HP Vertica® Analytics Platform

Configuring the HP DL380p Gen8 for HP Vertica

Contents

Using the HP ProLiant DL380p Gen8 25-SFF Server as an HP Vertica Node

The HP Vertica Analytics Platform software runs on a shared-nothing Massively Parallel Processing (MPP) cluster of peer nodes; each of the peer nodes is independent. An HP Vertica *node* is a hardware host configured to run an instance of HP Vertica.

This document provides recommendations for configuring an individual DL380p Gen8 25-SFF as an HP Vertica node.

The recommendations presented in this document are intended to help you create a cluster with the highest possible HP Vertica software performance.

This document includes a Bill of Materials (BOM) as a reference and to help with the DL380p Gen8 25-SFF.

Considering a Software Baseline for Linux and HP Vertica

This document assumes that your services, once you configure them, will be running the following minimum software versions.

> **HP Vertica 6.1.x Enterprise Edition.** The HP Vertica 6.1.x version implements a new I/O profile to best use the enhanced capabilities of a modern hardware platform like the Proliant Gen8. As such, HP Vertica recommends using the HP Vertica 6.1 (or later) Enterprise Edition.

> **Red Hat Enterprise Linux 6.x.** Note that, mention of Red Hat Linux is not an indication of HP Vertica's preference over other supported Linux distributions. The purpose of this recommendation is solely to enhance discussion points in this document.

Selecting a Server Model

The HP DL380 Gen8 product family includes several server models.

For maximum HP Vertica software performance, HP Vertica recommends the DL380p Gen8 25-SFF Server (part number 665554-B21).

Selecting a Processor

The HP DL380p Gen8 25-SSF Vertica nodes should include

> For Maximum Performance in high concurrency workloads: two (2) Intel® Xeon® E5-2697v2 2.7GHz/12-core DDR3-1866 130W processors.

> For Maximum Performance in moderate concurrency workloads: two (2) Intel® Xeon® E5-2690v2 3.0GHz/10-core DDR3-1866 130W processors.

> For Balanced Performance: two (2) Intel® Xeon® E5-2680v2 2.8GHz/10-core DDR3-1866 115W processors.

The processor recommendation is based on the fastest 10-12 core processors available for the DL380p Gen8 25-SFF platform. These processors allow HP Vertica to deliver the fastest response time across a wide spectrum of concurrent database workloads.

The processor's clock speed directly affects the HP Vertica database response time while additional cores enhance the cluster's ability to simultaneously execute multiple MPP queries and data loads.

Selecting Memory

The recommended memory configuration based on 16GB DDR3 1866MHz DIMMs and 128-256GB of RAM is intended to achieve the best memory performance along with the option of future expansion.

For maximum HP Vertica performance, DL380p Gen8 25-SFF each HP Vertica node should include 128 or 256GB of RAM, configured as follows:

> 8 x 16GB DDR3-1866 RDIMMs, 1DPC (16GB per channel)

> 16 x 16GB DDR3-1866 RDIMMs, 1DPC (32GB per channel)

This configuration is recommended for the best memory performance and because it allows for future expansion (if you chose the 128GB version). You can expand this configuration in the field to 192GB or 256GB by adding 8 x 8GB or 8 x 16GB DIMMs respectively.

A two-processor DL380p Gen8 25-SFF server has 8 memory channels with 3 DIMM slots in each channel, for a total of 24 slots. Follow these guidelines when configuring DL380p Gen8 25-SFF memory configuration for best performance:

a) Populate all memory channels. (No channel should be left completely blank.) Load all channels similarly.

b) To use the highest supported DIMM speed of 1866MHz, populate the maximum number of DIMMS per channel (DPC) to 2. . 3DPC DIMMs run at 1333MHz or lower.

The following table provides several additional memory configurations.

Sample Configuration	Total Memory	Considerations
16 x 8GB DDR3-1866 RDIMMs 2DPC (8GB + 8GB per channel)	128GB	This configuration is less expensive version of the recommended maximum performance configuration. However, note that you cannot expand 2DPC by adding additional modules without some memory speed degradation.
8 x 16GB DDR3-1866 RDIMMs + 8 x 8GB DDR3-1866 RDIMMs 2DPC (16GB + 8GB per channel)	192GB	A high-memory option that may be beneficial to support some database workloads.
16 x 16GB DDR3-1866 RDIMMs 2DPC (16GB + 16GB per channel)	256GB	A high-memory option that may be beneficial to support some database workloads.

Configuring Storage

Configure storage hardware for a maximum performance DL380p Gen8 25-SFF HP Vertica node as follows:

➢ 1x HP DL380p Gen8 25-SFF CTO Chassis with 24 Hot Plug SmartDrive SFF (2.5-inch) Drive Bays

➢ 1x Embedded HP P420i/2GB Flash Backed Cache SAS Array Controller (integrated on system board)

➢ 2x 300GB 6G SAS 10K 2.5in SC ENT drives (configured as RAID1 for the OS and the HP Vertica Catalog location)

➢ 22x 1.2TB 6G SAS 10K 2.5in SC ENT drives (configured as one RAID 1+0 device for the HP Vertica Data location, for approximately 13TB total formatted storage capacity per HP Vertica node)

You can configure an HP Vertica node with less storage capacity:

- Substitute 22x 1.2TB 6G SAS 10K 2.5in SC ENT drives with 22x HP 600GB 6G SAS 10K 2.5in SC ENT drives.

- Configure the drives as one RAID 1+0 device for the HP Vertica Data location, for approximately 6TB total data storage capacity per HP Vertica node.

You can configure the 23rd 1.2TB (or 600GB) data drive (for 25 drives in total) as a hot spare; however, such a configuration is unnecessary with RAID 1+0 configuration.

HP Vertica can operate on any storage type. For example, HP Vertica can run on internal storage, a SAN array, a NAS storage unit, or a DAS enclosure, as long as the storage is presented to the host as a file system, and is capable of providing sufficient I/O bandwidth. Internal storage in a RAID configuration offers the best price/performance/availability characteristics at the lowest TCO.

An HP Vertica installation recommends at least 2 storage locations in smaller clusters – one for the operating system and catalog, and the other for data. You should place the data location on a dedicated contiguous storage volume.

HP Vertica is a multithreaded application. While each worker thread issues predominantly large sequential I/O, in a typical concurrent multi-user environment the HP Vertica Data location I/O profile is best characterized as *large block random I/O.*

Drive Bay Population

The 25 drive bays on the DL380p Gen8 25-SFF servers are attached to the Smart Array P410i Controller over 2 internal SAS port connectors as follows:

port 1 – drive bays 1 through 13   (13 drives)

port 2 – drive bays 14 through 25   (12 drives)

👍 *For best storage performance, I/O should be distributed across the two SAS physical links as evenly as possible.*

For example, an ideal implementation of the HP Vertica-recommended 24 drive configuration:

➤ 300GB drives placed in bays 1 and 14; this spreads the OS RAID1 I/O evenly across the two SAS groups.

➤ 1.2TB (or 600GB) drives placed in all remaining vacant bays except bay 13; this spreads HP Vertica data RAID10 I/O evenly across the two SAS groups.

Keep in mind that, with 24 drives, pursuing a "perfectly" balanced physical placement of drives is desirable but not critical. However, if you procure DL380p Gen8 25-SFF servers with substantially fewer drives (e.g., 16 or less), loading the SAS links evenly is essential in avoiding a storage performance bottleneck.

In general, the following simple "bays population rule" works well for any number of available drives:

👍 *Populate bays with drives from "both ends" towards the "middle".* That is, insert drives into bays 1 – 25 – 2 – 24 – 3 – 23 – 4 – 22 – and so forth.

Hpsa Device Driver for HP Smart Array Controllers

Starting with Red Hat Enterprise Linux (RHEL) 6, the new Hpsa SCSI driver debuts as the default driver for newer HP Smart Array controllers such as Gen8's P420. The main user-visible change is that the logical RAID devices are named sda, sdb, sdc, and so forth.

Unlike its cciss driver predecessor, hpsa presents logical disk devices in the standard SCSI namespace. This change greatly improves compatibility with system tools and utilities designed to work with SCSI storage.

The latest versions of Linux include the hpsa driver.

More information about the SCSI drivers is available in the "Migrating to hpsa" white paper.

Required HP System Software

The following HP System Software should be installed on each DL380p Gen8 25-SFF HP Vertica server:

System Software	Minimum Required Version (Release Date)
System ROM (BIOS)	2012.08.20 (26 Oct 2012)
HP ProLiant Smart Array Controller Driver (also known as the "hpsa" driver)	3.1.0-7 (26 Oct 2012)
Smart Array P410i Controller Firmware	3.22 (26 Oct 2012)
HP Array Configuration Utility CLI for Linux 64-bit	9.30.15.0 (26 Oct 2012)

To check the version of the currently installed Smart Array Controller (hpsa) driver:

```
% modinfo hpsa | grep version
```

HP ACU utility is the RAID configuration utility for HP Smart Array Controllers. It is required to configure storage with non-default parameters. HP Array Configuration Utility (ACU) is available as a command line program, as well as with a web-based interface.

The following storage configuration examples use the HP ACU Command Line Interface (CLI).

To check whether HP ACU CLI utility is installed:

```
% rpm -q hpacucli
```

To check the version of the currently installed Smart Array Controller Firmware:

```
% hpacucli ctrl slot=0 show | grep Firmware
```

👍 *To obtain all available fixes and enhancements, HP recommends keeping the system software (i.e. drivers, firmware and management utilities) up-to-date with the latest release available.*

All system software is available for unrestricted download at HP Business Support Center. No registration is required. To obtain the software:

1. At www.hp.com, select Support & Drivers tab.

2. Enter the product name (e.g., "DL380p Gen8").

3. Select your product, and then select the operating system. All necessary software downloads are available from this web page.

Data RAID Configuration

RAID 10
RAID 0
RAID 1   RAID 1
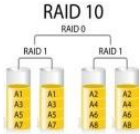
The 22 data drives should be configured as one RAID 1+0 device as follows:

➢ The recommended strip size for the data RAID 1+0 is 256K, which is the default setting for the P420i controller.

➢ The recommended Controller Cache (Accelerator) Ratio is 10/90, which is the default setting for the P420i controller.

➢ The logical drive should be partitioned with a single primary partition spanning the entire drive.

For more information on Vertica storage locations, see the "Before You Install" section in the HP Vertica Installation Guide.

HP Vertica *data location* should be placed on a dedicated physical storage volume and should not be co-located with the HP Vertica *catalog*. The recommended HP Vertica catalog location on a DL380p Gen8 25-SFF HP Vertica node is the OS drive. No special OS/HP Vertica catalog drive configuration/tuning actions are necessary.

✎ *Note: HP Vertica does not support storage configured with the Linux Logical Volume Manager in the I/O path. This applies to all HP Vertica storage locations including the catalog which is typically placed on the OS drive.*

Linux I/O Subsystem Tuning

To support the maximum performance DL380p Gen8 25-SFF node configuration, the following Linux I/O configuration settings are recommended for the HP Vertica Data location volumes:

The recommended Linux *file system* is `ext4`

The recommended Linux *I/O Scheduler* is `deadline`

The recommended Linux *Readahead* setting is `8192 512-byte sectors (4 MB)`

System administrators should configure the deadline scheduler and the readahead settings for the HP Vertica data volume so that these setting persist across server restarts.

✎ *A failure to follow the recommended Linux I/O subsystem settings will adversely affect performance of the HP Vertica software.*

Data RAID Configuration Example

All configuration/tuning instructions below pertain to the HP Vertica data storage location.

Example steps to configure data location RAID device:

(1) View the current storage configuration:

```
% hpacucli ctrl all show config detail
```

(2) The RAID1 OS drive should be configured as a logical drive 1 and the 22 data drives should either be "loose" or pre-configured into a temporary logical drive. Note that, if this is not your configuration, you may have to destroy the non-OS logical drives to create a new logical drive with the recommended parameters (note that content will be lost). To destroy a non-OS logical drive:

```
% hpacucli ctrl slot=0 ld 2 delete forced
```

(3) Create a new RAID10 data drive with 256K strip size:

```
% hpacucli ctrl slot=0 create type=ld raid=1+0 ss=256
drives=1I:1:2,1I:1:3,1I:1:4,1I:1:5,1I:1:6,1I:1:7,1I:1:8,1I:1:9,1I:1:10,1I:1:11,1I:1:12,2I:1:15,2
I:1:16,2I:1:17,2I:1:18,2I:1:19,2I:1:20,2I:1:21,2I:1:22,2I:1:23,2I:1:24,2I:1:25
```

(4) Partition and format the RAID10 data drive:

```
% parted -s /dev/sdb mklabel gpt mkpart primary ext4 0% 100%
% mkfs.ext4 /dev/sdb1
```

(5) Create a /data mount point, add a line to the /etc/fstab file, and mount the HP Vertica data volume:

```
% mkdir /data
[add line to /etc/fstab]: /dev/sdb1  /data  ext4  defaults  0 0
% mount /data
```

(6) Add the following lines to /etc/rc.local file so the Linux I/O scheduler, Linux Read-Ahead, and hugepage defragmentation settings persist across system restarts:

```
echo deadline > /sys/block/sdb/queue/scheduler
blockdev --setra 8192 /dev/sdb
echo never > /sys/kernel/mm/redhat_transparent_hugepage/enabled
echo never > /sys/kernel/mm/redhat_transparent_hugepage/defrag
echo no > /sys/kernel/mm/redhat_transparent_hugepage/khugepaged/defrag
```

Selecting a Network Adapter

To support the maximum performance MPP cluster operations, DL380p Gen8 25-SFF Vertica nodes should be provisioned with at least two (2) 10 Gigabit Ethernet ports (for full network redundancy):

➢ HP Ethernet 10Gb 2-port 530FLR-SFP+ Adapter in the FlexibleLOM slot

These ports should be bonded together in an active-active configuration, then divided into two VLAN's to provide independent private and public networks.

Alternatively if 1Gb networking is needed in addition to 10G, the DL380p Gen8 25-SFF HP Vertica nodes should be provisioned with at least two (2) 10 Gigabit Ethernet ports, and two (2) 1 Gigabit Ethernet ports.

➢ HP Ethernet 1Gb 4-port 366FLR Adapter

➢ HP Ethernet 10Gb 2-port 530SFP+ Adapter

A Vertica cluster is formed with DL380p Gen8 25-SFF servers, associated network switches, and the HP Vertica software.

Each HP Vertica DL380p Gen8 25-SFF node should be connected to 2 separate Ethernet networks:

✓ The *private* network (i.e., cluster interconnect, used exclusively for internal cluster communications) must be same subnet, with a dedicated VLAN. HP Vertica does TCP P2P communications and UDP broadcasts on this network. IP addresses for the *private* network interfaces must be assigned statically. No external traffic should be allowed over the *private* cluster network.

✓ The *public* network is used for database client (i.e., application) connectivity. HP Vertica has no rigid requirements around public network configuration, but recommends that you assign static IP addresses for the public network interfaces.

All network interconnect should be provisioned with Ethernet redundancy. Otherwise the interconnect (specifically the switch) would be a single point of a cluster-wide failure. Although in some larger cluster configurations where a full rack failure is tolerable a single switch is acceptable.
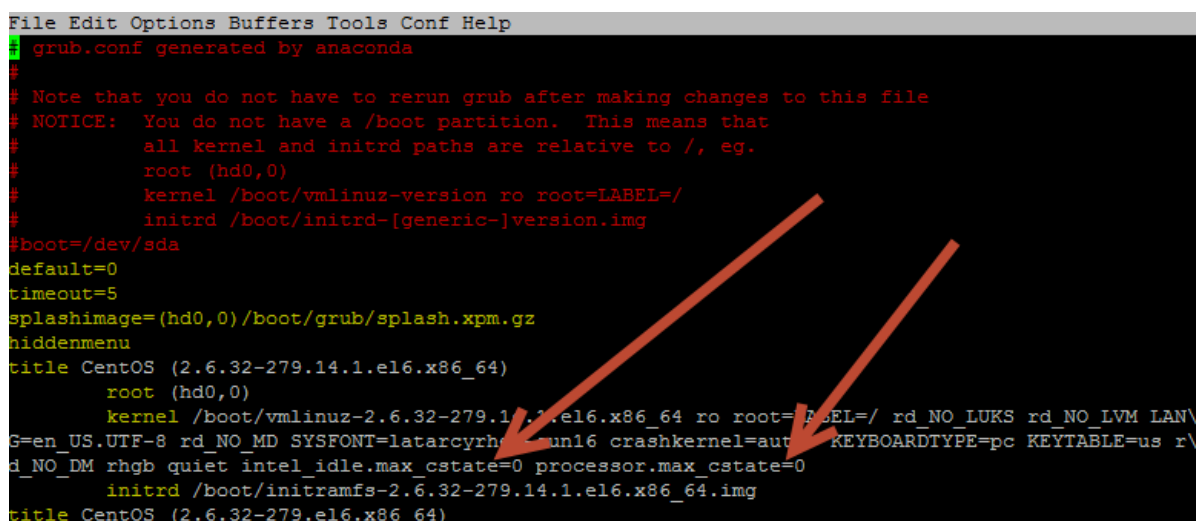
This generation of HP DL380p's is the first generation of the DL380p that does not include 1Gb networking integrated from the factory.

HP ROM–Based Setup Utility (BIOS) Settings

HP Vertica recommends changing the following Power Management Options from default values to ensure highest performance:

Option	Recommended Setting	Explanation
HP Power Profile	Maximum Performance	Enabling Maximum Performance disables all power management options that could have a negative impact on performance.
HP Power Regulator	HP Static High Performance Mode	Enabling the HP Power Regulator allows processors to run in the maximum power and performance state, regardless of the OS power management policy.
Advanced Power Management Options → Collaborative Power Control	Disable	HP Static Power Regulator Modes Are Ignored Under RHEL6 and SLES 11 on ProLiant Gen8 Server With Intel Processor For more info see the <u>HP CUSTOMER ADVISORY</u>.

Additionally a modification will have to be made to the kernel entry in the `/etc/grub.conf` file as shown below. (Append `intel_idle.max_cstate=0` and `processor.max_cstate=0` to the kernel line.)

HP ProLiant DL380p Gen8 25-SFF BOM

A sample BOM follows.
Notes:

- ✓ Part numbers are listed for reference only.
- ✓ Networking equipment and racks are not included.
- ✓ All hardware requests should come through the CSE team for review.

Quantity	Part Number	Description	Notes
1	665554-B21	HP DL380p Gen8 25-SFF CTO Server	
1	715214-L21	HP DL380p Gen8 E5 2690v2 FIO Kit	
1	715214-B21	HP DL380p Gen8 E5 2690v2 Kit	
16	708641-B21	HP 16GB 2Rx4 PC3 14900R 13 Kit	for 256GB of Memory
1	631681 B21	HP 2GB FBWC for P Series Smart Array	
2	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD	for OS and Catalog
23	697574-B21	HP 1.2TB 6G SAS 10K 2.5in SC ENT HDD	for approximately 13TB formatted storage capacity
1	684210-B21	HP Ethernet 10GbE 530FLR SFP+ FIO Adapter	for 10Gb networking
2	512327-B21	HP 750W CS Gold Ht Plg Pwr Supply Kit	
1	663478-B21	HP 2U SFF BB Rail Gen8 Kit	

Alternate Part Appendix

Processor

Quantity	Part Number	Description	Notes
1	715224-L21	HP DL380p Gen8 E5 2697v2 FIO Kit	12 cores at 2.7 GHz
1	715224-B21	HP DL380p Gen8 E5 2697v2 Kit	
1	715214-L21	HP DL380p Gen8 E5 2690v2 FIO Kit	10 cores at 3.0 GHz
1	715214-B21	HP DL380p Gen8 E5 2690v2 Kit	
1	715215-L21	HP DL380p Gen8 E5 2680v2 FIO Kit	10 cores at 2.8 GHz
1	715215-B21	HP DL380p Gen8 E5 2680v2 Kit	

Memory

Quantity	Part Number	Description	Notes
8	708641-B21	HP 16GB 2Rx4 PC3 14900R 13 Kit	for 128GB of Memory
16	708641-B21	HP 16GB 2Rx4 PC3 14900R 13 Kit	for 256GB of Memory

Disk Drives

Quantity	Part Number	Description	Notes
23	697574-B21	HP 1.2TB 6G SAS 10K 2.5in SC ENT HDD	for approximately 13TB formatted storage capacity
23	652589-B21	HP 900GB 6G SAS 10K 2.5in SC ENT HDD	for approximately 10TB formatted storage capacity
23	652583-B21	HP 600GB 6G SAS 10K 2.5in SC ENT HDD	for approximately 6TB formatted storage capacity
23	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD	for approximately 3TB formatted storage capacity (expandable)

Networking

Quantity	Part Number	Description	Notes
1	684210-B21	HP Ethernet 10GbE 530FLR SFP+ FIO Adapter	for 10Gb networking only
1	665240-B21	HP Ethernet 1Gb 4-port 366FLR Adapter	for 1Gb and 10Gb networking
1	652503 B21	HP Ethernet 10Gb 2P 530SFP+ Adapter	