



Hewlett Packard
Enterprise

HPE ConvergedSystem 300 for Microsoft Analytics Platform

Data Integration Platform – Gen9 reference guide for
the AU5 release

Contents

Executive summary.....	3
Overview of services for the Data Integration Platform.....	3
Updates.....	3
Solution overview.....	4
Backup and restore platform.....	4
Data Loading Staging Platform.....	5
Measured performance.....	6
Platform configuration.....	6
Database configuration.....	6
Test details.....	7
Performance results.....	7
Licensing and support considerations.....	8
Licensing considerations.....	8
Support considerations.....	8
General considerations.....	8
Recommended hardware for the Data Integration Platform.....	8
Recommended networking configurations for the Data Integration Platform.....	9
Recommended software for the Data Integration Platform.....	12
Caveats.....	12
Appendix A – Recommended configuration for a Data Integration Platform (with part numbers).....	13
Resources and additional links.....	18

Executive summary

HPE ConvergedSystem 300 for Microsoft Analytics Platform addresses the key needs of business insights through greater understanding of business data. It is pre-configured and the only purpose-built solution customized for Data Warehouse (DW) with the HPE Support Pack to meet a range of business needs. The HPE ConvergedSystem 300 for Microsoft Analytics Platform can be easily and rapidly deployed to support both large relational database solutions with Microsoft® SQL Server Parallel Data Warehouse. All required components, including servers, storage, and networking, are integrated into a single ConvergedSystem that is expandable to meet virtually any big data challenge.

The HPE ConvergedSystem 300 for Microsoft Analytics Platform comes as a self-contained solution. However, there are recommended functions that can be supported with a separate server (or set of servers) to provide optimal backup, restore, and data loading (Extract, Transform, and Load). This paper will review these external capabilities and make configuration recommendations.

It is important that the reader have a good understanding of the HPE ConvergedSystem 300 for Microsoft Analytics Platform solution prior to reading this paper. It is strongly recommended that the reader review the [HPE ConvergedSystem 300 for Microsoft Analytics Platform solution architecture](#) white paper or similar documentation before proceeding.

Target audience: The target audience for this document is IT directors, IT architects, database administrators, business intelligence (BI) administrators, and others wishing to learn more about extending the functionality of the HPE ConvergedSystem 300 for Microsoft Analytics Platform solution. This document assumes the reader has a basic understanding of several key data center technologies including, but not limited to: servers, storage, networking, power, solution management, virtualization, and hypervisors.

Overview of services for the Data Integration Platform

The Data Integration Platform (DIP) refers to one or more servers you may elect to accompany your purchase of an HPE ConvergedSystem 300 for Microsoft Analytics Platform solution. You may also choose to re-purpose an existing server or set of servers in your environment to provide this functionality. The Data Integration Platform server(s) will provide auxiliary support for your system, including support for the following capabilities:

- Intermediate location for backups (and therefore a staging area for restore operations)
- A staging location for Extract, Transform, and Load (ETL) or ELT operations

The Data Integration Platform must be ordered separately from your HPE ConvergedSystem 300 for Microsoft Analytics Platform solution.

Updates

This Data Integration Platform continues to be based on the HPE ProLiant DL380 Gen9 platform, and now features the Intel® Xeon® E5-2600 Version 4 processors.

The HPE ProLiant Gen9 portfolio is a major delivery milestone in HPE's compute strategy, which addresses IT demands with a vast pool of processing resources that can be located anywhere, scaled to any workload and available at all times. The servers are optimized for convergence, cloud and software-defined environments, and feature new technology innovations such as:

- HPE-unique PCIe accelerators and 2400MHz DDR4 SmartMemory that increase compute capacity
- Converged management across servers, storage and networking to enable a software defined enterprise
- Faster setup, monitoring and firmware maintenance with reliable, secure and innovative embedded management, including UEFI and RESTful APIs for hybrid cloud environments
- PCIe Accelerators, HPE SmartCache and HPE FlexFabric adapters that deliver improved performance

Solution overview

The HPE DIP server is a companion product for the HPE ConvergedSystem 300 for Microsoft Analytics Platform that is purchased separately. The DIP server can be customized to fit each customer environment and server configuration. This document provides a tested reference guide for CS300 appliance Backup/Restore/ and Load. The DIP server, while not part of the appliance, can be racked into the appliance rack (if space allows) and is designed to connect to the appliance network switches. This section details the tested configuration for this reference guide.

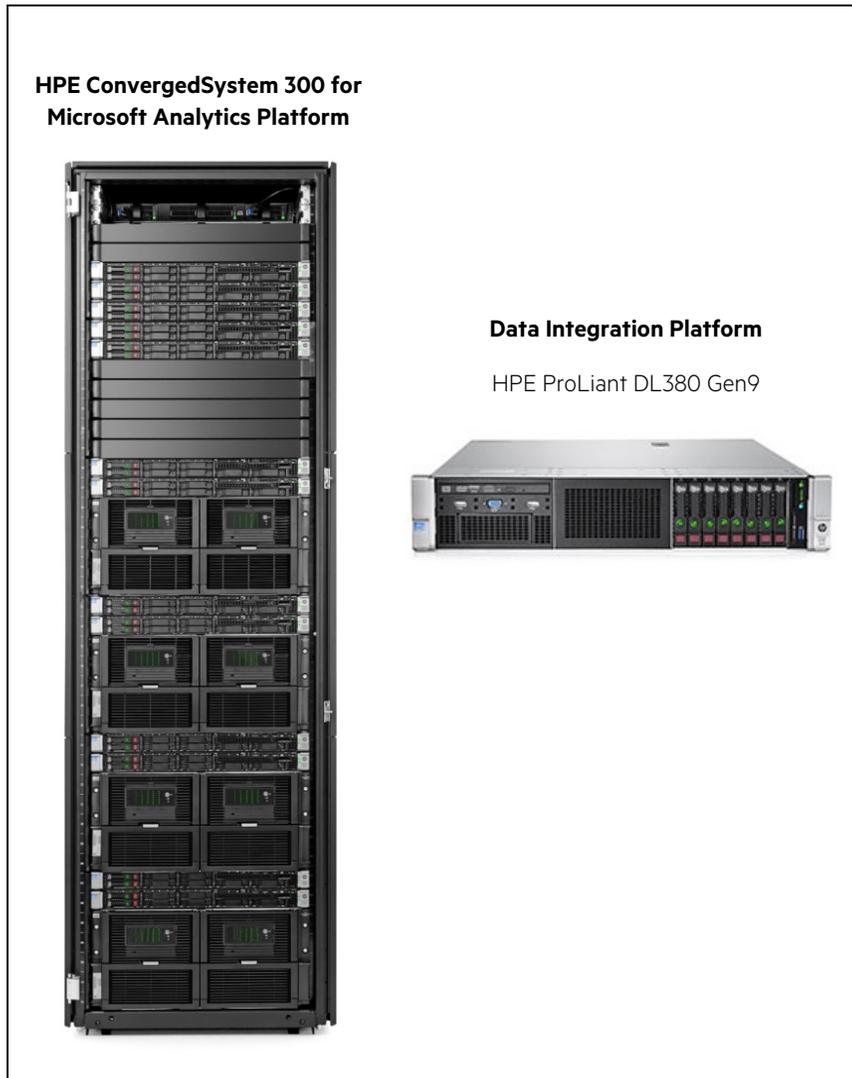


Figure 1. Overview

For additional product information you can read the QuickSpecs available at:
<http://h20195.www2.hp.com/V2/GetDocument.aspx?docname=c04126041>

Backup and restore platform

A key requirement of most database and analytics data storage systems is to have a backup of the data stored on the server. This backup location is typically going to be migrated off-site for extended storage and disaster recovery purposes. However, in the interim, a local copy of the backup is taken for efficiency and performance of the backup itself. The Data Integration Platform is the ideal location as the first storage location for Microsoft Analytics Platform System (APS) backups.

SQL Server Parallel Data Warehouse backup

SQL Server Parallel Data Warehouse (PDW) offers two backup types for your PDW databases. You can either take a full database backup (as it sounds, a copy of everything), or a differential backup, which will contain everything in your database that has changed since the last time a full database backup was taken. A differential backup will depend upon the last full backup for its baseline. For further details, review the APS.chm documentation file.

For each of these backup types, SQL Server Parallel Data Warehouse produces a set of files. These files together, rather than a single file, make up the database backup or differential backup. These backups will always use compression. The BACKUP DATABASE syntax assumes that the backup location will be on disk, via a UNC path to a specific directory to retain the backup (and must NOT be on the system itself). Therefore, a separate backup server is needed, with enough disk storage space to contain the full database backups and differential backups for each user database. Additionally, as with any local backup strategy an off-site or enterprise backup solution should be employed to make secondary (offsite) backups.

Note

Backup sets cannot be restored to a smaller system, so if a disaster recovery strategy involves using backups, the other system being restored to, must have at least the same number of compute nodes.

You should also back up the master database of the control node. Note, however, that a backup of the master database is a regular SQL Server backup of the master database, and must always be a full database backup. The master database will likely not be significant in size, but is an integral part of any recovery operation. It is strongly recommended that the master database backups use a parallel strategy to user database backups in terms of off-site storage, DR recovery planning, etc.

SQL Server Parallel Data Warehouse restore

To restore any individual database, a copy of the last full database backup must be available at an accessible UNC location. Once the full database backup is applied, the latest differential backup can then be applied to bring the database up to the latest point of recoverability. The last differential backup must also be accessible via a UNC share for the restore process.

If the entire system must be restored, the restore of the master database on the control node should be performed, and then the recovery as noted above for each individual user database (via the full and latest differential backup). Each of these backup file sets should be present on the Data Integration Platform for optimal recovery performance.

Using the Data Integration Platform

The recommendation for the backup and restore platform is that the server should be connected to the solution on the InfiniBand network if possible, to maximize the performance of the data movement, and use 10Gb Ethernet to connect to the customer corporate network. The Data Integration Platform as documented in this paper meets these requirements and will provide optimal networking and data transfer capabilities for the solution.

Data archival

It is strongly recommended that the database backups will be copied to a remote storage location (for example, a Disaster Recovery data center). The Data Integration Platform is an ideal location to have the "latest" backups (for faster recovery purposes), but typically does not have enough storage to have multiple copies of database backups. Additionally, for disaster recovery planning, a remote copy of the data is critical in the event of a severe event affecting the primary data center.

Data Loading Staging Platform

SQL Server Parallel Database Warehouse data will typically be loaded into the appropriate database using one of the following mechanisms:

- ETL Software solution, such as SQL Server Integration Services (SSIS)
- DWLoader utility (provided with the HPE ConvergedSystem 300 for Microsoft Analytics Platform solution)
- A custom data loading process incorporating one or more of the above tools
- A third party ETL solution

For most usages of data loading, an interim platform must be used to stage the data before it can be loaded into the SQL Server PDW database. Assuming, as above, that the Data Integration Platform is connected to the solution via the InfiniBand network, data can be transferred and loaded using the high bandwidth InfiniBand interconnects.

The Data Integration Platform is also an ideal location for any workflow or transformation engine(s) needed to support the data loading activities. Note that most data loading activities will not necessarily require significant CPU scalability, but the ETL software may be more CPU intensive.

If being used as a data staging location, focus on the storage capabilities and capacity of the Data Integration Platform storage system. If you are also planning on hosting your ETL engine (such as SSIS) on the server, consider adding additional CPU and memory capacity to the server.

Measured performance

Starting with the AU3 release, performance baselines have been created for the DIP server platform. These baselines will be used to compare performance on future releases of the appliance and the DIP platform. This section details the performance tests that were run as part of the baseline creation.

Platform configuration

Several different platform configurations were used for the performance testing, from 2, 4, 8 and 16 node systems running on HPE Generation 9 (Gen9) hardware with the AU4 software to an 8 node system running with the recently released AU5 software.

16 Node system – The 16 node system used 2 full racks with 16 compute nodes to test a multi-rack configuration.

8 Node system – The 8 node system was comprised of 8 compute nodes, and was used to establish a full rack performance baseline.

4 Node system – The 4 node system tested a ½ rack configuration.

2 Node system – The 2 node system used 2 compute nodes to create a quarter-rack baseline for Gen9/AU4.

Note

8 Node configurations and larger, now feature a “High Performance” drive option that substitutes 15K SAS drives for the standard SAS MDL drives. These drives support faster transfer rates at the expense of lower storage density.

Database configuration

For the SQL server performance tests, a 1TB TPCh database was created on the CS300 APS appliance. This database was set up to use both distributed and replicated tables. For tests that use indexes, indexes were created on the tables using clustered columnstore. For the SQL table scan throughput test, the indexes were deleted. Table 1 below details the settings for the various TPCh database tables.

Table 1. Database Setup

TABLE	DISTRIBUTED OR REPLICATED	INDEX
Lineitem	Distributed	Clustered/Columnstore
Orders	Distributed	Clustered/Columnstore
Partsupp	Distributed	
Part	Replicated	
Supplier	Replicated	
Customer	Replicated	
Region	Replicated	
Nation	Replicated	

Note

The test performed for this performance baseline DO NOT constitute a formal TPCh benchmark. The TPCh database was used to provide test data only.

Test details

Five different tests were performed for the performance baseline. These tests include SQL Server DB administration tasks as well as tests to measure throughput both in RAW form and through SQL table scans.

- **Database load rate:** This test measures the aggregate load performance of the DIP server and the appliance. For this test, unindexed heap tables were loaded simultaneously through both InfiniBand connections. The flat data files are created and stored on the DIP server.
- **Database backup rate:** Overall Backup rate is measured in this test. The 1TB database is backed up through a single InfiniBand connection to a file share on the DIP server. The Backup set size is divided by the time taken to back up the Database to determine the backup rate.
- **Database restore rate:** The Database restore rate is measured by taking the previous backup and restoring it through a single InfiniBand connection. The Backup set size is then divided by the time taken for the restore to complete.
- **SQL Server scan rate:** SQL Scan rate is measured by performing a table scan on a large table. In this case, the index on the lineitem table is deleted, and a “Select count_big (*) from lineitem” is performed to scan the table and count the records. Performance monitor is used to aggregate the I/O rates on each of the 16 LUNs on each HSA storage node.
- **RAW I/O throughput:** This is a non-SQL server test using the Microsoft SQLIO utility to measure raw I/O performance on all LUNs within a scale unit. A data file is created on each LUN, and this file is then read using connection parameters to determine the maximum I/O rate on each scale unit.

Performance results

The performance results are detailed in Table 2.

Table 2. Performance Metrics

PERFORMANCE METRIC	HPE 16 NODE	HPE 8 NODE (AUS)	HPE 4 NODE	HPE 2 NODE	COMMENTS
Database Load Rate (High Performance)	5.2TB/Hr 7.1TB/Hr	5.1TB/Hr 7.0TB/Hr	4TB/Hr N/A	2.1TB/Hr N/A	Loading TPCh data into 8 heap tables concurrently over 2 IB connections
Database Backup Rate (High Performance)	5.1TB/Hr 8.4TB/Hr	5.0TB/Hr 6.6TB/Hr	4.4TB/Hr N/A	2TB/Hr	
Database Restore Rate (High Performance)	4.5TB/Hr 6.6TB/Hr	4.0TB/Hr 5.0TB/Hr	3.7TB/Hr N/A	366GB/Hr N/A	Restoring 1TB TPCh data over IB connections
SQL Server Scan Rate	63GB/Sec	36.8GB/Sec	18.4GB/Sec	9.2GB/Sec	
RAW I/O Throughput	67.6GB/Sec	38.8GB/Sec	19.4GB/Sec	9.7GB/Sec	1 SQLIO Instance for each LUN in an HSA

Licensing and support considerations

The server used as the Data Integration Platform is not part of the HPE ConvergedSystem 300 for Microsoft Analytics Platform. It is essential to understand that the licensing and support considerations for this server and any other ancillary servers are separate from that of the ConvergedSystem.

Licensing considerations

The Data Integration Platform is a completely separate component from your HPE ConvergedSystem 300 for Microsoft Analytics Platform. Therefore, you must purchase licenses for all software that will be used on this server.

The list of software that needs to be licensed will include:

- HPE iLO Advanced 1-Server including 1 yr. 24x7 software support
 - For optimal management of the hardware, HPE iLO is strongly recommended
- Microsoft® Windows Server® 2012 R2 Standard Edition (x64 version)
 - SQL Server 2014 (If SSIS is to be used)

You must either provide your own licenses or purchase these licenses from your authorized reseller (or directly from HPE if ordering a new server).

Support considerations

The Data Integration Platform is **not** part of your HPE ConvergedSystem 300 for Microsoft Analytics Platform support contract with either HPE or Microsoft. You must purchase support separately for all hardware and software components that will be used on the Data Integration Platform.

General considerations

It is highly recommended (though not required) that the Data Integration Platform is connected to the InfiniBand fabric for connectivity to the ConvergedSystem. Using the high speed data interconnect will maximize the performance of the backup, restore, and data loading operations. Any modifications to the Ethernet or InfiniBand switch configurations must only be to the ports that are being used by the Data Integration Platform (e.g., setting VLANs) in the rack and not global switch configuration changes or changes to the ports used by the solution components.

Therefore, it is recommended that the Data Integration Platform is placed in the same rack or in an adjoining rack as the HPE ConvergedSystem 300 for Microsoft Analytics Platform. There is no requirement in terms of location, however if placed within the same rack, we highly recommend that you order a separate PDU power source for the Data Integration Platform. If this is not possible, then there is a risk that any change in power balancing, overloading of PDUs, etc. could impact the system and impact the power availability envelopes that have been designed in the rack. Secondly, there is a self-imposed limitation to upgrading within the HPE ConvergedSystem 300 for Microsoft Analytics Platform rack that could require removing the Data Integration Platform from the rack, if future expansion is desired or required.

Recommended hardware for the Data Integration Platform

HPE recommends the following hardware solution and configuration for a Data Integration Platform:

For a minimum configuration (1/4 rack which has 1 data scale unit):

- One (1) HPE ProLiant DL380 Gen9 (2x E5-2667 v4) Server including:
 - 64GB RAM (Eight (8) HPE 8GB (1R x4) PC4-2400R-40 Memory)
 - Twelve (12) HPE 1TB 6G SAS 7.2K 3.5in
 - HPE IB FDR/EN 10/40Gb 2P 544+QSFP Adapter
 - HPE P440ar/2GB FBWC 2P SAS Cntrl
 - HPE SAS Expander
 - 2x HPE IB FDR QSFP Copper Cables (to connect to the InfiniBand Switches)
 - Appropriate Ethernet Networking Components ([see below](#))

See [Appendix A](#) for complete information and part numbers.

If additional storage is required, you should select the HPE D6000 Disk Enclosure with the following options:

- For ½ rack with 2 data scale units or ¾ rack with 3 data scale units:
 - Twenty-four (24) HPE 1TB 6G SAS 7.2K 3.5in
 - 1x HPE Smart Array P841/4GB FBWC Controller
 - 1x HPE D6000 Disk Enclosure
 - 1x HPE D6000 Dual I/O Module Option Kit
 - 4x HPE Ext Mini SAS HD 1m Cable
- For full rack with 3 data scale units:
 - Forty-eight (48) HPE 1TB 6G SAS 7.2K 3.5in

OR

 - Forty-eight (48) HPE 600GB 6G SAS 15K 3.5in **(High Performance Option)**
 - 1x HPE Smart Array P841/4GB FBWC Controller
 - 1x HPE D6000 Disk Enclosure
 - 1x HPE D6000 Dual I/O Module Option Kit
 - 4x HPE Ext Mini SAS HD 1m Cable
- For Configurations larger than a full rack (4+ scale units):
 - Seventy (70) HPE 1TB 6G SAS 7.2K 3.5in

OR

 - Seventy (70) HPE 600GB 6G SAS 15K 3.5in **(High Performance Option)**
 - 1x HPE Smart Array P841/4GB FBWC Controller
 - 1x HPE D6000 Disk Enclosure
 - 1x HPE D6000 Dual I/O Module Option Kit
 - 4x HPE Ext Mini SAS HD 1m Cable

If the storage space required exceeds the estimated disk storage needs of current forecasts, consider ordering 2, 3, 4 or 6TB disks instead. If selecting larger drives, ensure that all the drives are of equal size to eliminate wasted drive space.

Recommended networking configurations for the Data Integration Platform

The default (and **strongly** recommended) configuration for connectivity to the HPE ConvergedSystem 300 for Microsoft Analytics Platform solution is to connect directly to the InfiniBand switches for high data throughput for backup, and loading performance.

If connecting directly to the InfiniBand network is not an option, the secondary strategy would be to connect via the 10Gb Ethernet network interface. Again, a direct connection to the switches is strongly recommended.

As a last (and not recommended) option, connectivity can be provided by connecting to the 1Gb management network of the system directly.

Table 3 below summarizes the connectivity options to the CS300 appliance.

Table 3. Connectivity overview

CONNECTIVITY TO THE HPE CONVERGEDSYSTEM 300 FOR MICROSOFT ANALYTICS PLATFORM

InfiniBand Connection x 2 (one to each switch)

This is the recommended configuration

10Gb Ethernet Connection x 2 (one to each switch)

1Gb Ethernet Connection x 2 (one to each switch)

This is the least recommended configuration

Additionally, direct connectivity to the customer network over either 10Gb or 1Gb Ethernet from the Data Integration Platform is the recommended configuration. This will provide the fastest data transfer, loading, and backup speeds.

The recommended InfiniBand connections (and cabling) are built into the parts list above. If the connectivity to the customer network is connecting via 10Gb Ethernet, then you would add a 10Gb Ethernet network card into the order (and the necessary cables).

Note that the cable length for the cables will depend upon the location of the solution and distance to customer connectivity endpoints. Table 4 below depicts the connection options to the customer’s network.

Table 4. Customer network Connectivity

CONNECTIVITY TO THE CUSTOMER NETWORK

Redundant (or LACP) connectivity via 10Gb Ethernet

This is the recommended configuration

A single 10Gb connection

Redundant (or LACP) connectivity via 1Gb Ethernet

A single 1Gb connection is also an option

Network connectivity specifics

For the InfiniBand network, Figure 2 shows the current network port usage. Any of the ports in green (19-30) are available to connect to the system. Ensure that you use IP addresses that are in the correct (customer-specified) ranges for each switch, and verify that the addresses are not currently in use.

Color key		IB cable lengths				Ethernet cable lengths																	
10 GbE Uplink		5M	2M			12M																	
Unused		1M	3M			4ft	7ft																
User available																							
Reserved																							
InfiniBand Switch #1 (R1BSW1) - U42																							
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35						
R1ETHSW1 U40	R1BSW2 U41	HST01 U38	HSA02 U7	HSA04 U14	HSA06 U21	HSA08 U28									R2BSW1 U42							R1ETHSW1 U40	
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36						
	HST03 U34	HST02 U37	HSA01 U6	HSA03 U13	HSA05 U20	HSA07 U27									R2BSW2 U41								
InfiniBand Switch #2 (R1BSW2) - U41																							
1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35						
R1ETHSW2 U39	R1BSW1 U42	HST01 U38	HSA02 U7	HSA04 U14	HSA06 U21	HSA08 U28									R2BSW1 U42							R1ETHSW2 U39	
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36						
	HST03 U34	HST02 U37	HSA01 U6	HSA03 U13	HSA05 U20	HSA07 U27									R2BSW2 U41								

Figure 2. The InfiniBand switches cable wiring in an HPE ConvergedSystem 300 for Microsoft Analytics Platform (Rack #1)

For the Ethernet network, Figure 3 shows the current 1/10/40Gb network port usage. The green port (#23) is available to connect to the system (1Gb Ethernet). Ensure that you use IP addresses that are in the correct (customer-specified) ranges for each switch, and verify that the addresses are not currently in use.

Color key		IB cable lengths			Ethernet cable lengths																	
10 GbE Uplink		.5M	2M		1.2M																	
Unused		1M	3M		4ft	7ft																
User available																						
Reserved																						
Ethernet Switch #1 (R2ETHSW1) - U40 HPE 5900-48G-4XG-2QSFP+ (JG510A)																						
1 GbE Ports (1-48)														10 GbE Ports (49-52)		40 GbE Ports (53-54)						
1	3	5	7	9	11	13	15	17	19	21	23	25 to 43	45	47	49	51	53					
HST03	HST02	HSA01	HSA03	HSA05	HSA07		HST03 iLO	HSA04 iLO	HSA08 iLO							R1BSW1 Port1 U42	R1ETHSW2 Port 53 U39					
U34	U37	U6	U13	U20	U27		U34	U14	U28													
2	4	6	8	10	12	14	16	18	20	22	24	26 to 44	46	48	50	52	54					
R1PDU1	HST01	HSA02	HSA04	HSA06	HSA08		HST01 iLO	HSA02 iLO	HSA06 iLO	R1BSW1 U42				R1ETHSW2 Port 48 U39			R1ETHSW2 Port 54 U39					
U38	U7	U14	U21	U28			U38	U7	U21													
Ethernet Switch #2 (R2ETHSW2) - U39 HPE 5900-48G-4XG-2QSFP+ (JG510A)																						
1 GbE Ports (1-48)														10 GbE Ports (49-52)		40 GbE Ports (53-54)						
1	3	5	7	9	11	13	15	17	19	21	23	25 to 43	45	47	49	51	53					
HST03	HST02	HSA01	HSA03	HSA05	HSA07			HSA03 iLO	HSA07 iLO							R1BSW2 Port1 U41	R1ETHSW1 Port 53 U40					
U34	U37	U6	U13	U20	U27			U13	U27													
2	4	6	8	10	12	14	16	18	20	22	24	26 to 44	46	48	50	52	54					
R1PDU2	HST01	HSA02	HSA04	HSA06	HSA08		HST02 iLO	HSA01 iLO	HSA05 iLO	R1BSW2 U41				R1ETHSW1 Port 48 U40			R1ETHSW1 Port 54 U40					
U38	U7	U14	U21	U28			U37	U6	U20													

Figure 3. The Ethernet switches cable wiring in an HPE ConvergedSystem 300 for Microsoft Analytics Platform [Rack #1]

For the 10Gb Ethernet network, Figure 4 shows the current Multi-Rack 10Gb connectivity. Again, the green “customer network” ports are available for 10Gb connectivity, if any aren’t being used for rack connectivity to the customer network. Ensure that you use IP addresses that are in the correct (customer-specified) ranges for each switch, and verify that the addresses are not currently in use.

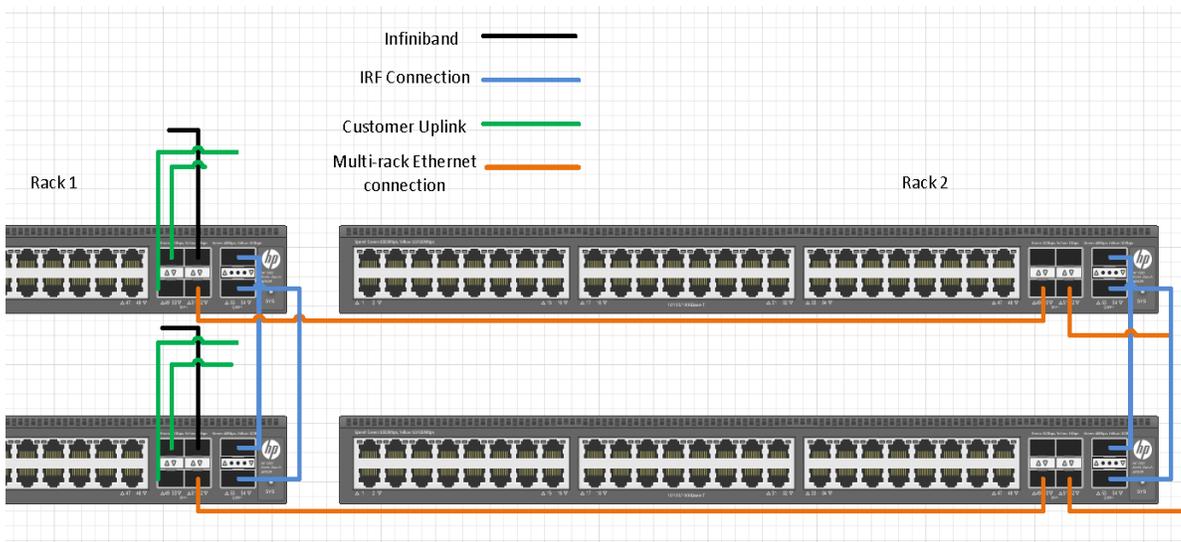


Figure 4. 10Gb Ethernet Connectivity for the HPE ConvergedSystem 300 for Microsoft Analytics Platform

Recommended software for the Data Integration Platform

HPE recommends the following software for a Data Integration Platform:

- HPE iLO Advanced 1-Server including 1 yr. 24x7 software support
 - For optimal management of the hardware, HPE iLO is strongly recommended
- Windows Server 2012 R2 Standard Edition (x64 version)
 - Windows Server should be purchased/licensed as part of the purchase from HPE. However, this version of Microsoft Windows Server is specifically requested because of the support for the InfiniBand firmware and drivers.
 - Enable the Application Server and File and Storage Services Roles
 - Enable Remote Desktop Services
- Version 4.85 or newer of the WinOF Driver for the HPE InfiniBand FDR/EN 10/40Gb Dual Port 544+QSFP Adapter
 - 5.10 is the most current version (as of press time) for the 544+QSFP adapter.
 - The driver is named: MLNX_VPI_WinOF-5_10_All_win2012R2_x64.exe
 - Any version from version 4.85 forward can be used.
- DWLoader.exe
 - This is the command-line data loading software that is included with your HPE ConvergedSystem 300 for Microsoft Analytics Platform purchase. The software is on the system in the ClientTools.msi file.
- SQL Server Data Tools
 - The SQL Server Data Tools are the primary SQL interface to the Parallel Data Warehouse software.
 - Only the Visual Studio 2012 version of the SQL Server Data Tools is supported with Parallel Data Warehouse.
 - The download is at: <http://msdn.microsoft.com/en-us/jj650015>

Note

The Data Integration Platform does not integrate into the domains on the HPE ConvergedSystem 300 for Microsoft Analytics Platform. The Data Integration Platform should either belong to a customer domain or be a stand-alone workgroup server.

Caveats

The hardware configuration for the HPE ConvergedSystem 300 for Microsoft Analytics Platform is preconfigured at the factory, and any topology changes need to be made by Microsoft or HPE-authorized personnel.

The customer may deploy the Data Integration Platform in the HPE ConvergedSystem 300 for Microsoft Analytics Platform rack provided it follows the recommendations in this white paper.

- Configurations outside of the recommendations in this white paper may impact core services such as performance, security, and high availability, and are not supported.
- The Data Integration Platform is not supported as part of the HPE ConvergedSystem 300 for Microsoft Analytics Platform, and needs to be supported with separate support and condition. Any issues would not flow through the standard HPE/Microsoft collaborative support model. HPE and Microsoft reserve the right to request that any additional equipment be shut down during troubleshooting if there is a possible correlation or potential issue that is causing problems with the HPE ConvergedSystem 300 for Microsoft Analytics Platform functionality.
- Any modifications to the Ethernet or IB switch configurations must only be to the ports that are being used by the additional equipment (e.g., setting VLANs) in the rack and not global switch configuration changes or changes to the ports used by appliance components.

Appendix A – Recommended configuration for a Data Integration Platform (with part numbers)

This appendix documents the specific parts list to order a “Data Integration Platform” for your HPE ConvergedSystem 300 for Microsoft Analytics Platform solution.

Note

This configuration does not ship with any preconfigured software or with an OS installed. Any software or configuration installation is performed on-site defined by the customer requirements. It is strongly recommended that you order the desired version of Windows Server with the hardware order.

Note

Part numbers are at time of publication and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your HPE Reseller or HPE Sales Representative for more details. hpe.com/us/en/services/consulting.html

Two BOMs are presented below. The first BOM is used when connecting the DIP server to a customer network through 1Gb Ethernet. The second BOM is required when connecting through 10Gb Ethernet.

Table 5. DIP Server BOM (1GbE Connectivity)

HPE Data Integration Platform (1GbE Customer Connectivity)

QTY	PART NUMBER	DESCRIPTION
1	719061-B21	HPE ProLiant DL380 Gen9 CTO Chassis (12LFF)
1	817947-L21	Eight-Core Intel Xeon Processor Model E5-2667v4 (3.2 GHz, 25MB Level 3 Cache, 135W)
1	817947-B21	Additional E5-2667v4 CPU Kit
1	749974-B21	HPE Smart Array P440ar 12Gb 2-port FIO adptr
1	727250-B21	HPE 12Gb SAS Expander Card w/cables
1	719073-B21	3 Slot Expansion
1	665240-B21	HPE 366FLR 4 port FlexLom Adapter
1	764284-B21	HPE IB FDR 2-port 544+QSFP Adapter
8	805347-B21	8 GB Reg DDR4-2400R CAS 17 Kit
2	720479-B21	HPE 800W Flexslot Platinum HPE Power Supply
1	512485-B21	One (1) HPE iLO Adv 1-Svr incl 1yr TS&U SW

OR (For customers requiring 10GbE customer connectivity instead of 1GbE)

Table 6. DIP Server BOM (10GbE Connectivity)

HPE Data Integration Platform (10GbE Customer Connectivity)

QTY	PART NUMBER	DESCRIPTION
1	719061-B21	HPE ProLiant DL380 Gen9 CTO Chassis (12LFF)
1	817947-L21	Eight-Core Intel Xeon Processor Model E5-2667v4 (3.2 GHz, 25MB Level 3 Cache, 135W)
1	817947-B21	Additional E5-2667v4 CPU Kit
1	749974-B21	HPE Smart Array P440ar 12Gb 2-port FIO adptr
1	727250-B21	HPE 12Gb SAS Expander Card w/cables
1	719073-B21	3 Slot Expansion
1	727054-B21	HPE 562FLR-SFP+ 2 port FlexLom Adapter
1	764284-B21	HPE IB FDR 2-port 544+QSFP Adapter
8	805347-B21	8 GB Reg DDR4-2400R CAS 17 Kit
2	720479-B21	HPE 800W Flexslot Platinum HPE Power Supply
1	512485-B21	One (1) HPE iLO Adv 1-Svr incl 1yr TS&U SW

Drives

If **¼ rack** (2 node) configuration: Twelve (**12**) HPE 6G SAS 7.2K 3.5in (Select 1, 2, 3, 4, or 6TB)

Table 7. Quarter Rack Internal Drives

QTY	PART NUMBER	DESCRIPTION
12	652753-B21	HPE 1TB 6G SAS 7.2K 3.5in SC MDL HDD
12	652757-B21	HPE 2TB 6G SAS 7.2K 3.5in SC MDL HDD
12	652766-B21	HPE 3TB 6G SAS 7.2K 3.5in SC MDL HDD
12	695510-B21	HPE 4TB 6G SAS 7.2K 3.5in SC MDL HDD
12	761477-B21	HPE 6TB 6G SAS 7.2K 3.5in SC MDL HDD

If the solution is **larger than ¼ rack** then add the following components **Plus disks (either (a) or (b) below)**:

Table 8. External Storage Hardware

QTY	PART NUMBER	DESCRIPTION
1	726903-B21	HPE Smart Array P841/4GB FBWC Controller
1	QQ695A	HPE D6000 Disk Enclosure
1	QQ696A	HPE D6000 Dual I/O Module Option Kit
4	716189-B21	HPE Ext Mini SAS HD 1m Cable

- a. If $\frac{1}{2}$ rack (4 nodes) or $\frac{3}{4}$ rack (6 nodes) configuration: **Twenty-four (24)** HPE 6G SAS 7.2K 3.5in (Select 1, 2, 3, 4 or 6TB)

Table 9. External Drives [Half Rack to Three Quarter Rack]

QTY	PART NUMBER	DESCRIPTION
24	507614-B21	HPE 1TB 6G SAS 7.2K 3.5in SC MDL HDD
24	507616-B21	HPE 2TB 6G SAS 7.2K 3.5in SC MDL HDD
24	625031-B21	HPE 3TB 6G SAS 7.2K 3.5in SC MDL HDD
24	693689-B21	HPE 4TB 6G SAS 7.2K 3.5in SC MDL HDD
24	782669-B21	HPE 6TB 6G SAS 7.2K 3.5in SC MDL HDD

- b. If **full rack** (8 nodes) configuration: **Forty-eight (48)** HPE 6G SAS 7.2K 3.5in (Select 1, 2, 3, 4 or 6TB)

Table 10. External Drives [Full Rack]

QTY	PART NUMBER	DESCRIPTION
48	507614-B21	HPE 1TB 6G SAS 7.2K 3.5in SC MDL HDD
48	507616-B21	HPE 2TB 6G SAS 7.2K 3.5in SC MDL HDD
48	625031-B21	HPE 3TB 6G SAS 7.2K 3.5in SC MDL HDD
48	693689-B21	HPE 4TB 6G SAS 7.2K 3.5in SC MDL HDD
48	782669-B21	HPE 6TB 6G SAS 7.2K 3.5in SC MDL HDD

OR

For the High Performance Storage Option: **Forty-eight (48)** HPE 6G SAS 15K 3.5in:

Table 11. External Drives [Full Rack - High Performance]

QTY	PART NUMBER	DESCRIPTION
48	516828-B21	HPE 600GB 6G SAS 15K 3.5in Ent HDD

- c. If Larger than a full **full rack** (10+ node) configuration: **Seventy (70)** HPE 6G SAS 7.2K 3.5in (Select 1, 2, 3, 4 or 6TB)

Table 12. External Drives [Full Rack+]

QTY	PART NUMBER	DESCRIPTION
70	507614-B21	HPE 1TB 6G SAS 7.2K 3.5in SC MDL HDD
70	507616-B21	HPE 2TB 6G SAS 7.2K 3.5in SC MDL HDD
70	625031-B21	HPE 3TB 6G SAS 7.2K 3.5in SC MDL HDD
70	693689-B21	HPE 4TB 6G SAS 7.2K 3.5in SC MDL HDD
70	782669-B21	HPE 6TB 6G SAS 7.2K 3.5in SC MDL HDD

OR

For the High Performance Storage Option: **Seventy (70)** HPE 6G SAS 15K 3.5in:

Table 13. External Drives (Full Rack+ - High Performance)

QTY	PART NUMBER	DESCRIPTION
70	516828-B21	HPE 600GB 6G SAS 15K 3.5in Ent HDD

Please be sure and add the necessary cables for network connectivity to the InfiniBand and Ethernet switches in the solution, and select the cables with the appropriate length. Use the following tables when selecting InfiniBand, 10Gb Ethernet and 1Gb Ethernet cables.

Table 14. InfiniBand Cables

InfiniBand cables

PART NUMBER	DESCRIPTION
670759-B21	HPE 0.5M IB FDR QSFP Copper Cable
670759-B22	HPE 1M IB FDR QSFP Copper Cable
670759-B23	HPE 15M IB FDR QSFP Copper Cable
670759-B24	HPE 2M IB FDR QSFP Copper Cable
670759-B25	HPE 3M IB FDR QSFP Copper Cable
670760-B22	HPE 5M IB FDR QSFP Optical Cable
670760-B23	HPE 7M IB FDR QSFP Optical Cable

Table 15. 10Gb Ethernet Cables

10Gb Ethernet cables

PART NUMBER	DESCRIPTION
JD096C	HPE X240 10G SFP+ SFP+ 1.2m DAC Cable
JD097C	HPE X240 10G SFP+ SFP+ 3m DAC Cable
JG081C	HPE X240 10G SFP+ SFP+ 5m DAC Cable
JC784C	HPE X240 10G SFP+ 7m DAC Cable

Table 16. 1Gb Ethernet Cables

1Gb Ethernet cables

PART NUMBER	DESCRIPTION
C7533A	HPE Ethernet 4ft CAT5e RJ45 M/M Cable
C7535A	HPE Ethernet 7ft CAT5e RJ45 M/M Cable
C7536A	HPE Ethernet 14ft CAT5e RJ45 M/M Cable
C7537A	HPE Ethernet 25ft CAT5e RJ45 M/M Cable

The transceivers must also be ordered for the connections to switches. You will need (at a minimum) two 10Gb Ethernet transceivers to connect to the 10Gb ports on the HPE 5900 Ethernet Switches.

Supported transceivers include:

Table 17. 10Gb Transceivers

HPE 5900 10Gb Ethernet transceivers

PART NUMBER	DESCRIPTION
JG234A	HPE X130 10G SFP+ LC ER 40km Transceiver
JD092B	HPE X130 SFP+ LC SR Transceiver
JD093B	HPE X130 SFP+ LC LRM Transceiver
JD094B	HPE X130 SFP+ LC LR Transceiver
JD095C	HPE X240 10G SFP+ SFP+ 0.65m DAC Cable
JD096C	HPE X240 10G SFP+ SFP+ 1.2m DAC Cable
JD097C	HPE X240 10G SFP+ SFP+ 3m DAC Cable
JC784C	HPE X240 10G SFP+ 7m DAC Cable
JG081C	HPE X240 10G SFP+ SFP+ 5m DAC Cable

Resources and additional links

HPE ConvergedSystem 300 for Microsoft Analytics Platform:
hpe.com/products/cs300aps

HPE ConvergedSystem 300 for Microsoft Analytics Platform System:
Solution architecture and reference guide for the AU5 release:
<http://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=4AA5-9215ENW>

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.



Sign up for updates

★ Rate this document



© Copyright 2014–2016 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for HPE 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. HPE shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft and Windows Server are trademarks of the Microsoft group of companies. Intel and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

4AA5-4660ENW, April 2016, Rev. 7