

Preparing the operating system

Prior to installing Cassandra, it is important to understand how Cassandra utilizes disk space. Disk space depends on usage. The database writes data to disk when appending data to the commit log for durability and when flushing memtables to SSTable data files for persistent storage. The commit log has a different access pattern (read/writes ratio) than the pattern for accessing data from SSTables. This is more important for spinning disks than for SSDs.

SSTables are periodically compacted. Compaction improves performance by merging and rewriting data as well as discarding old data. However, depending on the type and size of the compactions, disk utilization and data directory volume temporarily increases during compaction. For this reason, be sure to leave an adequate amount of free disk space available on a node.

Cassandra's data and commit logs should not, under any circumstances, be placed on the drive where the operating system is installed. Ideally, a server should have 3-4 drives or partitions. The root, /, or OS partition can be used as the target for the application. The /data partition should have adequate amounts of storage to accommodate your data. The /logs partition should hold your commit logs and (unless it is SSD) be on a different physical disk than the /data partition). The /backup partition should be allocated for backups.

For information about selecting hardware, see <http://cassandra.apache.org/doc/latest/operating/hardware.html>.

In order to achieve adequate performance, create separate partitions, ideally, on separate drives. This helps to avoid i/o contention. We recommend 3 separate block devices (disks). The first block device should contain the operating system as well as a mount for the programs (*/opt/local*). The second block device (preferably SSD) should contain a mount point at /data (the device must have high storage capacity for all the data). The third block device (preferably SSD but does not need to be of high capacity) should contain a mount point at /logs. If you use SSD, the device can be a partition on the same block device as the /data partition. All partitions should be formatted using the XFS file system, and there should not be a swap partition. The /backup partition can be a mount on a shared storage device but should not be on the same physical drive as the /data partition.

See an example of the contents of */etc/fstab* after partitioning where the partitions were created using LVM (without a mount for the /backup partition).

fstab

```
#
# /etc/fstab
# Created by anaconda on Tue May  2 16:31:05 2017
#
# Accessible filesystems, by reference, are maintained under '/dev/disk'
# See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info
#
/dev/mapper/cl_twccentos7-root / xfs defaults 0 0
/dev/mapper/cl_twccentos7-data /data xfs defaults 0 0
/dev/mapper/cl_twccentos7-logs /logs xfs defaults 0 0
/dev/mapper/cl_twccentos7-opt_local /opt/local xfs defaults 0 0
```

In the above example:

- Disk 1 contains the following partitions: */opt/local* (40GB) and / (rest of the drive capacity).
- Disk 2 (the disk with the highest capacity) contains the */data* partition (at least 250GB). Due to the way compactions are handled by Cassandra, up to 50% of headroom may be needed in a worst-case scenario.
- Disk 3 contains the */logs* partition (at least 10 GB).
- You should also create an additional mount for backups. Unlike the data and commit log partitions, which should be on SSD storage, this mount can be of any type (including centralized storage such as a SAN or NAS). It should have at least the same capacity as the /data partition.

The partitioning scheme above is an example. Internal security protocols in your organization may dictate that other directories may not be located in the main partition. During the installation, all applications will be installed in */opt/local*. By default, Cassandra will be installed in */var/lib*. Application logs will be written to */home/twcloud*.