



HP Vertica Hardware Planning Guide

This document provides information that helps you to appropriately size and recommend Hardware for HP Vertica.

Optimal Hardware for HP Vertica

Processor

For optimal performance, HP recommends running two socket servers with eight to fourteen core CPUs, clocked at or above 2.6 GHz for clusters over 10TB, and single socket servers with eight to twelve cores clocked at or above 2.6GHz for clusters under 10TB.

Memory

HP Vertica requires a minimum of 8 GB of memory per physical CPU core in the server. However in high performance applications, you should run 12-16 GB of memory per physical core. The memory should be at least DDR3-1600 (preferably DDR3-1866), and should be appropriately distributed across all memory channels in the server.

Storage

HP Vertica requires a minimum read/write speed of 40 MB/s per physical core of the CPU. However, you should have 60–80 MB/s per physical core for best performance. Each node should have 1–9 TB of storage post RAID. In a production setting, we recommend RAID 10; in some cases, RAID 50 will be acceptable,

SSDs are not required due to the heavy compression/encoding done by HP Vertica. In most cases, a RAID array of more, less expensive HDDs will work just as well as a RAID array of fewer SSDs to satisfy HP Vertica requirements.)

If you intend to use RAID 50 for your data partition, you should keep a spare node in every rack, allowing for manual failover of an HP Vertica node in the case of a drive failure. (An HP Vertica node recovery is faster than a RAID 50 rebuild). Also be sure to never put more than 10TB compressed on any node to keep node recovery times at an acceptable rate.

Network

HP recommends 10G networking in almost every case. With the introduction of 10G over cat6a (Ethernet) the cost difference is minimal.

Sizing Your Cluster

Consider these factors when sizing your cluster:

1. **Data Volume (Compression):** First look at the total raw data volume for the cluster, and then apply a reasonable compression number. In most cases 2:1 compression with high availability is a good start. You should use a previously-attained compression number or run some quick tests to determine a reasonable compression number for your data. To do this, install HP Vertica on an existing system and run Database Designer (DBD), or use a DDL you already have and load in several hundred gigabytes to 1 TB of data.
2. **Data Growth:** Once you have a good idea of the starting compressed data volume, look at the amount of data ingest (i.e. the amount of data you will be saving each day) and the retention policy (how long you intend to save the data). Consider your starting compressed data volume and your ingest rate multiplied by the retention period; this gives you a good idea of the total data volume for your cluster.
3. **Workload:** The next important factor to look into is the workload itself. Is it a high-concurrency workload? What type of workload management will you be doing? How much data will the average query be running against (that is, will you be using partition pruning?) These can all be major factors, but for the most part a high concurrency workload will require more disk speed (high spindle count) and more memory on each node, as well as more cores (which could also be more nodes).

Server Configuration

HP recommends a minimum of three nodes for high availability.

All configurations are based on raw data volumes assuming 2:1 compression. If you are getting better than 2:1 compression you may need fewer nodes. Always run your hardware BOM through the HP Vertica team before purchasing hardware for a production cluster. *Note that in a high concurrency or heavy workload environment, we may recommend to oversizing your cluster to achieve better performance.*

Raw Data Size	Recommended Server Configuration and Comments
Up to 5 TB	Three nodes with at least eight 10 k rpm spinning disks for the data partition (total of 1–2 TB per node or equivalent SSD), a single 8-12 core processor with clock speed or 2.6GHz or higher, and 96-128G of RAM
5-10 TB	Six nodes with at least eight 10 k rpm spinning disks for the data partition (total of 1–2 TB per node or equivalent SSD), a single 8-12 core processor with clock speed or 2.6GHz or higher, and 96-128G of RAM
10-40 TB	CS300 for Vertica (3-4 nodes each with 22 10k rpm drives, dual 10 core processors at 3.0 GHz, 256G of RAM)
40 TB – 1 PB	CS300 for Vertica (4-100 nodes each with 22 10k rpm drives, dual 10 core processors at 3.0 GHz, 256G of RAM)
Larger than 1 PB	In most cases the same configuration as listed above will work, but contact an HP Vertica Technical Representative for recommendations on sizing clusters over 1 PB.

Note: in 10TB+ cases where an appliance is not an option we recommend hardware configured to the HP Vertica DL380p Configuration guide, which is the same hardware spec as the CS300 for Vertica.

Get connected

hp.com/go/getconnected

Current HP driver, support, and security alerts delivered directly to your desktop.

© Copyright 2012, 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

