


Assignment 1

For all source and output files, PFA in the folder.

Q1

```
csgrads1.utdallas.edu - PuTTY
{csgrads1:~} hadoop jar MapReduce.jar MapReduce1 /Vedant/input/soc-LiveJournal1Adj.txt /Vedant/output1
```



Logged in as: dr.who

All Applications

Cluster

About Nodes Applications

NEW NEW SAVING SUBMITTED ACCEPTED RUNNING FINISHED FAILED KILLED

Scheduler

Tools

Cluster Metrics


Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
243	0	0	243	0	0 B	64 GB	0 B	8	0	0	0	0

Show 20 entries

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI
application_1579313296114_0263	vxk180003	MapReduce1	MAPREDUCE	default	Sat, 15 Feb 2020 20:27:55 GMT	Sat, 15 Feb 2020 20:28:17 GMT	FINISHED	SUCCEEDED		History

Q2

```
csgrads1.utdallas.edu - PuTTY
{csgrads1:~} hadoop jar MapReduce.jar MapReduce2 /Vedant/input/soc-LiveJournal1Adj.txt /Vedant/output2
20/02/15 14:32:09 INFO client.RMProxy: Connecting to ResourceManager at cshadoop1.utdallas.edu/10.176.92.71:8032
20/02/15 14:32:09 INFO input.FileInputFormat: Total input paths to process : 1
20/02/15 14:32:09 INFO mapreduce.JobSubmitter: number of splits:1
20/02/15 14:32:09 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1579313296114_0264
20/02/15 14:32:10 INFO impl.YarnClientImpl: Submitted application application_1579313296114_0264
20/02/15 14:32:10 INFO mapreduce.Job: The url to track the job: http://cshadoop1.utdallas.edu:8088/proxy/application_1579313296114_0264/
```



Logged in as: dr.who

All Applications

Cluster

About Nodes Applications

NEW NEW SAVING SUBMITTED ACCEPTED RUNNING FINISHED FAILED KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
245	0	0	245	0	0 B	64 GB	0 B	8	0	0	0	0

Show 20 entries

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI
application_1579313296114_0265	vxk180003	MapReduce2, Phase2	MAPREDUCE	default	Sat, 15 Feb 2020 20:32:14 GMT	Sat, 15 Feb 2020 20:32:29 GMT	FINISHED	SUCCEEDED		History
application_1579313296114_0264	vxk180003	MapReduce2, Phase1	MAPREDUCE	default	Sat, 15 Feb 2020 20:31:49 GMT	Sat, 15 Feb 2020 20:32:11 GMT	FINISHED	SUCCEEDED		History

Q3

```
csgrads1.utdallas.edu - PuTTY
{csgrads1:~} hadoop jar MapReduce.jar MapReduce3 /Vedant/input/soc-LiveJournal1Adj.txt /Vedant/input/userdata.txt /Vedant/output3
20/02/15 14:35:34 INFO client.RMProxy: Connecting to ResourceManager at cshadoop1.utdallas.edu/10.176.92.71:8032
20/02/15 14:35:34 INFO input.FileInputFormat: Total input paths to process : 1
20/02/15 14:35:34 INFO mapreduce.JobSubmitter: number of splits:1
20/02/15 14:35:35 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1579313296114_0266
20/02/15 14:35:35 INFO impl.YarnClientImpl: Submitted application application_1579313296114_0266
20/02/15 14:35:35 INFO mapreduce.Job: The url to track the job: http://cshadoop1.utdallas.edu:8088/proxy/application_1579313296114_0266/
20/02/15 14:35:35 INFO mapreduce.Job: Running job: job_1579313296114_0266
```

Cluster	Cluster Metrics												
<div> <div>About</div> <div>Nodes</div> <div>Applications</div> <div>NEW</div> <div>SAVING</div> <div>SUBMITTED</div> <div>ACCEPTED</div> <div>RUNNING</div> <div>FINISHED</div> <div>FAILED</div> <div>KILLED</div> </div>	Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
	246	0	0	246	0	0 B	64 GB	0 B	8	0	0	0	0
	<div> <div>Show 20 ▾ entries</div> <div>Search: <input type="text"/></div> </div>												
	ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI		
	application_1579313296114_0266	vrx180003	MapReduce3	MAPREDUCE	default	Sat, 15 Feb 2020 20:35:14 GMT	Sat, 15 Feb 2020 20:35:44 GMT	FINISHED	SUCCEEDED	<div></div>	History		

Q4

```
csgrads1.utdallas.edu - PuTTY
(csggrads1:~) hadoop jar MapReduce.jar MapReduce4 /Vedant/input/soc-LiveJournal1Adj.txt /Vedant/input/userdata.txt /Vedant/output4
20/02/15 14:38:29 INFO client.RMProxy: Connecting to ResourceManager at cshadoop1.utdallas.edu/10.176.92.71:8032
20/02/15 14:38:29 INFO input.FileInputFormat: Total input paths to process : 1
20/02/15 14:38:29 INFO input.FileInputFormat: Total input paths to process : 1
20/02/15 14:38:29 INFO mapreduce.JobSubmitter: number of splits:2
20/02/15 14:38:29 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1579313296114_0267
20/02/15 14:38:30 INFO impl.YarnClientImpl: Submitted application application_1579313296114_0267
20/02/15 14:38:30 INFO mapreduce.Job: The url to track the job: http://cshadoop1.utdallas.edu:8088/proxy/application_1579313296114_0267/
20/02/15 14:38:30 INFO mapreduce.Job: Running job: job_1579313296114_0267
```

Cluster

About Nodes

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
250	0	1	249	1	2 GB	64 GB	0 B	8	0	0	0	0
Show 20 entries <div>Search:</div>												
ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI		
application_1579313296114_0270	vxk180003	MapReduce4, Phase4	MAPREDUCE	default	Sat, 15 Feb 2020 20:39:09 GMT	Sat, 15 Feb 2020 20:39:23 GMT	FINISHED	SUCCEEDED	<div></div>	History		
application_1579313296114_0269	vxk180003	MapReduce4, Phase3	MAPREDUCE	default	Sat, 15 Feb 2020 20:38:51 GMT	Sat, 15 Feb 2020 20:39:08 GMT	FINISHED	SUCCEEDED	<div></div>	History		
application_1579313296114_0268	vxk180003	MapReduce4, Phase2	MAPREDUCE	default	Sat, 15 Feb 2020 20:38:30 GMT	Sat, 15 Feb 2020 20:38:50 GMT	FINISHED	SUCCEEDED	<div></div>	History		
application_1579313296114_0267	vxk180003	MapReduce4, Phase1	MAPREDUCE	default	Sat, 15 Feb 2020 20:38:09 GMT	Sat, 15 Feb 2020 20:38:29 GMT	FINISHED	SUCCEEDED	<div></div>	History		