Performance Analysis of SSL: (OSS Yarn)

A Hadoop cluster has been setup with 4 worker nodes, with each node having the below configuration.

Specification of each node:
- OS: buntu14 (x86_64)
- Cores (CPU): 4 (4)
- Disk: 1181.01GB
- Memory: 13.69GB

Teragen job of 100GB was run, followed by a consecutive execution of Terasort job. The following charts show the behavior of both these jobs with SSL on and off. From the chart, we can see that both Teragen and Terasort when run with increasing mapper and reducer configuration, the job completion time increased. At parallelism 4, all the machines were running at their maximum core capacity. Also, we can observe the job completion times for the runs with SSL on are 11-14% higher than usual. This is an expected behavior because all the communication through Hadoop IPC gets encrypted when SSL is on which adds onto to the additional latency in job completion times.