Mapper and Reducer API - LifeCycle

Mapper

org.apache.hadoop.mapreduce Class Mapper<KEYIN,VALUEIN,KEYOUT,VALUEOUT>

Method Summary	
All Methods Instance Methods	Concrete Methods
Modifier and Type	Method and Description
protected void	<pre>cleanup(org.apache.hadoop.mapreduce.Mapper.Context context) Called once at the end of the task.</pre>
protected void	<pre>map(KEYIN key, VALUEIN value, org.apache.hadoop.mapreduce.Mapper.Context context Called once for each key/value pair in the input split.</pre>
void	<pre>run(org.apache.hadoop.mapreduce.Mapper.Context context)</pre> <pre>Expert users can override this method for more complete control over the execution of the Mapper.</pre>
protected void	<pre>setup(org.apache.hadoop.mapreduce.Mapper.Context context) Called once at the beginning of the task.</pre>

org.apache.hadoop.mapreduce

Class Reducer<KEYIN, VALUEIN, KEYOUT, VALUEOUT>

- java.lang.Object
 - $\bullet \quad \text{org.apache.hadoop.mapreduce.Reducer} < \text{KEYIN,VALUEIN,KEYOUT,VALUEOUT}$



P.s. Map-Reduce with hadoop 1.0 is still continued as a different package at org.apache.hadoop.mapred

Note the difference between mapred (V1) vs mapreduce (V2).

Also note the difference between interface and abstract class usage.

Reference:

https://hadoop.apache.org/docs/r3.2.1/api/index.html

V2 org.apache.hadoop.mapreduce

Documented By:
Jigar M. Pandya
https://www.linkedin.com/in/jigar-pandya

Document Last updated: 30th July, 2020