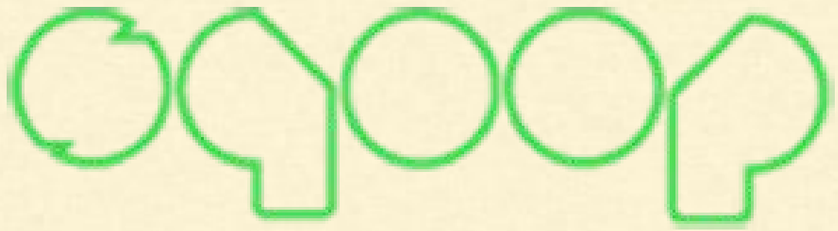
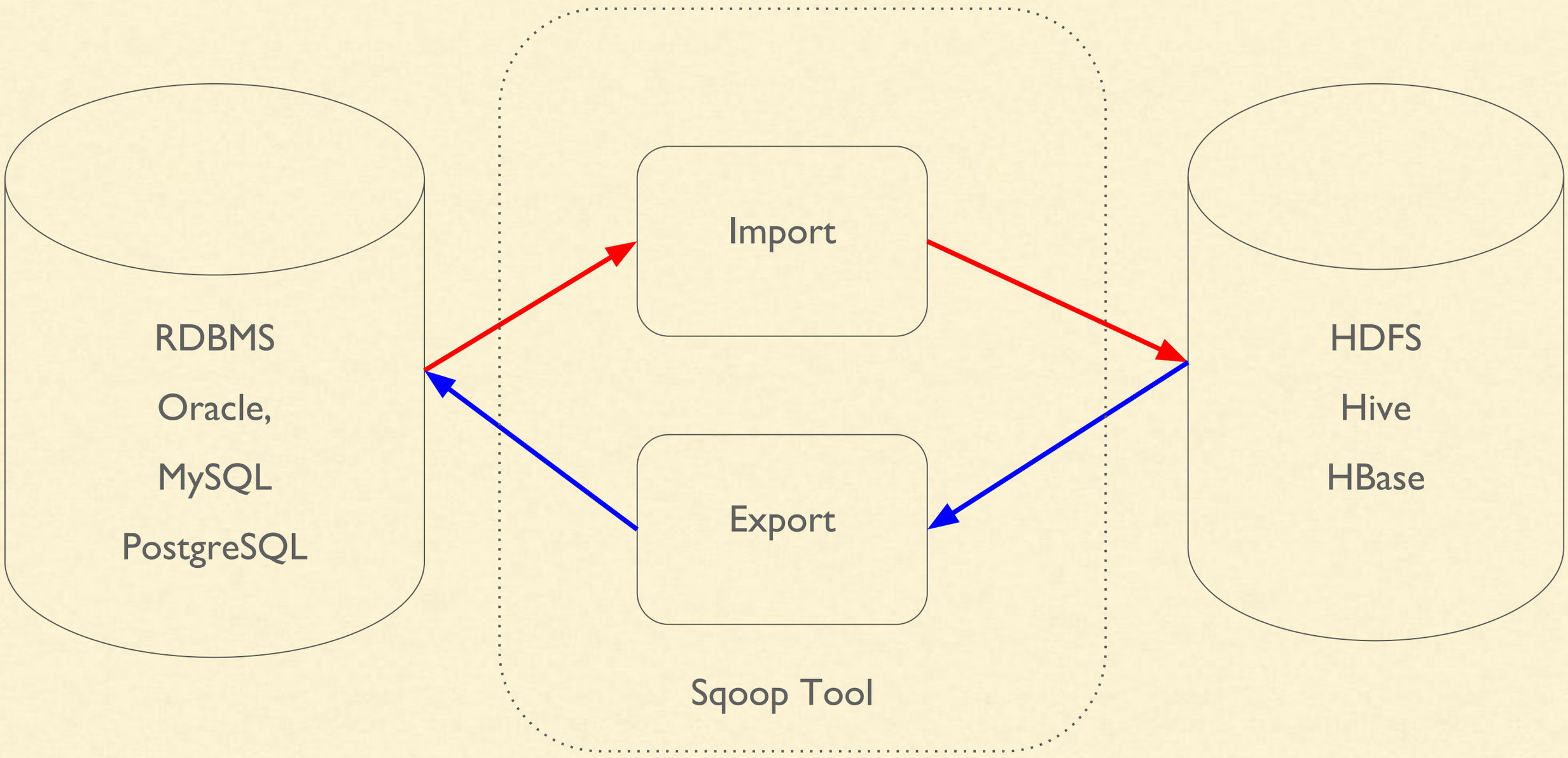


Welcome to Sqoop

Sqoop - Introduction

Open source tool to efficiently transferring bulk data between Hadoop and structured datastores such as MySQL, Oracle and HBase

Sqoop - Tools


RDBMS
Oracle,
MySQL
PostgreSQL
Import
Export
Sqoop Tool
HDFS
Hive
HBase

Sqoop - Connectors

Available Connectors:

Include MySQL, PostgreSQL, Oracle, SQL Server, DB2.
Generic JDBC Connector - any database that support jdbc
Third Party too - Netezza, Teradata

Sqoop - Help

Go to shell:

>sqoop help
Available commands:

codegen	Generate code to interact with database records
create-hive-table	Import a table definition into Hive
eval	Evaluate a SQL statement and display the results
export	Export an HDFS directory to a database table
help	List available commands
import	Import a table from a database to HDFS
import-all-tables	Import tables from a database to HDFS
job	Work with saved jobs
list-databases	List available databases on a server
list-tables	List available tables in a database
merge	Merge results of incremental imports
metastore	Run a standalone Sqoop metastore
version	Display version information

Sqoop Import - MySQL to HDFS

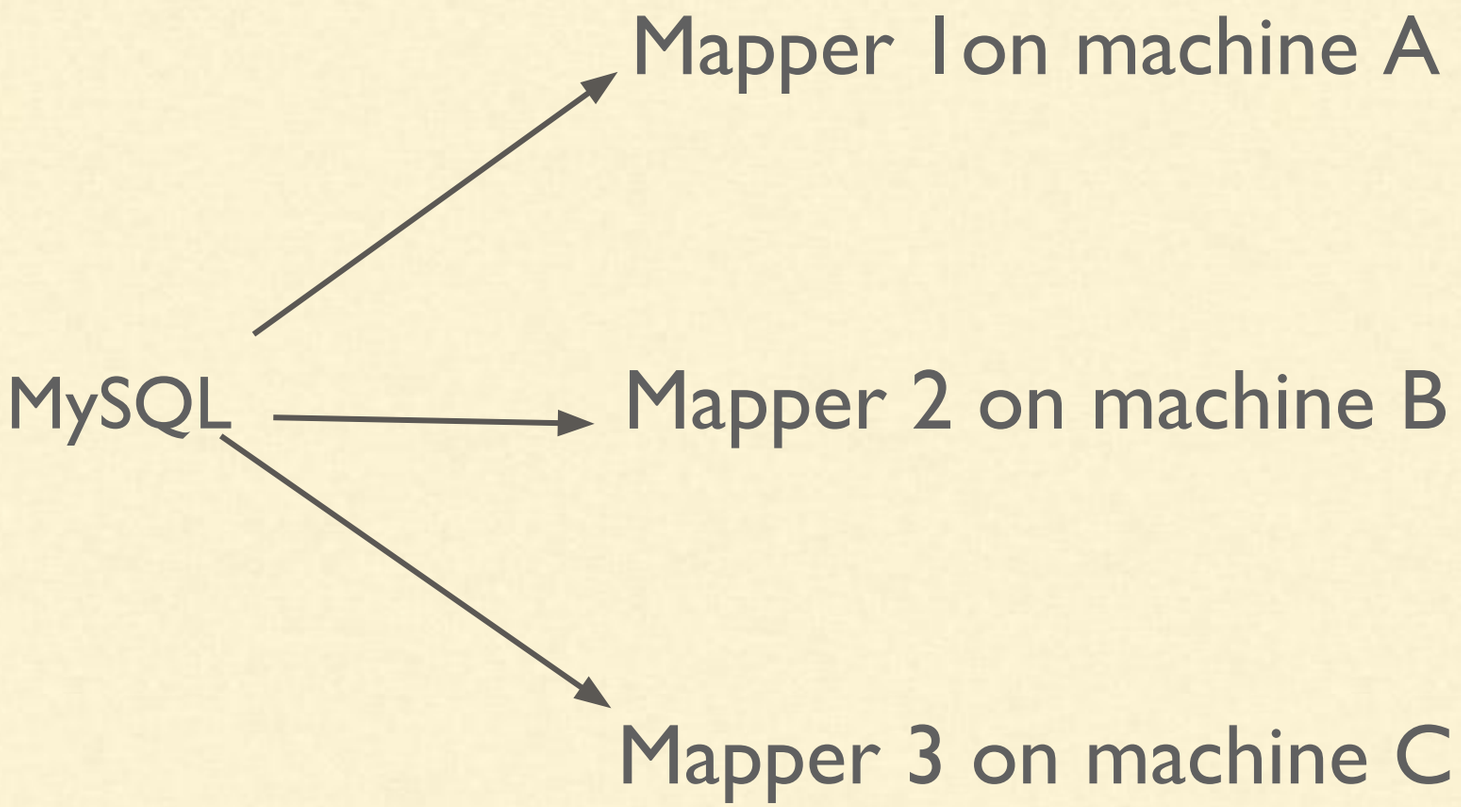
sqoop import --connect jdbc:mysql://ip-172-31-13-154/sqoopex --table widgets -m 2 --username sqoopuser -P --split-by id

Check the content of the imported File:

hadoop fs -cat widgets/part-m-00000

Also notice that widgets.java was created.

Sqoop - MySQL Connection



Sqoop Import - MySQL to Hive

sqoop import --connect jdbc:mysql://172.31.13.154/sqoopex --table widgets -m 2 --hive-import --username sqoopuser -P --hive-database sqoop_testing

Sqoop Import - MySQL to HBase

```
sqoop import --connect jdbc:mysql://172.31.13.154/sqoopex --table widgets
--hbase-table 'widgets' --column-family cf2 --username sqoopuser -P
--hbase-create-table --columns id,widget_name,price --hbase-row-key
'widget_name' -m 1
```

Sqoop Export - Hive to MySQL

Copy sales.log locally

hadoop fs -copyToLocal /data/hive/sales.log

Create Hive Table:

```
CREATE TABLE sales_test(widget_id INT, qty INT,
street STRING, city STRING, state STRING,
zip INT, sale_date STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
```

Load Data:

LOAD DATA LOCAL INPATH "sales.log" INTO TABLE sales_test;

Select rows to see data:

*select * from sales_test;*

Sqoop Export - Hive To MySQL

#Create MYSQL Table:

CREATE TABLE sales_test(widget_id INT, qty INT,
street varchar(100), city varchar(100), state varchar(100),
zip INT, sale_date varchar(100))

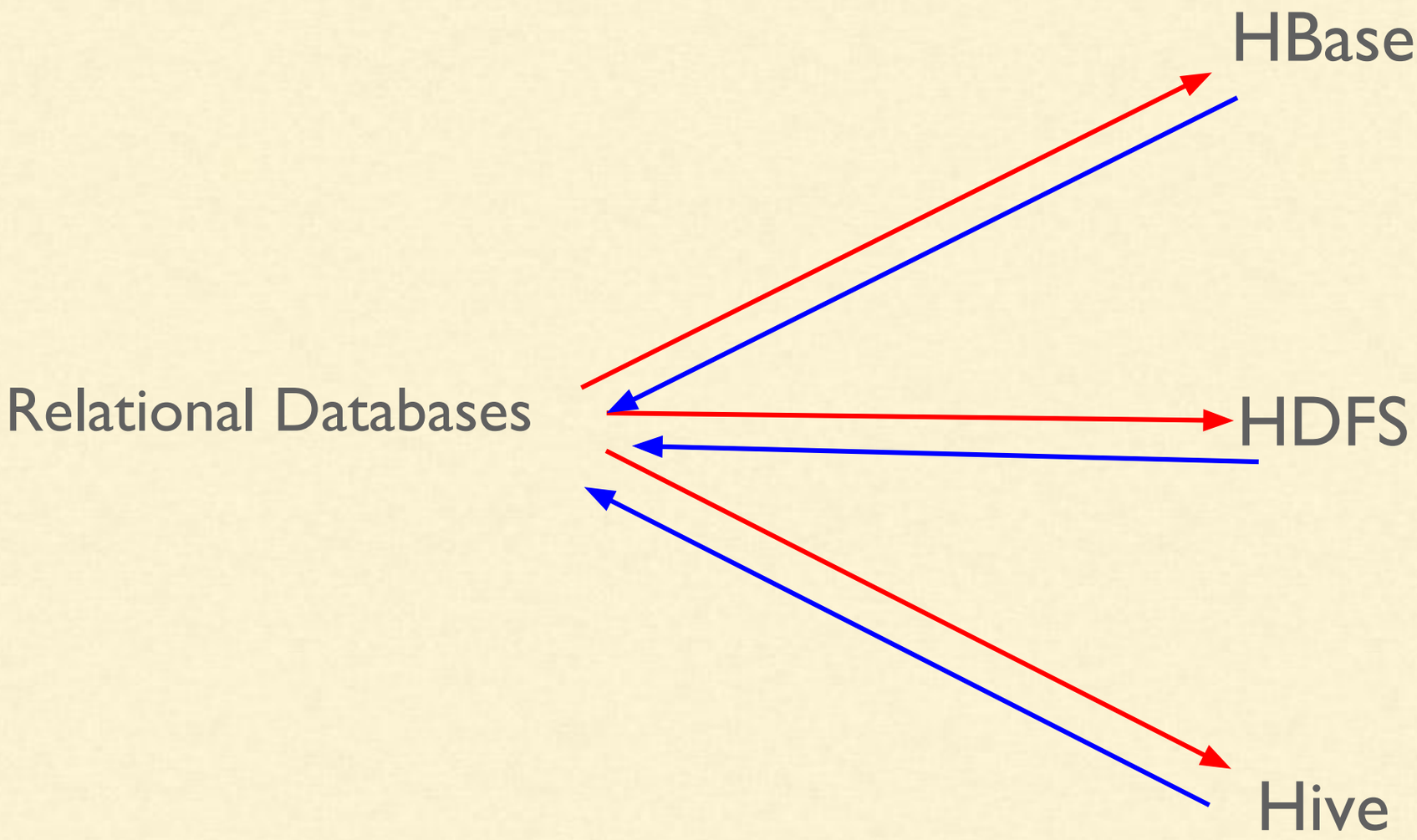
Sqoop Export:

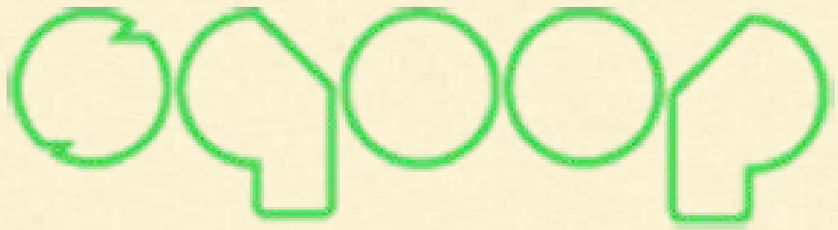
sqoop export --connect jdbc:mysql://172.31.13.154/sqoopex -m 1
--table sales_test --export-dir /apps/hive/warehouse/sales_test
--input-fields-terminated-by ',' --username sqoopuser -P

Sqoop - Summary

- Introduction
- Import
- Export

Sqoop - Introduction Contd.





Thank you!