* We will install Apache Cassandra DB on CentOS machine, then to benchmark this NoSQL database we install YCSB client on it or remotely. Then, we will test Cassandra machine with different workloads we have in YCSB. The workloads we use in YCSB to test database are A,B & C where each have different read, write and update ratios.

Workload A – 50% reads, 50% writes.

Workload B – 95% reads, 5% writes.

Workload C – 100% reads.

Workload D – Reads latest workload or new records inserted.

Workload E – short ranges of records are queried, instead of individual records.

Workload F – Read-modify-write.

* While testing we use different parameters such as operation type(load/run), workload type(a,b,c), record count(no. of records to load), operation count(no. of operations to run), thread count(no. of parallel threads), hosts IP.

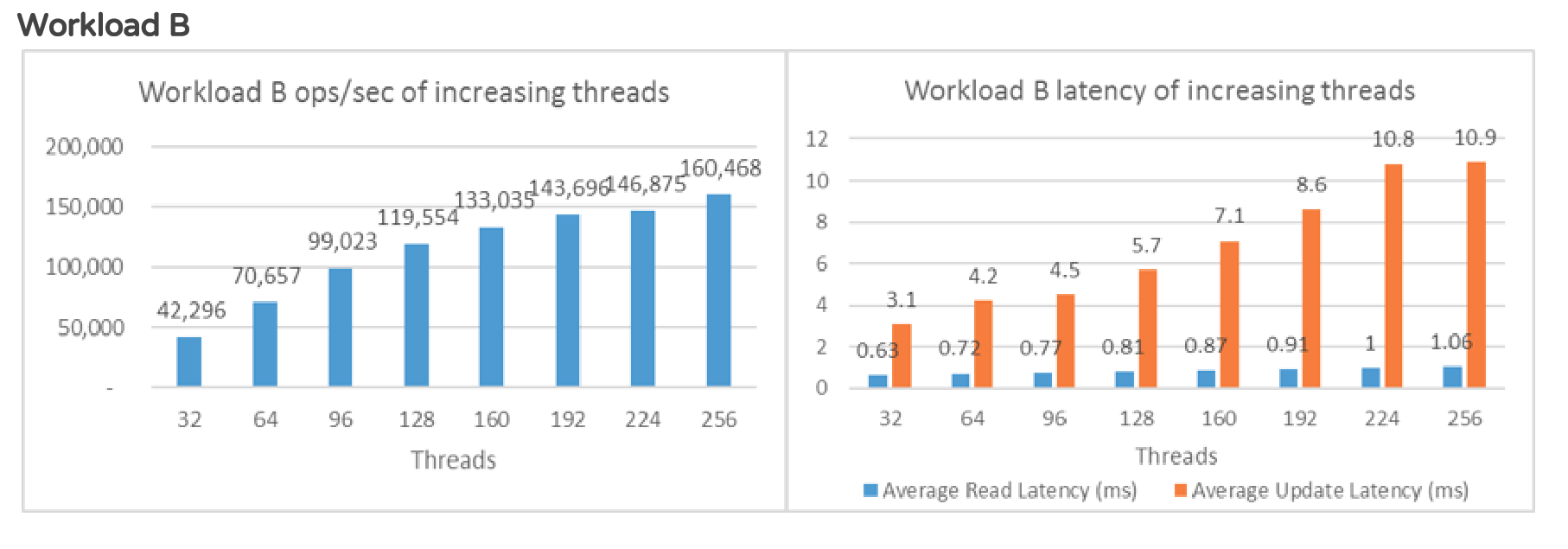
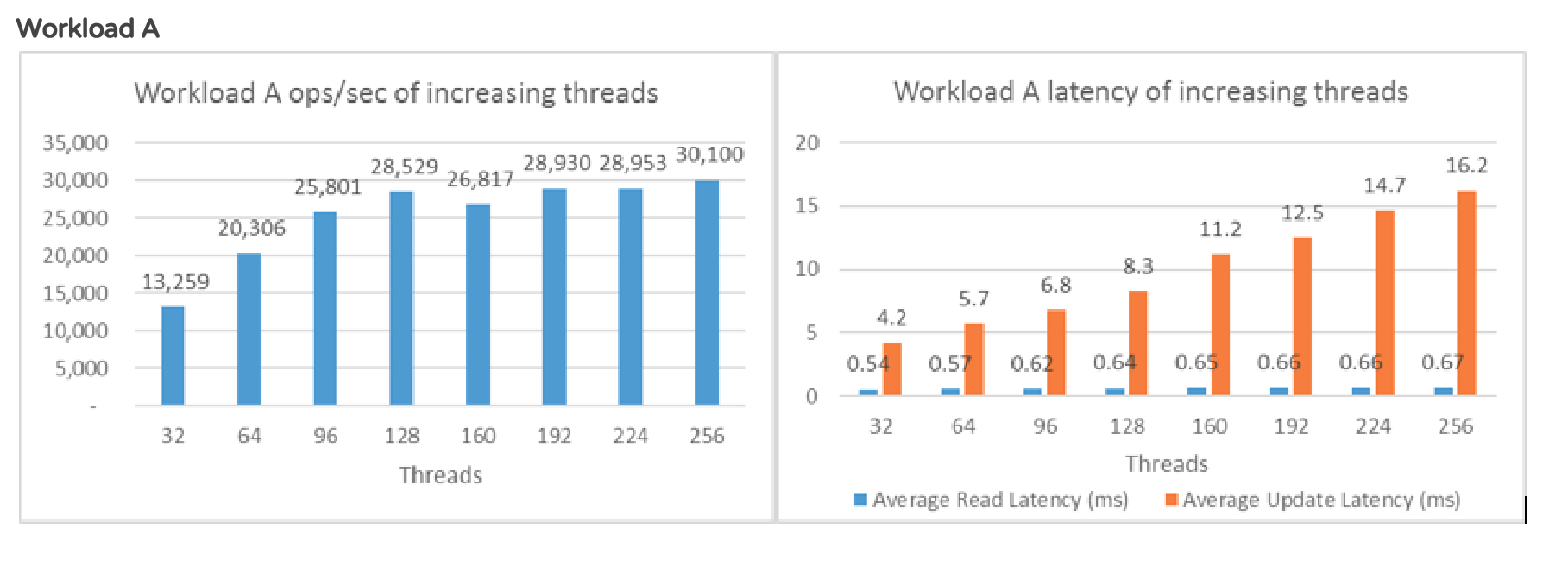
Test command will be similar to this.

*./bin/ycsb load cassandra-10 -s -P workloads/workloada -p recordcount=1000000 -threads 16 > outputLoadA.txt*

*./bin/ycsb run cassandra-10 -s -P workloads/workloada -p operationcount=1000000 -threads 2 > outputRunA.txt*

* From the above testing results, we gather details such as threads count, operations count, throughput operation/second, read & update latency, test load type to plot graphs to find point at which the machine is working with high performance with great accuracy and least data loss.

The results will look similar as below.

Changing Threads count:Changing Operations count: