Hortonworks

Apache Hadoop YARN and Tez

Future of Data Processing with Hadoop

Chris Harris
charris@hortonworks.com

Demo Benchmark Spec

- **The TPC-DS benchmark data+query set**

- **Query 27**
  - big table(store_sales) joins lots of small tables
  - A.K.A Star Schema Join

- **What does Query 27 do?**

  For all items sold in stores located in specified states during a given year, find the average quantity, average list price, average list sales price, average coupon amount for a given gender, marital status, education and customer demographic..

Query 27 - Star Schema Join

- **Derived from TPC-DS Query 27**

```
SELECT col5, avg(col6)
FROM store_sales_fact ssf
    join item_dim on (ssf.col1 = item_dim .col1)
    join date_dim on (ssf.col2 = date_dim.col2
    join custdmgrphcs_dim on (ssf.col3 =custdmgrphcs_dim.col3)
    join store_dim on (ssf.col4 = store_dim.col4)
GROUP BY col5
ORDER BY col5
LIMIT 100;
```

41 GB

58 MB

11MB

80MB

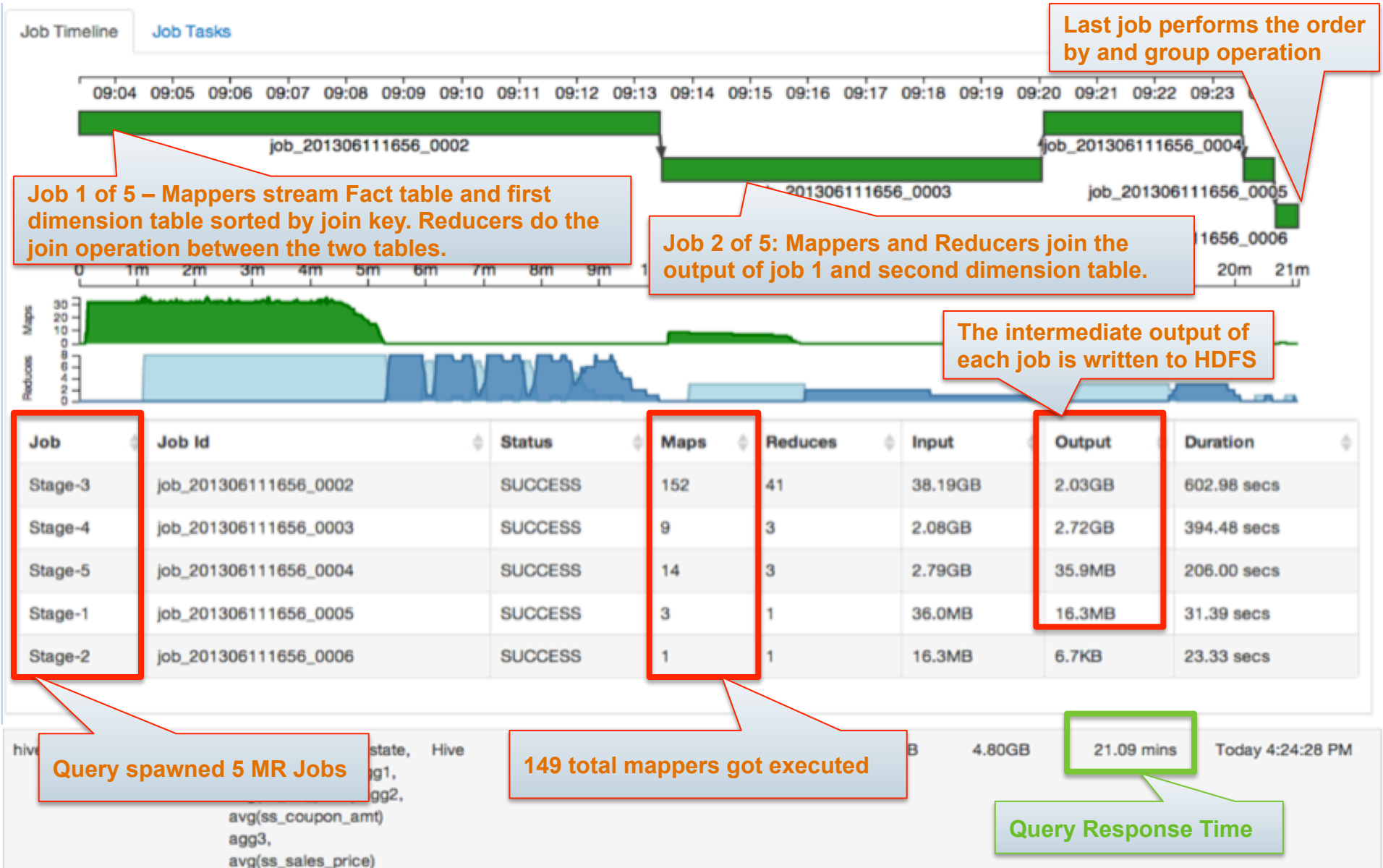
106 KB

Benchmark Cluster Specs

- **6 Node HDP 1.3 Cluster**
  - 2 Master Nodes
  - 4 Data/Slave Nodes

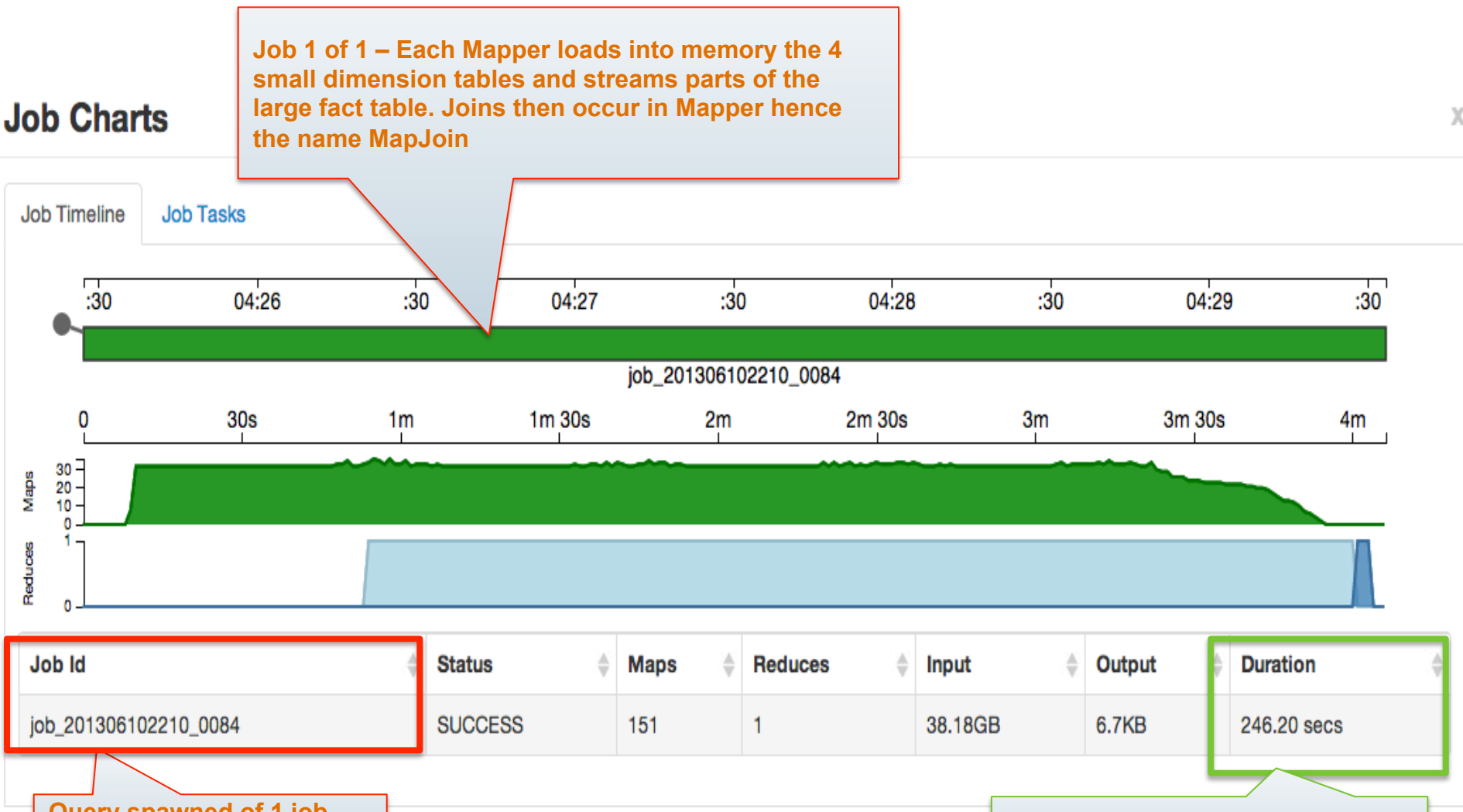
- **Slave Node Specs**
  - Dual Processors
  - 14 GB

Query27 Execution Before Hive 11-Text Format



Query27 Execution With Hive 11-Text Format

Job 1 of 1 – Each Mapper loads into memory the 4 small dimension tables and streams parts of the large fact table. Joins then occur in Mapper hence the name MapJoin

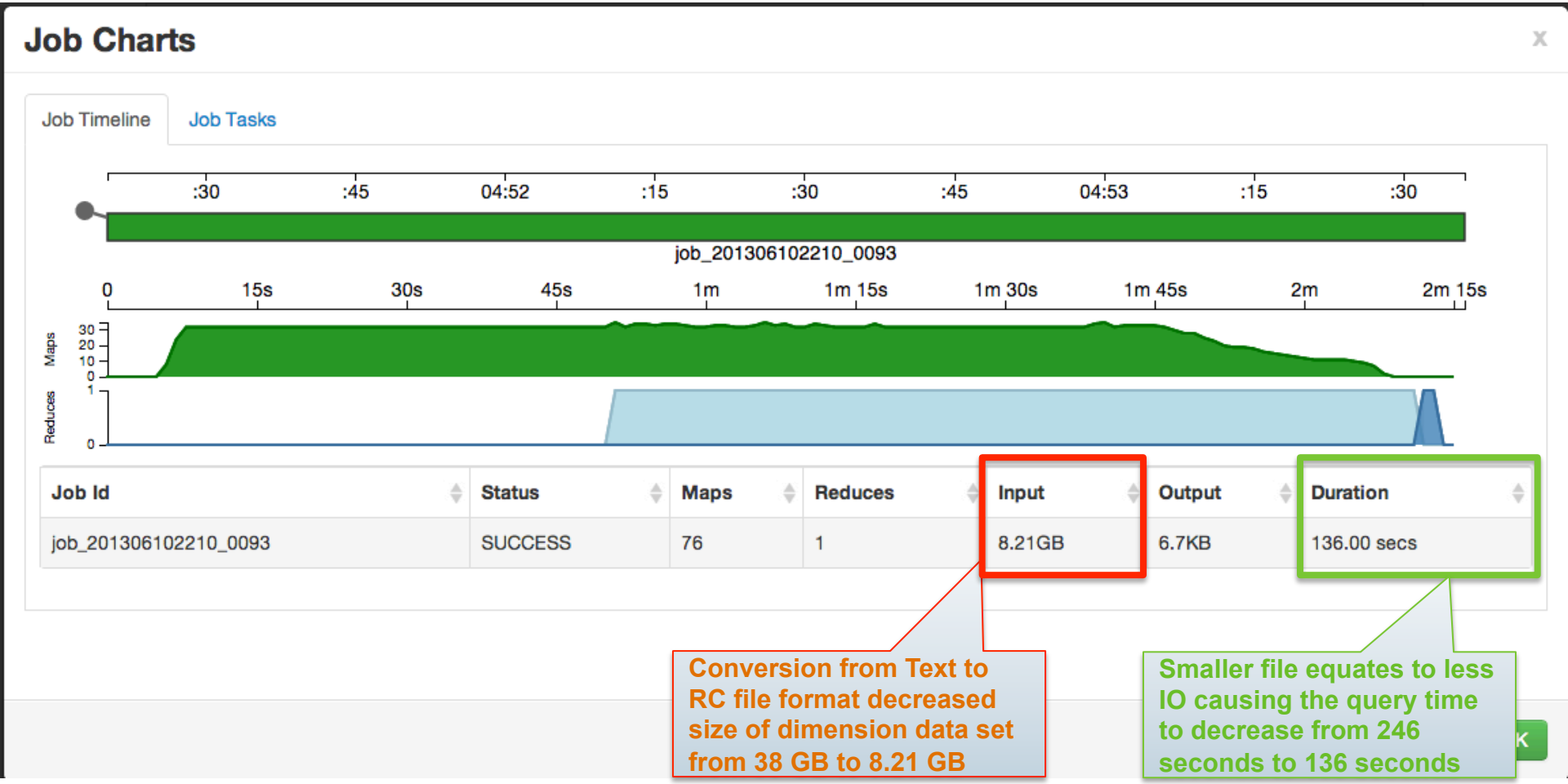


Job Id	Status	Maps	Reduces	Input	Output	Duration
job_201306102210_0084	SUCCESS	151	1	38.18GB	6.7KB	246.20 secs

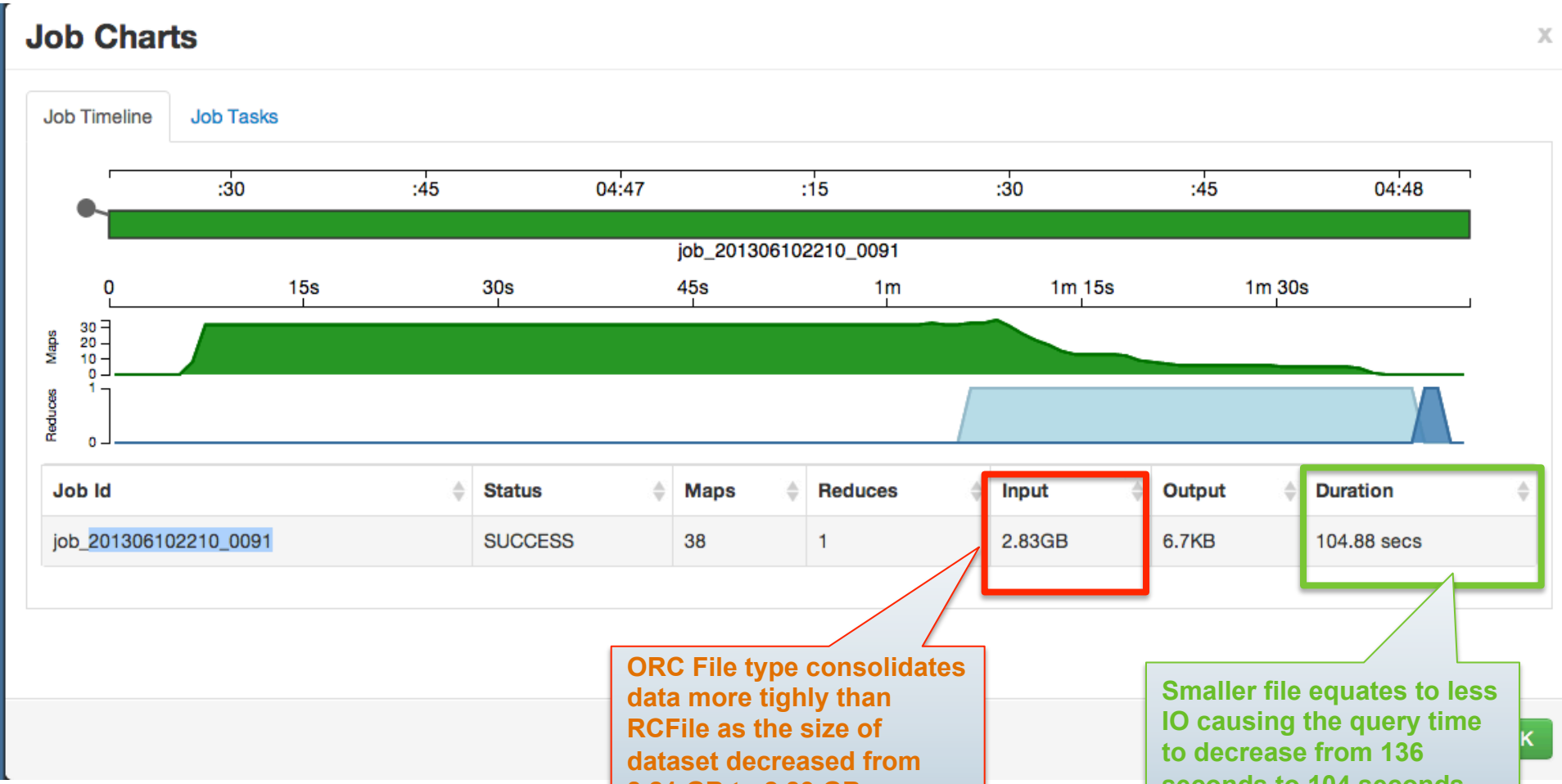
Query spawned of 1 job with Hive 11 compared to 5 MR Jobs with Hive 10

Increase in performance with Hive 11 as query time went down from 21 minutes to about 4 minutes

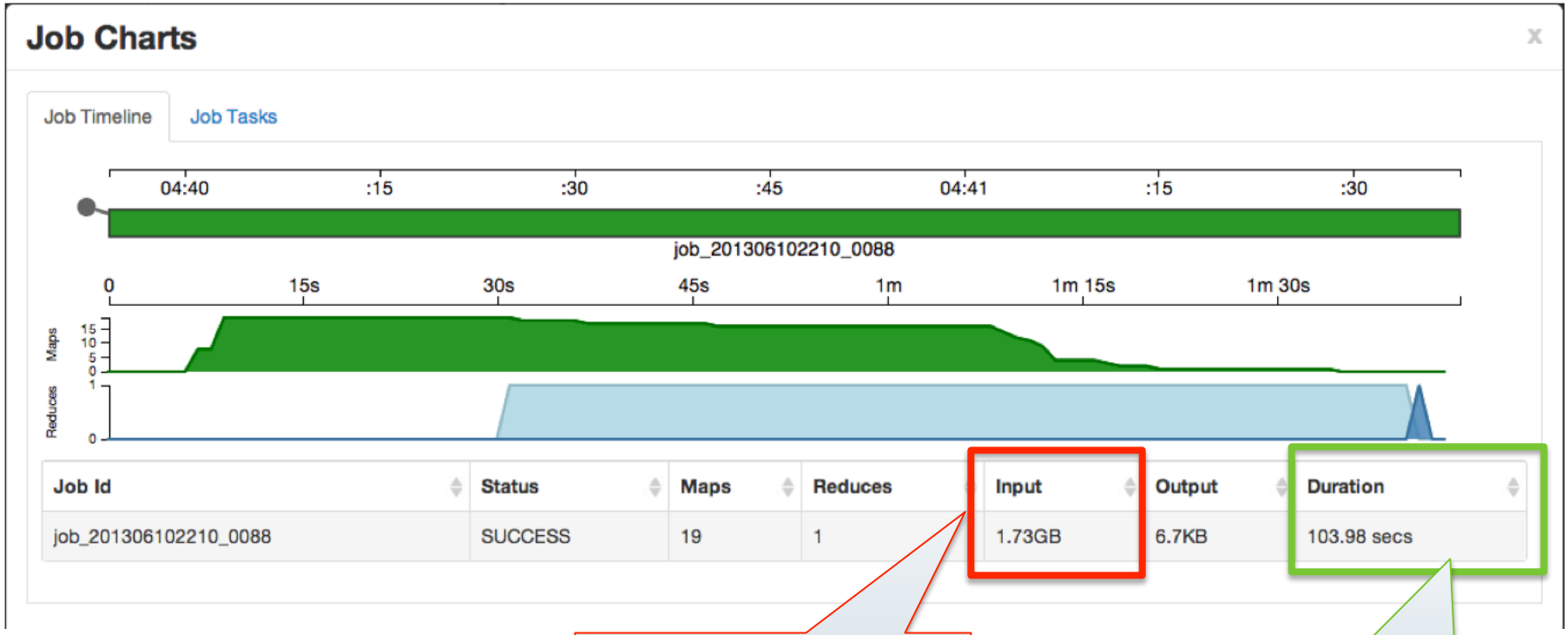
Query27 Execution With Hive 11- RC Format



Query27 Execution With Hive 11- ORC Format



Query27 Execution With Hive 11- RC Format/ Partitioned, Bucketed and Sorted

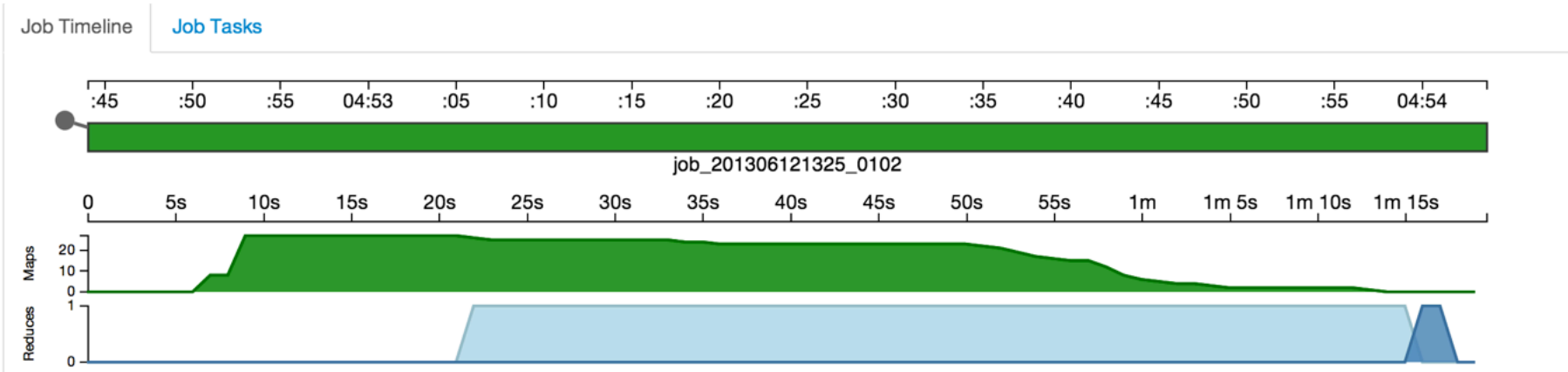


Job Id	Status	Maps	Reduces	Input	Output	Duration
job_201306102210_0088	SUCCESS	19	1	1.73GB	6.7KB	103.98 secs

Partitioned the data decreases the file input size to 1.73 GB

Smaller file equates to less IO causing the query time to decrease from 136 seconds to 104 seconds

Query27 Execution With Hive 11- ORCFormat/ Partitioned, Bucketed and Sorted



Job Id	Status	Maps	Reduces	Input	Output	Duration
job_201306121325_0102	SUCCESS	27	1	687.2MB	6.7KB	79.62 secs

Partitioned Data with ORC file format produces the smallest input size of 687 MB a decrease from 43 GB

Smallest Input Size allows us the fastest response time for the query: 79 seconds

Summary of Results

File Type	Number of MR Jobs	Input Size	Mappers	Time
Text/Hive 10	5	43.1 GB	179	21 minutes
Text/Hive 11	1	38 GB	151	246 seconds
RC/Hive 11	1	8.21 GB	76	136 seconds
ORC/Hive 11	1	2.83 GB	38	104 seconds
RC/Hive 11/ Partitioned/ Bucketed	1	1.73 GB	19	104 seconds
ORC/Hive 11/ Partitioned/ Bucketed	**1**	**687 MB**	**27**	**79.62**

Resources

hadoop-2.0.3-alpha:
http://hadoop.apache.org/common/releases.html

Release Documentation:
http://hadoop.apache.org/common/docs/r2.0.3-alpha/

Blogs:
http://hortonworks.com/blog/category/apache-hadoop/yarn/
http://hortonworks.com/blog/introducing-apache-hadoop-yarn/
http://hortonworks.com/blog/apache-hadoop-yarn-background-and-an-overview/
http://hortonworks.com/blog/apache-hadoop-yarn-background-and-an-overview/

Thank You!

Questions & Answers