Analysis for MapReduce Vs Spark

Tables of the iterations Vs Running time for each query (for each scenario/reason) are shown below.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  | **Carrier Delay** | |  |
|  | **HiveQL** | **Spark-SQL** |  |
|  | 37.632 | 13.26 |  |
|  | 17.239 | 14.64 |  |
|  | 3.624 | 1.02 |  |
|  | 19.707 | 13.4 |  |
|  | 13.673 | 0.93 |  |
|  | 14.481 | 0.83 |  |
|  |  |  |  |
|  |  |  |  |
| |  | | --- | |  | |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | **NAS Delay** | |  |  |  |
|  | **HiveQL** | **Spark-SQL** |  |  |  |
|  | 18.879 | 12.99 |  |  |  |
|  | 13.458 | 1.09 |  |  |  |
|  | 18.663 | 1.21 |  |  |  |
|  | 14.421 | 0.66 |  |  |  |
|  | 13.074 | 0.73 |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| |  | | --- | |  | |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  | **Weather Delay** | |  |  |  |
|  |  | **HiveQL** | **Spark-SQL** |  |  |  |
|  |  | 18.55 | 12.56 |  |  |  |
|  |  | 12.445 | 0.98 |  |  |  |
|  |  | 13.887 | 0.61 |  |  |  |
|  |  | 13.523 | 0.61 |  |  |  |
|  |  | 13.353 | 0.7 |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| |  | | --- | |  | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  | **Late Aircraft Delay** | |  |  |  |
|  |  |  |  | **HiveQL** | **Spark-SQL** |  |  |  |
|  |  |  |  | 18.17 | 13.11 |  |  |  |
|  |  |  |  | 18.22 | 0.71 |  |  |  |
|  |  |  |  | 13.05 | 0.63 |  |  |  |
|  |  |  |  | 13.795 | 0.85 |  |  |  |
|  |  |  |  | 3.099 | 0.84 |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| |  | | --- | |  | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  | **Security Delay** | |  |  |  |  |
|  |  |  | **HiveQL** | **Spark-SQL** |  |  |  |  |
|  |  |  | 13.738 | 13.46 |  |  |  |  |
|  |  |  | 12.856 | 1.01 |  |  |  |  |
|  |  |  | 12.272 | 0.77 |  |  |  |  |
|  |  |  | 12.526 | 0.71 |  |  |  |  |
|  |  |  | 13.648 | 0.64 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| |  | | --- | |  | |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

The final results of the whole assignment is shown as a graph below. Here the average time of 5 tests which used same query is considered as the execution time for each scenario.

|  |  |  |
| --- | --- | --- |
| **Average Time Taken By Query (s)** | **HiveQL** | **Spark-SQL** |
| Carrier Delay | 17.726 | 7.346 |
| NAS Delay | 15.699 | 3.336 |
| Weather Delay | 14.3516 | 3.092 |
| Late Aircraft Delay | 15.031 | 3.228 |
| Security Delay | 13.008 | 3.318 |