

Assignment No. 10

6CS371 : Advanced Database System Lab

Cassandra Clustering

Name : Jay Shirgupe

PRN: 21510026

Batch: T-7

TY CSE

Aim





To create a cluster of cassandra nodes running on windows machines.

Procedure

1. Perform Cassandra Installation on each node

a. Install JDK 8.0 (jdk-8u251-windows-x64.exe) from

<https://www.oracle.com/java/technologies/javase/javase8u211-later-archive-downloads.html>

Solaris x64 (SVR4 package)	133.64 MB	 jdk-8u251-solaris-x64.tar.Z
Solaris x64	91.9 MB	 jdk-8u251-solaris-x64.tar.gz
Windows x86	201.17 MB	 jdk-8u251-windows-i586.exe
Windows x64	211.54 MB	 jdk-8u251-windows-x64.exe

b. Install Python 2.7 from <https://www.python.org/downloads/release/python-2718/>

Python 2.7.18

Release Date: April 20, 2020

Python 2.7.18 is the last release of Python 2.

Files

Version	Operating System	Description	MD5 Sum	File Size	GPG
Gzipped source tarball	Source release		38c84292658ed4456157195f1c9bcbe1	16.7 MB	SIG
XZ compressed source tarball	Source release		fd6cc8ec0a78c44036f825e739f36e5a	12.3 MB	SIG
macOS 64-bit installer	macOS	for OS X 10.9 and later	ce98eeb7bdf806685adc265ec1444463	23.7 MB	SIG
Windows help file	Windows		b3b753dffe1c7930243c1c40ec3a72b1	6.0 MB	SIG
Windows debug information files	Windows		20b111cfe8d06d2fe8c77679a86113d	24.0 MB	SIG
Windows debug information files for 64-bit binaries	Windows		bb0897ea20fda343e5179d413d4a4a7c	24.8 MB	SIG
Windows x86 MSI installer	Windows		db6ad9195b3086c6b4cef9493d738d2	18.7 MB	SIG
Windows x86-64 MSI installer	Windows	for AMD64/EM64T/x64	a425c758d38f8e28b56f4724b499239a	19.6 MB	SIG

- c. Download the 3.11.6 version of cassandra from <https://archive.apache.org/dist/cassandra/3.11.6/>

Older Maintained Releases

Apache Cassandra 3.11

Latest release on 2023-08-20

Maintained until 5.0.0 release (Nov-Dec 2023)

3.11.16

(pgp, sha256, sha512)

(source: pgp, sha256, sha512)

Apache Cassandra 3.0

Latest release on 2023-05-15

Maintained until 5.0.0 release (Nov-Dec 2023)

3.0.29

(pgp, sha256, sha512)

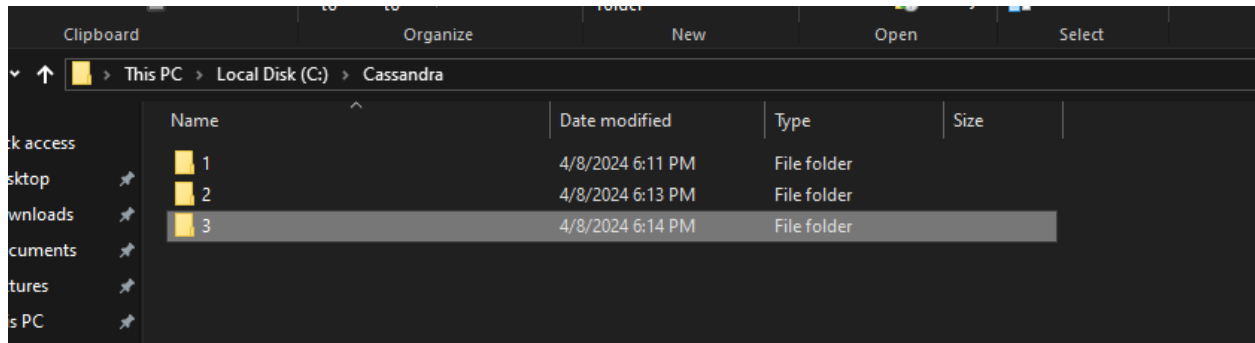
(source: pgp, sha256, sha512)

Unmaintained older versions of Cassandra are archived [here](#).

CVE fixes may be applied to unmaintained versions as decided on a case-by-case basis.

Extract this folder to a certain location. Add the /bin to the PATH variable. Add the location of the folder as CASSANDRA_HOME environment variable.

- Create 3 folders named 1, 2 and 3 in C:\Cassandra\. Copy the extracted cassandra directory into these 3 folders. These will act as nodes.



- Change the listen_address and rpc_address in the conf/cassandra.yaml to 127.0.0.2 for node 1, 127.0.0.2 for node 2 and 127.0.0.3 for node 3. Also edit the following entries.

```

10 # Setting listen_address to 0.0.0.0
11 #
12 listen_address: 127.0.0.2
13
14 # Set listen_address OR listen_interface

```

```

87 #
88 # For security reasons, you should
89 rpc_address: 127.0.0.2
90
91 # Set rpc_address OR rpc_interface
92 # to a single address. IP aliases

```

```

414 # any class that implements the SeedProvider interface and has a
415 # constructor that takes a Map<String, String> of parameters will do.
416 seed_provider:
417     # Addresses of hosts that are deemed contact points.
418     # Cassandra nodes use this list of hosts to find each other and learn
419     # the topology of the ring. You must change this if you are running
420     # multiple nodes!
421     - class_name: org.apache.cassandra.locator.SimpleSeedProvider
422       parameters:
423         # seeds is actually a comma-delimited list of addresses.
424         # Ex: "<ip1>,<ip2>,<ip3>"
425         - seeds: "127.0.0.2"
426
427 # For workloads with more data than can fit in memory, Cassandra's

```

```

# If not set, the default directory is $CASSANDRA_HOME/data
✓ data_file_directories:
| | - /opt/cassandra/3/data

# commit log. when running on magnetic HDD, this should
# separate spindle than the data directories.
# If not set, the default directory is $CASSANDRA_HOME/data/commitlog
94 # separate spindle than the data directories.
95 # If not set, the default directory is $CASSANDRA_HOME/data/commitlog
96 commitlog_directory: /opt/cassandra/1/commitlog
97
98 # Enable / disable CDC functionality on a per-node basis. This
99 # for write path allocation rejection (standard: never reject

# saved caches
# If not set, the default directory is $CASSANDRA_HOME/data/saved_caches
saved_caches_directory: /opt/cassandra/1/saved_caches

# commitlog_sync may be either "periodic" or "batch."
#

```

4. Now run each cassandra node.

```

C:\opt\Cassandra\1\bin>cassandra.bat -f
WARNING! Powershell script execution unavailable.
Please use 'powershell Set-ExecutionPolicy Unrestricted'
on this user-account to run cassandra with fully featured
functionality on this platform.
Starting with legacy startup options
Starting Cassandra Server
INFO [main] 2024-04-08 18:32:42,481 YamlConfigurationLoader.java:89 - Configuration location: file:/C:/opt/Cassandra/1/
conf/cassandra.yaml
INFO [main] 2024-04-08 18:32:42,951 Config.java:516 - Node configuration:[allocate_tokens_for_keyspace=null; authentica
tor=AllowAllAuthenticator; authorizer=AllowAllAuthorizer; auto_bootstrap=true; auto_snapshot=true; back_pressure_enabled
=false; back_pressure_strategy=org.apache.cassandra.net.RateBasedBackPressure{high_ratio=0.9, factor=5, flow=FAST}; batc

C:\opt\cassandra\2\bin>cassandra.bat -f
WARNING! Powershell script execution unavailable.
Please use 'powershell Set-ExecutionPolicy Unrestricted'
on this user-account to run cassandra with fully featured
functionality on this platform.
Starting with legacy startup options
Starting Cassandra Server

```

```
File Edit View Insert Format Tools Extensions Help
Command Prompt - cassandra.bat -f

C:\opt\cassandra\3\bin>cassandra.bat -f
WARNING! Powershell script execution unavailable.
Please use 'powershell Set-ExecutionPolicy Unrestricted'
on this user-account to run cassandra with fully featured
functionality on this platform.
Starting with legacy startup options
Starting Cassandra Server
```

```
datacenter1: datacenter1
=====
Status=Up/Down
-/ State=Normal/Leaving/Joining/Moving
-- Address Load Tokens Owns (effective) Host ID Rack
UN 127.0.0.2 269.34 KB 256 69.3% 14df64f7-82f2-4bf2-a4ec-add12d60d121 rack1
UN 127.0.0.3 233.42 KB 256 62.8% edd17770-58a9-4b2f-a8f7-8673d045a6a8 rack1
UN 127.0.0.4 107.59 KB 256 67.9% 1225ce9d-4de7-445c-9451-b415052045a3 rack1
```

5. Create keyspace on node1

```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19045.4046]
(c) Microsoft Corporation. All rights reserved.

C:\anurag\1\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.4 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> desc keyspaces;

system_traces system_schema system_auth system system_distributed

cqlsh> CREATE KEYSPACE IF NOT EXISTS my_keyspace
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
cqlsh> desc keyspaces;

system_schema system system_traces
system_auth system_distributed my_keyspace

cqlsh> CREATE KEYSPACE IF NOT EXISTS 21510016
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
SyntaxException: line 1:30 no viable alternative at input '21510016' (CREATE KEYSPACE IF NOT EXISTS [21510016]...)
cqlsh> CREATE KEYSPACE IF NOT EXISTS anurag
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
cqlsh>
cqlsh> desc keyspaces;

anurag system_auth system_distributed my_keyspace
system_schema system system_traces

cqlsh> _
```

See that it is visible from other nodes

```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19045.4046]
(c) Microsoft Corporation. All rights reserved.

C:\anurag\2\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.4 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> desc keyspaces;

anurag      system_auth  system_distributed  my_keyspace
system_schema  system      system_traces
cqlsh> _
```

6. Create a table at node 1 and check it's visibility at other nodes.

```
C:\Windows\System32\cmd.exe - cqlsh
cqlsh> desc keyspaces;

system_schema  system              system_traces
system_auth    system_distributed  my_keyspace

cqlsh> CREATE KEYSPACE IF NOT EXISTS 21510016
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
SyntaxException: line 1:30 no viable alternative at input '21510016' (CREATE KEYSPACE IF NOT EXISTS [21510016]...)
cqlsh> CREATE KEYSPACE IF NOT EXISTS anurag
... WITH replication = {'class': 'SimpleStrategy', 'replication_factor': 1};
cqlsh>
cqlsh> desc keyspaces;

anurag          system_auth  system_distributed  my_keyspace
system_schema  system              system_traces

cqlsh> use anurag;
cqlsh:anurag> CREATE TABLE my_keyspace.users (
...     user_id UUID PRIMARY KEY,
...     username TEXT,
...     email TEXT
... );
cqlsh:anurag>
cqlsh:anurag> INSERT INTO my_keyspace.users (user_id, username, email)
... VALUES (
...     uuid(),
...     'john_doe',
...     'john.doe@example.com'
... );
cqlsh:anurag>
```

```
C:\Windows\System32\cmd.exe - cqlsh
Microsoft Windows [Version 10.0.19045.4046]
(c) Microsoft Corporation. All rights reserved.

C:\anurag\2\bin>cqlsh

WARNING: console codepage must be set to cp65001 to support utf-8 encoding on Windows platforms.
If you experience encoding problems, change your console codepage with 'chcp 65001' before starting cqlsh.

Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.4 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
WARNING: pyreadline dependency missing. Install to enable tab completion.
cqlsh> desc keyspaces;

anurag          system_auth  system_distributed  my_keyspace
system_schema  system              system_traces

cqlsh> use anurag;
... ;
cqlsh:anurag> SELECT * FROM my_keyspace.users;

 user_id          | email          | username
-----+-----+-----
 837d68bb-f645-4892-9b2c-d7cca03409f | john.doe@example.com | john_doe

(1 rows)
cqlsh:anurag>
```

Conclusion

This assignment encapsulates the comprehensive process of establishing a multi-node Cassandra cluster. Delving into essential concepts such as Cassandra clustering and pivotal network configuration parameters including `rpc_address`, `listen_address`, and `seeds`. By engaging in practical tasks and hands-on exercises, participants not only grasp the fundamentals of distributed database systems but also delve into intricate facets of data modeling, proactive monitoring, and proficient cluster management.

References

<https://phoenixnap.com/kb/install-cassandra-on-windows>

<https://extendit.us/articles/steps-configure-multiple-nodes-cassandra-single-windows-machine>