Installing Single-node Vertica on RHEL

By Yiqun Zhang

1. Vertica supports both ext3 and ext4 file system, but ext4 is recommended for performance reasons. Normally we always choose ext4 during system installation.

2. We should not enable LVM (Logical Volume Manager) while partitioning the hard drive because it is not supported by Vertica.

3. If you are using this note right after a clean installation of RHEL, please refer to http://www2.cs.uh.edu/~yzhang/ref/network_config.pdf first to make sure the network is configured right.

4. The swap space should be **2GB regardless** the amount of installed RAM.
   Note that this is **NOT** a must, the guide said *larger swap space is acceptable, although unnecessary*.
   The first time when we install RedHat, we didn't know this requirement, so the swap partition was as large as 4GB, to make up for that we adopt swap file instead of the partition. And here is how:
   To see the swap usage summary, type `swapon -s`.
   To create a **2GB** file for swap usage, type:

   ```
   dd if=/dev/zero of=/swapfile bs=1024k count=2000
   ```

   `/dev/zero` is a special file in Unix-like operating systems that provides as many null characters (ASCII NUL, `0x00`) as are read from it.

   By calling `mkswap /swapfile`, we set up a swap area on the file `/swapfile`, and to put this swap file into use, type: `swapon /swapfile`.

   After that, we need to let the OS stop using current swap partition, type `df -h` to see the name of the partition that is currently used for swapping, then type `swapoff [name]` to stop the OS from using that partition as swap space.

   Finally, write the changes to the file `/etc/fstab`, thus it will take effect every time the server startups.
   Type `vim /etc/fstab` to edit it, comment out the line where the swap partition is set, and add a new line for our swap file:

   ```
   /swapfile    swap    swap    defaults    0    0
   ```

5. We need to change the scheduler Linux uses for I/O requests. To view the current configuration, type:

   ```
   cat /sys/block/sda/queue/scheduler
   ```

   "`sda`" in the command stands for the name of the disk, Normally you will get an output like:

   ```
   noop deadline [cfq]
   ```

   "`cfq`" stands for *completely fair queuing*, it is now selected because it is surrounded by a pair of brackets. To change it to `deadline`, type:

   ```
   echo deadline > /sys/block/sda/queue/scheduler
   ```

   This command should also be added to `/etc/rc.d/rc.local` to let it be executed automatically every time the server startups.

6. Note that the time zone configuration is in the guide page 22, currently we don't have to change it.

7. According to the guide, each host has a system setting for the Linux environment variable `LANG`. `LANG` determines the locale category for native language, local customs, and coded character set in the absence of the `LC_ALL` and other `LC_` environment variables.

I checked our system, there is no `LC_ALL` variable, and the value of `LANG` is same as recommendation which is `en_US.UTF-8`.

NOTE: To view the system environment variables, type `export`.

8. It is a **MUST** to make sure that the network time protocol (NTP) daemon is configured to run on startup. If the server does not have NTP running, the installation can fail with a spread configuration error as well as other potential errors. In RHEL, the NTP daemon is referred as `ntpd`. To check its status, type:

```
chkconfig --list ntpd
```

This service is **DISABLED** by default in RHEL. To enable it, type:

```
chkconfig ntpd on
```

Since this will only take effect after rebooting, we could manually start it now by calling:

```
/etc/init.d/ntpd start
```

9. 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:

```
127.0.0.1 localhost.localdomain localhost
192.168.13.128 host01
192.168.13.129 host02
192.168.13.130 host03
192.168.13.131 host04
```

NOTE: Unless you are installing Vertica on a single host machine with no plans to expand to a multiple-node cluster or import or export data to or from another database, **DO NOT** use the loopback address `127.0.0.1` or the name `localhost` in a node definition, and **DO NOT** assign the cluster node host name to `127.0.0.1`.

Here I adopt 127.0.0.1 for the simplest case.

NOTE: The host name could be changed by editing file `/etc/sysconfig/network`.

10. The firewall considerations is covered in the guide page 31, how to ensure port 5433 is enabled is in page 86.

11. Make sure that `/dev/pts` is mounted by running command `mount`.

12. Download the Vertica install package, I have put it in http://www2.cs.uh.edu/~dbms/vertica.rpm.

13. Install the package by the command:

```
rpm –Uvh [pathname]
```

14. Run `/opt/vertica/sbin/install_vertica`

The parameters are covered in the guide page 35, because I am installing a vertica with single node, so I didn't provide any parameters.

After Vertica is successfully installed for one single node, to begin with creating database, run:

```
/opt/vertica/bin/admintools
```

Then a GUI will appear asking for license file, for community edition, press "`OK`" directly.
Then a second screen will appear asking you to accept the EULA, choose "`Accept`" and press "`OK`".

After the two steps above, a main menu will appear for you to play with the server. On our server, I created an empty database called "`test`".