Programming Hadoop Map-Reduce Programming, Tuning & Debugging

Arun C Murthy
Yahoo! CCDI
acm@yahoo-inc.com
ApacheCon US 2008



Existential angst: Who am 1?

- Yahoo!
- Grid Team (CCDI)
- Apache Hadoop
- Developer since April 2006
- Core Committer (Map-Reduce)
- Member of the Hadoop PMC





Hadoop - Overview

- Hadoop includes:
- Distributed File System distributes data
- Map/Reduce distributes application
- Open source from Apache
- Written in Java
- Runs on
- Linux, Mac OS/X, Windows, and Solaris
- Commodity hardware



Distributed File System

- Designed to store large files
- Stores files as large blocks (64 to 128 MB)
- Each block stored on multiple servers
- Data is automatically re-replicated on need
- Accessed from command line, Java API, or C API
- bin/hadoop fs -put my-file hdfs://node1:50070/foo/bar
- Path p = new Path("hdfs://node1:50070/foo/bar");
 - FileSystem fs = p.getFileSystem(conf); DataOutputStream file = fs.create(p);
- file.writeUTF("hello\n");
- file.close();



Map-Reduce

- Map-Reduce is a programming model for efficient distributed computing
- It works like a Unix pipeline:
- cat > output - cat input | grep | sort | unique -c | cat > ot
 - Input | Map | Shuffle & Sort | Reduce | Output
- Efficiency from
- Streaming through data, reducing seeks
 - Pipelining
- A good fit for a lot of applications
- Log processing
- Web index building



Map/Reduce features

- Fine grained Map and Reduce tasks
- Improved load balancing
- Faster recovery from failed tasks
- Automatic re-execution on failure
- In a large cluster, some nodes are always slow or flaky
- Introduces long tails or failures in computation
- Framework re-executes failed tasks
- Locality optimizations
- With big data, bandwidth to data is a problem
- Map-Reduce + HDFS is a very effective solution
- Map-Reduce queries HDFS for locations of input data
- Map tasks are scheduled local to the inputs when possible

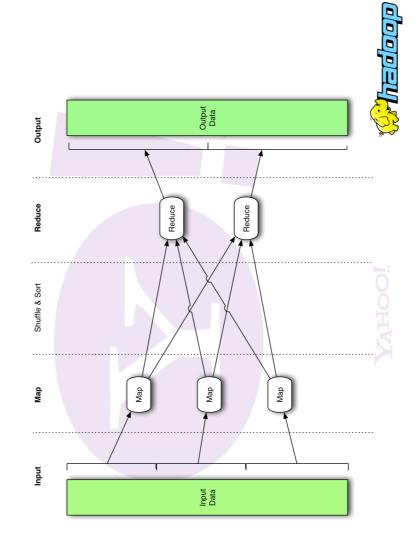


Mappers and Reducers

- Every Map/Reduce program must specify a Mapper and typically a Reducer
- The *Mapper* has a *map* method that transforms input (key, value) pairs into any number of intermediate (key', value') pairs
- The Reducer has a reduce method that transforms intermediate (key', value'*) aggregates into any number of output (key", value") pairs



Map/Reduce Dataflow



Example...

"45% of all Hadoop tutorials count words. 25% count sentences. 20% are about paragraphs. 10% are log parsers. The remainder are helpful."

jandersen @http://twitter.com/jandersen/statuses/ 926856631



Example: Wordcount Mapper

```
public static class MapClass extends MapReduceBase
implements Mapper<LongWritable, Text, Text, IntWritable> {
   private final static IntWritable one = new IntWritable(1);
   private Text word = new Text();
   public void map(LongWritable key, Text value,
        OutputCollector<Text, IntWritable> output,
        Reporter reporter) throws IOException {
        String line = value.toString();
        StringTokenizer itr = new StringTokenizer(line);
        word.set(itr.nextToken());
        output.collect(word, one);
    }
}
```



Example: Wordcount Reducer





Input and Output Formats

- A Map/Reduce may specify how it's input is to be read by specifying an *InputFormat* to be used
- InputSplit
- RecordReader
- A Map/Reduce may specify how it's output is to be written by specifying an OutputFormat to be used
- These default to TextInputFormat and TextOutputFormat, which process line-based text data
- SequenceFile: SequenceFileInputFormat and SequenceFileOutputFormat
- These are file-based, but they are not required to be



Configuring a Job

- Jobs are controlled by configuring JobConf
- JobConfs are maps from attribute names to string value
- The framework defines attributes to control how the job is executed.

```
conf.set("mapred.job.name", "MyApp");
```

- Applications can add arbitrary values to the JobConf conf.set("my.string", "foo");
 conf.setInteger("my.integer", 12);
- JobConf is available to all of the tasks



Putting it all together

- Create a launching program for your application
- The launching program configures:
- The Mapper and Reducer to use
- The output key and value types (input types are inferred from the *InputFormat*)
- The locations for your input and output
- Optionally the InputFormat and OutputFormat to use
- The launching program then submits the job and typically waits for it to complete



Putting it all together

```
public static void main(String[] args) throws IOException {
                                                                                                                                                                                                                                                                                                                                                                                                                            conf.setOutputValueClass(IntWritable.class);
                                                                                                                                         JobConf conf = new JobConf(WordCount.class);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         conf.setOutputPath(new Path(args[1]);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         conf.setInputPath(new Path(args[0]);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               conf.setMapperClass(MapClass.class);
                                                                                                                                                                                                                                                                                                                                  conf.setOutputKeyClass(Text.class);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            conf.setReducerClass(Reduce.class);
                                                                                                                                                                                                                                                                                                                                                                                    // the values are counts (ints)
                                                                                                                                                                                                                                                                                     // the keys are words (strings)
                                                                                                                                                                                        conf.setJobName("wordcount");
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JobClient.runJob(conf);
public class WordCount {
```



Non-Java Interfaces

- Streaming
- Pipes (C++)
- Pig Hive
- Jaql
- Cascading





Streaming

- What about Unix hacks?
- Can define Mapper and Reduce using Unix text filters
- Typically use grep, sed, python, or perl scripts
- Format for input and output is: key \t value \n
- Allows for easy debugging and experimentation
- Slower than Java programs

bin/hadoop jar hadoop-streaming jar -input in-dir -output out-dir -mapper streaming Mapper.sh -reducer streaming Reducer.sh

- Mapper: /bin/sed -e 's | $|\ln|g'|$ / |bin/grep .
- Reducer: /usr/bin/uniq -c | /bin/awk '{print \$2 "\t" \$1}'





Pipes (C++)

```
C++ API and library to link application with
```

C++ application is launched as a sub-process of the Java task

Keys and values are std::string with binary data

```
    Word count map looks like:
```

```
class WordCountMap: public HadoopPipes::Mapper {
   public:
```

```
WordCountMap(HadoopPipes::TaskContext& context){}
void map(HadoopPipes::MapContext& context) {
    std::vector<std::string> words =
        HadoopUtils::splitString(context.getInputValue(),"");
    for(unsigned int i=0; i < words.size(); ++i) {
        context.emit(words[i],"1");
};</pre>
```



Pipes (C++)

The reducer looks like:

class WordCountReduce: public HadoopPipes::Reducer {

```
public:
WordCountReduce(HadoopPipes::TaskContext& context){}

void reduce(HadoopPipes::ReduceContext& context) {
  int sum = 0;
  while (context.nextValue()) {
    sum += HadoopUtils::toInt(context.getInputValue());
  }

context.emit(context.getInputKey(),
  HadoopUtils::toString(sum));
}
```



Pipes (C++)

And define a main function to invoke the tasks:

```
int main(int argc, char *argv[]) {
   return HadoopPipes::runTask(
```

HadoopPipes::TemplateFactory<WordCountMap,

```
WordCountReduce, void, WordCountReduce>());
```



Pig - Hadoop Sub-project

- Scripting language that generates Map/Reduce jobs
- User uses higher level operations
- Group by
- Foreach
- Word Count:

input = LOAD 'in-dir' USING TextLoader();

words = FOREACH input GENERATE
FLATTEN(TOKENIZE(*));

grouped = GROUP words BY \$0;

counts = FOREACH grouped GENERATE group, COUNT(words);

STORE counts INTO 'out-dir';



Hive - Hadoop Sub-project

- SQL-like interface for querying tables stored as flat-files on HDFS, complete with a meta-data repository
- Developed at Facebook
- In the process of moving from Hadoop contrib to a stand-alone Hadoop sub-project





How many Maps and Reduces

Maps

- Usually as many as the number of HDFS blocks being processed, this is the default
- Else the number of maps can be specified as a hint
- The number of maps can also be controlled by specifying the minimum split size
- The actual sizes of the map inputs are computed by:
- max(min(block_size, data/#maps), min_split_size)

Reduces

- Unless the amount of data being processed is small
- 0.95*num_nodes*mapred.tasktracker.reduce.tasks.maximum



Performance Example

- Bob wants to count lines in text files totaling several terabytes
- He uses
- Identity Mapper (input: text, output: same text)
- A single Reducer that counts the lines and outputs the total
- What is he doing wrong?
- This happened, really!
- I am not kidding!



Some handy tools

- Partitioners
- Combiners
- Compression
- Counters
- Speculation
- Zero reduces
- Distributed File Cache
- Tool





Partitioners

- Partitioners are application code that define how keys are assigned to reduces
- Default partitioning spreads keys evenly, but randomly
- Uses key.hashCode() % num_reduces
- Custom partitioning is often required, for example, to produce a total order in the output
- Should implement Partitioner interface
- Set by calling conf.setPartitionerClass(MyPart.class)
- To get a total order, sample the map output keys and pick values to divide the keys into roughly equal buckets and use that in your partitioner



Combiners

- When maps produce many repeated keys
- It is often useful to do a local aggregation following the map
- Done by specifying a Combiner
- Goal is to decrease size of the transient data
- Combiners have the same interface as Reduces, and often are the same class.
- Combiners must not have side effects, because they run an indeterminate number of times.
- In WordCount, conf.setCombinerClass(Reduce.class);



Compression

- Compressing the outputs and intermediate data will often yield huge performance gains
- Can be specified via a configuration file or set programatically
- Set mapred.output.compress to true to compress job output
- Set mapred.compress.map.output to true to compress map outputs
- Compression Types (mapred.output.compression.type) for SequenceFiles
- "block" Group of keys and values are compressed together
- "record" Each value is compressed individually
- Block compression is almost always best
- Compression Codecs (mapred(.map)?.output.compression.codec)
- Default (zlib) slower, but more compression
- LZO faster, but less compression



Counters

- Often Map/Reduce applications have countable events
- For example, framework counts records in to and out of Mapper and Reducer
- To define user counters:

```
static enum Counter {EVENT1, EVENT2};
reporter.incrCounter(Counter.EVENT1, 1);
```

• Define nice names in a MyClass_Counter.properties file
CounterGroupName=My Counters

EVENT1.name=Event 1

EVENT2.name=Event 2



Speculative execution

- The framework can run multiple instances of slow tasks
- Output from instance that finishes first is used
- Controlled by the configuration variable mapred.speculative.execution
- Can dramatically bring in long tails on jobs





Zero Reduces

- Frequently, we only need to run a filter on the input data
- No sorting or shuffling required by the job
- Set the number of reduces to 0
- Output from maps will go directly to OutputFormat and disk





Distributed File Cache

- Sometimes need read-only copies of data on the local computer.
- Downloading 1GB of data for each Mapper is expensive
- Define list of files you need to download in JobConf
- Files are downloaded once per a computer
- Add to launching program:

DistributedCache.addCacheFile(new URI("hdfs://nn:8020/foo"), conf);

Add to task:

Path[] files = DistributedCache.getLocalCacheFiles(conf);



Tool

- Handle "standard" Hadoop command line options:
- -conf file load a configuration file named file
- -D prop=value define a single configuration property prop
- Class looks like:



Debugging & Diagnosis

- Run job with the Local Runner
- Set mapred.job.tracker to "local"
- Runs application in a single process and thread
- Run job on a small data set on a 1 node cluster
- Can be done on your local dev box
- Set keep.failed.task.files to true
- This will keep files from failed tasks that can be used for debugging
 - Use the IsolationRunner to run just the failed task
- Java Debugging hints
- Send a kill -QUIT to the Java process to get the call stack, locks held, deadlocks



Profiling

- Set mapred.task.profile to true
- Use mapred.task.profile.{maps|reduces}
- hprof support is built-in
- Use mapred.task.profile.params to set options for the debugger
- Possibly use DistributedCache for the profiler's agent



Jobtracker front page

Jobid User Name Map % complete Map total Maps completed Reduce % complete Reduces completed Job 0001 parthas quArray 100.00% 22000 22000 96.34% 10 8 Maps Reduces Tasks/Node Nodes Completed Jobs Failed Jobs Running Jobs none 2 kry1112 Hadoop Map/Reduce Administration Started: Mon Aug 27 18:39:15 UTC 2007 Version: 0.13.1, r558872 Compiled: Mon Jul 23 22:07:51 UTC 2007 by hadoopqa Log directory, Job Tracker History Cluster Summary Completed Jobs Running Jobs Failed Jobs Local logs Hadoop, 2006.

المناب المالية

Job counters

Hadoop job_0001 on kry1112

User: parthas
Job Name: quArray
Job File: /mapredsystem/kry1112/submit 3n1dpt/job.xml
Started at: Mon Aug 27 18:40:53 UTC 2007
Status: Running

Kind	% Complete	Num Tasks	Pending	Running	Complete	Killed	Failed/Killed Task Attempts
map	100.00%	22000	0	0	22000	0	0/0
reduce	97.19%	10	0	1	6	0	0/0

	Counter	Map	Reduce	Total
	Map input records	23,680,136,843	0	23,680,136,843
	Map output records	529,463,712	0	529,463,712
	Map input bytes	1,447,917,806,993	0	0 1,447,917,806,993
Map-Reduce Framework Map output bytes	Map output bytes	15,840,622,445	0	15,840,622,445
	Reduce input groups	0	64,042	64,042
	Reduce input records	0	0 474,566,962	474,566,962
	Reduce output records	0	64,040	64,040

Go back to JobTracker Hadoop, 2006.



Task status

Hadoop reduce task list for job_0001 on kry1112

Tasks

Task	Complete Status	Status	Start Time	Finish Time Errors Counters	Errors	Counters
tip_0001_r_000000	32.95%	$\frac{\text{tip}\ 0001\ \ r\ 0000000\ \ 32.95\%}{\text{reduce}} \ \ \ \text{reduce} > \text{copy}\ (21750\ \text{of}\ 220000\ \text{at}\ 0.80\ \text{MB/s}) > \ \ 27\text{-Aug}-2007\ 18.41.06$	27-Aug-2007 18:41:06			
tip_0001_r_000001	32.78%	$\frac{\text{tip}\ 0001\ \text{r}\ 000001}{\text{co}\ 1000001}\ \ 32.78\% \qquad \text{reduce} > \text{copy}\ (21640\ \text{of}\ 22000\ \text{at}\ 0.31\ \text{MB/s}) > \ \ 27\text{-Aug}-2007\ 18.41.06$	27-Aug-2007 18:41:06			
tip_0001_r_000002	32.83%	tip 0001 r 000002 32.83% reduce > copy (21671 of 22000 at 2.37 MB/s) > 27-Aug-2007 18:41:06	27-Aug-2007 18:41:06			
tip_0001_r_000003	32.84%	tip 0001 r 000003 32.84% reduce > copy (21675 of 22000 at 1.53 MB/s) > 27-Aug-2007 18:41:06	27-Aug-2007 18:41:06			
tip_0001_r_000004	32.83%	$\frac{\text{tip}\ 0001\ \ r\ 0000004}{\text{s}\ 19.83\%} \text{reduce} > \text{copy}\ (21674\ \text{of}\ 22000\ \text{at}\ 0.41\ \text{MB/s}) > \ \ 27\text{-Aug}-2007\ 18.41.06$	27-Aug-2007 18:41:06			7
tip_0001_r_000005	32.81%	$\frac{\text{tip }0001 \text{ r }000005}{\text{132.81\%}} \text{reduce} > \text{copy } (21658 \text{ of } 22000 \text{ at } 0.76 \text{ MB/s}) > 27-\text{Aug-}2007 18.41.06 18.4$	27-Aug-2007 18:41:06			
tip_0001_r_000006	32.76%	$\frac{\text{tip}\ 0001\ \text{r}\ 000006}{\text{c}}\ [32.76\%]\ \ \text{reduce} > \text{copy}\ (21627\ \text{of}\ 22000\ \text{at}\ 0.26\ \text{MB/s}) > \ \ 27\text{-Aug-}2007\ 18.41.06$	27-Aug-2007 18:41:06			
tip_0001_r_000007	32.81%	$\frac{\text{tip}\ 0001\ \text{r}\ 000007}{\text{co}}\ [32.81\%]\ \text{reduce} > \text{copy}\ (21656\ \text{of}\ 22000\ \text{at}\ 0.19\ \text{MB/s}) > \ \ 27\text{-Aug-}2007\ 18.41.06$	27-Aug-2007 18:41:06			
tip_0001_r_000008	32.69%	$\frac{\text{tip}\ 0001\ \text{r}\ 0000008}{\text{co}}\ [32.69\%]\ \text{reduce} > \text{copy}\ (21578\ \text{of}\ 22000\ \text{at}\ 0.85\ \text{MB/s}) > \ \ 27-\text{Aug-}2007\ 18.41.06$	27-Aug-2007 18:41:06			
tip_00001_r_000009	32.70%	$\frac{\text{tip 00001 T 0000009}}{\text{22.70\%}} \text{reduce} > \text{copy (21585 of 22000 at 0.63 MB/s)} > 27-\text{Aug-2007 18.41.06} 27-\text{Aug-2007 18.06} 2$	27-Aug-2007 18:41:06			

Go back to JobTracker Hadoop, 2006.





Drilling down

Job job_0001

All Task Attempts

Task Attempts	Machine	Status	Progress	Progress Start Time	Shuffle Finished	Sort Finished	Finish Time	Errors Task Logs	Logs	Counters
sk_0001_r_000000_0	kry1110 inktomisearch.com	SUCCEEDED	100.00%	27-Aug-2007 18:41:06	27-Aug-2007 19:21:09 (40mins, 2sec)	27-Aug-2007 19:21:10 (1sec)	27-Aug-2007 19:29:09 (48mins, 2sec)		Last 4KB Last 8KB All	.co

Go back to the job Go back to JobTracker





Drilling down -- logs

Task Logs: 'task_0001_r_000000_0'

STDOUT logs

STDERR logs

199229 05 66 64 11870 0199229 05 66 64 11870 0199229 05 66 65 11870 0199229 05 66 65 11870 0199229 05 66 65 11870 0199229 05 66 67 11870 0199229 05 66 67 11870 0199229 05 66 67 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 11870 0199229 05 67 07 07 07 07 07 07 07 07 07 07 07 07 07		19:29:05,663	INFO	org.apache.hadoop.streaming.PipeMapRed:	.hadoop.	.streaming	1. PipeMapR		W/S=518830	R/W/S-51883001/6545/0	in:18	8033-51	in:18033-51883001/2877	77 [rec/s]		out:2=6545/2877	7 [rec/	/8]	
19:2916; 665 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 88320165455 (in its 18033-5) 188301/2877 recol. 19:2916; 665 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 88320165455 (in its 18033-5) 188301/2877 recol. 19:2916; 665 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188301016545 (in its 18033-5) 188301/2877 recol. 19:2916; 665 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188301016545 (in its 18033-5) 188301/2877 recol. 19:2916; 668 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188301016545 (in its 18033-5) 188301/2877 recol. 19:2916; 667 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188301016545 (in its 18033-5) 188301/2877 recol. 19:2916; 673 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 188401/2877 recol. 19:2916; 673 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 188401/2877 recol. 19:2916; 673 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 188401/2877 recol. 19:2916; 673 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 188401/2877 recol. 19:2916; 673 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 188401/2877 recol. 19:2916; 674 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 18840101/2877 recol. 19:2916; 675 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 18840101/2877 recol. 19:2916; 676 INCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 18840101/2877 recol. 19:2916; 678 1NCO Org. appeche haddoop streaming PipMakpRed R (W.S-5) 188401016545 (in its 18034-5) 18840101/2877 recol. 19:2916; 678 1NCO Org. appeche haddoop streaming PipMakpRed R (W.S-5)	-08-27 19	:29:05,664	INFO	org.apache	.hadoop.	streaming.	. PipeMapR		W/S=51883)	101/6545/0	in:18	8033=51	883101/28	_	_	t:2=6545/287	_		
19:29:16, 665 1800 org.appeche haddoop streaming PippAmpRed N/M/S=51883010.6545/0 in:18031=5188301/2877 rece/s out.2=6545/2877 rece/s 19:29:16, 667 1800 org.appeche haddoop streaming PippAmpRed N/M/S=51883010.6445/0 in:18031=5188301/2877 rece/s out.2=6545/2877 rece/s 19:29:16, 667 r		:29:05,664		org.apache	.hadoop.	.streaming	PipeMapR		W/S-518831	201/6545/0	in:18	8033-51	883201/28	_	<u>_</u>	t:2-6545/287	_	(8)	
19:3916, 665 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883401.6545 0 int 18031-51883601.2277 recof 10:2916, 667 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883601.6545 0 int 18031-51883601.2277 recof 10:2916, 668 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883601.6545 0 int 18031-51883601.2277 recof 10:2916, 668 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883601.6545 0 int 18031-51883601.2277 recof 10:2916, 668 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883001.6545 0 int 18031-51883001.2277 recof 10:2916, 671 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1883001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 672 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 673 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1884001.6545 0 int 18034-51884001.2277 recof 10:2916, 674 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1885001.6545 0 int 18034-51886001.2277 recof 10:2916, 675 INVO Org. appeche haddoop streaming PipManpRed N.W.S=1885001.6545 0 int 18034-51886001.2277 recof 10:2916, 675 INVO Org. appeche haddoop streaming PipM		:29:05,665		org.apache	.hadoop.	streaming	1. PipeMapR	ď	W/S=51883	301/6545/0	in:18	8033=51	883301/28	_	_	t:2=6545/287	_		
19:3916, 667 1NRO Org. appeche haddoop streaming. PipMakpRed 1 R/M-5-1883001/5545 0 in 18033-51883001/2877 recof soutt-2-655552877 recof soutt-2-65552877 recof soutt		:29:05,665		org.apache	.hadoop.	streaming	. PipeMapR		W/S=51883	401/6545/0	in:18	8033-51	883401/28	_	_	t:2=6545/287	_	[8]	
19:3916; 668 NNO org. appeche haddoop streaming PipMakpRed N.W.5=188360156450 inis 1803=518836012277 recof out 2=6555242877 recof out 2=65552487 recof out		:29:05,667		org.apache	.hadoop.	streaming	. PipeMapR		W/S=518835	501/6545/0	lini (8033=51	883501/28	_	_	t:2=6545/287	_	[8]	
19:3916, 669 1800 org.apeche hadoop streaming PipaMapRed R NWS-51883001/5545 0 int 18031-51883001/2277 rece's out 2-6555/2877 (rece) 19:3916, 671 1800 org.apeche hadoop streaming PipaMapRed R NWS-51883001/5545 0 int 18031-51883001/2877 rece's out 2-6555/2877 (rece) 19:3916, 672 1800 org.apeche hadoop streaming PipaMapRed R NWS-51883001/5545 0 int 18034-51883001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51883001/5645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/5645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/5645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rece) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rec) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51884001/2877 rece's out 2-6555/2877 (rec) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51884001/6645 0 int 18034-51886001/2877 rece's out 2-6555/2877 (rec) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51885001/6645 0 int 18034-51886001/2877 rece's out 2-6555/2877 (rec) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed R NWS-51885001/6645 0 int 18034-51886001/2877 rece's out 2-6555/5877 (rec) 19:3916, 673 1800 org.apeche hadoop streaming PipaMapRed		:29:05,668		org.apache	.hadoop.	streaming	. PipeMapRe		W/S=518836	601/6545/0	in:18	8033-51	883601/28	_	_	t:2=6545/287	_	[8]	
19:3916, 671 1800 org. appeche haddoop streaming PipMakpRed R (W.S=5) 883001/5645 0 inis 18034-51888001/2277 recef sort.2=5545/2877 recef 19:3916, 673 recef sort.2=5545/2877 recef sort.2=5554/2877 recef sort.2=5554/2877 recef sort.2=5555/2877 recef sort.2=5554/2877 recef sort.2=5555/2877 recef sort.2=55		:29:05,669		org.apache	.hadoop.	streaming	. PipeMapR	ď	W/S=51883	701/6545/0		8033=51	883701/28	_	_	t:2=6545/287	_	[8]	
19:3916, 672 INOO org.apeche hadoop streaming PipaMapRed R W/S=51883001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.65450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.66450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.66450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.66450 init 8034=51884001.2277 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.66450 init 8034=51884001.2877 rece's ort.2=6545.2877 rece's org.apeche hadoop streaming PipaMapRed R W/S=51884001.66450 init 8034=51884001.2877 rece's ort.2=6545.2877 rece's ort.2		:29:05,671		org.apache	.hadoop.	streaming	. PipeMapR		W/S=518838	801/6545/0		8033-51	883801/28	_	_	t:2=6545/287	_	[8]	
19:39:16, 673 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884001/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884001/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51884010/545/50 in 18034-51884010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-5555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-51885010/545/50 in 18034-51885010/2877 rece/s out.2-555/2877 1NP		:29:05,672	INFO	org.apache	.hadoop.	streaming	PipeMapR		W/S=51883	901/6545/0		8034=51	883901/28	_	w	t:2=6545/287	_	(8)	
19:39:16, 673 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/12877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/12877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/12877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51884010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51886010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51884010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885010/65450 in:18034-51885010/2877 [rec/s] out.2-6555/2877 INNO org.apatche Andoop streaming FighWapRed: R/W5-51885000/65450 in:18034-51885010/2877 [rec/s] out.2-6555/		:29:05,673		org.apache	.hadoop.	streaming	. PipeMapR		W/S=518840	001/6545/0	in:18	8034=51	884001/28	rec/	·	t:2=6545/287	_	[8]	
19:39:16, 675 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184401/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184401/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184401/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188401/65450 in 18034-5184601/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-6555/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-655/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-655/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-655/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5186501/2877 rece/s out.2-655/2877 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in 18034-5		:29:05,673	INFO	org.apache	.hadoop.	streaming	PipeMapR		W/S=51884)	101/6545/0	in:18	8034=51	884101/28	_	_	t:2=6545/287	_	(8)	
19:3916, 675 INO org.apenche haddoop streaming PipaMapRed R NWS-51884001/65450 init 18034-51884010/2277 rec6's] outt-2-65555287 rec6/19:2916, 675 INO org.apenche haddoop streaming PipaMapRed R NWS-51884001/65450 init 18034-51884001/2277 rec6's] outt-2-65555287 rec6/19:2916, 677 rec6/19:2916, 677		:29:05,675	INFO	org.apache	.hadoop.	streaming	. PipeMapR		W/S=518841	201/6545/0	in:18	8034=51	884201/28	_	_	t:2=6545/287	_	[8]	
19:39:16, 667 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184401/2577 rece's out. 2-65552877 19:29:05, 667 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184661/2577 rece's out. 2-65552877 19:09:05, 667 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184661/2577 rece's out. 2-65552877 19:09:05, 661 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184601/2577 rece's out. 2-65552877 19:09:05, 661 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184601/2577 rece's out. 2-65552877 19:09:05, 661 1NPO org. apache hadoop streaming FighWapRed R/W/S-1884601/65450 in: 8034-5184601/2577 rece's out. 2-65552877 19:09:05, 661 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in: 8034-5186501/2577 rece's out. 2-65552877 19:09:05, 662 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in: 8034-5186501/2577 rece's out. 2-65552877 19:09:05, 662 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in: 8034-5186501/2577 rece's out. 2-65552877 19:09:05, 663 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in: 8034-5186501/2577 rece's out. 2-6555287 19:09:05, 663 1NPO org. apache hadoop streaming FighWapRed R/W/S-188501/65450 in: 8034-5186501/2577 rece's out. 2-6555287 19:09:05, 663 1NPO org. apache hadoop streaming FighWapRed R/W/S-1885001/65450 in: 8034-5186501/2577 rece's out. 2-6555287 19:09:05, 666 1NPO org. apache hadoop streaming FighWapRed R/W/S-1885001/65450 in: 8034-5186501/2577 rece's out. 2-655287 19:09:05, 666 1NPO org. apache hadoop streaming FighWapRed R/W/S-1885001/65450 in: 8034-5186501/2577 rece's out. 2-655287 19:09:05, 666 1NPO org. apache hadoop streaming FighWapRed R/W/S-1885001/65450 in: 8034-5186501/2577 rece's out. 2-655287 19:09:00.0000000000000000000000000000000		:29:05,675	INFO	org.apache	.hadoop.	streaming	. PipeMapRe		W/S=518843	301/6545/0	lini (8034=51	884301/28	_	_	t:2=6545/287	_	[8]	
19:39:105,671 WIND org apatche hadoop streaming FighWakped R /W/S=1884601/64550 in:18034-518846501/2877 [rec/] out:1=6555/2877 [rec/] streaming FighWakped R /W/S=1884001/64550 in:18034-51884601/2877 [rec/] out:1=6555/2877 [rec/] streaming FighWakped R /W/S=1884001/64550 in:18034-51884601/2877 [rec/] out:1=6555/2877 [rec/] streaming FighWakped R /W/S=1884001/6450 in:18034-51884001/2877 [rec/] out:1=6555/2877 [rec/] streaming FighWakped R /W/S=1884001/6450 in:18034-51884001/2877 [rec/] out:1=6555/2877 [rec/] streaming FighWakped R /W/S=1885001/6450 in:18034-51885001/2877 [rec/] streaming FighWakped R /W/S=1885001/6450 in:18034-51885001/2877 [rec/] streaming FighWakped R /W/S=51885001/6450 [rec/] streaming FighWakped R /W/S=5		:29:05,676		org.apache	.hadoop.	streaming	. PipeMapR		W/S=518844	401/6545/0	in:18	8034=51	884401/28	_	_	t:2=6545/287	_	[8]	
19:2916, 5678 INO org.appeche haddoop streaming PippMapReda R/MS-51884601/55450 inis 18034-51884601/2277 rece's out:2-655512/2877 recell 19:2916, 5678 INO org.appeche haddoop streaming PippMapReda R/MS-51884601/55450 inis 18034-51884601/2877 rece's out:2-655512/877 rece's out:2-655512/877 recell 19:2916, 5601 INO org.appeche haddoop streaming PipPMapReda R/MS-51884601/55450 inis 18034-51884601/2877 rece's out:2-655512/877 recell 19:2916, 5611 INO org.appeche haddoop streaming PipPMapReda R/MS-51885001/55450 inis 18034-51884601/2877 rece's out:2-655512/877 recell 19:2916, 5611 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/55450 inis 18034-51886101/2877 rece's out:2-655512/877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/55450 inis 18034-51886501/2877 rece's out:2-655512/877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/55450 inis 18034-51886501/2877 rece's out:2-655512/877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/55450 inis 18034-51886501/2877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/55450 inis 18034-51886501/2877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/54550 inis 18034-51886501/2877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/54550 inis 18034-51886501/2877 recell 19:2916, 5631 INO org.appeche haddoop streaming PipPMapReda R/MS-51885010/54550 inis 18034-51886501/2877 recell 19:2916, 5631 Rec		:29:05,677	INFO	org.apache	.hadoop.	.streaming	. PipeMapR		W/S=518845	501/6545/0	in:18	8034=51	884501/28	_	_	t:2=6545/287	_	(8)	
19:39:05,679 1NO org.apench hadoop streaming.Pipakaphede 1, N/S-51884701/5545/ 0.in18034-51884701/2277 reco's outt-2-6555/2877 [reco'] 19:39:05,660 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-51884701/5545/ 0.in18034-51884701/2277 reco's outt-2-6555/2877 [reco'] 19:39:05,661 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-51884501/645/ 0.in18034-51884501/2277 reco's outt-2-6555/2877 [reco'] 19:39:05,661 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-5188501/2277 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-5188501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-51885011/645/ 0.in18034-5188501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-51885011/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-51885011/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,682 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco'] 19:39:05,687 1NO org.apench hadoop streaming.Pipakaphed 1, N/S-5188501/645/ 0.in18034-51886501/2877 reco's outt-2-6555/2877 [reco']		:29:05,678	INFO	org.apache	.hadoop.	streaming	1. PipeMapR		W/S=51884	601/6545/0		8034-51	884601/28	_	_	t:2=6545/287	_	[8]	
19:29:16, 660 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-5188400116456/0 in: 18034-5184808112877 ree/s out: 2-655512877 119:29:16, 661 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5186500112877 ree/s out: 2-655512877 119:29:16, 661 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5186500112877 ree/s out: 2-655512877 119:29:16, 662 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5186510112877 ree/s out: 2-655512877 119:29:16, 663 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 663 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:16, 666 1NNO org.apache hadoop streaming FlyewhopRed R /R/S-518850011645/0 in: 18034-5188500112877 ree/s out: 2-655512877 119:29:10 10:12-65551287 10:12-65501287 ree/s out: 2-65521287 10:12-655201287 ree/s out: 2-6552201287 ree/s out: 2		:29:05,679		org.apache	.hadoop.	.streaming	. PipeMapR		W/S=51884)	701/6545/0	in:18	8034=51	884701/28	_	<u></u>	t:2=6545/287	_		
19:39:05,681 INPO org.apatche.hadoop streaming.pipeMkpged R/Wis-1884001/6545/0 in:18034-51880901/2877 rec/s] out:2-6555/2877 19:39:05,681 INPO org.apatche.hadoop streaming.pipeMkpged R/Wis-1885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,682 INPO org.apatche.hadoop streaming.pipeMkpged R/Wis-1885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,683 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-1885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,683 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-1885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,683 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-1885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,686 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,686 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,686 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:39:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:30:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-6555/2877 19:30:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-655/2877 19:30:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-655/2877 19:30:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-655/2877 19:30:05,687 INPO org.apatche.hadoops.treaming.pipeMkpged R/Wis-51885001/6545/0 in:18034-51885001/2877 rec/s] out:2-655/2877 10:30:05,000 10:30:05,000 10:30:05,000 10:30:05,000 10:30:05,000 10:30:05,000 10:30:05,000 10:30:05,000		:29:05,680		org.apache	.hadoop.	streaming	1. PipeMapR		W/S=518848	801/6545/0	in:18	8034=51	884801/28	_	8	t:2=6545/287	_	[8]	
19:3916, 661 INTO org.appeche haddoop streaming PippMapRed A NWS-51885001.56450 inii 8034-51886010.12377 rece's outt.2-65451.2877 (rece) 19:2916, 662 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885010.66450 inii 8034-51885010.12377 rece's outt.2-65451.2877 (rece) 19:3916, 663 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.465450 inii 8034-51885010.12377 rece's outt.2-65451.2877 (rece) 19:3916, 663 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.465450 inii 8034-5188501.2877 rece's outt.2-65451.2877 (rece) 19:3916, 664 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.465450 inii 8034-5188503.2877 rece's outt.2-65451.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188503.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188503.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188503.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188603.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188603.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188603.2877 rece's outt.2-65545.2877 (rece) 19:3916, 666 INTO org.appeche haddoop streaming PipPMapRed A NWS-51885030.455450 inii 8034-5188603.2877 rece's outt.2-65545.2877 (rece) 19:3916.2877 (rece) 10:3916.2877 (rece) 10:3916.28		:29:05,681		org.apache	.hadoop.	.streaming	. PipeMapR		W/S=518845	901/6545/0	in:18	8034=51	884901/28	_	_	t:2=6545/287	_	(8)	
19:39:16, 662 1NRO org.apeche hadoop streaming.PipeMapRed R.W.S=51885101/54545 in in 18034=51885101/2877 rece/s ort.2=65455/2877 rece/s org.apeche hadoop streaming.PipeMapRed R.W.S=51885101/54545 in in 18034=51885101/2877 rece/s ort.2=65455/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885101/64545 in 18034=51885101/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 18034=51885101/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 118034=51886510/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 118034=51886510/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 118034=5188610/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 118034=5188610/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 in 118034=5188610/2877 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 rece/s org.apeche.hadoop.streaming.PipeMapRed R.W.S=51885010/64545 rece/s org.apeche.hadoop.streaming.PipeRed R.W.S=51885010/64545 r		:29:05,681		org.apache	.hadoop.	streaming.	. PipeMapR		W/S=51885(001/6545/0	in:18	8034=51	885001/28	_	_	t:2=6545/287	_	[8]	
1942:016,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=5188520164545/0 in:180044-51888510.12877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=518853016455/0 in:180044-51888510.12877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=518854016456/0 in:180044-51888510.12877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=518855016456/0 in:180044-518885012877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=518855015465/0 in:180044-5188855012877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop, streaming. PippambRed; R/WiS=518855015465/0 in:180044-5188855012877 [recc/s] out.2-654512877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=5188550154070 [recc/s] out.2-654512877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=5180501545/0 in:180044-5188857012877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] out.2-654512877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] out.2-654512877 [recc/s] 1940:015,5811NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] out.2-654512877 [recc/s] 1940:015,58111NPO ovg. pacheb. haddop. streaming. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] out.2-6545142877 [recc/s] 1940:015,58111NPO ovg. pacheb. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] 1940:015,58111NPO ovg. pacheb. PippambRed; R/WiS=51805010545/0 in:180044-5188657014877 [recc/s] 1940:015,58111NPO ovg. pacheb. PippambRed; R/WiS=51805010545/0 in:180044-5188857012877 [recc/s] 1940:015,58111NPO ovg. pach		:29:05,682		org.apache	.hadoop.	.streaming	.PipeMapR		W/S=51885	101/6545/0	ini	8034-51	885101/28	_	<u>_</u>	t:2=6545/287	_	[8]	
1912:915; 881 INNO org. apache, hadoop, streaming. PipeMaphedt R/W/S=1885301/6545/0 in:18014-518853012871 (rec/s) outt2-6545/2871 (rec/) 1912:915; 881 INNO org. apache, hadoop, streaming. PipeMaphedt R/W/S=1885401/645/0 in:18034-51886501/2877 (rec/s) outt2-6545/2877 (rec/) 1912:916; 886 INNO org. apache, hadoop, streaming. PipeMaphedt R/W/S=1885601/645/0 in:18034-51886501/2877 (rec/s) outt2-6545/2877 (rec/) 1912:916; 886 INNO org. apache, hadoop, streaming. PipeMaphedt R/W/S=1885601/645/0 in:18034-51888501/2877 (rec/s) outt2-6545/2877 (rec/) 1912:916; 887 INNO org. apache, hadoop, streaming. PipeMaphedt R/W/S=1885601/654/0 in:18034-5188868701/2877 (rec/s) outt2-6545/2877 (rec/)		:29:05,683		org.apache	.hadoop.	streaming	1. PipeMapR		W/S=518851	201/6545/0	in:18	8034=51	885201/28	[rec/	8	t:2=6545/287	_	[8]	
1923:016, 584 XNPO Oxy, apache, hadoop, streaming.Pippampled.RN/N-518854016/55/0 in:18034-518864012877 [rec/] 1923:016, 586 XNPO Oxy, apache, hadoop, streaming.Pippampled.RN/N-51885501655/0 in:18034-518865012877 [rec/] 1923:016, 586 XNPO Oxy, apache, hadoop, streaming.Pippampled.RN/N-51885501765/0 in:18034-518885012877 [rec/] 1923:016, 587 XNPO Oxy, apache, hadoop, streaming.Pippampled.RN/N-51885501765/0 in:18034-518885012877 [rec/] 1923:016, 587 XNPO Oxy, apache, hadoop, streaming.Pippampled.RN/N-5188501765/0 in:18034-5188857012877 [rec/]		:29:05,683		org.apache	.hadoop.	.streaming	. PipeMapRe		W/S=51885	301/6545/0	in:18	8034=51	885301/28	_	_	t:2=6545/287	_	(8)	
19:29:05.686 INPO org.apache.hadoop.streaming.PipeMapRed: R/W/S=51885501/645/0 in:18034=51885611/2877 [rec/s] out.2-6555/2877 [rec/] 19:29:05.886 INPO org.apache.hadoop.streaming.PipeMapRed: R/W-S=51885601/655/0 in:18034=51885601/2877 [rec/] out.2-6555/2877 [rec/] 19:29:05.887 [rec/] out.2-6555/2877 [rec/] 19:29:05.887 [rec/] Out.2-6555/2877 [rec/] Out.2-6557/2877		:29:05,684		org.apache	.hadoop.	streaming	1. PipeMapR		W/S=51885	401/6545/0	in:18	8034=51	885401/28	_	_	t:2=6545/287	_	[8]	
1972:916, 566 INRO Org. nameche Andoop, streaming.pienkampheda N. (%)=5188560165540 in 118034=518856012897 [recy/s] out 12-2655452897		:29:05,686		org.apache	.hadoop.	streaming	. PipeMapR		W/S=51885	501/6545/0	in:18	8034-51	885501/28	_	_	t:2=6545/287	_	[8]	
19:29:05,687 INFO org.apache.hadoop.streaming.PipeMapRed: R/W/S=51885701/6545/0 in:18034=51885701/2877 [rec/s] out:2=6545/2877 [:29:05,686	INFO	org.apache	.hadoop.	streaming	. PipeMapR		W/S=518856	601/6545/0	lini (8034=51	885601/28	_	_	t:2=6545/287	_	[8]	
	-08-27 19	:29:05,687	INFO	org.apache	.hadoop.	.streaming	.PipeMapR		W/S-51885	701/6545/0	in: 18	8034=51	885701/28	_	_	t:2-6545/287	_	[8]	

Performance

- Is your input splittable?
- Gzipped files are NOT splittable
- Use compressed SequenceFiles
- Are partitioners uniform?
- Buffering sizes (especially io.sort.mb)
- Can you avoid Reduce step?
- Only use singleton reduces for very small data
- Use Partitioners and cat to get a total order
- Memory usage
- Please do not load all of your inputs into memory!



Q&A

- For more information:
- Website: http://hadoop.apache.org/core
- Mailing lists:
- general@hadoop.apache.org
- · core-dev@hadoop.apache.org
- core-user@hadoop.apache.org
- IRC: #hadoop on irc.freenode.org



