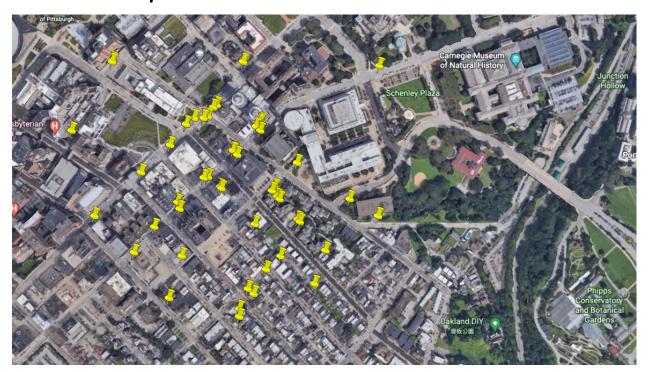# Project5

UserID: student100
Password: Sum73326Aborig123!

Part1Task7 Output:

Part2 Code:

```java
import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;

import java.util.Arrays;
import java.util.Scanner;


/**
 * Distributed systems
 * Analyzing a txt file.
 *
 * Andrew ID: hungfanl
 * @author  Leo Lin
 */

public class TempestAnalytics {
    static SparkConf sparkConf = new
SparkConf().setMaster("local").setAppName("JD Word Counter");
    static JavaSparkContext sparkContext = new JavaSparkContext(sparkConf);
```

```java
    /**
     * Calculating the number of lines in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void lineCount(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("\n")));

        System.out.println("Number of lines: " + wordsFromFile.count());
    }

    /**
     * Calculating the number of words in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void wordCount(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("[^a-zA-Z]+")));

        Function<String, Boolean> filter = k -> (!k.isEmpty());
        wordsFromFile = wordsFromFile.filter(filter);

        System.out.println("Number of words: " + wordsFromFile.count());
    }

    /**
     * Calculating the number of distinct words in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void distinctWordCount(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("[^a-zA-Z]+")));

        Function<String, Boolean> filter = k -> (!k.isEmpty());
        wordsFromFile = wordsFromFile.filter(filter);

        System.out.println("Number of distinct words: " +
wordsFromFile.distinct().count());
    }

    /**
     * Calculating the number of symbols in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void symbols(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);
```

```java
        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("")));

        Function<String, Boolean> filter = k -> (!k.isEmpty());
        wordsFromFile = wordsFromFile.filter(filter);

        System.out.println("Number of symbols: " + wordsFromFile.count());
    }

    /**
     * Calculating the number of distinct symbols in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void distinctSymbols(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("")));

        Function<String, Boolean> filter = k -> (!k.isEmpty());
        wordsFromFile = wordsFromFile.filter(filter);

        System.out.println("Number of symbols: " +
wordsFromFile.distinct().count());
    }

    /**
     * Calculating the number of distinct english letters in the txt file.
     * @param fileName The filename of the txt file
     */
    private static void distinctLetters(String fileName) {

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("")));

        Function<String, Boolean> filter = k -> (!k.isEmpty());
        wordsFromFile = wordsFromFile.filter(filter);
        Function<String, Boolean> f = k -> ((k.charAt(0) >= 'a' &&
k.charAt(0) <= 'z') || (k.charAt(0) >= 'A' && k.charAt(0) <= 'Z'));

        wordsFromFile = wordsFromFile.filter(f);

        System.out.println("Number of distinct Letters: " +
wordsFromFile.distinct().count());
    }

    /**
     * Have the user insert a word and print out all the lines that include
that particular word.
     * @param fileName The filename of the txt file
     */
    private static void search(String fileName) {
        System.out.println("Insert the word you want to search");
        Scanner readInput = new Scanner(System.in);
```

```java
        String key = readInput.nextLine();

        JavaRDD<String> inputFile = sparkContext.textFile(fileName);

        JavaRDD<String> wordsFromFile = inputFile.flatMap(content ->
Arrays.asList(content.split("\n")));

        Function<String, Boolean> filter = k -> (k.contains(key));
        wordsFromFile = wordsFromFile.filter(filter);
        for(int i = 0; i < wordsFromFile.count(); i++) {
            System.out.println(wordsFromFile.collect().get(i));
        }
    }

    public static void main(String[] args) {

        if (args.length == 0) {
            System.out.println("No files provided.");
            System.exit(0);
        }
        lineCount(args[0]);
        wordCount(args[0]);
        distinctWordCount(args[0]);
        symbols(args[0]);
        distinctSymbols(args[0]);
        distinctLetters(args[0]);
        search(args[0]);
    }
}
```

Part2 Output:

/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/bin/java -
javaagent:/Applications/IntelliJ IDEA.app/Contents/lib/idea_rt.jar=57099:/Applications/IntelliJ
IDEA.app/Contents/bin -Dfile.encoding=UTF-8 -classpath
/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/charsets.jar:/Library/Java/J
avaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/deploy.jar:/Library/Java/JavaVirtualMachi
nes/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/cldrdata.jar:/Library/Java/JavaVirtualMachines/jdk1.8.
0_341.jdk/Contents/Home/jre/lib/ext/dnsns.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Co
ntents/Home/jre/lib/ext/jaccess.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Hom
e/jre/lib/ext/jfxrt.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/lo
caledata.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/nashorn.ja
r:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/sunec.jar:/Library/Jav
a/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/sunjce_provider.jar:/Library/Java/J
avaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/sunpkcs11.jar:/Library/Java/JavaVirtu
alMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/ext/zipfs.jar:/Library/Java/JavaVirtualMachines/jd
k1.8.0_341.jdk/Contents/Home/jre/lib/javaws.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/C
ontents/Home/jre/lib/jce.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/li
b/jfr.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/jfxswt.jar:/Library/
Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/jsse.jar:/Library/Java/JavaVirtualMa
chines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/management-

agent.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/plugin.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/resources.jar:/Library/Java/JavaVirtualMachines/jdk1.8.0_341.jdk/Contents/Home/jre/lib/rt.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/out/production/Part2:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-catalyst_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/scala-reflect-2.10.5.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/scala-library-2.10.5.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-core_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/avro-mapred-1.7.7-hadoop2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/avro-ipc-1.7.7.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/avro-1.7.7.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/avro-ipc-1.7.7-tests.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-core-asl-1.9.13.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-mapper-asl-1.9.13.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/chill_2.10-0.5.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/kryo-2.21.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/reflectasm-1.07-shaded.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/minlog-1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/objenesis-1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/chill-java-0.5.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/xbean-asm5-shaded-4.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-client-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-common-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-cli-1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-math-2.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/xmlenc-0.52.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-configuration-1.6.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-collections-3.2.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-digester-1.8.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-beanutils-1.7.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-beanutils-core-1.8.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-auth-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-compress-1.4.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/xz-1.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-hdfs-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jetty-util-6.1.26.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-mapreduce-client-app-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-mapreduce-client-common-

2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-yarn-client-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/guice-3.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/javax.inject-1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/aopalliance-1.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-test-framework-grizzly2-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-test-framework-core-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/javax.servlet-api-3.0.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-client-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-grizzly2-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/grizzly-http-2.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/grizzly-framework-2.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/gmbal-api-only-3.0.0-b023.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/management-api-3.0.0-b012.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/grizzly-http-server-2.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/grizzly-rcm-2.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/grizzly-http-servlet-2.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/javax.servlet-3.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-json-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jettison-1.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/stax-api-1.0.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jaxb-impl-2.2.3-1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jaxb-api-2.2.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/activation-1.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-jaxrs-1.8.3.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-xc-1.8.3.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-guice-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-yarn-server-common-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-mapreduce-client-shuffle-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-yarn-api-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-mapreduce-client-core-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-yarn-common-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-mapreduce-client-jobclient-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/hadoop-annotations-2.2.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-launcher_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-network-common_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-network-shuffle_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/leveldbjni-all-1.8.jar:/Users/linhungfan/Desktop/CMU/Semester

2/Distributed System/Project5/Part2/lib/jackson-annotations-2.4.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jets3t-0.7.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-codec-1.3.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-httpclient-3.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/curator-recipes-2.4.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/curator-framework-2.4.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/curator-client-2.4.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/zookeeper-3.4.5.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jline-0.9.94.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/guava-14.0.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/javax.servlet-3.0.0.v201112011016.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-lang3-3.3.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-math3-3.4.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jsr305-1.3.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/slf4j-api-1.7.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jul-to-slf4j-1.7.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jcl-over-slf4j-1.7.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/log4j-1.2.17.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/slf4j-log4j12-1.7.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/compress-lzf-1.0.3.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/snappy-java-1.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/lz4-1.3.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/RoaringBitmap-0.5.11.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-net-2.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/akka-remote_2.10-2.3.11.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/akka-actor_2.10-2.3.11.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/config-1.2.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/netty-3.8.0.Final.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/protobuf-java-2.5.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/uncommons-maths-1.2.2a.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/akka-slf4j_2.10-2.3.11.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/json4s-jackson_2.10-3.2.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/json4s-core_2.10-3.2.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/json4s-ast_2.10-3.2.10.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/scalap-2.10.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/scala-compiler-2.10.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-server-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/asm-

3.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jersey-core-1.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/mesos-0.21.1-shaded-protobuf.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/netty-all-4.0.29.Final.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/stream-2.7.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/metrics-core-3.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/metrics-jvm-3.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/metrics-json-3.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/metrics-graphite-3.1.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-databind-2.4.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-core-2.4.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/jackson-module-scala_2.10-2.4.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/paranamer-2.6.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/ivy-2.4.0.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/oro-2.0.8.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/tachyon-client-0.8.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-lang-2.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-io-2.4.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/tachyon-underfs-hdfs-0.8.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/tachyon-underfs-s3-0.8.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/tachyon-underfs-local-0.8.2.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/pyrolite-4.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/py4j-0.9.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/spark-unsafe_2.10-1.6.1.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/janino-2.7.8.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/commons-compiler-2.7.8.jar:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/lib/unused-1.0.0.jar TempestAnalytics TheTempest.txt

Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties

22/12/08 16:05:20 INFO SparkContext: Running Spark version 1.6.1

22/12/08 16:05:20 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

22/12/08 16:05:20 INFO SecurityManager: Changing view acls to: linhungfan

22/12/08 16:05:20 INFO SecurityManager: Changing modify acls to: linhungfan

22/12/08 16:05:20 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(linhungfan); users with modify permissions: Set(linhungfan)

22/12/08 16:05:21 INFO Utils: Successfully started service 'sparkDriver' on port 57105.

22/12/08 16:05:21 INFO Slf4jLogger: Slf4jLogger started

22/12/08 16:05:21 INFO Remoting: Starting remoting

22/12/08 16:05:22 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://sparkDriverActorSystem@10.37.129.2:57106]

22/12/08 16:05:22 INFO Utils: Successfully started service 'sparkDriverActorSystem' on port 57106.

22/12/08 16:05:22 INFO SparkEnv: Registering MapOutputTracker

22/12/08 16:05:22 INFO SparkEnv: Registering BlockManagerMaster

22/12/08 16:05:22 INFO DiskBlockManager: Created local directory at /private/var/folders/55/rnjlg8zj0yng2nbxj9dmfjn80000gn/T/blockmgr-9d01e924-e9ed-4c70-a8c9-ca02016458dc

22/12/08 16:05:22 INFO MemoryStore: MemoryStore started with capacity 2.4 GB

22/12/08 16:05:22 INFO SparkEnv: Registering OutputCommitCoordinator

22/12/08 16:05:22 INFO Utils: Successfully started service 'SparkUI' on port 4040.

22/12/08 16:05:22 INFO SparkUI: Started SparkUI at http://10.37.129.2:4040

22/12/08 16:05:22 INFO Executor: Starting executor ID driver on host localhost

22/12/08 16:05:22 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 57107.

22/12/08 16:05:22 INFO NettyBlockTransferService: Server created on 57107

22/12/08 16:05:22 INFO BlockManagerMaster: Trying to register BlockManager

22/12/08 16:05:22 INFO BlockManagerMasterEndpoint: Registering block manager localhost:57107 with 2.4 GB RAM, BlockManagerId(driver, localhost, 57107)

22/12/08 16:05:22 INFO BlockManagerMaster: Registered BlockManager

22/12/08 16:05:23 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 107.7 KB, free 107.7 KB)

22/12/08 16:05:23 INFO MemoryStore: Block broadcast_0_piece0 stored as bytes in memory (estimated size 9.8 KB, free 117.5 KB)

22/12/08 16:05:23 INFO BlockManagerInfo: Added broadcast_0_piece0 in memory on localhost:57107 (size: 9.8 KB, free: 2.4 GB)

22/12/08 16:05:23 INFO SparkContext: Created broadcast 0 from textFile at TempestAnalytics.java:16

22/12/08 16:05:24 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:24 INFO SparkContext: Starting job: count at TempestAnalytics.java:20

22/12/08 16:05:24 INFO DAGScheduler: Got job 0 (count at TempestAnalytics.java:20) with 1 output partitions

22/12/08 16:05:24 INFO DAGScheduler: Final stage: ResultStage 0 (count at TempestAnalytics.java:20)

22/12/08 16:05:24 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:24 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:24 INFO DAGScheduler: Submitting ResultStage 0 (MapPartitionsRDD[2] at flatMap at TempestAnalytics.java:18), which has no missing parents

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_1 stored as values in memory (estimated size 4.1 KB, free 121.6 KB)

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_1_piece0 stored as bytes in memory (estimated size 2.4 KB, free 124.0 KB)

22/12/08 16:05:25 INFO BlockManagerInfo: Added broadcast_1_piece0 in memory on localhost:57107 (size: 2.4 KB, free: 2.4 GB)

22/12/08 16:05:25 INFO SparkContext: Created broadcast 1 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:25 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 0 (MapPartitionsRDD[2] at flatMap at TempestAnalytics.java:18)

22/12/08 16:05:25 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks

22/12/08 16:05:25 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:25 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)

22/12/08 16:05:25 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:25 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id

22/12/08 16:05:25 INFO deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id

22/12/08 16:05:25 INFO deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap

22/12/08 16:05:25 INFO deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition

22/12/08 16:05:25 INFO deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id

22/12/08 16:05:25 INFO Executor: Finished task 0.0 in stage 0.0 (TID 0). 2082 bytes result sent to driver

22/12/08 16:05:25 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 175 ms on localhost (1/1)

22/12/08 16:05:25 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool

22/12/08 16:05:25 INFO DAGScheduler: ResultStage 0 (count at TempestAnalytics.java:20) finished in 0.190 s

22/12/08 16:05:25 INFO DAGScheduler: Job 0 finished: count at TempestAnalytics.java:20, took 0.292136 s

Number of lines: 3466

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 59.6 KB, free 183.5 KB)

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 13.8 KB, free 197.3 KB)

22/12/08 16:05:25 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:25 INFO SparkContext: Created broadcast 2 from textFile at TempestAnalytics.java:25

22/12/08 16:05:25 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:25 INFO SparkContext: Starting job: count at TempestAnalytics.java:32

22/12/08 16:05:25 INFO DAGScheduler: Got job 1 (count at TempestAnalytics.java:32) with 1 output partitions

22/12/08 16:05:25 INFO DAGScheduler: Final stage: ResultStage 1 (count at TempestAnalytics.java:32)

22/12/08 16:05:25 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:25 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:25 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at filter at TempestAnalytics.java:30), which has no missing parents

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_3 stored as values in memory (estimated size 4.6 KB, free 201.9 KB)

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_3_piece0 stored as bytes in memory (estimated size 2.5 KB, free 204.4 KB)

22/12/08 16:05:25 INFO BlockManagerInfo: Added broadcast_3_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:25 INFO SparkContext: Created broadcast 3 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:25 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at filter at TempestAnalytics.java:30)

22/12/08 16:05:25 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks

22/12/08 16:05:25 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:25 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)

22/12/08 16:05:25 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:25 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 2082 bytes result sent to driver

22/12/08 16:05:25 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 90 ms on localhost (1/1)

22/12/08 16:05:25 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool

22/12/08 16:05:25 INFO DAGScheduler: ResultStage 1 (count at TempestAnalytics.java:32) finished in 0.092 s

22/12/08 16:05:25 INFO DAGScheduler: Job 1 finished: count at TempestAnalytics.java:32, took 0.115771 s

Number of words: 18167

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_4 stored as values in memory (estimated size 127.2 KB, free 331.6 KB)

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_4_piece0 stored as bytes in memory (estimated size 13.8 KB, free 345.4 KB)

22/12/08 16:05:25 INFO BlockManagerInfo: Added broadcast_4_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:25 INFO SparkContext: Created broadcast 4 from textFile at TempestAnalytics.java:37

22/12/08 16:05:25 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:25 INFO SparkContext: Starting job: count at TempestAnalytics.java:44

22/12/08 16:05:25 INFO DAGScheduler: Registering RDD 11 (distinct at TempestAnalytics.java:44)

22/12/08 16:05:25 INFO DAGScheduler: Got job 2 (count at TempestAnalytics.java:44) with 1 output partitions

22/12/08 16:05:25 INFO DAGScheduler: Final stage: ResultStage 3 (count at TempestAnalytics.java:44)

22/12/08 16:05:25 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 2)

22/12/08 16:05:25 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 2)

22/12/08 16:05:25 INFO DAGScheduler: Submitting ShuffleMapStage 2 (MapPartitionsRDD[11] at distinct at TempestAnalytics.java:44), which has no missing parents

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_5 stored as values in memory (estimated size 5.6 KB, free 350.9 KB)

22/12/08 16:05:25 INFO MemoryStore: Block broadcast_5_piece0 stored as bytes in memory (estimated size 3.0 KB, free 354.0 KB)

22/12/08 16:05:25 INFO BlockManagerInfo: Added broadcast_5_piece0 in memory on localhost:57107 (size: 3.0 KB, free: 2.4 GB)

22/12/08 16:05:25 INFO SparkContext: Created broadcast 5 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:25 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 2 (MapPartitionsRDD[11] at distinct at TempestAnalytics.java:44)

22/12/08 16:05:25 INFO TaskSchedulerImpl: Adding task set 2.0 with 1 tasks

22/12/08 16:05:25 INFO TaskSetManager: Starting task 0.0 in stage 2.0 (TID 2, localhost, partition 0,PROCESS_LOCAL, 2185 bytes)

22/12/08 16:05:25 INFO Executor: Running task 0.0 in stage 2.0 (TID 2)

22/12/08 16:05:25 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:25 INFO Executor: Finished task 0.0 in stage 2.0 (TID 2). 2253 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 2.0 (TID 2) in 331 ms on localhost (1/1)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 2.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: ShuffleMapStage 2 (distinct at TempestAnalytics.java:44) finished in 0.335 s

22/12/08 16:05:26 INFO DAGScheduler: looking for newly runnable stages

22/12/08 16:05:26 INFO DAGScheduler: running: Set()

22/12/08 16:05:26 INFO DAGScheduler: waiting: Set(ResultStage 3)

22/12/08 16:05:26 INFO DAGScheduler: failed: Set()

22/12/08 16:05:26 INFO DAGScheduler: Submitting ResultStage 3 (MapPartitionsRDD[13] at distinct at TempestAnalytics.java:44), which has no missing parents

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_6 stored as values in memory (estimated size 3.2 KB, free 357.2 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_6_piece0 stored as bytes in memory (estimated size 1904.0 B, free 359.0 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_6_piece0 in memory on localhost:57107 (size: 1904.0 B, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 6 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 3 (MapPartitionsRDD[13] at distinct at TempestAnalytics.java:44)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 3.0 with 1 tasks

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 3.0 (TID 3, localhost, partition 0,NODE_LOCAL, 1894 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 3.0 (TID 3)

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 5 ms

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 3.0 (TID 3). 1203 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 3.0 (TID 3) in 149 ms on localhost (1/1)

22/12/08 16:05:26 INFO DAGScheduler: ResultStage 3 (count at TempestAnalytics.java:44) finished in 0.151 s

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 3.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: Job 2 finished: count at TempestAnalytics.java:44, took 0.582486 s

Number of distinct words: 3595

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_7 stored as values in memory (estimated size 127.2 KB, free 486.2 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_7_piece0 stored as bytes in memory (estimated size 13.8 KB, free 500.0 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_7_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 7 from textFile at TempestAnalytics.java:49

22/12/08 16:05:26 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:26 INFO SparkContext: Starting job: count at TempestAnalytics.java:56

22/12/08 16:05:26 INFO DAGScheduler: Got job 3 (count at TempestAnalytics.java:56) with 1 output partitions

22/12/08 16:05:26 INFO DAGScheduler: Final stage: ResultStage 4 (count at TempestAnalytics.java:56)

22/12/08 16:05:26 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:26 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:26 INFO DAGScheduler: Submitting ResultStage 4 (MapPartitionsRDD[17] at filter at TempestAnalytics.java:54), which has no missing parents

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_8 stored as values in memory (estimated size 4.6 KB, free 504.5 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_8_piece0 stored as bytes in memory (estimated size 2.5 KB, free 507.0 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_8_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 8 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 4 (MapPartitionsRDD[17] at filter at TempestAnalytics.java:54)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 4.0 with 1 tasks

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 4.0 (TID 4, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 4.0 (TID 4)

22/12/08 16:05:26 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 4.0 (TID 4). 2082 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 4.0 (TID 4) in 72 ms on localhost (1/1)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 4.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: ResultStage 4 (count at TempestAnalytics.java:56) finished in 0.076 s

22/12/08 16:05:26 INFO DAGScheduler: Job 3 finished: count at TempestAnalytics.java:56, took 0.107209 s

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_9 stored as values in memory (estimated size 127.2 KB, free 634.2 KB)

Number of symbols: 96381

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_9_piece0 stored as bytes in memory (estimated size 13.8 KB, free 648.0 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_9_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 9 from textFile at TempestAnalytics.java:61

22/12/08 16:05:26 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:26 INFO SparkContext: Starting job: count at TempestAnalytics.java:68

22/12/08 16:05:26 INFO DAGScheduler: Registering RDD 22 (distinct at TempestAnalytics.java:68)

22/12/08 16:05:26 INFO DAGScheduler: Got job 4 (count at TempestAnalytics.java:68) with 1 output partitions

22/12/08 16:05:26 INFO DAGScheduler: Final stage: ResultStage 6 (count at TempestAnalytics.java:68)

22/12/08 16:05:26 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 5)

22/12/08 16:05:26 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 5)

22/12/08 16:05:26 INFO DAGScheduler: Submitting ShuffleMapStage 5 (MapPartitionsRDD[22] at distinct at TempestAnalytics.java:68), which has no missing parents

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_10 stored as values in memory (estimated size 5.6 KB, free 653.6 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_10_piece0 stored as bytes in memory (estimated size 3.0 KB, free 656.6 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_10_piece0 in memory on localhost:57107 (size: 3.0 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 10 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 5 (MapPartitionsRDD[22] at distinct at TempestAnalytics.java:68)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 5.0 with 1 tasks

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 5.0 (TID 5, localhost, partition 0,PROCESS_LOCAL, 2185 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 5.0 (TID 5)

22/12/08 16:05:26 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 5.0 (TID 5). 2253 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 5.0 (TID 5) in 71 ms on localhost (1/1)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 5.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: ShuffleMapStage 5 (distinct at TempestAnalytics.java:68) finished in 0.072 s

22/12/08 16:05:26 INFO DAGScheduler: looking for newly runnable stages

22/12/08 16:05:26 INFO DAGScheduler: running: Set()

22/12/08 16:05:26 INFO DAGScheduler: waiting: Set(ResultStage 6)

22/12/08 16:05:26 INFO DAGScheduler: failed: Set()

22/12/08 16:05:26 INFO DAGScheduler: Submitting ResultStage 6 (MapPartitionsRDD[24] at distinct at TempestAnalytics.java:68), which has no missing parents

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_11 stored as values in memory (estimated size 3.2 KB, free 659.8 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_11_piece0 stored as bytes in memory (estimated size 1891.0 B, free 661.6 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_11_piece0 in memory on localhost:57107 (size: 1891.0 B, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 11 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 6 (MapPartitionsRDD[24] at distinct at TempestAnalytics.java:68)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 6.0 with 1 tasks

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 6.0 (TID 6, localhost, partition 0,NODE_LOCAL, 1894 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 6.0 (TID 6)

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 0 ms

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 6.0 (TID 6). 1203 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 6.0 (TID 6) in 17 ms on localhost (1/1)

22/12/08 16:05:26 INFO DAGScheduler: ResultStage 6 (count at TempestAnalytics.java:68) finished in 0.017 s

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 6.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: Job 4 finished: count at TempestAnalytics.java:68, took 0.120058 s

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_12 stored as values in memory (estimated size 127.2 KB, free 788.8 KB)

Number of symbols: 76

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_12_piece0 stored as bytes in memory (estimated size 13.8 KB, free 802.6 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_12_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 12 from textFile at TempestAnalytics.java:73

22/12/08 16:05:26 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:26 INFO SparkContext: Starting job: count at TempestAnalytics.java:83

22/12/08 16:05:26 INFO DAGScheduler: Registering RDD 30 (distinct at TempestAnalytics.java:83)

22/12/08 16:05:26 INFO DAGScheduler: Got job 5 (count at TempestAnalytics.java:83) with 1 output partitions

22/12/08 16:05:26 INFO DAGScheduler: Final stage: ResultStage 8 (count at TempestAnalytics.java:83)

22/12/08 16:05:26 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 7)

22/12/08 16:05:26 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 7)

22/12/08 16:05:26 INFO DAGScheduler: Submitting ShuffleMapStage 7 (MapPartitionsRDD[30] at distinct at TempestAnalytics.java:83), which has no missing parents

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_13 stored as values in memory (estimated size 5.7 KB, free 808.3 KB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_13_piece0 stored as bytes in memory (estimated size 3.1 KB, free 811.4 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_13_piece0 in memory on localhost:57107 (size: 3.1 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 13 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 7 (MapPartitionsRDD[30] at distinct at TempestAnalytics.java:83)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 7.0 with 1 tasks

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 7.0 (TID 7, localhost, partition 0,PROCESS_LOCAL, 2185 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 7.0 (TID 7)

22/12/08 16:05:26 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 7.0 (TID 7). 2253 bytes result sent to driver

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_11_piece0 on localhost:57107 in memory (size: 1891.0 B, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 7

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_10_piece0 on localhost:57107 in memory (size: 3.0 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 6

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 7.0 (TID 7) in 132 ms on localhost (1/1)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 7.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: ShuffleMapStage 7 (distinct at TempestAnalytics.java:83) finished in 0.132 s

22/12/08 16:05:26 INFO DAGScheduler: looking for newly runnable stages

22/12/08 16:05:26 INFO DAGScheduler: running: Set()

22/12/08 16:05:26 INFO DAGScheduler: waiting: Set(ResultStage 8)

22/12/08 16:05:26 INFO DAGScheduler: failed: Set()

22/12/08 16:05:26 INFO DAGScheduler: Submitting ResultStage 8 (MapPartitionsRDD[32] at distinct at TempestAnalytics.java:83), which has no missing parents

22/12/08 16:05:26 INFO ContextCleaner: Cleaned shuffle 1

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_9_piece0 on localhost:57107 in memory (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_14 stored as values in memory (estimated size 3.2 KB, free 660.0 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_8_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO MemoryStore: Block broadcast_14_piece0 stored as bytes in memory (estimated size 1894.0 B, free 659.3 KB)

22/12/08 16:05:26 INFO BlockManagerInfo: Added broadcast_14_piece0 in memory on localhost:57107 (size: 1894.0 B, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 5

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_7_piece0 on localhost:57107 in memory (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO SparkContext: Created broadcast 14 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_6_piece0 on localhost:57107 in memory (size: 1904.0 B, free: 2.4 GB)

22/12/08 16:05:26 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 8 (MapPartitionsRDD[32] at distinct at TempestAnalytics.java:83)

22/12/08 16:05:26 INFO TaskSchedulerImpl: Adding task set 8.0 with 1 tasks

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 4

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_5_piece0 on localhost:57107 in memory (size: 3.0 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 3

22/12/08 16:05:26 INFO TaskSetManager: Starting task 0.0 in stage 8.0 (TID 8, localhost, partition 0,NODE_LOCAL, 1894 bytes)

22/12/08 16:05:26 INFO Executor: Running task 0.0 in stage 8.0 (TID 8)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned shuffle 0

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_4_piece0 on localhost:57107 in memory (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_3_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 2

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_2_piece0 on localhost:57107 in memory (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks

22/12/08 16:05:26 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 1 ms

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_1_piece0 on localhost:57107 in memory (size: 2.4 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO ContextCleaner: Cleaned accumulator 1

22/12/08 16:05:26 INFO BlockManagerInfo: Removed broadcast_0_piece0 on localhost:57107 in memory (size: 9.8 KB, free: 2.4 GB)

22/12/08 16:05:26 INFO Executor: Finished task 0.0 in stage 8.0 (TID 8). 1203 bytes result sent to driver

22/12/08 16:05:26 INFO TaskSetManager: Finished task 0.0 in stage 8.0 (TID 8) in 29 ms on localhost (1/1)

22/12/08 16:05:26 INFO DAGScheduler: ResultStage 8 (count at TempestAnalytics.java:83) finished in 0.031 s

22/12/08 16:05:26 INFO TaskSchedulerImpl: Removed TaskSet 8.0, whose tasks have all completed, from pool

22/12/08 16:05:26 INFO DAGScheduler: Job 5 finished: count at TempestAnalytics.java:83, took 0.209650 s

Number of distinct Letters: 51

Insert the word you want to search

love

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_15 stored as values in memory (estimated size 127.2 KB, free 282.0 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_15_piece0 stored as bytes in memory (estimated size 13.8 KB, free 295.8 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_15_piece0 in memory on localhost:57107 (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 15 from textFile at TempestAnalytics.java:91

22/12/08 16:05:37 INFO FileInputFormat: Total input paths to process : 1

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 6 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 9 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 9 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_16 stored as values in memory (estimated size 4.6 KB, free 300.4 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_16_piece0 stored as bytes in memory (estimated size 2.5 KB, free 302.9 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_16_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 16 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 9 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 9.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 9.0 (TID 9, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 9.0 (TID 9)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 9.0 (TID 9). 2082 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 9.0 (TID 9) in 21 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 9.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 9 (count at TempestAnalytics.java:97) finished in 0.022 s

22/12/08 16:05:37 INFO DAGScheduler: Job 6 finished: count at TempestAnalytics.java:97, took 0.037083 s

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 7 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 10 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 10 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_17 stored as values in memory (estimated size 4.8 KB, free 307.7 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_17_piece0 stored as bytes in memory (estimated size 2.6 KB, free 310.3 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_17_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 17 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 10 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 10.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 10.0 (TID 10, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 10.0 (TID 10)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 10.0 (TID 10). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 10.0 (TID 10) in 13 ms on localhost (1/1)

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 10 (collect at TempestAnalytics.java:98) finished in 0.014 s

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 10.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: Job 7 finished: collect at TempestAnalytics.java:98, took 0.022307 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 8 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 11 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 11 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

BOATSWAIN None that I more love than myself. You are

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_18 stored as values in memory (estimated size 4.6 KB, free 314.9 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_18_piece0 stored as bytes in memory (estimated size 2.5 KB, free 317.4 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_18_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 18 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 11 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 11.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 11.0 (TID 11, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 11.0 (TID 11)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 11.0 (TID 11). 2082 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 11.0 (TID 11) in 10 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 11.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 11 (count at TempestAnalytics.java:97) finished in 0.010 s

22/12/08 16:05:37 INFO DAGScheduler: Job 8 finished: count at TempestAnalytics.java:97, took 0.019500 s

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 9 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 12 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 12 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_19 stored as values in memory (estimated size 4.8 KB, free 322.2 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_19_piece0 stored as bytes in memory (estimated size 2.6 KB, free 324.7 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_19_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 19 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 12 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 12.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 12.0 (TID 12, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 12.0 (TID 12)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 12.0 (TID 12). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 12.0 (TID 12) in 6 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 12.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 12 (collect at TempestAnalytics.java:98) finished in 0.007 s

22/12/08 16:05:37 INFO DAGScheduler: Job 9 finished: collect at TempestAnalytics.java:98, took 0.014258 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 10 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 13 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 13 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_20 stored as values in memory (estimated size 4.6 KB, free 329.4 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_20_piece0 stored as bytes in memory (estimated size 2.5 KB, free 331.9 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_20_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 20 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 13 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 13.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 13.0 (TID 13, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 13.0 (TID 13)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

Of all the world I loved, and to him put

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 13.0 (TID 13). 2082 bytes result sent to driver

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 13 (count at TempestAnalytics.java:97) finished in 0.017 s

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 13.0 (TID 13) in 16 ms on localhost (1/1)

22/12/08 16:05:37 INFO DAGScheduler: Job 10 finished: count at TempestAnalytics.java:97, took 0.028053 s

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 13.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 11 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 14 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 14 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_21 stored as values in memory (estimated size 4.8 KB, free 336.7 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_21_piece0 stored as bytes in memory (estimated size 2.6 KB, free 339.2 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_21_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 21 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 14 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 14.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 14.0 (TID 14, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 14.0 (TID 14)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 14.0 (TID 14). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 14.0 (TID 14) in 10 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 14.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 14 (collect at TempestAnalytics.java:98) finished in 0.011 s

22/12/08 16:05:37 INFO DAGScheduler: Job 11 finished: collect at TempestAnalytics.java:98, took 0.016317 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 12 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 15 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 15 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_22 stored as values in memory (estimated size 4.6 KB, free 343.9 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_22_piece0 stored as bytes in memory (estimated size 2.5 KB, free 346.4 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_22_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 22 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 15 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 15.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 15.0 (TID 15, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 15.0 (TID 15)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

So dear the love my people bore me, nor set

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 15.0 (TID 15). 2082 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 15.0 (TID 15) in 9 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 15.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 15 (count at TempestAnalytics.java:97) finished in 0.009 s

22/12/08 16:05:37 INFO DAGScheduler: Job 12 finished: count at TempestAnalytics.java:97, took 0.028541 s

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 13 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 16 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 16 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_23 stored as values in memory (estimated size 4.8 KB, free 351.2 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_23_piece0 stored as bytes in memory (estimated size 2.6 KB, free 353.7 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_23_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 23 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 16 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 16.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 16.0 (TID 16, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 16.0 (TID 16)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 16.0 (TID 16). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 16.0 (TID 16) in 7 ms on localhost (1/1)

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 16 (collect at TempestAnalytics.java:98) finished in 0.007 s

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 16.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: Job 13 finished: collect at TempestAnalytics.java:98, took 0.012377 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 14 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 17 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 17 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_24 stored as values in memory (estimated size 4.6 KB, free 358.4 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_24_piece0 stored as bytes in memory (estimated size 2.5 KB, free 360.9 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_24_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 24 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 17 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 17.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 17.0 (TID 17, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 17.0 (TID 17)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 17.0 (TID 17). 2082 bytes result sent to driver

Knowing I loved my books, he furnished me

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 17.0 (TID 17) in 8 ms on localhost (1/1)

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 17 (count at TempestAnalytics.java:97) finished in 0.009 s

22/12/08 16:05:37 INFO DAGScheduler: Job 14 finished: count at TempestAnalytics.java:97, took 0.015983 s

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 17.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 15 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 18 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 18 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_25 stored as values in memory (estimated size 4.8 KB, free 365.7 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_25_piece0 stored as bytes in memory (estimated size 2.6 KB, free 368.2 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_25_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 25 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 18 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 18.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 18.0 (TID 18, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 18.0 (TID 18)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 18.0 (TID 18). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 18.0 (TID 18) in 8 ms on localhost (1/1)

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 18 (collect at TempestAnalytics.java:98) finished in 0.008 s

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 18.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: Job 15 finished: collect at TempestAnalytics.java:98, took 0.016195 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 16 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 19 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 19 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_26 stored as values in memory (estimated size 4.6 KB, free 372.8 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_26_piece0 stored as bytes in memory (estimated size 2.5 KB, free 375.4 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_26_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 26 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 19 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 19.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 19.0 (TID 19, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 19.0 (TID 19)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

Into a cloven pine, within which rift

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 19.0 (TID 19). 2082 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 19.0 (TID 19) in 10 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 19.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 19 (count at TempestAnalytics.java:97) finished in 0.011 s

22/12/08 16:05:37 INFO DAGScheduler: Job 16 finished: count at TempestAnalytics.java:97, took 0.015963 s

22/12/08 16:05:37 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:37 INFO DAGScheduler: Got job 17 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 20 (collect at TempestAnalytics.java:98)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 20 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_27 stored as values in memory (estimated size 4.8 KB, free 380.2 KB)

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_27_piece0 stored as bytes in memory (estimated size 2.6 KB, free 382.7 KB)

22/12/08 16:05:37 INFO BlockManagerInfo: Added broadcast_27_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:37 INFO SparkContext: Created broadcast 27 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:37 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 20 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Adding task set 20.0 with 1 tasks

22/12/08 16:05:37 INFO TaskSetManager: Starting task 0.0 in stage 20.0 (TID 20, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:37 INFO Executor: Running task 0.0 in stage 20.0 (TID 20)

22/12/08 16:05:37 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:37 INFO Executor: Finished task 0.0 in stage 20.0 (TID 20). 2945 bytes result sent to driver

22/12/08 16:05:37 INFO TaskSetManager: Finished task 0.0 in stage 20.0 (TID 20) in 9 ms on localhost (1/1)

22/12/08 16:05:37 INFO TaskSchedulerImpl: Removed TaskSet 20.0, whose tasks have all completed, from pool

22/12/08 16:05:37 INFO DAGScheduler: ResultStage 20 (collect at TempestAnalytics.java:98) finished in 0.009 s

22/12/08 16:05:37 INFO DAGScheduler: Job 17 finished: collect at TempestAnalytics.java:98, took 0.014751 s

22/12/08 16:05:37 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:37 INFO DAGScheduler: Got job 18 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:37 INFO DAGScheduler: Final stage: ResultStage 21 (count at TempestAnalytics.java:97)

22/12/08 16:05:37 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:37 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:37 INFO DAGScheduler: Submitting ResultStage 21 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:37 INFO MemoryStore: Block broadcast_28 stored as values in memory (estimated size 4.6 KB, free 387.3 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_28_piece0 stored as bytes in memory (estimated size 2.5 KB, free 389.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_28_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 28 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 21 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 21.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 21.0 (TID 21, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

I do not love to look on.

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 21.0 (TID 21)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 21.0 (TID 21). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 21.0 (TID 21) in 14 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 21.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 21 (count at TempestAnalytics.java:97) finished in 0.015 s

22/12/08 16:05:38 INFO DAGScheduler: Job 18 finished: count at TempestAnalytics.java:97, took 0.020169 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 19 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 22 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 22 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_29 stored as values in memory (estimated size 4.8 KB, free 394.7 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_29_piece0 stored as bytes in memory (estimated size 2.6 KB, free 397.2 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_29_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 29 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 22 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 22.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 22.0 (TID 22, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 22.0 (TID 22)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 22.0 (TID 22). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 22.0 (TID 22) in 14 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 22 (collect at TempestAnalytics.java:98) finished in 0.015 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 22.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 19 finished: collect at TempestAnalytics.java:98, took 0.027052 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 20 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 23 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 23 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_30 stored as values in memory (estimated size 4.6 KB, free 401.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_30_piece0 stored as bytes in memory (estimated size 2.5 KB, free 404.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_30_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 30 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 23 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 23.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 23.0 (TID 23, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 23.0 (TID 23)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

That burn by day and night. And then I loved thee,

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 23.0 (TID 23). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 23.0 (TID 23) in 10 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 23.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 23 (count at TempestAnalytics.java:97) finished in 0.010 s

22/12/08 16:05:38 INFO DAGScheduler: Job 20 finished: count at TempestAnalytics.java:97, took 0.023471 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 21 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 24 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 24 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_31 stored as values in memory (estimated size 4.8 KB, free 409.1 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_31_piece0 stored as bytes in memory (estimated size 2.6 KB, free 411.7 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_31_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 31 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 24 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 24.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 24.0 (TID 24, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 24.0 (TID 24)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 24.0 (TID 24). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 24.0 (TID 24) in 13 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 24.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 24 (collect at TempestAnalytics.java:98) finished in 0.013 s

22/12/08 16:05:38 INFO DAGScheduler: Job 21 finished: collect at TempestAnalytics.java:98, took 0.020516 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 22 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 25 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 25 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_32 stored as values in memory (estimated size 4.6 KB, free 416.3 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_32_piece0 stored as bytes in memory (estimated size 2.5 KB, free 418.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_32_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 32 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 25 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 25.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 25.0 (TID 25, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 25.0 (TID 25)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 25.0 (TID 25). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 25.0 (TID 25) in 10 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 25 (count at TempestAnalytics.java:97) finished in 0.012 s

22/12/08 16:05:38 INFO DAGScheduler: Job 22 finished: count at TempestAnalytics.java:97, took 0.018028 s

And I the King shall love thee.

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 25.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 23 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 26 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 26 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_33 stored as values in memory (estimated size 4.8 KB, free 423.6 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_33_piece0 stored as bytes in memory (estimated size 2.6 KB, free 426.2 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_33_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 33 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 26 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 26.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 26.0 (TID 26, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 26.0 (TID 26)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 26.0 (TID 26). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 26.0 (TID 26) in 11 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 26.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 26 (collect at TempestAnalytics.java:98) finished in 0.011 s

22/12/08 16:05:38 INFO DAGScheduler: Job 23 finished: collect at TempestAnalytics.java:98, took 0.026665 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 24 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 27 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 27 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_34 stored as values in memory (estimated size 4.6 KB, free 430.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_34_piece0 stored as bytes in memory (estimated size 2.5 KB, free 433.4 KB)

All wound with adders, who with cloven tongues

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_34_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 34 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 27 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 27.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 27.0 (TID 27, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 27.0 (TID 27)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 27.0 (TID 27). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 27 (count at TempestAnalytics.java:97) finished in 0.008 s

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 27.0 (TID 27) in 8 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: Job 24 finished: count at TempestAnalytics.java:97, took 0.023845 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 27.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 25 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 28 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 28 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_35 stored as values in memory (estimated size 4.8 KB, free 438.1 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_35_piece0 stored as bytes in memory (estimated size 2.6 KB, free 440.7 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_35_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 35 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 28 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 28.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 28.0 (TID 28, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 28.0 (TID 28)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 28.0 (TID 28). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 28.0 (TID 28) in 9 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 28.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 28 (collect at TempestAnalytics.java:98) finished in 0.010 s

22/12/08 16:05:38 INFO DAGScheduler: Job 25 finished: collect at TempestAnalytics.java:98, took 0.018058 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 26 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 29 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 29 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_36 stored as values in memory (estimated size 4.6 KB, free 445.3 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_36_piece0 stored as bytes in memory (estimated size 2.5 KB, free 447.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_36_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

        She loved not the savor of tar nor of pitch,

22/12/08 16:05:38 INFO SparkContext: Created broadcast 36 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 29 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 29.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 29.0 (TID 29, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 29.0 (TID 29)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 29.0 (TID 29). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 29.0 (TID 29) in 6 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 29.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 29 (count at TempestAnalytics.java:97) finished in 0.009 s

22/12/08 16:05:38 INFO DAGScheduler: Job 26 finished: count at TempestAnalytics.java:97, took 0.012553 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 27 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 30 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 30 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_37 stored as values in memory (estimated size 4.8 KB, free 452.6 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_37_piece0 stored as bytes in memory (estimated size 2.6 KB, free 455.2 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_37_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 37 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 30 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 30.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 30.0 (TID 30, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 30.0 (TID 30)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 30.0 (TID 30). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 30.0 (TID 30) in 15 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 30.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 30 (collect at TempestAnalytics.java:98) finished in 0.018 s

22/12/08 16:05:38 INFO DAGScheduler: Job 27 finished: collect at TempestAnalytics.java:98, took 0.023072 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 28 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 31 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 31 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_38 stored as values in memory (estimated size 4.6 KB, free 459.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_38_piece0 stored as bytes in memory (estimated size 2.5 KB, free 462.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_38_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 38 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 31 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 31.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 31.0 (TID 31, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 31.0 (TID 31)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 31.0 (TID 31). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 31.0 (TID 31) in 8 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 31.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 31 (count at TempestAnalytics.java:97) finished in 0.008 s

22/12/08 16:05:38 INFO DAGScheduler: Job 28 finished: count at TempestAnalytics.java:97, took 0.014095 s

MIRANDA Do you love me?

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 29 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 32 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 32 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_39 stored as values in memory (estimated size 4.8 KB, free 467.1 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_39_piece0 stored as bytes in memory (estimated size 2.6 KB, free 469.7 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_39_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 39 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 32 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 32.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 32.0 (TID 32, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 32.0 (TID 32)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 32.0 (TID 32). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 32.0 (TID 32) in 8 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 32 (collect at TempestAnalytics.java:98) finished in 0.010 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 32.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 29 finished: collect at TempestAnalytics.java:98, took 0.019945 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 30 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 33 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 33 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

Do love, prize, honor you.

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_40 stored as values in memory (estimated size 4.6 KB, free 474.3 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_40_piece0 stored as bytes in memory (estimated size 2.5 KB, free 476.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_40_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 40 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 33 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 33.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 33.0 (TID 33, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 33.0 (TID 33)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 33.0 (TID 33). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 33.0 (TID 33) in 19 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 33.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 33 (count at TempestAnalytics.java:97) finished in 0.019 s

22/12/08 16:05:38 INFO DAGScheduler: Job 30 finished: count at TempestAnalytics.java:97, took 0.035138 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 31 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 34 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 34 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_41 stored as values in memory (estimated size 4.8 KB, free 481.6 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_41_piece0 stored as bytes in memory (estimated size 2.6 KB, free 484.2 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_41_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 41 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 34 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 34.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 34.0 (TID 34, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 34.0 (TID 34)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 34.0 (TID 34). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 34.0 (TID 34) in 7 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 34.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 34 (collect at TempestAnalytics.java:98) finished in 0.008 s

22/12/08 16:05:38 INFO DAGScheduler: Job 31 finished: collect at TempestAnalytics.java:98, took 0.014912 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 32 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 35 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 35 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_42 stored as values in memory (estimated size 4.6 KB, free 488.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_42_piece0 stored as bytes in memory (estimated size 2.5 KB, free 491.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_42_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 42 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 35 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 35.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 35.0 (TID 35, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 35.0 (TID 35)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

And his and mine loved darling. [He exits, above.]

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 35.0 (TID 35). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 35.0 (TID 35) in 10 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 35.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 35 (count at TempestAnalytics.java:97) finished in 0.011 s

22/12/08 16:05:38 INFO DAGScheduler: Job 32 finished: count at TempestAnalytics.java:97, took 0.016511 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 33 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 36 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 36 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_43 stored as values in memory (estimated size 4.8 KB, free 496.1 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_43_piece0 stored as bytes in memory (estimated size 2.6 KB, free 498.7 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_43_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 43 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 36 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 36.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 36.0 (TID 36, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 36.0 (TID 36)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 36.0 (TID 36). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 36.0 (TID 36) in 14 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 36 (collect at TempestAnalytics.java:98) finished in 0.014 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 36.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 33 finished: collect at TempestAnalytics.java:98, took 0.024600 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 34 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 37 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 37 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_44 stored as values in memory (estimated size 4.6 KB, free 503.3 KB)

Were but my trials of thy love, and thou

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_44_piece0 stored as bytes in memory (estimated size 2.5 KB, free 505.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_44_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_43_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 44 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 37

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 37 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 37.0 with 1 tasks

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_42_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 36

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_41_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 35

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_40_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 34

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_39_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 33

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 37.0 (TID 37, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 37.0 (TID 37)

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_38_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 32

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_37_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 31

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_36_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 30

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_35_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 29

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_34_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 28

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_33_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 27

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_32_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 26

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_31_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 25

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_30_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 24

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_29_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 37.0 (TID 37). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 23

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_28_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 22

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_27_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 37.0 (TID 37) in 34 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 37.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 21

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 37 (count at TempestAnalytics.java:97) finished in 0.036 s

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_26_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 20

22/12/08 16:05:38 INFO DAGScheduler: Job 34 finished: count at TempestAnalytics.java:97, took 0.110636 s

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_25_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 19

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_24_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 18

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_23_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 17

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_22_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO DAGScheduler: Got job 35 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 38 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 16

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 38 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_21_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 15

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_45 stored as values in memory (estimated size 4.8 KB, free 343.8 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_20_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 14

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_45_piece0 stored as bytes in memory (estimated size 2.6 KB, free 339.2 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_45_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_19_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 13

22/12/08 16:05:38 INFO SparkContext: Created broadcast 45 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_18_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 38 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 38.0 with 1 tasks

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 12

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_17_piece0 on localhost:57107 in memory (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 11

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_16_piece0 on localhost:57107 in memory (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 38.0 (TID 38, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 10

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 38.0 (TID 38)

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_14_piece0 on localhost:57107 in memory (size: 1894.0 B, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 9

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_13_piece0 on localhost:57107 in memory (size: 3.1 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO ContextCleaner: Cleaned accumulator 8

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO ContextCleaner: Cleaned shuffle 2

22/12/08 16:05:38 INFO BlockManagerInfo: Removed broadcast_12_piece0 on localhost:57107 in memory (size: 13.8 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 38.0 (TID 38). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 38 (collect at TempestAnalytics.java:98) finished in 0.018 s

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 38.0 (TID 38) in 18 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 38.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 35 finished: collect at TempestAnalytics.java:98, took 0.053451 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 36 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 39 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 39 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_46 stored as values in memory (estimated size 4.6 KB, free 160.1 KB)

With such love as 'tis now, the murkiest den,

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_46_piece0 stored as bytes in memory (estimated size 2.5 KB, free 162.6 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_46_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 46 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 39 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 39.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 39.0 (TID 39, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 39.0 (TID 39)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 39.0 (TID 39). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 39.0 (TID 39) in 18 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 39.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 39 (count at TempestAnalytics.java:97) finished in 0.019 s

22/12/08 16:05:38 INFO DAGScheduler: Job 36 finished: count at TempestAnalytics.java:97, took 0.029319 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 37 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 40 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 40 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_47 stored as values in memory (estimated size 4.8 KB, free 167.4 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_47_piece0 stored as bytes in memory (estimated size 2.6 KB, free 169.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_47_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 47 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 40 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 40.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 40.0 (TID 40, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 40.0 (TID 40)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 40.0 (TID 40). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 40.0 (TID 40) in 14 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 40.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 40 (collect at TempestAnalytics.java:98) finished in 0.017 s

22/12/08 16:05:38 INFO DAGScheduler: Job 37 finished: collect at TempestAnalytics.java:98, took 0.025622 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

     Do you love me, master? No?

22/12/08 16:05:38 INFO DAGScheduler: Got job 38 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 41 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 41 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_48 stored as values in memory (estimated size 4.6 KB, free 174.5 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_48_piece0 stored as bytes in memory (estimated size 2.5 KB, free 177.1 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_48_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 48 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 41 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 41.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 41.0 (TID 41, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 41.0 (TID 41)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 41.0 (TID 41). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 41.0 (TID 41) in 11 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 41.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 41 (count at TempestAnalytics.java:97) finished in 0.011 s

22/12/08 16:05:38 INFO DAGScheduler: Job 38 finished: count at TempestAnalytics.java:97, took 0.025756 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 39 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 42 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 42 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_49 stored as values in memory (estimated size 4.8 KB, free 181.9 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_49_piece0 stored as bytes in memory (estimated size 2.6 KB, free 184.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_49_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 49 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 42 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 42.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 42.0 (TID 42, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 42.0 (TID 42)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 42.0 (TID 42). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 42.0 (TID 42) in 18 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 42.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 42 (collect at TempestAnalytics.java:98) finished in 0.019 s

22/12/08 16:05:38 INFO DAGScheduler: Job 39 finished: collect at TempestAnalytics.java:98, took 0.032633 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

Whose shadow the dismissed bachelor loves,

22/12/08 16:05:38 INFO DAGScheduler: Got job 40 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 43 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 43 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_50 stored as values in memory (estimated size 4.6 KB, free 189.0 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_50_piece0 stored as bytes in memory (estimated size 2.5 KB, free 191.6 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_50_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 50 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 43 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 43.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 43.0 (TID 43, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 43.0 (TID 43)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 43.0 (TID 43). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 43.0 (TID 43) in 17 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 43 (count at TempestAnalytics.java:97) finished in 0.019 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 43.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 40 finished: count at TempestAnalytics.java:97, took 0.038268 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 41 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 44 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 44 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_51 stored as values in memory (estimated size 4.8 KB, free 196.4 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_51_piece0 stored as bytes in memory (estimated size 2.6 KB, free 198.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_51_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 51 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 44 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 44.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 44.0 (TID 44, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 44.0 (TID 44)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 44.0 (TID 44). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 44.0 (TID 44) in 16 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 44 (collect at TempestAnalytics.java:98) finished in 0.017 s

22/12/08 16:05:38 INFO DAGScheduler: Job 41 finished: collect at TempestAnalytics.java:98, took 0.028359 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 44.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 42 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 45 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 45 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

A contract of true love to celebrate,

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_52 stored as values in memory (estimated size 4.6 KB, free 203.5 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_52_piece0 stored as bytes in memory (estimated size 2.5 KB, free 206.1 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_52_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 52 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 45 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 45.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 45.0 (TID 45, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 45.0 (TID 45)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 45.0 (TID 45). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 45.0 (TID 45) in 22 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 45.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 45 (count at TempestAnalytics.java:97) finished in 0.023 s

22/12/08 16:05:38 INFO DAGScheduler: Job 42 finished: count at TempestAnalytics.java:97, took 0.034360 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 43 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 46 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 46 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_53 stored as values in memory (estimated size 4.8 KB, free 210.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_53_piece0 stored as bytes in memory (estimated size 2.6 KB, free 213.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_53_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 53 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 46 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 46.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 46.0 (TID 46, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 46.0 (TID 46)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 46.0 (TID 46). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 46.0 (TID 46) in 13 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 46 (collect at TempestAnalytics.java:98) finished in 0.013 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 46.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 43 finished: collect at TempestAnalytics.java:98, took 0.019277 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 44 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 47 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 47 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_54 stored as values in memory (estimated size 4.6 KB, free 218.0 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_54_piece0 stored as bytes in memory (estimated size 2.5 KB, free 220.6 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_54_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 54 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 47 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 47.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 47.0 (TID 47, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 47.0 (TID 47)

On the blest lovers.

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 47.0 (TID 47). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 47.0 (TID 47) in 11 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 47.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 47 (count at TempestAnalytics.java:97) finished in 0.011 s

22/12/08 16:05:38 INFO DAGScheduler: Job 44 finished: count at TempestAnalytics.java:97, took 0.017433 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 45 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 48 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 48 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_55 stored as values in memory (estimated size 4.8 KB, free 225.3 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_55_piece0 stored as bytes in memory (estimated size 2.6 KB, free 227.9 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_55_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 55 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 48 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 48.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 48.0 (TID 48, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 48.0 (TID 48)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 48.0 (TID 48). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 48.0 (TID 48) in 6 ms on localhost (1/1)

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 48 (collect at TempestAnalytics.java:98) finished in 0.007 s

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 48.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: Job 45 finished: collect at TempestAnalytics.java:98, took 0.014369 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 46 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 49 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 49 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_56 stored as values in memory (estimated size 4.6 KB, free 232.5 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_56_piece0 stored as bytes in memory (estimated size 2.5 KB, free 235.1 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_56_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 56 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 49 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 49.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 49.0 (TID 49, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 49.0 (TID 49)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

A contract of true love. Be not too late.

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 49.0 (TID 49). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 49.0 (TID 49) in 6 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 49.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 49 (count at TempestAnalytics.java:97) finished in 0.006 s

22/12/08 16:05:38 INFO DAGScheduler: Job 46 finished: count at TempestAnalytics.java:97, took 0.019464 s

22/12/08 16:05:38 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:38 INFO DAGScheduler: Got job 47 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 50 (collect at TempestAnalytics.java:98)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 50 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_57 stored as values in memory (estimated size 4.8 KB, free 239.8 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_57_piece0 stored as bytes in memory (estimated size 2.6 KB, free 242.4 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_57_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 57 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 50 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 50.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 50.0 (TID 50, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 50.0 (TID 50)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 50.0 (TID 50). 2945 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 50.0 (TID 50) in 14 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 50.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 50 (collect at TempestAnalytics.java:98) finished in 0.014 s

22/12/08 16:05:38 INFO DAGScheduler: Job 47 finished: collect at TempestAnalytics.java:98, took 0.018670 s

22/12/08 16:05:38 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:38 INFO DAGScheduler: Got job 48 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:38 INFO DAGScheduler: Final stage: ResultStage 51 (count at TempestAnalytics.java:97)

22/12/08 16:05:38 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:38 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:38 INFO DAGScheduler: Submitting ResultStage 51 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_58 stored as values in memory (estimated size 4.6 KB, free 247.0 KB)

22/12/08 16:05:38 INFO MemoryStore: Block broadcast_58_piece0 stored as bytes in memory (estimated size 2.5 KB, free 249.6 KB)

22/12/08 16:05:38 INFO BlockManagerInfo: Added broadcast_58_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:38 INFO SparkContext: Created broadcast 58 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:38 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 51 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Adding task set 51.0 with 1 tasks

22/12/08 16:05:38 INFO TaskSetManager: Starting task 0.0 in stage 51.0 (TID 51, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:38 INFO Executor: Running task 0.0 in stage 51.0 (TID 51)

22/12/08 16:05:38 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:38 INFO Executor: Finished task 0.0 in stage 51.0 (TID 51). 2082 bytes result sent to driver

22/12/08 16:05:38 INFO TaskSetManager: Finished task 0.0 in stage 51.0 (TID 51) in 8 ms on localhost (1/1)

22/12/08 16:05:38 INFO TaskSchedulerImpl: Removed TaskSet 51.0, whose tasks have all completed, from pool

22/12/08 16:05:38 INFO DAGScheduler: ResultStage 51 (count at TempestAnalytics.java:97) finished in 0.009 s

22/12/08 16:05:38 INFO DAGScheduler: Job 48 finished: count at TempestAnalytics.java:97, took 0.016193 s

FERDINAND No, my dearest love,

22/12/08 16:05:39 INFO SparkContext: Starting job: collect at TempestAnalytics.java:98

22/12/08 16:05:39 INFO DAGScheduler: Got job 49 (collect at TempestAnalytics.java:98) with 1 output partitions

22/12/08 16:05:39 INFO DAGScheduler: Final stage: ResultStage 52 (collect at TempestAnalytics.java:98)

22/12/08 16:05:39 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:39 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:39 INFO DAGScheduler: Submitting ResultStage 52 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:39 INFO MemoryStore: Block broadcast_59 stored as values in memory (estimated size 4.8 KB, free 254.3 KB)

22/12/08 16:05:39 INFO MemoryStore: Block broadcast_59_piece0 stored as bytes in memory (estimated size 2.6 KB, free 256.9 KB)

22/12/08 16:05:39 INFO BlockManagerInfo: Added broadcast_59_piece0 in memory on localhost:57107 (size: 2.6 KB, free: 2.4 GB)

22/12/08 16:05:39 INFO SparkContext: Created broadcast 59 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:39 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 52 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:39 INFO TaskSchedulerImpl: Adding task set 52.0 with 1 tasks

22/12/08 16:05:39 INFO TaskSetManager: Starting task 0.0 in stage 52.0 (TID 52, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:39 INFO Executor: Running task 0.0 in stage 52.0 (TID 52)

22/12/08 16:05:39 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:39 INFO Executor: Finished task 0.0 in stage 52.0 (TID 52). 2945 bytes result sent to driver

22/12/08 16:05:39 INFO TaskSetManager: Finished task 0.0 in stage 52.0 (TID 52) in 6 ms on localhost (1/1)

22/12/08 16:05:39 INFO TaskSchedulerImpl: Removed TaskSet 52.0, whose tasks have all completed, from pool

22/12/08 16:05:39 INFO DAGScheduler: ResultStage 52 (collect at TempestAnalytics.java:98) finished in 0.006 s

22/12/08 16:05:39 INFO DAGScheduler: Job 49 finished: collect at TempestAnalytics.java:98, took 0.012813 s

22/12/08 16:05:39 INFO SparkContext: Starting job: count at TempestAnalytics.java:97

22/12/08 16:05:39 INFO DAGScheduler: Got job 50 (count at TempestAnalytics.java:97) with 1 output partitions

22/12/08 16:05:39 INFO DAGScheduler: Final stage: ResultStage 53 (count at TempestAnalytics.java:97)

22/12/08 16:05:39 INFO DAGScheduler: Parents of final stage: List()

22/12/08 16:05:39 INFO DAGScheduler: Missing parents: List()

22/12/08 16:05:39 INFO DAGScheduler: Submitting ResultStage 53 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96), which has no missing parents

22/12/08 16:05:39 INFO MemoryStore: Block broadcast_60 stored as values in memory (estimated size 4.6 KB, free 261.5 KB)

22/12/08 16:05:39 INFO MemoryStore: Block broadcast_60_piece0 stored as bytes in memory (estimated size 2.5 KB, free 264.1 KB)

22/12/08 16:05:39 INFO BlockManagerInfo: Added broadcast_60_piece0 in memory on localhost:57107 (size: 2.5 KB, free: 2.4 GB)

22/12/08 16:05:39 INFO SparkContext: Created broadcast 60 from broadcast at DAGScheduler.scala:1006

22/12/08 16:05:39 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 53 (MapPartitionsRDD[36] at filter at TempestAnalytics.java:96)

22/12/08 16:05:39 INFO TaskSchedulerImpl: Adding task set 53.0 with 1 tasks

22/12/08 16:05:39 INFO TaskSetManager: Starting task 0.0 in stage 53.0 (TID 53, localhost, partition 0,PROCESS_LOCAL, 2196 bytes)

22/12/08 16:05:39 INFO Executor: Running task 0.0 in stage 53.0 (TID 53)

Of these our dear-beloved solemnized,

22/12/08 16:05:39 INFO HadoopRDD: Input split: file:/Users/linhungfan/Desktop/CMU/Semester 2/Distributed System/Project5/Part2/TheTempest.txt:0+99846

22/12/08 16:05:39 INFO Executor: Finished task 0.0 in stage 53.0 (TID 53). 2082 bytes result sent to driver

22/12/08 16:05:39 INFO TaskSetManager: Finished task 0.0 in stage 53.0 (TID 53) in 24 ms on localhost (1/1)

22/12/08 16:05:39 INFO DAGScheduler: ResultStage 53 (count at TempestAnalytics.java:97) finished in 0.024 s

22/12/08 16:05:39 INFO TaskSchedulerImpl: Removed TaskSet 53.0, whose tasks have all completed, from pool

22/12/08 16:05:39 INFO DAGScheduler: Job 50 finished: count at TempestAnalytics.java:97, took 0.029099 s

22/12/08 16:05:39 INFO SparkContext: Invoking stop() from shutdown hook

22/12/08 16:05:39 INFO SparkUI: Stopped Spark web UI at http://10.37.129.2:4040

22/12/08 16:05:39 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!

22/12/08 16:05:39 INFO MemoryStore: MemoryStore cleared

22/12/08 16:05:39 INFO BlockManager: BlockManager stopped

22/12/08 16:05:39 INFO BlockManagerMaster: BlockManagerMaster stopped

22/12/08 16:05:39 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!

22/12/08 16:05:39 INFO SparkContext: Successfully stopped SparkContext

22/12/08 16:05:39 INFO ShutdownHookManager: Shutdown hook called

22/12/08 16:05:39 INFO ShutdownHookManager: Deleting directory /private/var/folders/55/rnjlg8zj0yng2nbxj9dmfjn80000gn/T/spark-0f488ab5-56e2-41dd-a812-4997428cba22

Process finished with exit code 0