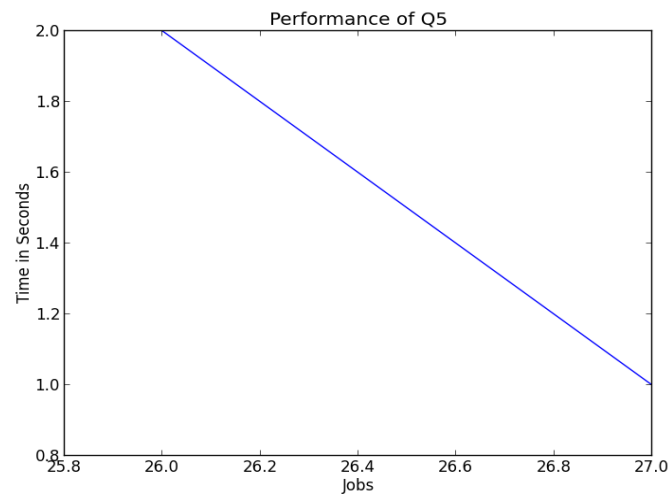
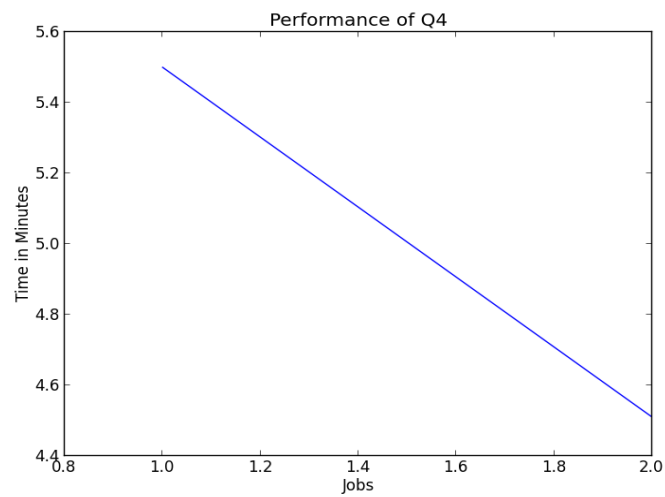


ASSIGNMENT -2: HADOOP MAPREDUCE

RAJIV MADESWARAN

1. Performance Graph:



Performance Parameters Used:

- `mapred.reduce.tasks = 2`
- `mapred.compress.map.output = true`.

2. Screenshot of the Console Output:

```

17/03/07 04:40:17 INFO streaming.StreamJob: map 81% reduce 0%
17/03/07 04:40:20 INFO streaming.StreamJob: map 82% reduce 0%
17/03/07 04:40:26 INFO streaming.StreamJob: map 83% reduce 0%
17/03/07 04:40:29 INFO streaming.StreamJob: map 84% reduce 0%
17/03/07 04:40:32 INFO streaming.StreamJob: map 85% reduce 0%
17/03/07 04:40:35 INFO streaming.StreamJob: map 86% reduce 0%
17/03/07 04:40:39 INFO streaming.StreamJob: map 87% reduce 0%
17/03/07 04:40:42 INFO streaming.StreamJob: map 88% reduce 0%
17/03/07 04:40:45 INFO streaming.StreamJob: map 89% reduce 0%
17/03/07 04:40:50 INFO streaming.StreamJob: map 90% reduce 0%
17/03/07 04:40:56 INFO streaming.StreamJob: map 91% reduce 0%
17/03/07 04:40:59 INFO streaming.StreamJob: map 92% reduce 0%
17/03/07 04:41:02 INFO streaming.StreamJob: map 93% reduce 0%
17/03/07 04:41:08 INFO streaming.StreamJob: map 94% reduce 0%
17/03/07 04:41:17 INFO streaming.StreamJob: map 95% reduce 0%
17/03/07 04:41:18 INFO streaming.StreamJob: map 95% reduce 8%
17/03/07 04:41:24 INFO streaming.StreamJob: map 96% reduce 8%
17/03/07 04:41:27 INFO streaming.StreamJob: map 96% reduce 17%
17/03/07 04:41:30 INFO streaming.StreamJob: map 97% reduce 17%
17/03/07 04:41:42 INFO streaming.StreamJob: map 98% reduce 17%
17/03/07 04:41:51 INFO streaming.StreamJob: map 99% reduce 17%
17/03/07 04:42:00 INFO streaming.StreamJob: map 100% reduce 17%
17/03/07 04:42:03 INFO streaming.StreamJob: map 100% reduce 21%
17/03/07 04:42:07 INFO streaming.StreamJob: map 100% reduce 25%
17/03/07 04:42:13 INFO streaming.StreamJob: map 100% reduce 29%
17/03/07 04:42:16 INFO streaming.StreamJob: map 100% reduce 50%
17/03/07 04:42:19 INFO streaming.StreamJob: map 100% reduce 84%
17/03/07 04:42:22 INFO streaming.StreamJob: map 100% reduce 90%
17/03/07 04:42:25 INFO streaming.StreamJob: map 100% reduce 95%
17/03/07 04:42:28 INFO streaming.StreamJob: map 100% reduce 98%
17/03/07 04:42:31 INFO streaming.StreamJob: map 100% reduce 100%
17/03/07 04:42:35 INFO streaming.StreamJob: Job complete: job_201703070437_0001
17/03/07 04:42:35 INFO streaming.StreamJob: Output: /hw2-output

```

3. Screenshot of the Job Tracker:

The screenshot shows the Hadoop Job Tracker web interface in a browser. The address bar shows the URL `129.115.26.130:50030/jobtracker.jsp`. The interface includes a search bar, a filter input field, and a table of jobs. The 'Completed Jobs' section shows a single job with the ID `job_201703070437_0001`, which is in a 'Completed' state. The job details show it was started on Tue Mar 07 at 04:37:43 UTC 2017, with a priority of 'NORMAL' and user 'hduser'. The job name is 'streamjob7816347577910532155.jar'. The progress bar for the job is at 100.00%. The 'Retired Jobs' section is empty.

Jobid	Started	Priority	User	Name	Map % Complete	Map Total	Maps Completed	Reduce % Complete	Reduce Total	Reduces Completed	Job Sched Inform
job_201703070437_0001	Tue Mar 07 04:37:43 UTC 2017	NORMAL	hduser	streamjob7816347577910532155.jar	100.00%	4	4	100.00%	2	2	NA