HPE ProLiant DL580 Gen9 and HPE PCIe LE Workload Accelerator 120TB Data Warehouse Fast Track Reference Architecture

Based on the SQL Server 2016 Data Warehouse Fast Track (DWFT) Reference Architecture

SanDisk®
a Western Digital brand

Table of Contents

Executive Summary

This guide details the server, storage and software configurations for the HPE ProLiant DL580 Gen9 with Fusion ioMemory™ SX350 PCIe Application Accelerator devices.

This document is for individuals (BI Architects, DBA's, Report-Developers, and IT Directors) involved in decision making who are looking for guidance when designing enterprise, business-intelligence applications.

The Microsoft® SQL Server® Data Warehouse Fast Track (DWFT) reference architecture is designed to eliminate the complexity of properly sizing hardware, which helps reduce unnecessary scale-out of storage and servers. The sizing techniques used in the SQL Server DWFT will properly size servers, based on I/O and CPU consumption. This consumption-based approach ensures your data warehouse can fully take advantage of your hardware investment.

Fusion ioMemory products from SanDisk® provide the following significant benefits for data warehouses deployed on SQL Server 2016:

- I/O performance and resiliency
- Lower capital cost
- Lower operational costs
- Simplified management
- Predictable maintenance

Fusion ioMemory devices use flash as if it were memory instead of like disk. This gives applications native access to data, delivering the lightning-fast response times that businesses depend on today.

Industries such as banking, social media, retail, transportation, healthcare, security, entertainment, and research and development, all benefit greatly from this solution. Service levels are increased, while complex and expensive storage sprawl is significantly reduced.

In addition, the Fusion ioMemory platform reduces energy consumption and total cost of ownership. The Fusion ioMemory platform integrates hardware and software to overcome the limitations of legacy architectures and specialized hardware.

Fusion ioMemory solutions significantly increase datacenter efficiency, while delivering enterprise-grade performance, reliability, availability, and manageability.

IT managers, database architects, and CTOs looking to explore and deploy data warehouses and BI applications now have added leverage – they can take advantage of Fusion ioMemory engineering, integration and optimization to quickly build and deploy their next data warehouses.

About the HP Enterprise ProLiant DL580 Gen9 Server

The HPE ProLiant DL580 Gen9 Server is the HPE four-socket (4S) enterprise standard x86 server, offering commanding performance, rock-solid reliability and availability, and compelling consolidation and virtualization efficiencies. Supporting Intel® Xeon® E7-4800/8800 v3 processors, the DL580 Gen9 Server offers enhanced processor performance, up to 6TB of memory, greater IO bandwidth (9 PCIe Gen 3.0 slots), and 12Gb/s performance for SAS devices.

The DL580 Gen9 Server has security and data protection features for system resiliency that your business can depend on. This makes it ideal for mission-critical enterprise, business intelligence, and database applications.

The DL580 Gen9 Server incorporates the intelligence and simplicity of automated management with HPE OneView and HPE iLO 4. This enables your business to improve agility and lower the cost of infrastructure management, for highly virtualized or cloud-based deployments.

What's New

- 64GB LRDIMM support for up to 6TB max of memory

- Support for up to five HPE NVMe SSDs, up to 1.6TB each

- Improved GPU support with the new NVIDIA Quadro M6000 GPU Module

- Support for HPE OneView for HPE ProLiant (DL) Rack Gen9 Servers, which provides infrastructure management that reduces complexity with automation simplicity.

Fusion ioMemory SX350 PCIe Application Accelerator
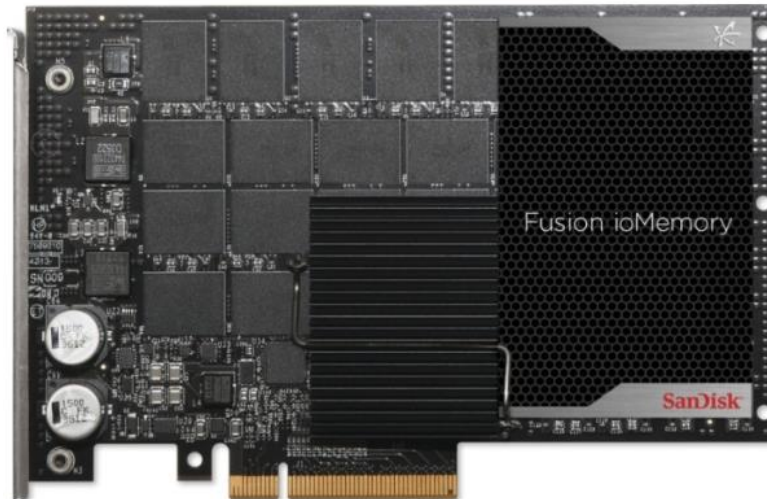
The Fusion ioMemory SX350 PCIe application accelerator is the scalable capacity leader for PCIe flash solutions. The Fusion ioMemory SX350 PCIe series provides a cost-effective solution for read-intensive application workloads that include web hosting, data mining, seismic data processing, content caching, 3D animation, and CAD/CAM.

The Fusion ioMemory SX350 PCIe application accelerator is available in capacities from 1.25TB–6.4TB, with ultra-low 75μs/15μs read/write data access latency, superior reliability, and outstanding random read/write performance of up to 225K/385K IOPS. It also provides updated VSL™ (Virtual Storage Layer) software that delivers direct memory access, minimizes latency, and maximizes application throughput.

With this significant performance improvement, customers can reduce infrastructure and reduce power and cooling costs over a traditional hard disk drive infrastructure, for a lower total cost of ownership (TCO). With over 7,000 customers and over 250,000 units sold, this latest generation of PCIe application accelerators is designed to provide customers with the peace of mind that these products will perform in the field as intended.

About the Data Warehouse Fast Track Reference Architecture

The SQL Server Data Warehouse Fast Track reference architecture provides a scalable framework, based on balancing I/O, to achieve maximum performance from SMP-based servers. SQL Server Data Warehouse Fast Track eliminates the complexity of sizing servers with data warehouses by providing a set of data consumption rates that properly balances performance between the disk subsystem, CPU, and memory.

More information on SQL Server DWFT can be found here:

http://www.microsoft.com/en-us/server-cloud/data-warehouse-fast-track.aspx

Microsoft SQL Server 2016: Data Warehousing with Improved Column Store Technology

Microsoft SQL Server 2016 has made significant improvements in data warehousing technologies and performance, including column-store features as well as many other improvements.

Column-store indices offer great advantages over traditional row stores for analytics and data warehousing queries. They are ideally suited for the star schemas, and tables with billions of rows which are commonly seen. Among their advantages for analytics are:

Up to 10X compression in data size
Data warehouses are very large by nature, and the compression offered by column store index technologies offers both space and cost savings, but also significantly increased performance, due to the dramatically reduced IO requirements given by the compression, coupled by the ability to only scan the specific columns required by each query. This compression also reduces the amount of memory required to hold a given number of rows from the source data warehouse.

Additional Indices
SQL Server 2016 adds the capability to add additional (B-Tree) indices to column store-based tables, which enables efficient single-row lookup.

In addition to these architectural features, we have further optimized the processing of queries in column-store indices in the following ways:

Operator Pushdown
Pushdown refers to moving both filter and aggregation query operations closer to the data, so that many of the filters and calculations can be done in the scan operators, dramatically reducing the volume of data which needs to be handled further on in query processing.

Batch Mode Processing
SQL Server 2016 includes enhancements in batch-mode processing which processes many rows at a time rather than serially doing calculations on each individual row. These batch operations are further optimized by leveraging Single Instruction Multiple Data (SIMD) vector processing CPU instructions in the Intel® architectures.

Reference Architecture

The following shows the configuration details for the HP Enterprise ProLiant DL580 Gen9 with Lightning Ascend™ Gen. II SAS SSDs, as part of the Fast Track Data Warehouse architecture.

Server	HPE ProLiant DL580 Gen9
Operating System	Microsoft Windows Server 2012 R2 Standard Edition
CPU	Intel Xeon E7-8890 v3 @ 2.5 GHz (4S/72C/144T)
PCI-E Slots	Nine FL/FH PCIe 3.0 slots
Drives	2 x 1.2TB SAS (OS)
RAM	2048GB

Server UEFI Configuration

- Hyper-Threading was enabled.
- Operating mode was changed to "Maximum Performance".
- Fan Offset was set to "Increased Cooling".

General Settings

Operating System Settings

The operating system used for this Fast Track Data Warehouse test was Microsoft Windows Server 2012 R2 Standard Edition. Standard installation steps were used to install the operating system with default values, followed by service packs and update patches.

Windows Configuration – Power Settings

The High Performance plan was chosen to reduce CPU throttling.

Fusion ioMemory SX350 Settings

Each device was formatted to its default capacity with a block size of 512B. The raw disk was then used as a mount point, and an NTFS file system was applied using the default space and geometry, with an allocation unit size of 4096KB.

Power Override

Enabling the power override setting on the Fusion ioMemory SX350 product line allows the device to draw up to 55 watts of power under heavy workloads and is required to achieve the performance results below. All cards must be configured at the same time, and a server reboot is required for the setting to be active and persist.

Example:

```
fio-config –p FIO_EXTERNAL_POWER_OVERRIDE SN:MW
```

where $<$ `SN` $>$ is the serial number of the card obtained from `fio-status`, and $<$ `MW` $>$ is the power in milliwatts. The following example configures the device with the given serial number to 55 W:

```
fio-config –p FIO_EXTERNAL_POWER_OVERRIDE 1410G0092:55000
```

Multiple cards must be configured with the same command:

```
fio-config -p FIO_EXTERNAL_POWER_OVERRIDE 1407G0327:55000,1504G0154:55000
```

Storage Configuration

PCIe	CPU	Device	Capacity	Mount Point	Allocation	Notes
1	4	Fusion ioMemory SX350	6.4TB	C:\DB\DATA1	Data Files \ TempDB	JBOD
2	4	Fusion ioMemory SX350	6.4TB	C:\DB\DATA2	Data Files \ TempDB	JBOD
3	3	Fusion ioMemory SX350	6.4TB	C:\DB\DATA3	Data Files \ TempDB	JBOD
4	3	Fusion ioMemory SX350	6.4TB	C:\DB\DATA4	Data Files \ TempDB	JBOD
5	3	Fusion ioMemory SX350	6.4TB	C:\DB\DATA5	Data Files \ TempDB	JBOD
6	2	Fusion ioMemory SX350	6.4TB	C:\DB\DATA6	Data Files \ TempDB	JBOD
7	2	Fusion ioMemory SX350	6.4TB	C:\DB\DATA7	Data Files \ TempDB	JBOD
9	1	Fusion ioMemory SX350	6.4TB	C:\DB\DATA8	Data Files \ TempDB	JBOD
N/A	N/A	2 x Lightning II Eco SAS SSD	1.6TB	C:\DB\LOGS	SQL Logs	Mirrored
N/A	N/A	2 x 10K SAS	1TB	C:\	OS	Mirrored

SQL SERVER SETTINGS

Database Configuration

A 2TB data warehouse schema was created for benchmarking, using the Fast Track toolkit. The schema used a master file group with 8 additional file groups.

TempDB Configuration

In total, eight 20GB tempdb files were created and stored on the volumes designated for data files. The tempdb transaction log file was stored on the volume designated for log files.

Log File Configuration

Two Lightning Eco Gen. II 12Gb/s SAS SSDs were used in a mirrored configuration for all SQL log files. These drives offer solid I/O cost-to-performance benefits with low-latency data access for high quality of service (QoS). For more information, see the following datasheet:

https://www.sandisk.com/content/dam/sandisk-main/en_us/assets/resources/enterprise/data-sheets/lightning-eco-genII-sas-ssd-datasheet.pdf

Memory Allocation

SQL Server was allocated 236GB of the available server memory (2048GB). This amount is allocated as part of the Fast Track test criteria to drive peak backend disk activity during the Row Store query runs.

Local Security Policy

The SQL Server maintenance account was granted the following privileges:

- Enable Lock Pages in Memory - prevents SQL Server buffer pool pages from paging out.
- Perform Volume Maintenance Tasks - enables Instant File Initialization.

SQL Server 2016 Configuration Parameters

Parameter	Setting	Description
Memory Allocation	236GB	This is the Fast Track-required value for a 4-socket, 2TB database. Memory is deliberately constrained to enforce I/O pressure on the subsystem.
Max Degree of Parallelism **Row Store**	72	When SQL Server runs on a computer with more than one microprocessor or CPU, it detects the best degree of parallelism (the number of processors that are used in the execution of a parallel plan).
Max Degree of Parallelism **Column Store**	144	
Resource Governor Memory Allocation **Row Store**	12%	The default is 25%. This is reduced to 12% for Row Store to reduce the maximum memory consumed per query.
Resource Governor Memory Allocation **Column Store**	25%	This is set at the default value.
Fast Track Required Startup Parameters	-T1117	This trace flag ensures even growth of all files in a file group in case `autogrow` is enabled. The standard FTDW recommendation for database file growth is to pre-allocate rather than use `autogrow` (with the exception of tempdb).
Optional Startup Parameters	-T1118	This flag helps alleviate allocation bit map contention in tempdb by switching allocations to full extents (8 physically contiguous pages, or 64KB).
Startup Parameters **Row Store only**	-T834	When this flag is set, SQL Server uses Windows large-page memory allocations for the buffer pool. This trace flag can improve throughput rates for many data warehousing workloads. This value is disabled for Column Store runs.

Measured Performance

During Fast Track Database Validation, Microsoft's Reference Point tool drives multiple concurrent query workloads designed to identify bottlenecks. The tool establishes the key performance metrics in the table below.

Scan Rate Type	Metric	Description
Rated User Capacity (TB)	120TB	Represents the optimal Fast Track-certified data capacity of the configuration.
		Allows for 5:1 compression with 10% recommended free space. 25% capacity is reserved for tempdb, while some memory and throughput based limits are also applied.
Row Store Relative Throughput	468	Percentage throughput of this configuration in comparison to the FTDW reference configuration. This result almost matched the reference configuration while using only a single socket. The reference architecture is a 25TB dual-socket configuration.
Column Store Relative Throughput	569	Percentage throughput of this configuration in comparison to the FTDW reference configuration.
Maximum User Data Capacity (TB)	178TB	Calculated, based on total disk capacity. This allows 5:1 compression. It factors recommended free space but ignores the throughput limits that are applied to the Rated User Capacity.
RS Measured Throughput (Q/Hr/TB)	468	Number of Row Store combined benchmark queries completed during the measurement interval. This is normalized to the 1TB database.
CS Measured Throughput (Q/Hr/TB)	3696	Number of Column Store combined benchmark queries completed during the measurement interval. This is normalized to the 1TB database.
Row Store Measured I/O Throughput (MB/S)	12,591	Average of the Physical and Logical scan rate, which demonstrates the throughput capability of the drives. *During maximum load, the physical read throughput peaked at 20,000 MB/sec.*

SQL Server Data Warehouse Fast Track Certification

DWFT Certification #2016-004	HP Proliant DL580 (Gen9) with Fusion io-Memory SX350 DWFT Reference Architecture for Microsoft SQL Server 2016 **DWFT Reference Architecture**		Report Date: 6/13/2016
DWFT Rev. 5.4			

System Provider	System Name	Processor Type	Memory
Hewlett Packard Enterprise	HP Proliant DL580 (Gen9)	Intel Xeon E7-8890 v3 2.5 GHz (4s/72c/144t)	2048 GB

Operating System	SQL Server Edition
Windows Server 2012 R2	SQL Server 2016 Enterprise Edition

Storage Provider	Storage Information
SanDisk a Western Digital brand	8 x SanDisk Fusion ioMemory SX350-6400 (JBOD) 51200GB allocated to Data and TempDB 2 x Lightning II Eco 1.6TB - 1600GB allocated to LOG (Mirror)

Primary Metrics

Rated User Data Capacity[1] (TB)	Row Store Relative Throughput[2]	Column Store Relative Throughput[3]	Maximum User Data Capacity[1] (TB)
120	427	569	178

Row Store

Relative Throughput[2] (Queries/Hr/TB)	Measured Throughput	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
427	468	11,356	13,826	12,591	80

Column Store

Relative Throughput[2] (Queries/Hr/TB)	Measured Throughput	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
569	3,696	2,611	N/A	N/A	93

The reference configuration is a 2 socket system rated for 25TB using SQL Server 2014 and the DWFT V4 methodology

[1] Assumes a data compression ratio of 5:1

[2] Percent ratio of the throughput to the row store throughput of the reference configuration.

[3] Percent ratio of the throughput to the column store throughput of the reference configuration.

[*] Reported metrics are based on the qualification configuration which specifies database size and SQL Server memory.

Summary

Together, Hewlett Packard Enterprise and Western Digital teams dedicated hundreds of hours of testing to engineer the SQL Server DWFT solution to provide the most optimal reliability and performance. These series of tests pushed the HPE ProLiant DL580 Gen9 to peak performance without hardware failure. The reliability and performance experienced during testing is what can be expected in production environments.

The same configuration meets the need of both Row Store and Column Store configurations. It delivers high physical read throughput in the Row Store configuration at a measured average of 12.5GB/s, and high query rates in the Column Store configuration at 3696 Q/Hr/TB.

Bill of Materials

HP Enterprise ProLiant DL580 Gen9 – Non-High Availability Option

Qty	SKU	Description
1	793161-B21	HPE DL580 Gen9 CTO Server
4	788319-B21	HPE DL580 Gen9 E7-8880 v3 1P Kit
64	726719-B21	HPE 16GB 2Rx4 PC4-2133P-R Kit
8	788360-B21	HPE DL580 Gen9 12 DIMMs Memory Cartridge
2	781518-B21	HPE 1.2TB 12G SAS 10K 2.5-in. SC ENT HDD
2	846436-B21	HPE 1.6TB 12G SAS Mixed Use-1 SFF (2.5-in.) SC 3-yr. Warranty Solid State Drive
8	831739-B21	HPE 6.4TB Read Intensive-2 FH/HL PCIe Workload Accelerator
1	629135-B22	HPE 1Gb Ethernet 4P 331FLR Adapter
4	684532-B21	HPE 1500W Hot Plug Power Supply Kit