**Session 21**

**Assignment 2**

**Problem Statement:** Explain the different methods of

* Mapper class
* Reducer class
* **Mapper Class:**

1. **Setup:**

protected void setup (org.apache.hadoop.mapreduce.Mapper.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException{ .. }](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

The setup method is the first of Mapper class methods and is called once at the beginning of the task.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Map:**

protected void map([KEYIN](https://hadoop.apache.org/docs/current/api/org/apache/hadoop/mapreduce/Mapper.html) key, [VALUEIN](https://hadoop.apache.org/docs/current/api/org/apache/hadoop/mapreduce/Mapper.html) value, org.apache.hadoop.mapreduce.Mapper.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

This method is called once for each key or value pair in the input split. Most applications should override this, but the default is the identity function.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Cleanup:**

protected void cleanup(org.apache.hadoop.mapreduce.Mapper.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

This method is called at the end of the task.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Run:**

public void run(org.apache.hadoop.mapreduce.Mapper.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

Expert users can override this method for more complete control over the execution of the Mapper.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)
* **Reducer Class:**

1. **Setup:**

protected void [setup](https://hadoop.apache.org/docs/r2.7.0/api/src-html/org/apache/hadoop/mapreduce/Reducer.html#line.136)(org.apache.hadoop.mapreduce.Reducer.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

This method is called once at the start of the task.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Reduce:**

protected void [reduce](https://hadoop.apache.org/docs/r2.7.0/api/src-html/org/apache/hadoop/mapreduce/Reducer.html#line.147)([KEYIN](https://hadoop.apache.org/docs/r2.7.0/api/org/apache/hadoop/mapreduce/Reducer.html) key, [Iterable](http://docs.oracle.com/javase/7/docs/api/java/lang/Iterable.html?is-external=true)<[VALUEIN](https://hadoop.apache.org/docs/r2.7.0/api/org/apache/hadoop/mapreduce/Reducer.html)> value,

org.apache.hadoop.mapreduce.Reducer.Context context)

throws [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true),[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

This method is called once for each key. Most applications will define their reduce class by overriding this method. The default implementation is an identity function.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Cleanup:**

protected void [cleanup](https://hadoop.apache.org/docs/r2.7.0/api/src-html/org/apache/hadoop/mapreduce/Reducer.html#line.157)(org.apache.hadoop.mapreduce.Reducer.Context context)

throws IOException,[InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true){ .. }

This method is called once at the end of the task.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)

1. **Run:**

public void run(org.apache.hadoop.mapreduce.Reducer.Context context)

throws IOException,InterruptedException{ .. }

Advanced application writers can use the run (org.apache.hadoop.mapreduce.Reducer.Context) method to control how the reduce task works.

**Throws:**

* [IOException](http://docs.oracle.com/javase/7/docs/api/java/io/IOException.html?is-external=true)
* [InterruptedException](http://docs.oracle.com/javase/7/docs/api/java/lang/InterruptedException.html?is-external=true)