

Detailed instructions for setting up Cassandra, DynamoDB, and MongoDB can be found in “Cassandra Set up Documentation”, “Dynamo Instructions”, and “MongoDB set-up info”, respectively.

YCSB was taken from <https://github.com/brianfrankcooper/YCSB> , but instructions for installing it can be found in each respective documentation file.

Final graphs of the results of loads and runs can be found in “Result Graphs”

```
''' Our github Repo with the same material + log files of our runs of YCSB
    Documentation for each database can be found in each respective folder
    Result graphs and graph plotting sheet provided
'''
```

Data Preparation and Setup

1. Which database system(s) and version(s) are you using? How do we install it/them? (providing a link to official documentation will be enough)

The databases we used were:

DynamoDB

Apache Cassandra (version debian 41x):

https://cassandra.apache.org/doc/stable/cassandra/getting_started/installing.html#installing-the-debian-packages (follow instructions for Debian installation)

MongoDB (version 7.0)

<https://www.mongodb.com/docs/manual/installation/>

2. How do we download the data you used for your project? Please do NOT submit ALL the data with your code (Submitting a very small portion (< 5 MB) so that we can run the demo might be okay)

All data is dynamically generated by YCSB so no dataset was used.

Follow the instructions for setting up YCSB in each environment.

3. How do we load this data into the database system?

Data is loaded into the system through YCSB commands listed in the instructions for each individual database.

1. Do you have some scripts to do that? If yes, how do we execute them?

The scripts we used consisted of just the load and run commands indicated in the individual database instructions.

2. Did you use some tools for loading? If yes, what are they? Provide appropriate details and links.

No other tools were used for loading.

4. If you are benchmarking different database systems, did you make changes to the configurations? If yes, what are they?

Yes, the original benchmarking scripts assumed that each database was running on the same machine as ycsb. To ensure that all resources in a given benchmark were dedicated wholly to the database, ycsb was run on a separate node. Necessary changes are documented in each set-up file.

Cassandra and MongoDB each have their own configurations to set their configurations

5. If you are generating your own data, how do we generate it?

The data is generated during the load phase of the benchmark. Instructions for specifically generating the data for each database can be found in the individual instruction files provided.

Application and code

1. Which programming language(s) and version(s) are you using (Python, Java 8, C++, etc.)?

We are using YSCB to generate our data.

2. List the third-party libraries needed to execute your code and how do we install them (For ex. MySQL/neo4j connector for Python)

No other outside libraries are required.

3. If you have a GUI, how do we run it?

No GUI was created

4. Anything else we need to know about running your application?

There is a script in the Cassandra folder to install Cassandra that we wrote to streamline the process.

Code Documentation and References

1. Did you use some code from GitHub or other sources? If yes provide the link.

Yes, we used Yahoo Cloud Serving Benchmark found here
<https://github.com/brianfrankcooper/YCSB>

2. If you used some online code, what changes did you make to the code?

The focus of our project was setting up each database in a distributed fashion and benchmarking them against each other.

We don't have any raw code for these libraries in our submission.

3. Give a list of files in your submission which are written by you.

We don't have any raw code for these libraries in our submission.

- The logs for our trials can be found here: <https://github.com/ldyim/6400-temp/tree/main>.
- Instructions for setting up for each database (for each Cassandra, MongoDB, DynamoDB)
- Excel sheet with data plotted and graphs generated for each workload
- Document of all of our graphs compiled for each workload
- Script for installing Cassandra
- And changes we made to the benchmark can be found in the individual setup instructions

4. Feel free to include images of your application's working in Readme/Instructions file.

An example output running the benchmark on MongoDB with 6 nodes:

```
[OVERALL], RunTime(ms), 6840
[OVERALL], Throughput(ops/sec), 146.19883040935673
[TOTAL_GCS_G1_Young_Generation], Count, 4
[TOTAL_GC_TIME_G1_Young_Generation], Time(ms), 42
[TOTAL_GC_TIME_%_G1_Young_Generation], Time(%), 0.6140350877192983
[TOTAL_GCS_G1_Old_Generation], Count, 0
[TOTAL_GC_TIME_G1_Old_Generation], Time(ms), 0
[TOTAL_GC_TIME_%_G1_Old_Generation], Time(%), 0.0
[TOTAL_GCs], Count, 4
[TOTAL_GC_TIME], Time(ms), 42
[TOTAL_GC_TIME_%], Time(%), 0.6140350877192983
[CLEANUP], Operations, 1
[CLEANUP], AverageLatency(us), 4194.0
[CLEANUP], MinLatency(us), 4192
[CLEANUP], MaxLatency(us), 4195
[CLEANUP], 95thPercentileLatency(us), 4195
[CLEANUP], 99thPercentileLatency(us), 4195
[INSERT], Operations, 1000
[INSERT], AverageLatency(us), 5659.91
[INSERT], MinLatency(us), 3102
[INSERT], MaxLatency(us), 273407
[INSERT], 95thPercentileLatency(us), 7479
[INSERT], 99thPercentileLatency(us), 9023
[INSERT], Return=OK, 1000
```

VM instances of Cassandra

VM instances

Filter

Enter property name or value

?

III

<input type="checkbox"/> Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect
<input type="checkbox"/>	cassandra1	us-east1-b			10.142.0.2 (nic0)	35.237.31.88 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra2	us-east1-b			10.142.0.4 (nic0)	35.229.89.234 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra3	us-east1-b			10.142.0.5 (nic0)	35.229.44.8 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra4	us-east1-b			10.142.0.6 (nic0)	35.231.44.120 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra5	us-east1-b			10.142.0.7 (nic0)	35.227.10.198 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra6	us-east1-b			10.142.0.8 (nic0)	34.74.25.80 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra7	us-east1-b			10.142.0.9 (nic0)	34.23.209.201 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	ycsb	us-east1-b			10.142.0.3 (nic0)	34.73.219.221 (nic0)	SSH <div>▼</div> <div>⋮</div>

Related actions

^

HIDE

Filter

Enter property name or value

🔍

III

<input type="checkbox"/> Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect
<input type="checkbox"/>	cassandra8	us-east1-b			11.0.0.6 (nic0)		SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra8-1	us-east1-b			11.0.0.5 (nic0)	34.75.154.52 (nic0)	SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra8-2	us-east1-b			11.0.0.8 (nic0)		SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra9	us-east1-b			11.0.0.3 (nic0)		SSH <div>▼</div> <div>⋮</div>
<input type="checkbox"/>	cassandra9-1	us-east1-b			11.0.0.7 (nic0)	34.139.108.161 (nic0)	SSH <div>▼</div> <div>⋮</div>

9 node cluster for cassandra

```
jodi — Jodie@cassandra1: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
Datacenter: datacenter1
=====
Status=Up/Down
|/ State=Normal/Leaving/Joining/Moving
-- Address      Load      Tokens    Owns (effective)  Host ID
-- Address      Load      Tokens    Owns (effective)  Host ID
UN  10.142.0.2   3.48 MiB   16        32.0%             45574ae1-ff30-4f
cb-9334-e1680dc2de30 rack1
UN  10.142.0.5   1.31 MiB   16        33.3%             1719e405-1cf9-4d
b9-b309-9452b7953d0d rack1
UN  10.142.0.6   927.93 KiB 16        34.4%             98cd0262-9343-44
f9-9526-7b964756ed28 rack1
UN  10.142.0.8   815.24 KiB 16        33.4%             82d38d47-eded-4a
aa-8673-59bb087abaf0 rack1
UN  11.0.0.5     562.06 KiB 16        34.2%             3bcd6e9-4140-47
00-92f7-5472b9e1dbad rack1
UN  10.142.0.9   642.91 KiB 16        33.1%             44de5010-0447-4c
a7-9443-facbe996cc52 rack1

jodi — Jodie@cassandra2: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
UN  10.142.0.5   1.31 MiB   16        33.3%             1719e405-1cf9-4db
9-b309-9452b7953d0d rack1
UN  10.142.0.6   927.93 KiB 16        34.4%             98cd0262-9343-44f
9-9526-7b964756ed28 rack1
UN  10.142.0.8   815.24 KiB 16        33.4%             82d38d47-eded-4aa
8-8673-59bb087abaf0 rack1
UN  10.142.0.9   642.91 KiB 16        33.1%             44de5010-0447-4ca
0-92f7-5472b9e1dbad rack1
UN  11.0.0.5     562.06 KiB 16        33.5%             5e8c2ea0-f304-41b
8-815b-a817857f0e65 rack1
UN  10.142.0.4   1.33 MiB   16        33.2%             31ca91bf-1bc4-40f
6-9698-396135b587e4 rack1
UN  10.142.0.7   778.29 KiB 16        32.8%             32474c14-88f7-42e
a-8ea9-32c925ce47f8 rack1

jodi — Jodie@cassandra3: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 19:33:12 2024 from 128.61.30.26
Jodie@cassandra1:~$ sudo systemctl start cassandra
Jodie@cassandra1:~$

jodi — Jodie@cassandra4: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 20:05:29 2024 from 128.61.30.26
Jodie@cassandra1:~$ sudo systemctl start cassandra
Jodie@cassandra1:~$

jodi — Jodie@cassandra5: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 20:05:18 2024 from 128.61.30.26
Jodie@cassandra1:~$ sudo systemctl restart cassandra
Jodie@cassandra5:~$

jodi — Jodie@cassandra6: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 19:33:08 2024 from 128.61.30.26
Jodie@cassandra6:~$ sudo systemctl start cassandra
Jodie@cassandra6:~$ sudo systemctl start cassandra
Jodie@cassandra6:~$ sudo systemctl start cassandra
Jodie@cassandra6:~$

jodi — Jodie@cassandra7: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 19:33:18 2024 from 128.61.30.26
Jodie@cassandra7:~$ sudo systemctl start cassandra
Jodie@cassandra7:~$

jodi — Jodie@cassandra8: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 19:33:25 2024 from 128.61.30.26
Jodie@cassandra8:~$ sudo systemctl start cassandra
Jodie@cassandra8:~$

jodi — Jodie@cassandra9: ~ — ssh -i ~/.ssh/CassandraKey Jodie...
croK8s
just raised the bar for easy, resilient and secure K8s cluster deployment.
https://ubuntu.com/engage/secure-kubernetes-at-the-edge
Expanded Security Maintenance for Applications is not enabled.
10 updates can be applied immediately.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
Last login: Tue Apr 23 19:33:05 2024 from 128.61.30.26
Jodie@cassandra9:~$ sudo systemctl start cassandra
Jodie@cassandra9:~$
```

```
jodi — Jodie@cassandra1: ~ — ssh -i ~/.ssh/CassandraKey Jodie...

Last login: Tue Apr 23 19:33:15 2024 from 128.61.30.26
Jodie@cassandra1:~$ sudo systemctl start cassandra
Jodie@cassandra1:~$ nodetool status
Datacenter: datacenter1
=====
Status=Up/Down
|/ State=Normal/Leaving/Joining/Moving
-- Address      Load      Tokens    Owns (effective)  Host ID
-- Address      Load      Tokens    Owns (effective)  Host ID
UN  10.142.0.2   3.48 MiB   16        32.0%             45574ae1-ff30-4f
cb-9334-e1680dc2de30 rack1
UN  10.142.0.5   1.31 MiB   16        33.3%             1719e405-1cf9-4d
b9-b309-9452b7953d0d rack1
UN  10.142.0.6   927.93 KiB 16        34.4%             98cd0262-9343-44
f9-9526-7b964756ed28 rack1
UN  10.142.0.8   815.24 KiB 16        33.4%             82d38d47-eded-4a
aa-8673-59bb087abaf0 rack1
UN  11.0.0.5     562.06 KiB 16        34.2%             3bcd6e9-4140-47
00-92f7-5472b9e1dbad rack1
UN  10.142.0.9   642.91 KiB 16        33.1%             44de5010-0447-4c
a7-9443-facbe996cc52 rack1
UN  11.0.0.7     654.03 KiB 16        33.5%             5e8c2ea0-f304-41
08-815b-a817857f0e65 rack1
UN  10.142.0.4   1.33 MiB   16        33.2%             31ca91bf-1bc4-40
f6-9698-396135b587e4 rack1
UN  10.142.0.7   778.29 KiB 16        32.8%             32474c14-88f7-42
ea-8ea9-32c925ce47f8 rack1

Jodie@cassandra1:~$
```