Hadoop Admin Notes for PD Environment

ReD Shield Engineering

Exported on 06/05/2020

Table of Contents

19 May 2020

Yellowbrick is confirmed and will be taking the majority of **load off** Hive/LLAP. So the Performance Hadoop cluster reconfigured and give more resources to our Spark nodes group.

Proposed changes in PERF
llap_nodes (LLAP0 as is currently is configured, single instance shared across all applications) – 4 Nodes **spark_nodes** – 6 Nodes

1 Steps:

Stop the Hive LLAP services of llap0 and llap1

Uninstall the LLAP1 services from the masternode4 (cov3lphdpmst04)

Delete the config group as llap_app_name

(Make sure all the config changes are deleted related to llap1).

Remove the node lables - spark_llap

yarn rmadmin -removeFromClusterNodeLabels "spark_llap"

Error:

20/05/19 00:41:22 WARN client.RequestHedgingRMFailoverProxyProvider: Invocation returned exception: org.apache.hadoop.yarn.exceptions.YarnException: java.io[1].IOException: Cannot remove label=spark_llap, because queue=spark_llap is using this label. Please remove label on queue before remove the label

Removed the spark_llap queue as part of llap1

Added that resource 28% to the Applications queue (27+28=55%)

Check the list of node labels make sure it deleted.

yarn cluster --list-node-labels

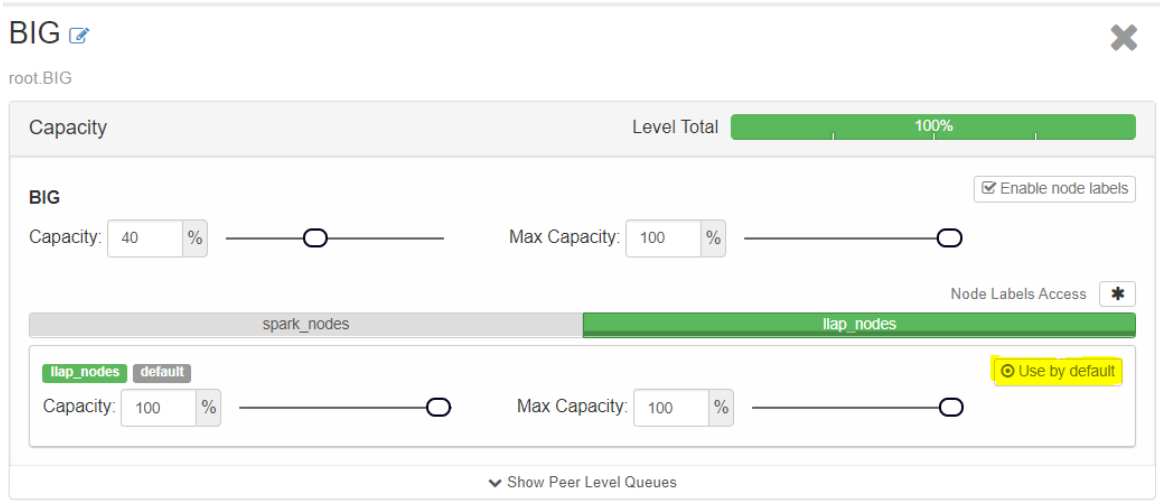
Add llap1 3 nodes to the node lables ***spark_nodes***

*yarn rmadmin -replaceLabelsOnNode "*cov3lphdpllap01am.tsacorp.com[2]*=spark_nodes* cov3lphdpwrk01am.tsacorp.com[3]*=spark_nodes* cov3lphdpwrk02.am.tsacorp.com[4]*=spark_nodes* cov3lphdpwrk03.am.tsacorp.com[5]*=spark_nodes* cov3lphdpwrk04.am.tsacorp.com[6]*=spark_nodes* cov3lphdpwrk05.am.tsacorp.com[7]*=spark_nodes"*

Restart all in YARN to remove the spark_llap entities from the cluster.

Choose the node labels **Use by default** in each queues according to the node labels.

1 http://java.io
2 http://cov3lphdpwrk02.am.tsacorp.com
3 http://cov3lphdpwrk03.am.tsacorp.com
4 http://cov3lphdpwrk04.am.tsacorp.com
5 http://cov3lphdpwrk05.am.tsacorp.com
6 http://cov3lphdpwrk05.am.tsacorp.com
7 http://cov3lphdpwrk05.am.tsacorp.com



In spark:

Redirects Spark to use llap0 instead of llap1

spark.hadoop.hive.llap.daemon.service.hosts = @llap0

spark.sql.hive.hiveserver2.jdbc.url=jdbc:hive2://cov3lphdpmst03.am.tsacorp.com:10501/default;httpPath=cliservice;principal=hive/_HOST@IPA.AM.TSACORP.COM;transportMode=http(see page 3)

13 Aug 2019

2 To separate the Spark/IBI load, created one more llap as llap1

For IBI front end loads connected to llap0 → llap_nodes → 4 nodes

For Spark backend loads connected to llap1 → spark_llap → 3 nodes

Node labels configuration in PERF
llap_nodes (LLAP0 used for IBI) – 4 Nodes **spark_llap** (LLAP1 used by spark jobs) – 3 Nodes **spark_nodes** – 3 Nodes

Step 1:

Created a separate node labels as 'spark_llap'

Removed the worker node 01 from the existing spark_nodes label (now we have 4 nodes alone)

Added the cov3lphdpwrk01 into spark_llap

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Alloca
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:4096, vCores:1>	<memory:354304, vCores:25>

Show 20 entries

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers
spark_llap	/default-rack	RUNNING	cov3lphdpwrk01.am.tsacorp.com:45454	cov3lphdpwrk01.am.tsacorp.com:8042	Tue Aug 13 04:31:48 -0500 2019		1

Show 20 entries

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers
spark_nodes	/default-rack	RUNNING	cov3lphdpwrk03.am.tsacorp.com:45454	cov3lphdpwrk03.am.tsacorp.com:8042	Tue Aug 13 04:31:53 -0500 2019		0
spark_nodes	/default-rack	RUNNING	cov3lphdpwrk05.am.tsacorp.com:45454	cov3lphdpwrk05.am.tsacorp.com:8042	Tue Aug 13 04:31:48 -0500 2019		0
spark_nodes	/default-rack	RUNNING	cov3lphdpwrk02.am.tsacorp.com:45454	cov3lphdpwrk02.am.tsacorp.com:8042	Tue Aug 13 04:31:53 -0500 2019		0
spark_nodes	/default-rack	RUNNING	cov3lphdpwrk04.am.tsacorp.com:45454	cov3lphdpwrk04.am.tsacorp.com:8042	Tue Aug 13 04:31:48 -0500 2019		1

Showing 1 to 4 of 4 entries

Step 2:

Removed the data_science team from the Teams parent queue

Added under the root leaf

Added the new node label – spark_llap

Configured the spark_llap queue with 100% resource of spark_llap node label.

Step 3:

Created separate config group called “llap_app_name” in Ambari Hive

Installed new interactive service as llap1 in cov3lphdpmst04 node

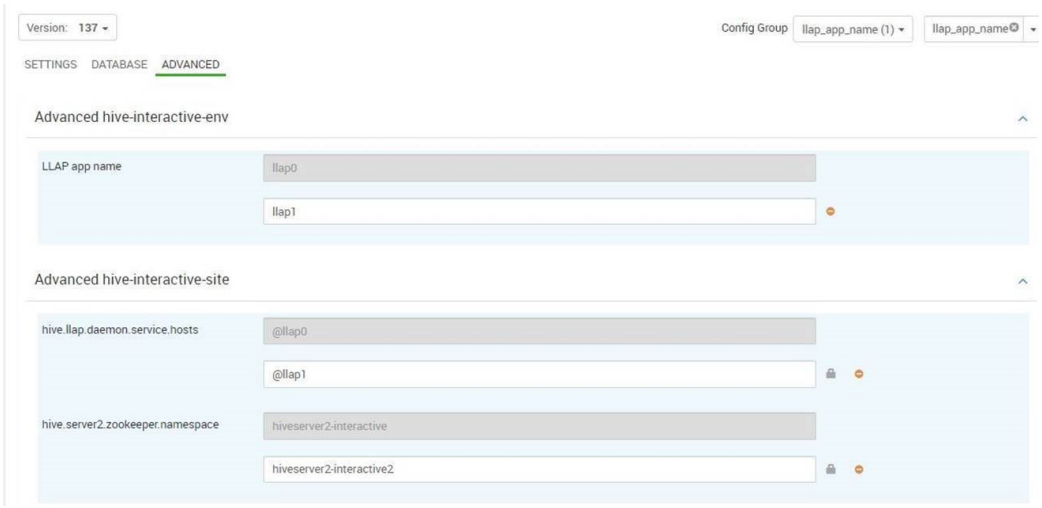
Added the configurations like:

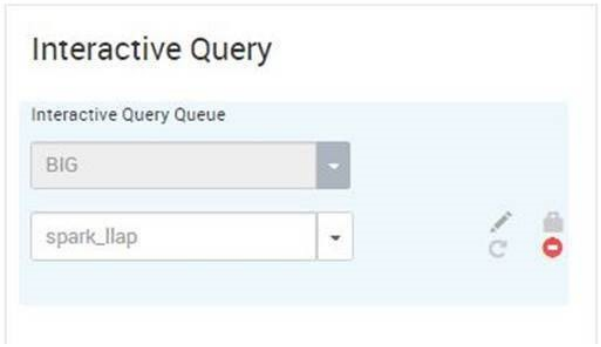
llap_app_name=llap1

hive.llap.daemon.service.hosts=@llap1[8]

hive.server2.zookeeper.namespace=hiveserver2-interactive2

Interactive Query Queue=spark_llap

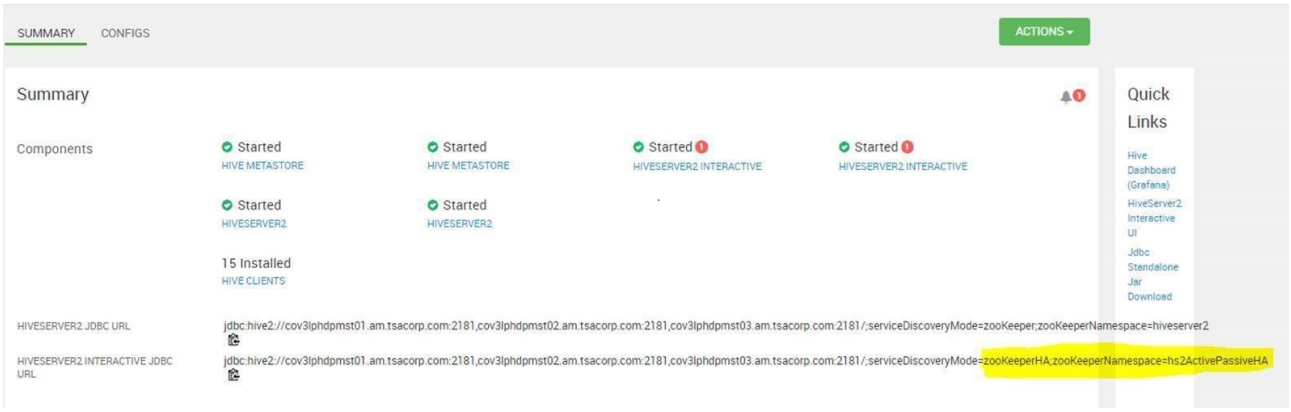




Post configurations in Hive’s Summary page we get:

I have done all the steps which Pablo mentioned in the comments to achieve @llap0 and @llap1 and get 3 JDBC URL to use (Normal hiveserver and 2 interactive namespaces).

8 mailto:hive.llap.daemon.service.hosts=@llap1


SUMMARY
CONFIGS
ACTIONS
Summary
Components
Started
HIVE METASTORE
Started
HIVE METASTORE
Started
HIVESERVER2 INTERACTIVE
Started
HIVESERVER2 INTERACTIVE
Started
HIVESERVER2
Started
HIVESERVER2
15 Installed
HIVE CLIENTS
Quick Links
Hive Dashboard (Grafana)
HiveServer2 Interactive UI
Jdbc Standalone Jar Download
HIVESERVER2 JDBC URL
jdbc:hive2://cov3lphdpmst01.am.tsacorp.com:2181,cov3lphdpmst02.am.tsacorp.com:2181,cov3lphdpmst03.am.tsacorp.com:2181/;serviceDiscoveryMode=zooKeeper;zooKeeperNamespace=hiveserver2
HIVESERVER2 INTERACTIVE JDBC URL
jdbc:hive2://cov3lphdpmst01.am.tsacorp.com:2181,cov3lphdpmst02.am.tsacorp.com:2181,cov3lphdpmst03.am.tsacorp.com:2181/;serviceDiscoveryMode=zooKeeperHA;zooKeeperNamespace=hs2ActivePassiveHA

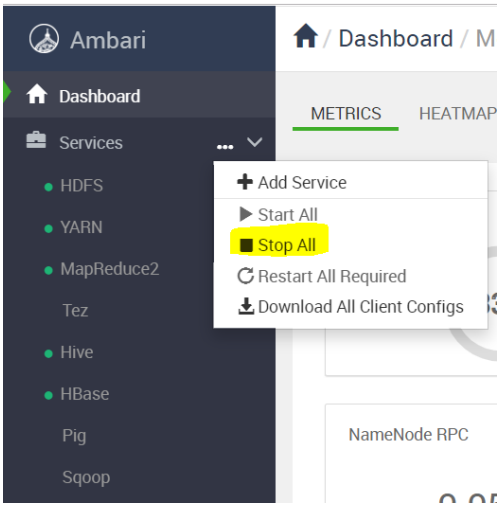
3 Restart guide for Hadoop Environments

We may face issues with the OS or Network in our environments. If in case, planned to bring down the clusters follow the steps to avoid the data loss.

1. Stop all the services in Ambari:

   → Require Admin access to stop those services

2. Locate the Services tab in the left side panel, click the "Stop All" to stop the services manually.



3. Takes sometime to stop each Hadoop components in a while and make sure that stopped successfully. If it doesnt stop any particular service, need to login to the node/service to identify the issue through logs. Clear the issue to proceed further.

4. Login to the Ambari node to stop the Ambari service,

   → Require 'root' access to stop the service

   $ ambari-server stop

5. Take the dump file for the Ambari server:

   $ pg_dump -U ranger ranger >/tmp/ranger.sql

6. Take the dump file for Ranger, Hive and Oozie

   $ pg_dump -U hive hive > /tmp/hive.sql

   $ pg_dump -U ranger ranger > /tmp/ranger.sql

   $ pg_dump -U oozie oozie > /tmp/oozie.sql

7. Post backup the dump files stop the Postgres DB.

   $ service postgresql status

   $ service postgres stop

8. Restart the nodes now to avoid the services or data get crashed now

   $ reboot

9. Post restart, bring up the services:

$ systemctl start postgres → Start the Postgres services.

$ systemctl status postgres → Check the status of the Postgres DB

9.1 Successful restart, the following steps to validate the DB

[root@cov3lphdpmst04 ~]# su - postgres → From root change the postgres user
[postgres@cov3lphdpmst04 ~]$ psql → Enter into postgres prompt
psql (9.6.11, server 9.2.23)
Type "help" for help.

postgres=# → Default postgres prompt

postgres=# \l → List the schemas available the postgres BD

List of databases
Name	Owner	Encoding	Collate	Ctype	Access privileges
ambari	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	=Tc/postgres +
				postgres=CTc/postgres+
				ambari=CTc/postgres
hive	hive	UTF8	en_US.UTF-8	en_US.UTF-8
oozie	oozie	UTF8	en_US.UTF-8	en_US.UTF-8
postgres	postgres	UTF8	en_US.UTF-8	en_US.UTF-8
ranger	ranger	UTF8	en_US.UTF-8	en_US.UTF-8	=Tc/ranger +
				ranger=CTc/ranger
ranger_db	ranger	UTF8	en_US.UTF-8	en_US.UTF-8	=Tc/ranger +
				ranger=CTc/ranger
rangerkms	rangerkms	UTF8	en_US.UTF-8	en_US.UTF-8
rangerkms_db	rangerkms	UTF8	en_US.UTF-8	en_US.UTF-8
template0	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	=c/postgres +
				postgres=CTc/postgres
template1	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	=c/postgres +
				postgres=CTc/postgres
(10 rows)

9.2 If in case, the data got corrupted the dump file has to be restored with the following steps:

postgres=# CREATE ROLE ranger superuser; → Create a user account for 'ranger' DB

postgres=# ALTER ROLE ranger WITH LOGIN; → Alter role for 'ranger' user

postgres=# \password ranger; → Password to be captured for the 'ranger' user

Enter password:

Enter password again: → Note down the same password to update in Ambari configs

postgres=# CREATE DATABASE rangerkms OWNER ranger; → Create a new database schema for 'ranger'

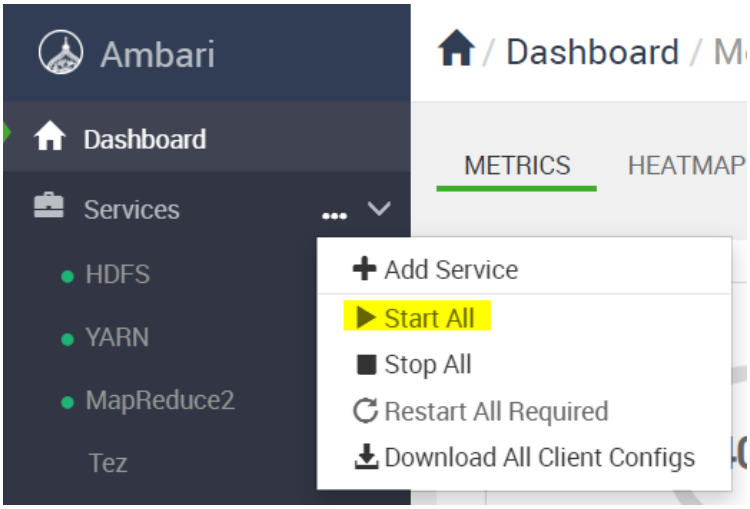
[postgres@cov3lphdpmst04 ~]$ psql ranger_db < /tmp/ranger_metastore-before-upgrade.sql → Restore the dump file into the 'ranger' database

10. Start the ambari server

$ ambari-server start

11. Start all the Hadoop services in Ambari UI:

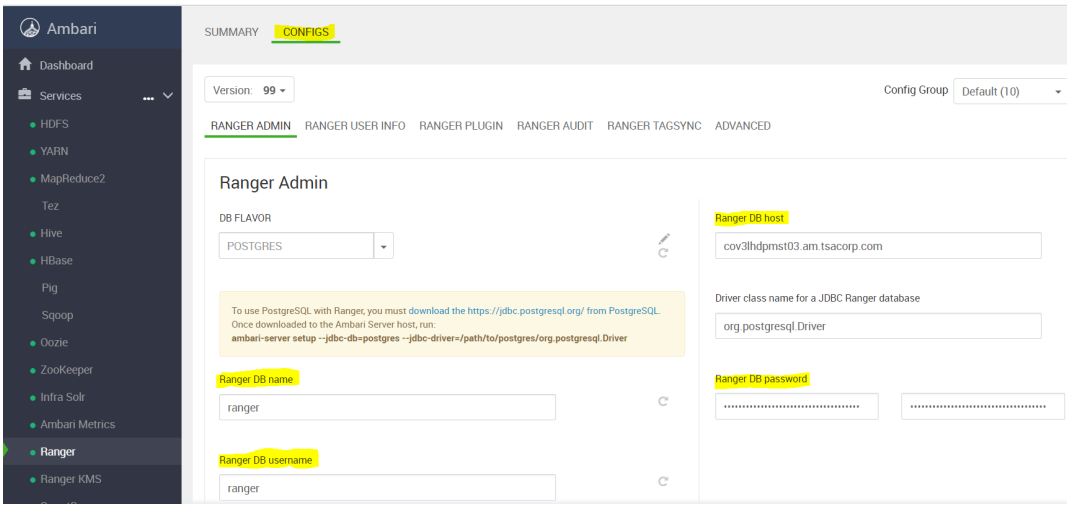
Services → Start All



12. If we restore the dump file, need to do the following steps in the Ambari to update the DB.

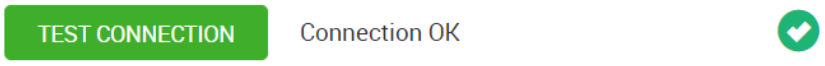
For ranger,

Ranger → Configs:



Update the Ranger DB Name, Ranger DB User Name, Ranger DB Host and Ranger DB Password:

Post update, click the "Test Connection" to validate the connectivity to the DB from Ambari:



If not, need to check the DB connectivity through the logs.

13. Make sure that, all the services are UP and running in Ambari.

13.1. If all the services are UP need to do smoke tests, in Hive

Edgenode → Beeline → show tables → select count (*) from <db.table_name>;

16 Mar 2020

Before;

[root@cov3lphdpmst03 ~]# systemctl status postgresql.service
● postgresql.service - PostgreSQL database server
Loaded: loaded (/usr/lib/systemd/system/postgresql.service; enabled; vendor preset: disabled)
Active: failed (Result: exit-code) since Mon 2020-03-16 00:22:49 CDT; 22s ago
Process: 155243 ExecStart=/usr/bin/pg_ctl start -D ${PGDATA} -s -o -p ${PGPORT} -w -t 300 (code=exited, status=1/FAILURE)
Process: 155237 ExecStartPre=/usr/bin/postgresql-check-db-dir ${PGDATA} (code=exited, status=0/SUCCESS)

Mar 16 00:22:48 cov3lphdpmst03.am.tsacorp.com[9] systemd[1]: Starting PostgreSQL database server...
Mar 16 00:22:48 cov3lphdpmst03.am.tsacorp.com[10] pg_ctl[155243]: FATAL: could not create lock file "/var/run/postgresql/.s.PGSQL.5432....ctory
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[11] pg_ctl[155243]: pg_ctl: could not start server
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[12] pg_ctl[155243]: Examine the log output.
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[13] systemd[1]: postgresql.service: control process exited, code=exited status=1
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[14] systemd[1]: Failed to start PostgreSQL database server.
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[15] systemd[1]: Unit postgresql.service entered failed state.
Mar 16 00:22:49 cov3lphdpmst03.am.tsacorp.com[16] systemd[1]: postgresql.service failed.
Hint: Some lines were ellipsized, use -l to show in full.

It shows /var/run/postgresql* is missing there, need to create it and change the ownership to postgres.

[root@cov3lphdpmst03 ~]# mkdir -p /var/run/postgresql
[root@cov3lphdpmst03 ~]# chown -R postgres:postgres /var/run/postgresql

Then start the service as,

[root@cov3lphdpmst03 ~]# systemctl start postgresql.service

and check the status as,

[root@cov3lphdpmst03 ~]# systemctl status postgresql.service
● postgresql.service - PostgreSQL database server

9 http://cov3lphdpmst03.am.tsacorp.com/
10 http://cov3lphdpmst03.am.tsacorp.com/
11 http://cov3lphdpmst03.am.tsacorp.com/
12 http://cov3lphdpmst03.am.tsacorp.com/
13 http://cov3lphdpmst03.am.tsacorp.com/
14 http://cov3lphdpmst03.am.tsacorp.com/
15 http://cov3lphdpmst03.am.tsacorp.com/
16 http://cov3lphdpmst03.am.tsacorp.com/

```
Loaded: loaded (/usr/lib/systemd/system/postgresql.service; enabled; vendor preset: disabled)
Active: active (running) since Mon 2020-03-16 00:23:18 CDT; 3s ago
Process: 155749 ExecStart=/usr/bin/pg_ctl start -D ${PGDATA} -s -o -p ${PGPORT} -w -t 300 (code=exited, status=0/
SUCCESS)
Process: 155743 ExecStartPre=/usr/bin/postgresql-check-db-dir ${PGDATA} (code=exited, status=0/SUCCESS)
Main PID: 155752 (postgres)
CGroup: /system.slice/postgresql.service
├─155752 /usr/bin/postgres -D /var/lib/pgsql/data -p 5432
├─155753 postgres: logger process
├─155757 postgres: checkpointer process
├─155758 postgres: writer process
├─155759 postgres: wal writer process
├─155760 postgres: autovacuum launcher process
└─155761 postgres: stats collector process

Mar 16 00:23:17 cov3lphdpmst03.am.tsacorp.com[17] systemd[1]: Starting PostgreSQL database server...
Mar 16 00:23:18 cov3lphdpmst03.am.tsacorp.com[18] systemd[1]: Started PostgreSQL database server.
[root@cov3lphdpmst03 ~]#
```

17 http://cov3lphdpmst03.am.tsacorp.com/
18 http://cov3lphdpmst03.am.tsacorp.com/