LEARNING

vertica

#vertica

Table of Contents

About

You can share this PDF with anyone you feel could benefit from it, downloaded the latest version from: vertica

It is an unofficial and free vertica ebook created for educational purposes. All the content is extracted from Stack Overflow Documentation, which is written by many hardworking individuals at Stack Overflow. It is neither affiliated with Stack Overflow nor official vertica.

The content is released under Creative Commons BY-SA, and the list of contributors to each chapter are provided in the credits section at the end of this book. Images may be copyright of their respective owners unless otherwise specified. All trademarks and registered trademarks are the property of their respective company owners.

Use the content presented in this book at your own risk; it is not guaranteed to be correct nor accurate, please send your feedback and corrections to info@zzzprojects.com

Chapter 1: Getting started with vertica

Remarks

This section provides an overview of what vertica is, and why a developer might want to use it.

It should also mention any large subjects within vertica, and link out to the related topics. Since the Documentation for vertica is new, you may need to create initial versions of those related topics.

Examples

Installation or Setup

Operational System Setup (Linux-CentOS x64 or RedHat x64)

1-Increase the swap space to a minimum of 2 Gb

2-Space and CPU requirements:

-Vertica requires at least 1 GB per CPU.

-disk utilization per node should no more than sixty percent (60%). Disk space is temporarily required by certain query execution operators, such as hash joins and sorts, in the case when they have to spill to disk.

-configure TEMP SPACE separate from data disk space.

3-Install the prerequisite for Vertica Cluster

```
yum install  rsync  python*  telnet ruby*  java* sudo  openssh-server openssh-clients
chkconfig sshd on
service sshd start
```

4-Edit the /etc/pam.d/su file

```
vi  /etc/pam.d/su
#add the line
session required pam_limits.so
```

5-Verify that the NTP Daemon is Running

```
chkconfig --list ntpd
#if is not on use the commands
chkconfig ntpd on
#start ntp service
/etc/init.d/ntpd start
```

6-Remove Nonessential Applications

> For optimal performance, Vertica is designed to use all available resources on each host
> machine. Vertica recommends that you remove or disable all non-essential applications from
> cluster hosts.

7-Configuring the Network

Alter the **/etc/hosts** file. Make sure file exists and that it contains the loopback address 127.0.0.1

7.1-Setting Up Cluster Hosts

Make sure that the /etc/hosts file includes all of the hosts that become part of the cluster. For example, if the hosts are named host01, host02, host03, and host04, the **/etc/hosts** file on each host looks like this:

```
# Do not remove the following line, or various programs
# that require network functionality will fail.
127.0.0.1               localhost.localdomain localhost
192.xxx.13.128          host01
192.xxx.13.129          host02
192.xxx.13.130          host03
192.xxx.13.131          host04
```

This should be done in all hosts(nodes)

7.2- Edit the /etc/sysconfig/network file :

```
vim /etc/ sysconfig/network
Alter the hostname and set it to the desired name :
HOSTNAME=host01
```

7.3 Setting the HOSTNAME Environment Variable

```
vim /etc/profile or /etc/bashrc
Add the line
export HOSTNAME=hostname
```

7.4-Verify that the hostname resolution works correctly Verify this with the command

```
/bin/hostname
Hostname
```

Restart the hosts(nodes) Make sure you do all this steps in all Hosts(nodes)as root user.

7.5 Disable the firewall

Firewalls not recommended for database hosts SELinux (Security-Enhanced Linux) Iptables

7.6 Provide Root and for dbadmin user to SSH Access to the Cluster

Steps to do it for root or dbadmin user:

```
[root@Vertica_Master1 ~]# ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
/root/.ssh/id_rsa already exists.
Overwrite (y/n) y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
7c:b5:11:48:d3:c1:e6:f5:80:b3:4a:4a:93:ed:16:99
 root@Vertica_Master1
The key's randomart image is:
+--[ RSA 2048]----+
|          .o+oo   |
|           ..*.o  |
|          o =o+ o |
|        .+ E.oo  .|
|        .S=.o.    |
|          ..+     |
|           .      |
|                  |
|                  |
+-----------------+
[root@Vertica_Master1 ~]# cd ~
[root@Vertica_Master1 ~]# chmod 700 .ssh
[root@Vertica_Master1 ~]# cd .ssh/
[root@Vertica_Master1 .ssh]# cp id_rsa.pub authorized_keys
```

Do on all hosts the steps show upper and then follow the next steps

```
[root@Vertica_Master1 .ssh]# cat id_rsa.pub
ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAQEArX26Pgsyvkw+o0Vimm26FSOAtTh9C9
mZ+tS7LfO92y7RDKsSm38tSQO/p1f2NWP6UzAam8dG77Zo8W+wjwF6bEJbFU9Gq+S/j
hETD8bMTKh6odZNhXmZanddvH4qnA0eKngAPe9Y93udA6kEYGpA0sCWMFcbrtvwraz7@Vertica_Master1
[root@Vertica_Master1 .ssh]# ssh root@Vertica_Master2
The authenticity of host 'vertica_master2 (10.xxx.1.224)' can't be established.
RSA key fingerprint is ff:9c:48:27:7d:6b:a1:39:5a:17:d0:a3:a3:9d:f0:48.
Are you sure you want to continue connecting (yes/no) yes
Warning: Permanently added 'vertica_master2,10.xxx.1.224' (RSA) to the list of known hosts.
root@vertica_master2's password:xxxxxx- this is the password for the root user
Last login: Tue Sep  4 15:11:35 2012 from e05347
-bash: Vertica_Master2: command not found
[root@Vertica_Master2 ~]# hostname-check to see that you are on the Vertica_Master2
Vertica_Master2

[root@Vertica_Master1 .ssh]# vim authorized_keys
```

And copy the the content of the id_rsa.pub into authorized_keys and save it .

Do this for all hosts so that they all have the keys form all hosts in their authorized_keys file.

- Host1 will hold host1,host2,host3 public keys inside
- Host2 will hold host1,host2,host3 public keys inside
- Host3 will hold host1,host2,host3 public keys inside

-and so on if you have more hosts in your cluster.

8 -Download and install Vertica software on Master Node(where you will run the Administrative jobs of the cluster)

As root :

```
#rpm –ihv  vertica-<version>.x86_64.RHEL5.rpm
After entering the command, a progress indicator appears:
Preparing...    ################################# [100%]
1:vertica       ################################# [100%]
Vertica 6.0.xx successfully installed on host hostname.
```

Normally by default vertica will be installed into **/opt/vertica** dir.

8.1- Run the Install Script

On the master node run the following command(The master node would be the node you most access for admin work)

```
/opt/vertica/sbin/install_vertica –s host_list –r rpm_package –u dba_username
```

Where options are : -s host_list comma-separated list of hostnames or IP addresses to include in the cluster; do not include space characters in the list. Example :

```
–s host01,host02,host03
```

or

```
–s 192.xxx.233.101,192.xxx.233.102,192.xxx.233.103
```

-r rpm_package The pathname of the Vertica RPM package.

Example:

-r "vertica_6.0.x.x86_64.RHEL5.rpm" **-u** dba_username -this will be the name of the user how will run the admintools(only)

- If you omit the parameter, the default database administrator account name is dbadmin.

Example of full command for 3 nodes cluster :

```
# /opt/vertica/sbin/install_vertica –s 10.xxx.1.216,10.xxx.1.224,10.xxx.1.225 –r
 /home/user/Downloads/vertica-6.0.0-3.x86_64.RHEL5.rpm –u dbadmin
```

9 – Create the Vertica Data and Catalog directories on each node:

```
Directory names are totally up to you. Remember that database user must have owner rights over
them.
```

Note: this refers to extra directories added after installation

---

```
mkdir /vertica_db/data
mkdir /vertica_db/data
chown  dbadmin:dbadmin  vertica_db/
```

10- Add the /opt/vertica/bin to you dbadmin user path so you can access the vertica tools without the full path.

11- Access the adminTool and you will be asked to introduce the licence.dat file provided by Vertica.(this applies only for Enterprise Edition).

12- Create the database and chose on which nodes it will reside !! Follow the next steps to see how to create a database that will reside on the Vertica Cluster .

Read Getting started with vertica online: https://riptutorial.com/vertica/topic/8889/getting-started-with-vertica

Chapter 2: TroubleShooting

Examples

ERROR: Too many ROS containers exist for the following projections:

it caused by your batch size too small, which lead to a lot of ROS Containers created and reach the limitation(1024 default). you should do defragment using TupleMover task(mergeout) before the error raised.

To do troubleshooting:

1. ROS Containers viewed from the projections.

```
select * from STORAGE_CONTAINERS where projection_name like '%DATASET_TABLE%';
```

2. check ContainersPerProjectionLimit settings Views

```
SELECT *
FROM CONFIGURATION_PARAMETERS
WHERE parameter_name = 'ContainersPerProjectionLimit' ;
```

3. ROS Container Number Lookup

```
select count(*) from STORAGE_CONTAINERS where projection_name like '%DATASET_TABLE%';
```

4. Solutions

```
-- change ContainersPerProjectionLimit settings
SELECT SET_CONFIG_PARAMETER('ContainersPerProjectionLimit', 2048);
```

or

```
-- change mergeout frequency
SELECT SET_CONFIG_PARAMETER('MergeOutInterval', 30);
```

or

```
-- Do TupleMover Task manually
select do_tm_task('mergeout','projection_name')
```

COPY Rejected Data and Exception Files

it often happened some rows with format issue, data type issue rejected by copy command while try load it by copy command. the query return succeed but some of data rejected.

To do troubleshooting

1. save Rejected Data and Exceptions

```
COPY large_tbl FROM :file1 ON site01,
               :file2 ON site01,
               :file3 ON site02,
               :file4 ON site02
               DELIMITER '|'
               REJECTED DATA :reject_s1 ON site01, :reject_s2 ON site02
               EXCEPTIONS :except_s1 ON site01, :except_s2 ON site02;
```

2. check errors in exceptions file, and fix it one by one

Read TroubleShooting online: https://riptutorial.com/vertica/topic/9778/troubleshooting

Credits

S. No	Chapters	Contributors
1	Getting started with vertica	Community, Up_One
2	TroubleShooting	suiwenfeng