

Configuring the HP DL380 Gen9 24-SFF CTO Server as an HP Vertica Node

HP Vertica Analytic Database

HP Big Data Foundations
Document Release Date: February, 2015



Contents

Using the HP DL380 Gen9 24-SFF Server as an HP Vertica Node	3
Recommended Software	3
Selecting a Server Model	3
Selecting a Processor	3
Selecting Memory	3
Selecting and Configuring Storage.....	4
Drive Bay Population	5
Protecting Data on Bulk Storage.....	5
Hpsa Device Driver for HP Smart Array Controllers.....	5
Data RAID Configuration.....	5
Linux I/O Subsystem Tuning.....	6
Data RAID Configuration Example.....	6
Selecting a Network Adapter.....	7
HP ROM-Based Setup Utility (BIOS) Settings	8
Alternate Part Appendix	10
Processor	10
Memory	10
Disk Drives	10
Networking	11

Using the HP DL380 Gen9 24-SFF Server as an HP Vertica Node

The HP Vertica Analytics Platform software runs on a shared-nothing MPP cluster of peer nodes. Each peer node is independent, and processing is massively parallel. An *HP Vertica node* is a hardware host configured to run an instance of HP Vertica.

This document provides recommendations for configuring an individual DL380 Gen9 24-SFF CTO Server 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 provide more information about the DL380 Gen9 24-SFF CTO Server.

Recommended Software

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

- **Vertica 7.1 Enterprise Edition.** The Vertica 7.1 version implements a new I/O profile to best use the enhanced capabilities of a modern hardware platform like the ProLiant Gen9. As such, Hewlett-Packard recommends using HP Vertica 7.1 (or later) Enterprise Edition.
- **Red Hat Enterprise Linux 6.x.** Recommending Red Hat Linux is not an indication of that Hewlett-Packard has a preference for this product 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 Gen9 product family includes several server models. The model that Hewlett-Packard recommends for maximum HP Vertica software performance is the DL380 Gen9 24-SFF CTO Server (part number 767032-B21).

Selecting a Processor

For maximum price/performance advantage on your HP Vertica database, the DL380 Gen9 24-SFF servers used for HP Vertica nodes should include two (2) Intel Xeon E5-2690v3 2.6 GHz/12-core DDR4-2133 135W processors.

This processor recommendation is based on the fastest 12-core processors (or optionally 14-core processors) available for the DL380 Gen9 24-SFF platform at the time of this writing. These processors allow HP Vertica to deliver the fastest possible response time across a wide spectrum of concurrent database workloads.

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

Selecting Memory

For maximum HP Vertica performance, DL380 Gen9 24-SFF servers used as HP Vertica nodes should include 256 GB of RAM. Configure this memory as follows:

- 8 x 32 GB DDR4-2133 RDIMMs, 1DPC (32 GB per channel)
- In the field, you can expand this configuration to 512 GB by adding 8 x 32 GB DIMMs.

A two-processor DL380 Gen9 24-SFF server has 8 memory channels with 3 DIMM slots in each channel, for a total of 24 slots. DL380 Gen9 24-SFF memory configuration should comply with DIMM population rules and guidelines:

- Do not leave any channel completely blank. Load all channels similarly.
- Populate the maximum number of DIMMs per channel (DPC) to 2. Doing so allows you to use the highest supported DIMM speed of 2133 MHz. DPCs with 3 DIMMs run at 1866 MHz or lower.

Note: Hewlett-Packard recommends that you follow these guidelines to avoid a reduction in memory speed that could adversely affect the performance of your HP Vertica database.

The preceding recommended memory configuration is based on 32 GB DDR4 2133 MHz DIMMs and 256 GB of RAM. That configuration is intended to achieve the best memory performance while providing the option of future expansion. The following table provides several alternate memory configurations:

Sample Configuration	Total Memory	Considerations
8 x 16 GB DDR4-2133 RDIMMs 1 DPC (8 GB per channel)	128 GB	A low-memory option for systems with less concurrency and slower speed requirements.
16 x 16 GB DDR4-2133 RDIMMs 2 DPC (16 GB, 16 GB per channel)	256 GB	A slightly less expensive option for 256 GB of RAM that does not allow expansion in the field.
8 x 32 GB DDR4-2133 RDIMMs 1 DPC (32 GB per channel)	256 GB	The standard memory recommendation for HP Vertica.
24 x 16 GB DDR4-2133 RDIMMs 3 DPC (16 GB, 16 GB, 16 GB per channel)	384 GB	A higher memory option, but with lower memory speed of 1866 MHz. This selection does not allow for expansion in the future.
16 x 32 GB DDR4-2133 RDIMMs 2 DPC (32 GB + 32 GB per channel)	512 GB	A high-memory option that may be beneficial to support some database workloads.
24 x 32 GB DDR4-2133 LRDIMMs 3 DPC (32 GB, 32 GB, 32 GB per channel) (LR-DIMM)	768 GB	A LR-DIMM memory option, a high memory option but with lower memory speed. This selection does not allow for expansion in the future.

Selecting and Configuring Storage

Configure the storage hardware as follows for maximum performance of the DL380 Gen9 24-SFF server used as an HP Vertica node:

- 1x HP DL380 Gen9 24-SFF CTO Chassis with 24 Hot Plug SmartDrive SFF (2.5-inch) Drive Bays
- 1x HP DL380 Gen9 2-SFF Kit with 2 Hot Plug SmartDrive SFF (2.5 inch) Drives Bays (on the back of the server)
- 1x Embedded HP P440/2 GB Flash Backed Cache SAS Array Controller (integrated on system board)
- 2x 300 GB 6G SAS 10 K 2.5 inch SC ENT drives (configured as RAID1 for the OS and the Vertica Catalog location)
- 24x 1.2 TB 6G SAS 10 K 2.5 inch SC ENT drives (configured as one RAID 1+0 device for the Vertica Data location, for approximately 13 TB total formatted storage capacity per HP Vertica node)

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

- Substitute 24x 12 TB 6G SAS 10 K 2.5 inch SC ENT drives with 24x HP 600 GB 6G SAS 10 K 2.5 inch SC ENT drives.
- Configure the drives as one RAID 1+0 device for the HP Vertica data location, for approximately 6 TB of total data storage capacity per HP Vertica node.

Alternatively, you can configure the 23rd and 24th 1.2 TB (or 600 GB) data drives (for 22 active drives in total) as hot spares. However, such configuration is unnecessary with a 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. In each case, the storage appears 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 requires at least two storage locations—one for the operating system and catalog, and the other for data. Place these data locations on a dedicated, contiguous-storage volume.

HP Vertica is a multithreaded application. The HP Vertica data location I/O profile is best characterized as *large block random I/O*.

Drive Bay Population

The 26 drive bays on the DL380 Gen9 24-SFF servers are attached to the Smart Array P440 Controller over 4 internal SAS port connectors (through a 12 G SAS Expander Card) as follows:

- Ports 1, 2, 3—Drive bays 1–8, 9–16, 17–24 (8 drives each)
- Port 4—Drive bays 25 and 26 (2 drives)

Important: For best storage performance, distribute I/O distributed across the three SAS physical links used for data as evenly as possible (the 8 drive ports).

For example, an ideal implementation of the recommended 26-drive configuration is:

- 300 GB drives placed in bays 25 and 26, with the 2SFF drive expander in the rear of the server.
- 1.2 TB (or 600 GB) drives placed in all bays on the front of the server. This approach spreads the Vertica data RAID10 I/O evenly across the three SAS groups.

Admittedly, with 24 drives for data, pursuing a perfectly balanced physical placement of drives may be desirable but not critical. However, if DL380 Gen9 24-SFF servers have substantially fewer drives (18 or less), loading the SAS links evenly is essential to avoid a storage performance bottleneck.

Protecting Data on Bulk Storage

The HP Smart Array P440ar Controller offers the optional HP Secure Encryption capability that protects data at rest on any bulk storage attached to the controller. (Additional software, hardware, and licenses may be required.) For more information, see [HP Secure Encryption product details](#).

Hpsa Device Driver for HP Smart Array Controllers

Note: The information in this section is a best practices recommendation.

Starting with Red Hat Enterprise Linux (RHEL) 6, the new Hpsa SCSI driver is the default driver for newer HP Smart Array controllers like the Gen9 P440. The main change visible to users is that the logical RAID devices are named sda, sdb, sdc, and so on.

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 includes the hpsa driver.

More information about the SCSI drivers is available in the white paper, "[Migrating to hpsa](#)".

Data RAID Configuration

The 24 data drives should be configured as one RAID 1+0 device as follows:

- The recommended strip size for the data RAID 1+0 is 256 K, which is the default setting for the P440 controller.
- The recommended Controller Cache (Accelerator) Ratio is 10/90, which is the default setting for the P440 controller.
- The logical drive should be partitioned with a single primary partition spanning the entire drive.

Place the HP Vertica data location on a dedicated physical storage volume. Do not co-locate the HP Vertica data location with the HP Vertica catalog location. Hewlett-Packard recommends that the HP Vertica catalog location on an HP Vertica node on a DL380 Gen9 24-SFF server be the operating system drive. You do not need to perform any special operating system/HP Vertica catalog drive configuration or tuning.

For more information, read [Before You Install HP Vertica](#) in the product documentation, particularly the discussion of HP Vertica storage locations.

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

Linux I/O Subsystem Tuning

To support the maximum performance DL380 Gen9 24-SFF node configuration, Hewlett-Packard recommends the following Linux I/O configuration settings 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).

The current configuration recommendations differ from the previously issued guidance due to the changes in the Vertica I/O profile implemented in HP Vertica 7.

System administrators should durably configure the deadline scheduler and the read-ahead settings for the Vertica data volume so that these settings persist across server restarts.

Caution: Failing to use the recommended Linux I/O subsystem settings will adversely affect performance of the HP Vertica software.

Data RAID Configuration Example

The following configuration and tuning instructions pertain to the HP Vertica data storage location.

Note: The following steps are provided as an example, and may not be correct for your machine. Verify the drive numbers and population for your machine before running these commands.

1. View the current storage configuration:

```
% hpacucli ctrl all show config detail
```
2. Assume that the RAID1 OS drive is configured as a logical drive 1 and the 24 data drives are either “loose” or pre-configured into a temporary logical drive. To create a new logical drive with recommended parameters, any non-OS logical drives must be destroyed, and 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 256 K strip size:

```
% hpacucli ctrl slot=0 create type=ld raid=1+0 ss=256
drives=1I:1:1,1I:1:2,1I:1:3,1I:1:4,1I:1:5,1I:1:6,1I:1:7,1I:1:8,2I:1:9,2I:1:10,
2I:1:11,2I:1:12,2I:1:13,2I:1:14,2I:1:15,2I:1:16,3I:1:17,3I:1:18,3I:1:19,3I:1:2
0,3I:1:21,3I:1:22,3I:1:23,3I:1:24
```
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. So that the Linux I/O scheduler Linux Read-Ahead, and hugepage defragmentation settings persist across system restarts, add the following five lines to /etc/rc.local file:

```
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
```
7. After you have configured the storage, run vioperf to understand the baseline I/O performance.

The minimum required I/O is 20 MB/s read and write per physical processor core on each node. This value is based on running in full duplex, reading and writing at this rate simultaneously, concurrently on all nodes of the cluster. The

recommended I/O is 40 MB/s per physical core on each node. For example, the I/O rate for a server node with 2 hyper-threaded 6-core CPUs is a required minimum of 240 MB/s. Hewlett-Packard recommends 480 MB/s.

For more information, see [vioperf](#) in the product documentation.

Selecting a Network Adapter

To support the maximum performance MPP cluster operations, DL380 Gen9 24-SFF servers used as HP Vertica nodes should include at least two (2) 10 Gigabit Ethernet (RJ-45) ports:

- HP Ethernet 10 Gb 2-port 561FLR-T Adapter in the FlexibleLOM slot

Alternatively, if you need SFP+ networking, DL380 Gen9 24-SFF servers used as HP Vertica nodes should include at least two (2) 10 Gigabit Ethernet (SFP+) ports.

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

An *HP Vertica cluster* is formed with DL380 Gen9 24-SFF servers, associated network switches and HP Vertica software.

When used as an HP Vertica node, each DL380 Gen9 24-SFF server should be connected to two separate Ethernet networks:

- The *private network*, such as a cluster interconnect, used exclusively for internal cluster communications. This network must be the same subnet, dedicated switch or VLAN, 10 Gb Ethernet. 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, should be 1 Gb Ethernet or better. A 10-Gb Ethernet is suggested for workloads including high-volume data transfers between the database server cluster and connecting applications. HP Vertica has no rigid requirements for public network configuration. However, Hewlett-Packard recommends that you assign static IP addresses for the public network interfaces.

The private network interconnect should have Ethernet redundancy. Otherwise, the interconnect (specifically the switch) would be a single point of a cluster-wide failure.

Cluster operations are not affected even in the event of a complete failure of a public network. Thus, public network redundancy is not technically required. However, if a failure occurs, the application connectivity to the database is affected. Therefore, consider public network redundancy for continuous availability of the entire environment.

To achieve redundancy on both the private and public networks:

1. Take the two ports from the Ethernet card on the server and run one to each of the two top of rack switches (which are bonded together in an IRF).
2. Bond the links together using LACP.
3. Divide the links into public and private networks, using VLANs.

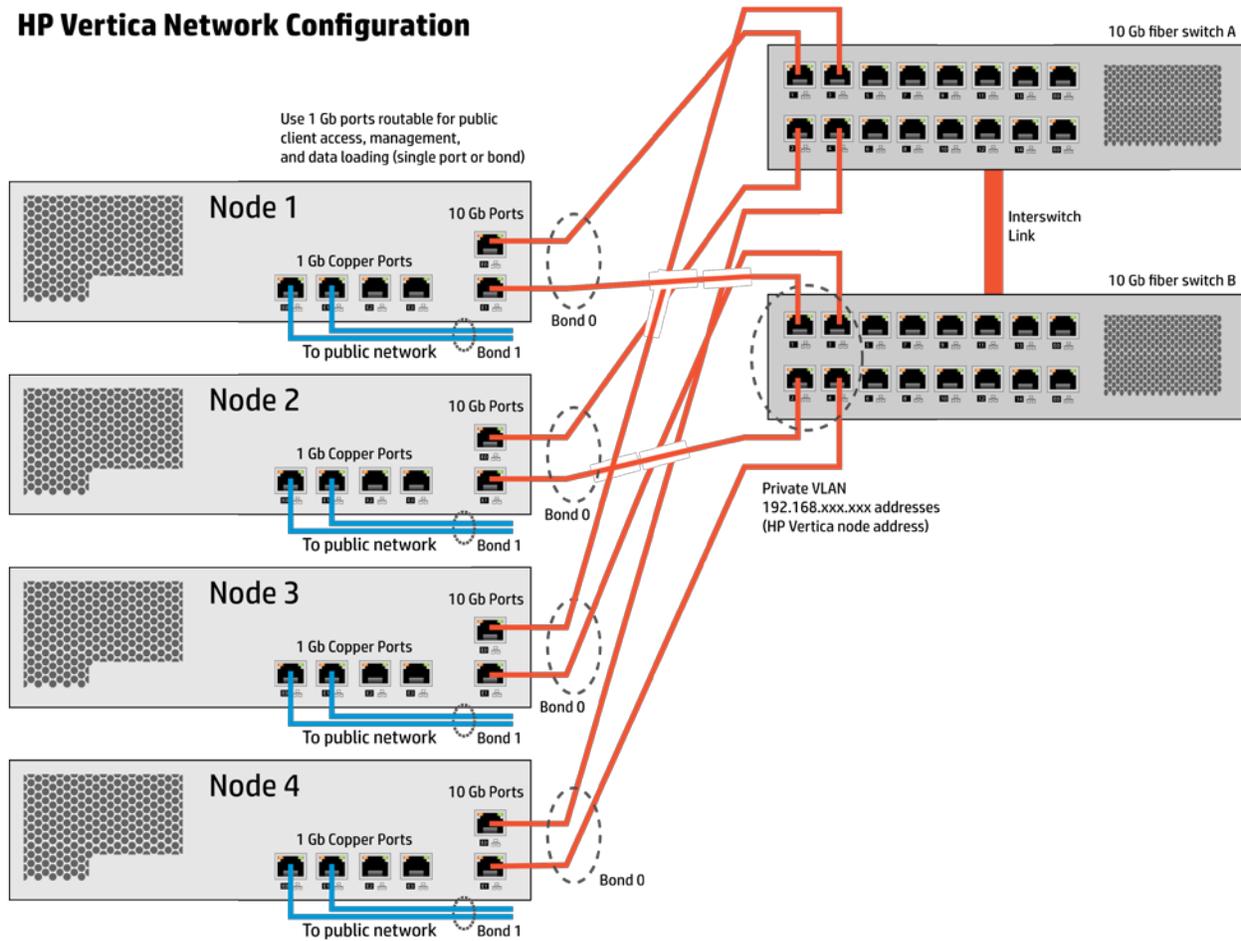
Configuring the Network

The following figure illustrates a typical network setup that achieves high throughput and high availability. (This figure is for demonstration purposes only.)

This figure shows that the bonding of the adapters allows for one adapter to fail without the connection failing. This bonding provides high availability of the network ports. Bonding the adaptors also doubles the throughput.

Furthermore, this configuration allows for high availability with respect to the switch. If a switch fails, the cluster does not go down. However, it may reduce the network throughput by 50%.

HP Vertica Network Configuration



HP ROM-Based Setup Utility (BIOS) Settings

Hewlett-Packard recommends changing the following Power Management Options from the default values for best performance:

Option	Recommended Setting	Explanation
HP Power Profile	Maximum Performance	Enabling Maximum Performance disables all power management options that could negatively affect 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 operating system management policy.
Advanced Power Management Options Collaborative Power Control	Disable	HP Static Power Regulator modes are ignored under RHEL6 and SLES11 on ProLiant Gen8 Servers with an Intel Processor. This is a known Linux issue. For more information, see the HP Customer Advisory .

Additionally, you must modify the kernel entry in the `/etc/grub.conf` file, as the following figure shows. (Append `intel_idle.max_cstate=0` and `processor.max_cstate=0` to the kernel line.)

```
File Edit Options Buffers Tools Conf Help
# grub.conf generated by anaconda
#
# Note that you do not have to rerun grub after making changes to this file
# NOTICE: You do not have a /boot partition. This means that
#         all kernel and initrd paths are relative to /, eg.
#         root (hd0,0)
#         kernel /boot/vmlinuz-version ro root=LABEL=/
#         initrd /boot/initrd-[generic-]version.img
#boot=/dev/sda
default=0
timeout=5
splashimage=(hd0,0)/boot/grub/splash.xpm.gz
hiddenmenu
title CentOS (2.6.32-279.14.1.el6.x86_64)
    root (hd0,0)
    kernel /boot/vmlinuz-2.6.32-279.14.1.el6.x86_64 ro root=LABEL=/ rd_NO_LUKS rd_NO_LVM LAN\
G=en_US.UTF-8 rd_NO_MD SYSFONT=latarcyrh\un16 crashkernel=auto KEYBOARDTYPE=pc KEYTABLE=us r\
d_NO_DM rhgb quiet intel_idle.max_cstate=0 processor.max_cstate=0
    initrd /boot/initramfs-2.6.32-279.14.1.el6.x86_64.img
title CentOS (2.6.32-279.el6.x86_64)
```



ProLiant DL380 Gen9 24-SFF Bill of Materials

The following table contains a sample Bill of Materials (BOM). Note the following about this BOM:

- 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	767032-B21	HP DL380 Gen9 24-SFF CTO Server	
1	719044-L21	HP DL380 Gen9 E5 2690v3 FIO Kit	For 12 cores
1	719044-B21	HP DL380 Gen9 E5 2690v3 Kit	For 12 cores
8	26722-B21	HP 32GB 4Rx4 PC42133P-L Kit	For 256 GB of memory
1	727250-B21	HP 12GB SAS Expander Card	
1	727258-B21	HP DL/ML/SL 96W Smart Storage Battery	
1	724864-B21	HP DL380 Gen9 2SFF Kit	
1	SG506A	HP C13-C14 2.5ft Sgl Data Special	
1	SG508A	HP C13-C14 4.5 Sgl Data Special	
1	768896-B21	HP DL380 Gen9 Rear Serial Cable Kit	
1	726736-B21	HP Smart Array P440ar/2G Controller	

2	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD	For operating system and database catalog
24	718162-B21	HP 1.2TB 6G SAS 10K 2.5in SC ENT HDD	For approximately 13 TB formatted storage capacity
1	647581-B21	HP Ethernet 10GbE 530FLR SFP+ Adapter	For SFP+ networking
2	720479-B21	HP 800W FS Play Ht Plg Pwr Supply Kit	
1	733660-B21	HP 2U SFF Easy Rail Kit	

Alternate Part Appendix

Processor

Quantity	Part Number	Description	Notes
1	719045-L21	HP DL380 Gen9 E5 2697v3 FIO Kit	14 cores at 2.6 GHz
1	719045-B21	HP DL380 Gen9 E5 2697v3 Kit	
1	719044-L21	HP DL380 Gen9 E5 2690v3 FIO Kit	12 cores at 2.6 GHz
1	719044-B21	HP DL380 Gen9 E5 2690v3 Kit	

Memory

Quantity	Part Number	Description	Notes
8	726722-B21	HP 32GB 4Rx4 PC4 2133P-L Kit	For 256 GB of memory
16	726722-B21	HP 32GB 4Rx4 PC4 2133P-L Kit	For 512 GB of memory

Disk Drives

Quantity	Part Number	Description	Notes
24	718162-B21	HP 1.2TB 6G SAS 10K 2.5in SC ENT HDD	For approximately 13 TB formatted storage capacity
24	652583-B21	HP 600GB 6G SAS 10K 2.5in SC ENT HDD	For approximately 6 TB formatted storage capacity

Networking

Quantity	Part Number	Description	Notes
1	647581-B21	HP Ethernet 10GbE 530FLR SFP+ FIO Adapter	For SFP+ networking
1	700699-B21	HP Ethernet 10Gb 2P 561FLR-T Adapter	For RJ-45 networking

Get connected

hp.com/go/getconnected

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

© Copyright 2012, 2013, 2014, 2015 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.