Main(String args[]){

Job job =;

Job.setMapperClass();

Job.setCombinerClass();

Job.setReducerClass();

Job.setNumReduceTasks();

Job.setInputFormatClass();

Job.setOutputFormatClass();

Job.setInputPath(args[0]);

Job.setOutputPath(args[1])

}

Public class MyCombiner extends Reducer{

Public void reduce(KEYIN key, Iterable<VALUEIN> values, Context context){

}

}

Custom writable class

Class Stock implements Writable{

String symbol:

Float opening;

Float closing;

}

Every value used in MapReduce needs to implement the writable interface

Classes that implement Writable Interface:

* IntWritable
* ArrayWritable
* …..

Override write and readFields

Every Key used in MapReduce needs to implement the writableComparable interface

Classes that implement Writable Interface:

* BooleanWritable
* ByteWritable

Class CustomPartitioner extends Partitioner{

String k = key.toString();

If(k.substring(4).equals(k1))

Return 0;

…….

}

Class MyPartitioner extends Partitioner{

int getPartition(KEY key, VALUE value, int numPartitions){

if(key == K1){

return 0;

}else if(key == K2 && 0<val <4K){

return 0;

}else if(key == K2 && 5<val <10<){

return 1;

}

//return hash(key) % numPartitions;

}

}

Prog in slide

1 comment1 timestamp

1 comment2 timestamp

1 comment3 timestamp

map(){

emit(userid, “firstdate:lastdate:1”)

}

public class MinMaxCountTuple implements Writable {

Date minDate;

Date maxDate;

int date;

}

**Distinct Pattern:**

map(){

emit(key,NullWritable.get())

}

reduce(){

emit(key,NullWritable.get())

}

**HBASE:**

Download

Enter java\_home in hbase\_env.sh

Create localhost/dir – dir in Hadoop

Create zookeeper dir in laptop.

run start-hbase.sh

run hbase shell

* create ‘usertable’, ‘cfuser’
* list ‘usertable’
* put 'usertable', '123', 'cfuser:first', 'Sumanth'
* put 'usertable', '123', 'cfuser:last, Gopal
* get ‘usertable’, ‘123’
* list ‘usertable’

**PIG:**

log = LOAD ‘access.log’;

$0 AS showName, $51 AS userAge………