

## HBASE Performance test by YCSB

---

### Bulid YCSB From github

We use thumbtack modify version <https://github.com/thumbtack-technology/ycsb>

Modify pom.xml – using CDH platform version :

```
<?xml version="1.0" encoding="UTF-8"?>
<project xmlns="http://maven.apache.org/POM/4.0.0" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/maven-v4_0_0.xsd">
  <modelVersion>4.0.0</modelVersion>
  <parent>
    <groupId>com.yahoo.ycsb</groupId>
    <artifactId>root</artifactId>
    <version>0.1.4</version>
  </parent>

  <artifactId>hbase-binding</artifactId>
```

```
<name>HBase DB Binding</name>
```

```
<repositories>
```

```
<repository>
```

```
<id>central</id>
```

```
<name>Maven Repository Switchboard</name>
```

```
<url>http://repo1.maven.org/maven2</url>
```

```
</repository>
```

```
<repository>
```

```
<id>maven-hadoop</id>
```

```
<name>Hadoop Releases</name>
```

```
<url>https://repository.cloudera.com/content/repositories/releases/</url>
```

```
</repository>
```

```
<repository>
```

```
<id>cloudera-repos</id>
```

```
<name>Cloudera Repos</name>
```

```
<url>https://repository.cloudera.com/artifactory/cloudera-repos/</url>
```

```
</repository>
```

```
</repositories>
```

```
<dependencies>
```

```
<dependency>
```

```
<groupId>org.apache.hbase</groupId>
```

```
<artifactId>hbase-client</artifactId>
```

```
<version>0.96.1.1-cdh5.0.0</version>
```

```
<!-- <version>0.96.1.1-hadoop2</version> -->
```

```
</dependency>
<dependency>
  <groupId>org.apache.hadoop</groupId>
  <artifactId>hadoop-hdfs</artifactId>
  <version>2.3.0-cdh5.0.0</version>
</dependency>
<dependency>
  <groupId>com.yahoo.ycsb</groupId>
  <artifactId>core</artifactId>
  <version>${project.version}</version>
</dependency>
</dependencies>

<build>
  <plugins>
    <plugin>
      <groupId>org.apache.maven.plugins</groupId>
      <artifactId>maven-assembly-plugin</artifactId>
      <version>${maven.assembly.version}</version>
      <configuration>
        <descriptorRefs>
          <descriptorRef>jar-with-dependencies</descriptorRef>
        </descriptorRefs>
        <appendAssemblyId>>false</appendAssemblyId>
      </configuration>
    </plugin>
  </plugins>
</build>
```

```
<execution>
  <phase>package</phase>
  <goals>
    <goal>single</goal>
  </goals>
</execution>
</executions>
</plugin>
</plugins>
</build>

</project>
```

Run mvn clean package – build HBase DB Binding successfully

Copy core-site.xml hbase-site.xml hdfs-site.xml to /home/hadoop/ycsb/hbase/src/main/conf

Load 100 million data to usertable:

*YCSB Client 0.1*

*Command line: -db com.yahoo.ycsb.db.HBaseClient -P workloads/workloada -p columnfamily=family -p recordcount=100000000 -s -threads 10 -p measurementtype=timeseries -p timeseries.granularity=2000 -load*

Run test scripts:

*\$ls -l |grep .sh*

```
-rwxrwxr-x 1 hadoop hadoop 3535 Sep  3 09:58 test_workloada.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:09 test_workloadb.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:10 test_workloadc.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:10 test_workloadd.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:11 test_workloade.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:11 test_workloadf.sh
-rwxr-xr-x 1 root  root  3535 Sep  3 10:11 test_workloadg.sh
```

```
[hadoop@a1-ops-hdclient /home/hadoop/ygsb]
```

```
$cat test_workloada.sh |more
```

```
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 1 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t1
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 2 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t2
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 4 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t4
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 8 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t8
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 16 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t16
/home/hadoop/ygsb/bin/ygsb run hbase -P workloads/workloada -p columnfamily=family -s -threads 32 -p
measurementtype=timeseries -p timeseries.granularity=2000|tee workloada_t32
```

```
.....
```

**Load 100 million data:**

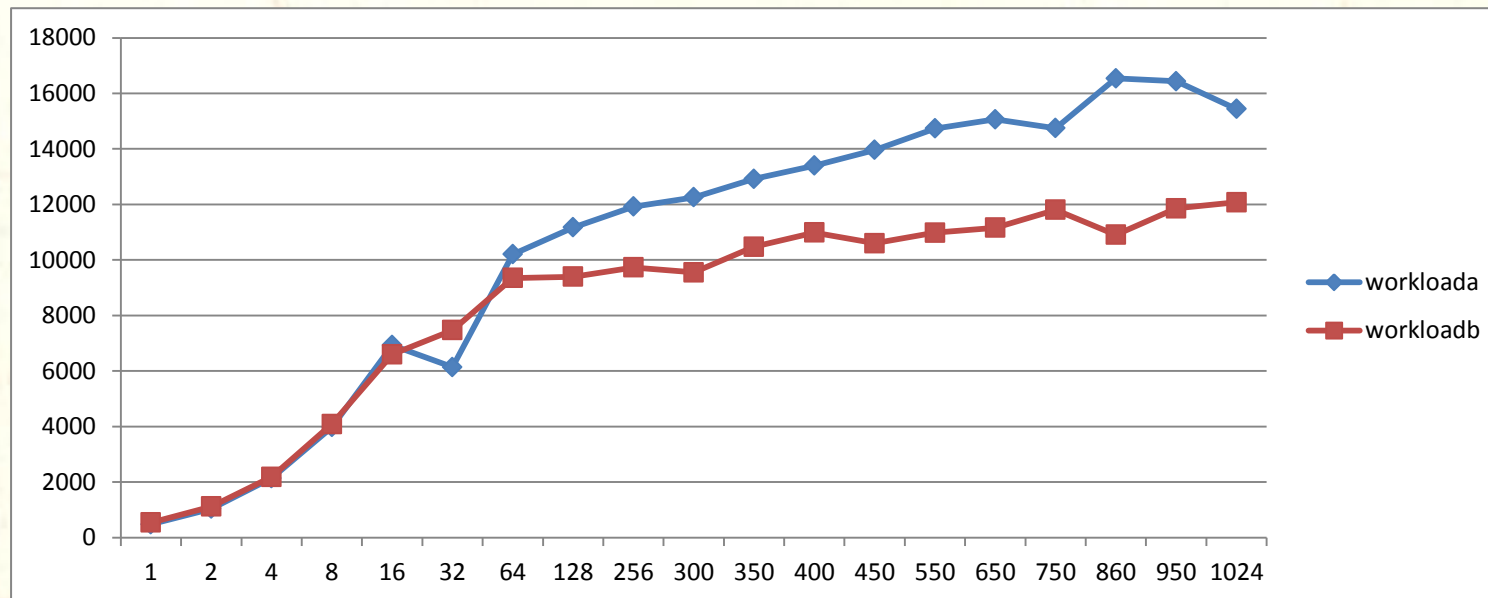
***[OVERALL], Reconnections, 0.0***

***[OVERALL], RunTime(ms), 3197006.0***

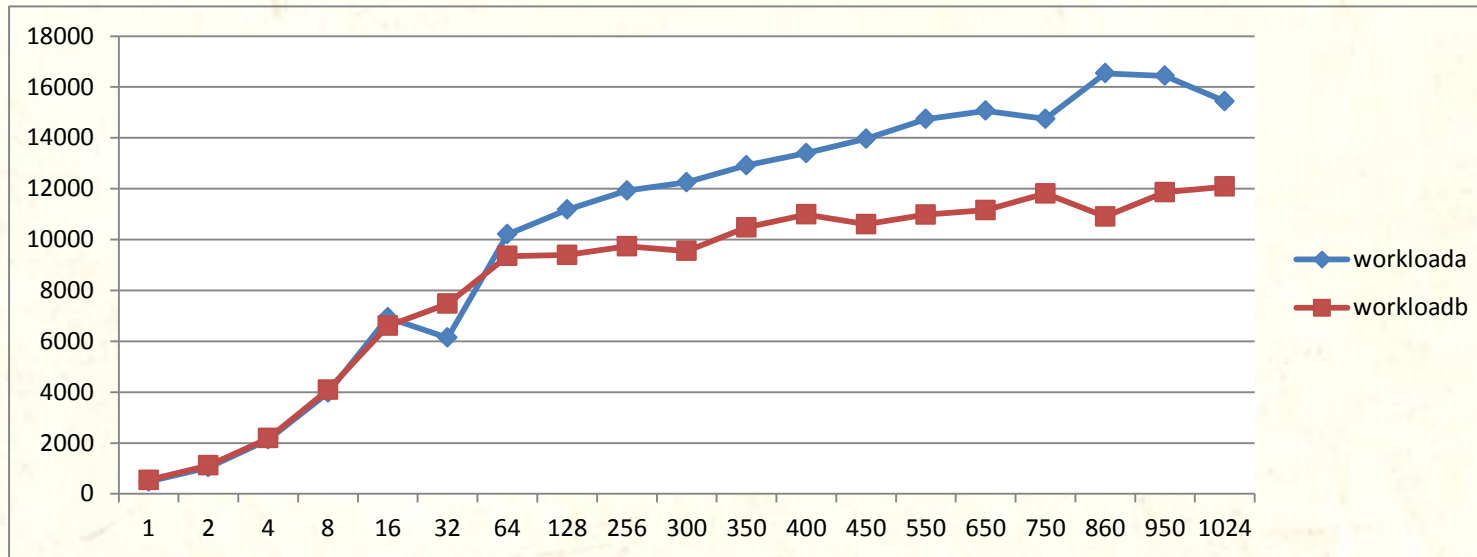
***[OVERALL], Operations, 1.0E8***

***[OVERALL], Throughput(ops/sec), 31279.265662935883***

### Workload A AverageLatency(us) : read 50% update 50%

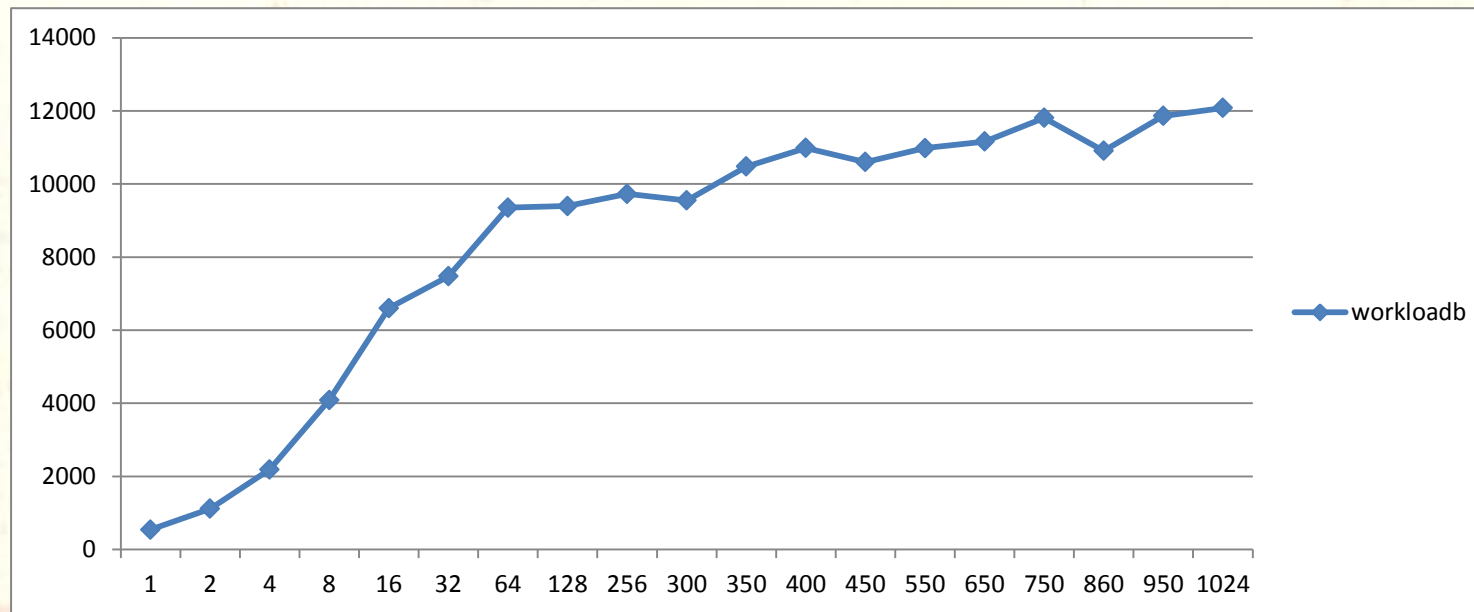


### Workload B AverageLatency(us): read 95% update 5%

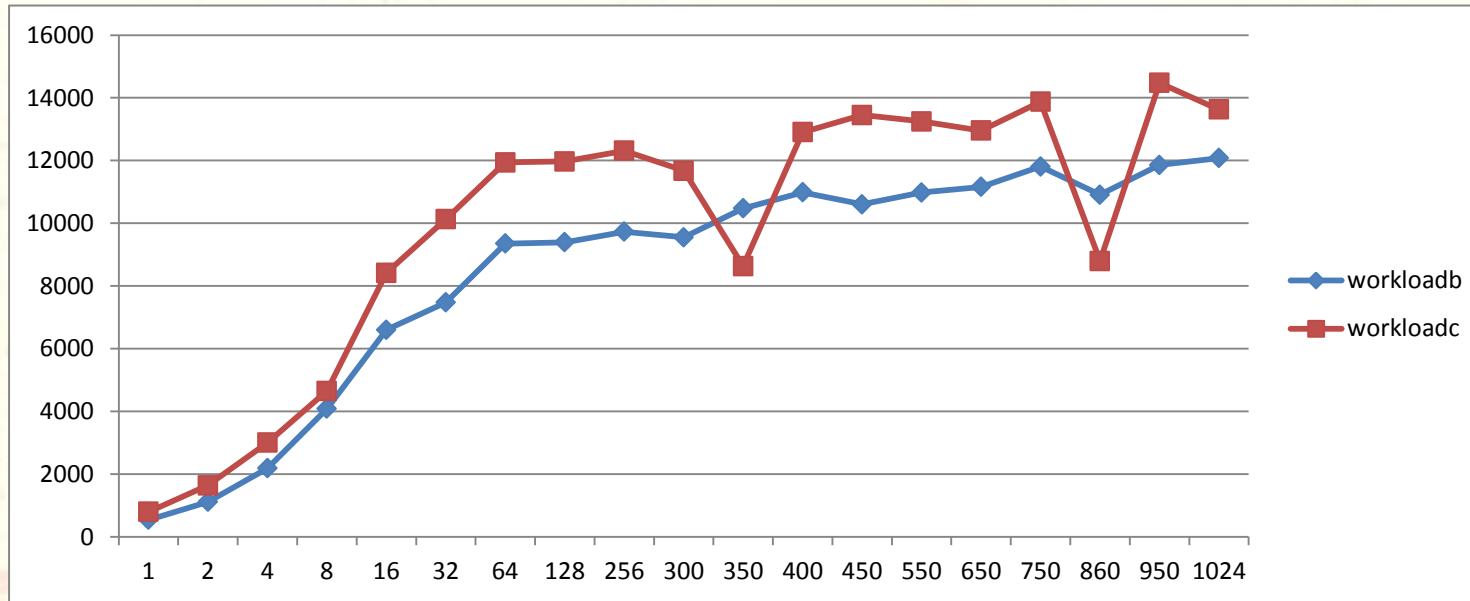




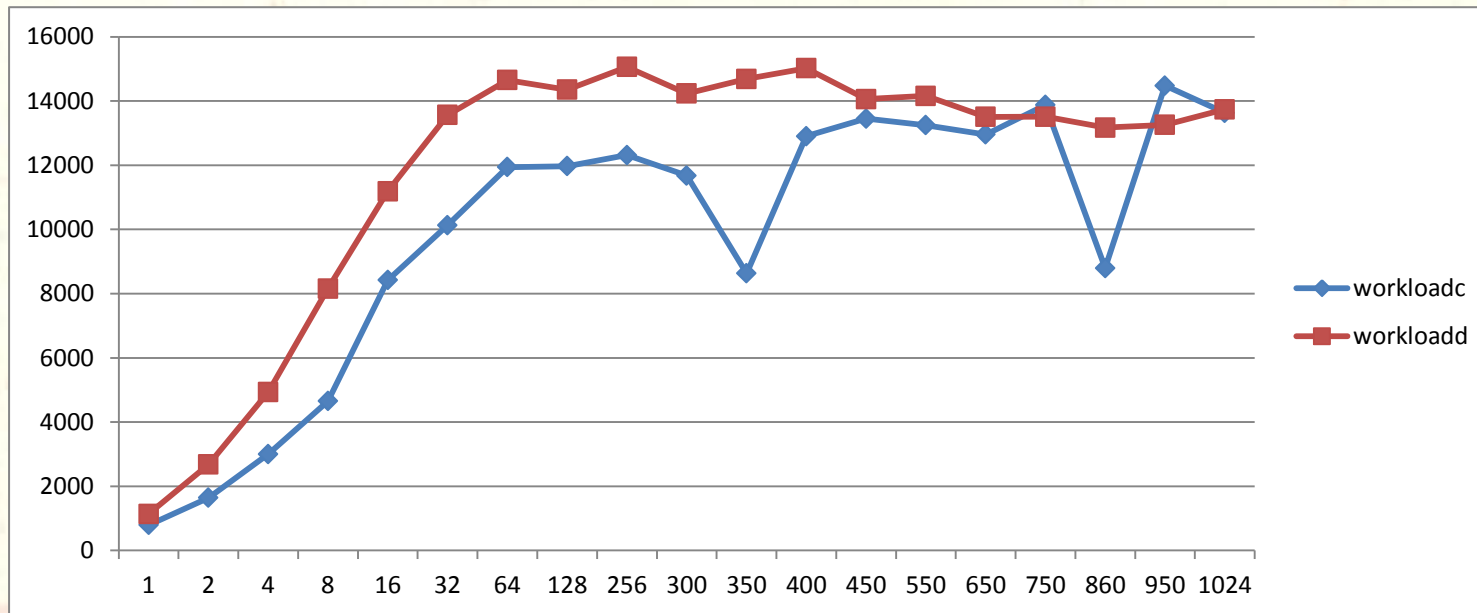
### Workload C AverageLatency(us): read 100%



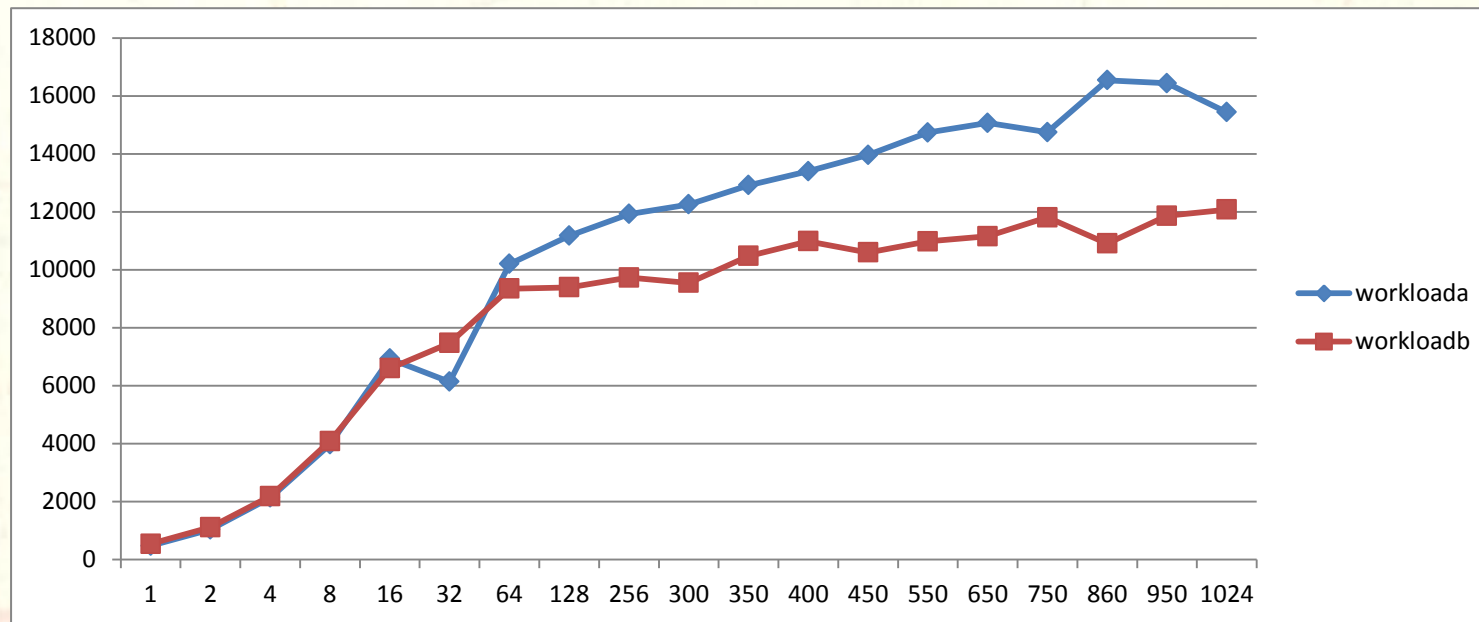
**Workload D AverageLatency(us): read 95% insert 5%**



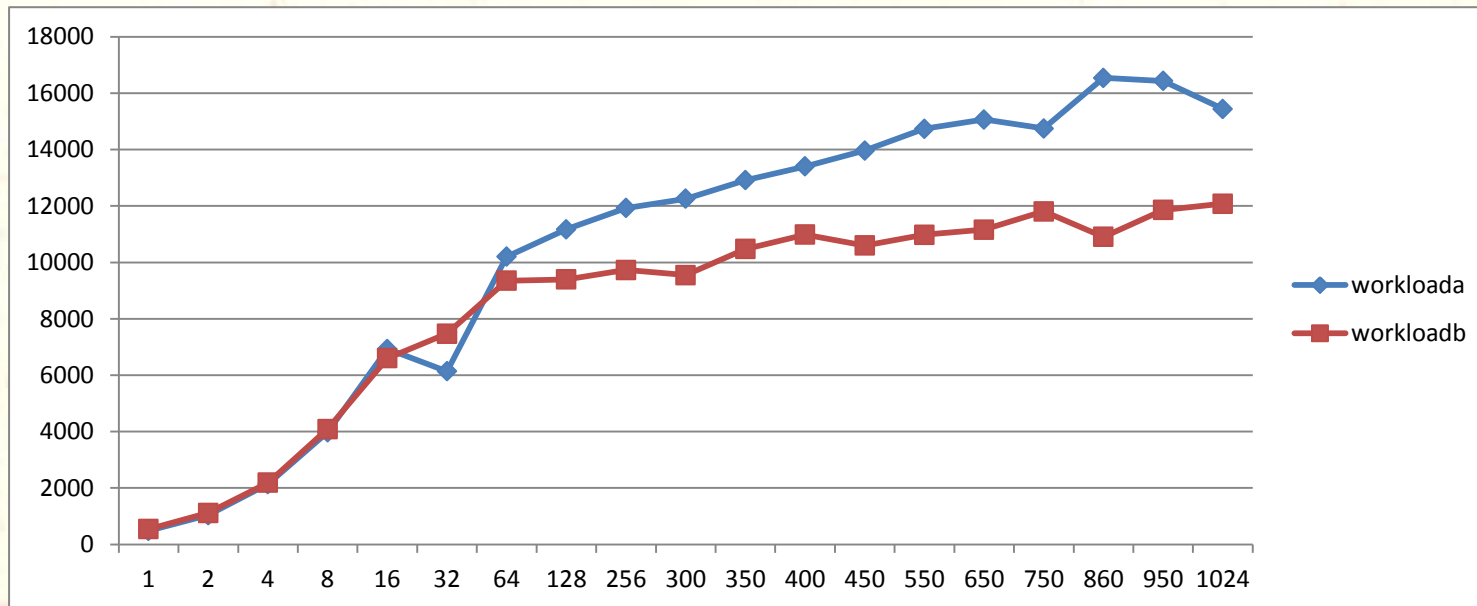
**Workload E AverageLatency(us): scanread 95 insert 5%**



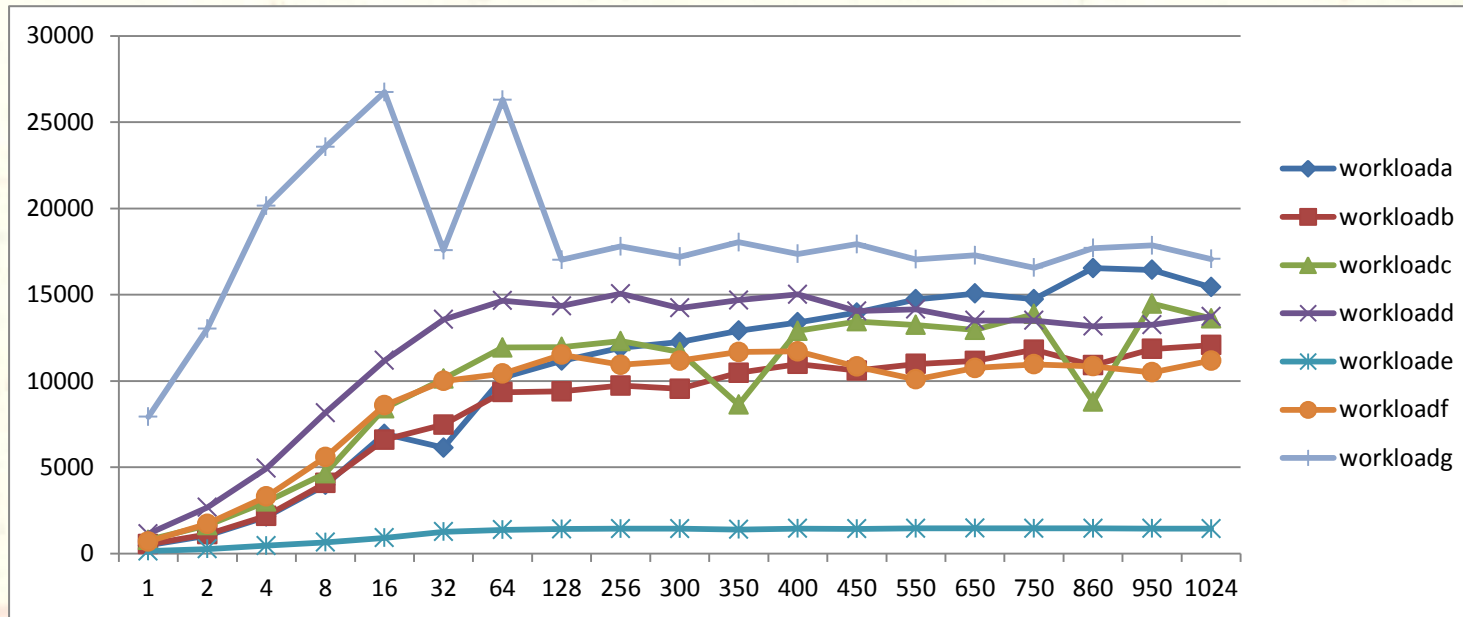
**Workload F AverageLatency(us): read 50% readmodify 50%**



**Workload G AverageLatency(us): read 5% update 95%**



### Workload all avg ops/s



## **HBASE ENV:**

```
#hbase version
```

```
2014-09-03 22:46:34,973 INFO [main] util.VersionInfo: HBase 0.96.1.1-cdh5.0.0
```

```
2014-09-03 22:46:34,973 INFO [main] util.VersionInfo: Subversion file:
```

```
///var/lib/jenkins/workspace/generic-binary-tarball-and-maven-deploy/CDH5.0.0-Packaging-HBase-2014-03-27_22-53-27/hbase-0.9  
6.1.1-cdh5.0.0 -r Unknown
```

```
2014-09-03 22:46:34,973 INFO [main] util.VersionInfo: Compiled by jenkins on Thu Mar 27 23:02:13 PDT 2014
```

## **Datanode OS system information :**

```
#cat /etc/issue
```

```
CentOS release 6.5 (Final)
```

```
#cat /proc/cpuinfo |grep processor |wc -l
```

```
24
```

```
#cat /proc/meminfo |more
```

```
MemTotal: 99035868 kB
```

*Mount system info:*

*/dev/sdb1 on /data1 type ext4 (rw)*  
*/dev/sdc1 on /data2 type ext4 (rw)*  
*/dev/sdd1 on /data3 type ext4 (rw)*  
*/dev/sde1 on /data4 type ext4 (rw)*  
*/dev/sdf1 on /data5 type ext4 (rw)*  
*/dev/sdg1 on /data6 type ext4 (rw)*  
*/dev/sdh1 on /data7 type ext4 (rw)*  
*/dev/sdi1 on /data8 type ext4 (rw)*  
*/dev/sdj1 on /data9 type ext4 (rw)*  
*/dev/sdk1 on /data10 type ext4 (rw)*  
*/dev/sdl1 on /data11 type ext4 (rw)*  
*/dev/sdm1 on /data12 type ext4 (rw)*

*#sg\_vpd /dev/sde --page=0xb1*

*Block device characteristics VPD page (SBC):*

*Nominal rotation rate: 7200 rpm*  
*: 3.5 inch*

**== START OF INFORMATION SECTION ==**

*Device Model: ST2000NM0033-9ZM175*  
*Serial Number: Z1X1YV4Q*  
*LU WWN Device Id: 5 000c50 06672b0c2*  
*Firmware Version: SN03*  
*User Capacity: 2,000,398,934,016 bytes [2.00 TB]*  
*Sector Size: 512 bytes logical/physical*



*Device is:* Not in smartctl database [for details use: -P showall]

*ATA Version is:* 8

*ATA Standard is:* ACS-2 (revision not indicated)

*Local Time is:* Wed Sep 3 15:53:55 2014 CST

*SMART support is:* Available - device has SMART capability.

*SMART support is:* Enabled