Readings:

|  |  |  |
| --- | --- | --- |
| **Metric** | **Spark Core** | **GraphX** |
| Tasks | 1251 | 2351 |
| Shuffle Read | 15.8 GB | 16.5 GB |
| Shuffle Write | 11.8 GB | 8.4 GB |
| Total IO Read | 2.0 MB | 2.1 MB |
| Total IO Write | 12.38 MB | 7.7 MB |
| Total Network Receive | 22.23 GB | 17.47 GB |
| Total Network Send | 11.32 MB | 7.7 MB |
| Time | 438 secs | 246 secs |

PageRank on GraphX performs 1.78X better than Spark core.

Reasons:

GraphX spawns more tasks leading to better parallelism and hence, faster completion time.

Also, GraphX’s default partitioning is better than Spark Core and hence, Shuffle Write, IO Writes and Network IO are better in GraphX than Spark Core.

Resources allocated to Spark Core and GraphX:

|  |  |
| --- | --- |
| Driver memory | 4GB |
| Executor memory | 4GB |
| Executor cores | 4 |
| Task cpus | 1 |