
NEC Express5800 A2040b 45TB Data Warehouse Fast Track

Reference Architecture with HGST FlashMAX[®] III

Based on Microsoft[®] SQL Server[®] 2014

Data Warehouse Fast Track (DWFT) Reference Architecture

Contents

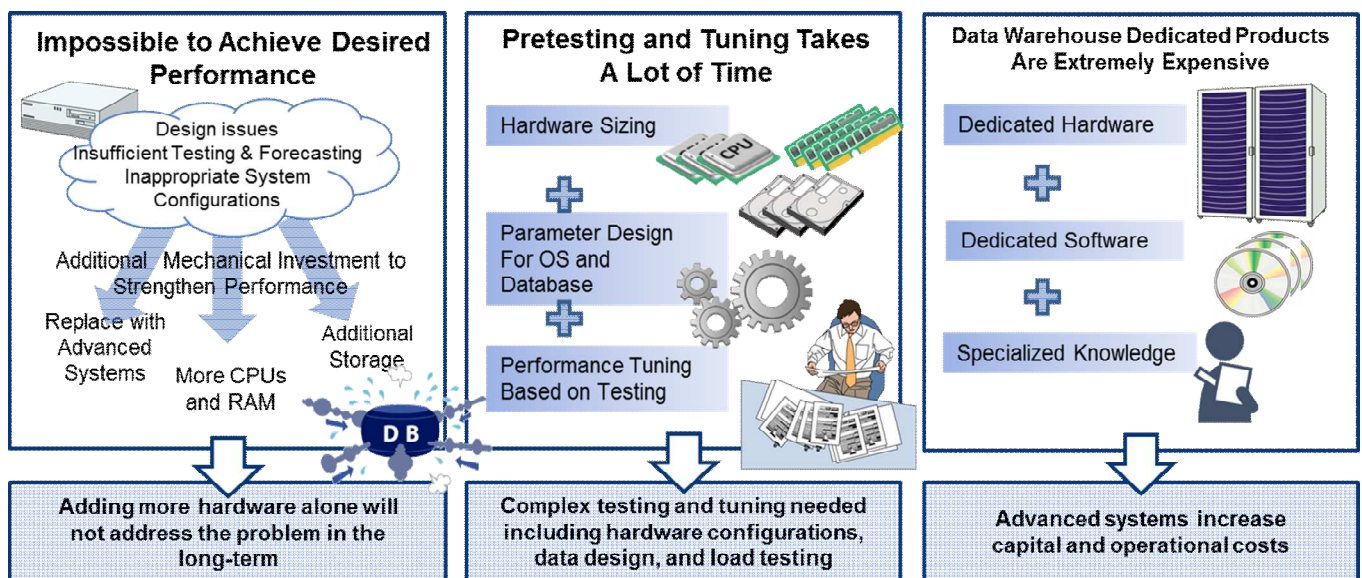
1. Preface.....	2
2. Data Warehouse Fast Track Reference Architecture.....	4
3. Express5800/A2040b DWFT Reference Architecture.....	6
3-1. HARDWARE CONFIGURATION	6
3-2. HARDWARE LAYOUT	9
3-3. STORAGE LAYOUT	10
3-4. FLASHMAX [®] III SETTINGS	11
3-5. SOFTWARE CONFIGURATION	12
4. Pre-verification using Express5800/A2040b DWFT Reference Architecture.....	16
4-1. SQL SERVER [®] DWFT BENCHMARK.....	16
4-2. VERIFICATION RESULTS.....	18
5. Summary	20
Appendix 1. List of models used to configure DWFT Reference Architecture	21
Appendix 2. Related information.....	22
Appendix 3. Inquiries	22

1. Preface

A data warehouse is a large data store integrated from various sources. The integrated data is used for trend analysis, business intelligence (BI), forecast analysis, and other analyses. As the volume of data and its complexity increases on a daily basis, so does the challenge of balancing the capacity of data warehouses and their performance.

Furthermore, hardware must be sized and performance measured whenever a data warehouse is constructed, requiring a great deal of time and money (Figure 1).

Figure 1. Data Warehouse construction problems



To solve these problems, the Data Warehouse Fast Track (DWFT) Reference Architecture uses a configuration that enables standard hardware/software to provide expected performance, and presents the best practices based on pre-verification. It also reduces the cost of such design and construction work.

NEC has worked in conjunction with Microsoft® to develop the Express5800/A2040b DWFT Reference Architecture, a hardware configuration, software configuration, and storage layout designed specifically for use as a data warehouse. Compliant with Microsoft® SQL Server® DWFT Reference Guide for SQL 2014, the Express5800/A2040b DWFT Reference Architecture offers the following advantages (Figure 2).

- Predictable and stable performance
- Fast installation of a pre-verified configuration
- Reasonable pricing and reduced TCO

Figure 2. Advantages of the Express5800/A2040b DWFT Reference Architecture



This whitepaper describes the Express5800/A2040b DWFT Reference Architecture and the pre-verification results of this configuration as executed by NEC.

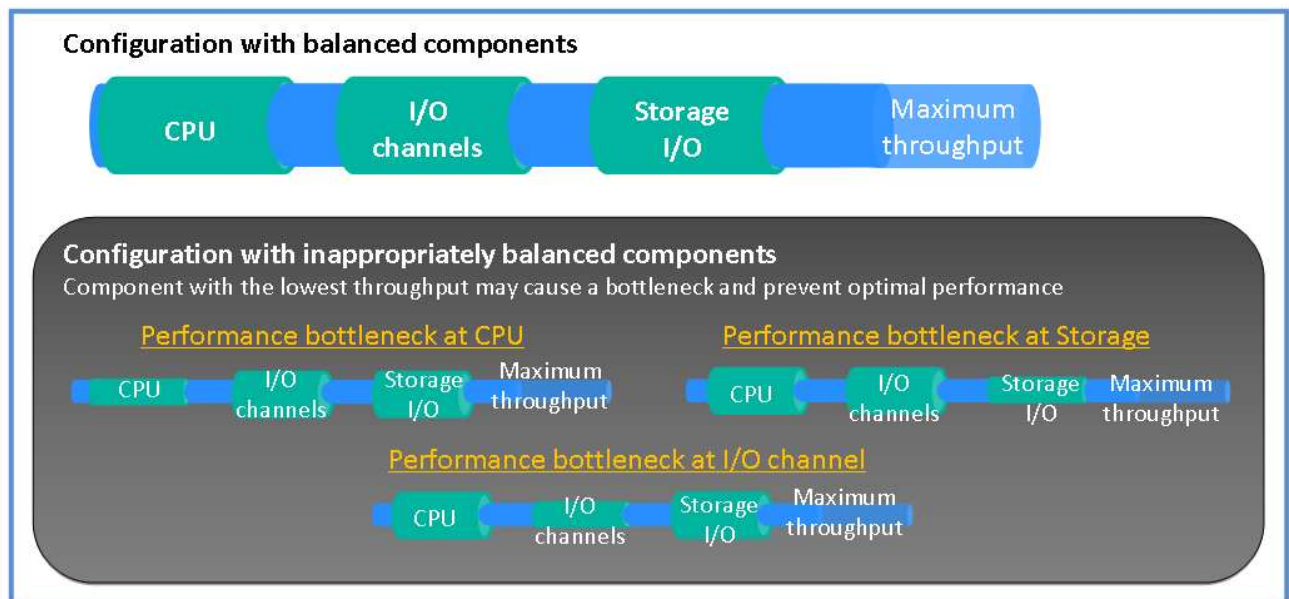
2. Data Warehouse Fast Track Reference Architecture

DWFT Reference Architecture is an architecture that specializes in increasing the speed of sequential I/O performance using a core-balanced architecture and workload optimization.

Core-balanced architecture

A core-balanced architecture adjusts the balance of the I/O channels and sequential I/O capacity of storage based on the CPU cores (Figure. 3). In a data warehouse tasked mainly with reading data sequentially, this architecture achieves high-speed throughput and an optimum configuration.

Figure 3. DWFT core-balanced architecture



Optimizing for the data warehouse workload

A database using DWFT Reference Architecture must be optimized for the data warehouse workload using mainly sequential I/O. Establishing and maintaining data placement to attain an optimum configuration for sequential I/O is important and can be achieved by partitioning.

New features in SQL Server® 2014 : "clustered column store index"

In DWFT Reference Architecture using SQL Server® 2014, users can take advantage of the "clustered column store index (CCI)", which is designed to decrease the query response time and provide a deeper level of data compression.

CCI is optimized for query performance and works well for data warehouse workloads which execute many bulk load and read-only queries. Thus, compared with read performance with a conventional row store index, the query performance is improved up to 10 times, and the data compression of the uncompressed data size is increased up to 7 times.

Further, by using CCI as the default configuration for a table, the compression ratio of DWFT will increase from 3.5 to 5.

3. Express5800/A2040b DWFT Reference Architecture

3-1. Hardware configuration

The following hardware configuration is suitable for a core-balanced architecture:

Express5800/A2040b

Well-suited for large-scale IT platforms, the Express5800 Scalable Enterprise Server features high levels of performance, reliability, and availability demanded by enterprise mission critical and social infrastructure systems. Scalable up to four of the latest Intel® Xeon® processor E7 family v2, this enterprise-class server achieves roughly twice the performance of NEC's previous Express5800/A1080a-S server. It is also ideal for in-memory databases that utilize the large memory size to process data at high speeds. With up to 16 of the latest PCI-Express 3.0 I/O slots available, the server can be used as a platform for real-time analysis, simultaneously incorporating multiple network interface cards, storage interface cards, and flash storage.

The RAS technology developed by NEC enhances the availability of CPU and memory to detect potential failures in advance. System level availability can be maximized by de-allocating resources before potential failures lead to system downtime, and recovering system operations automatically. This enterprise server seeks to set high standards for extensibility, availability, and reliability. Building on NEC's years of experience designing supercomputers and mainframes, this Express5800 server offers businesses superior performance with simple and flexible resource scalability.



CPU

The Express5800/A2040b DWFT Reference Architecture has Intel® Xeon® E7-8893v2 processors (6 core, 37.5M Cache, 3.40 GHz, 8.00 GT/s Intel® QPI) in 4 sockets.

Memory

Total memory is 512GB (DDR3-1600 8GB DIMM x 64), which is the capacity recommended by Microsoft® SQL Server® DWFT Reference Guide for SQL 2014.

FlashMAX® III

The high-performance HGST FlashMAX III server mounted PCIe flash storage, incorporating a hardware RAID mechanism optimized for flash memory, guarantees enterprise-class reliability, as well as delivering linear performance improvement in OLTP, analytics, and other processing applications.



Table 1 and 2 detail the server and PCIe-SSD storage hardware.

Table 1. Server hardware details

Item	Description	Qty
Server	Express5800/A2040b	1
CPU	Intel® Xeon® processor E7-8893v2 @3.4GHz (HT Enabled)	4
- Number of sockets		4
- Number of cores		24
Memory	DDR3L-1600 8GB DIMM	64
- Memory size		512 GB
Hard Drives	2.5-inch 10krpm SAS-HDD 600GB	7

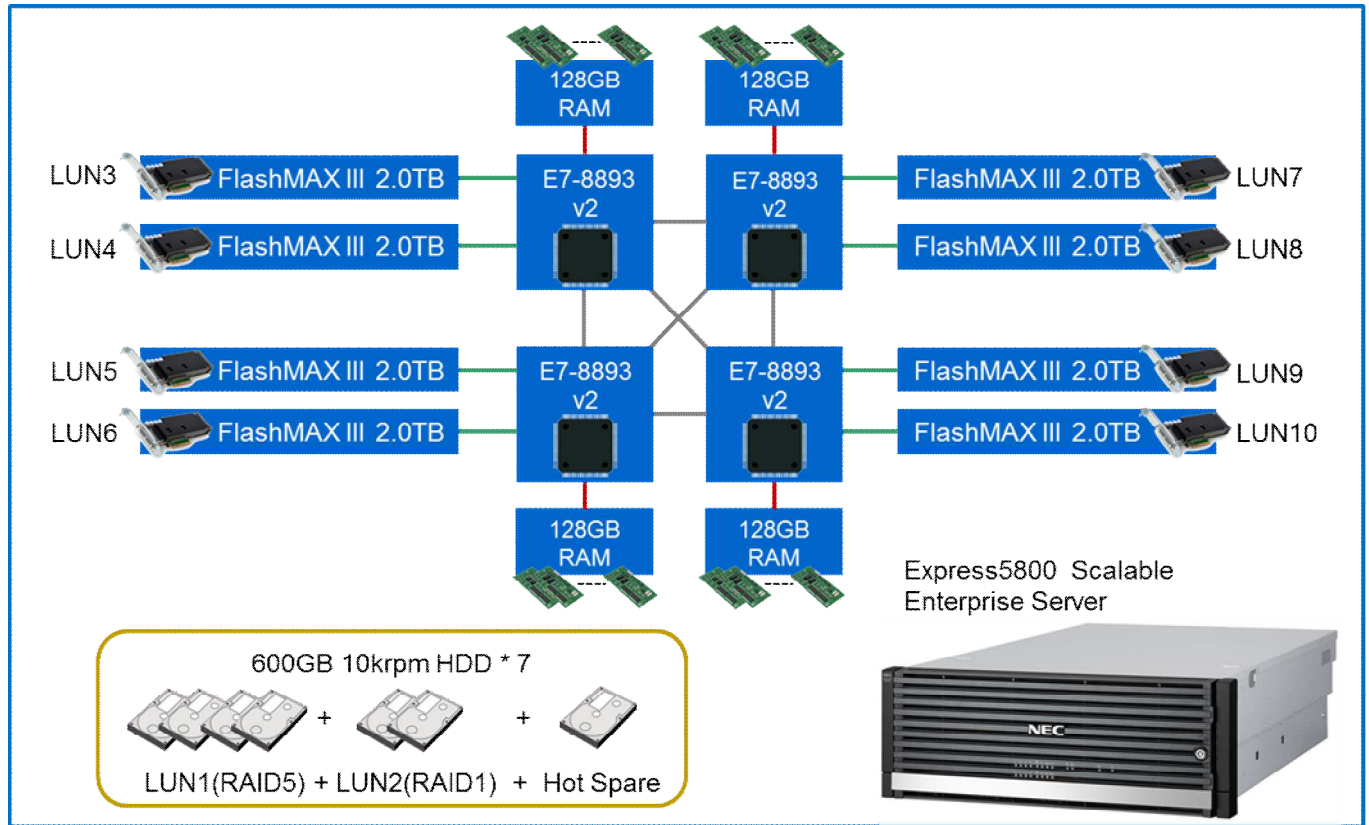
Table 2. PCIe-SSD Storage hardware details

Item	Description	Qty
PCIe-SSD Storage	FlashMAX® III (2.2TB)	8

3-2. Hardware layout

The following figures show the storage layout optimized for the data warehouse workload (Figure 4)

Figure 4. Hardware layout



3-3. Storage layout

The following figures show the storage layout optimized for the data warehouse workload (Figure 5) and area uses (Table 3).

Figure 5. Storage layout

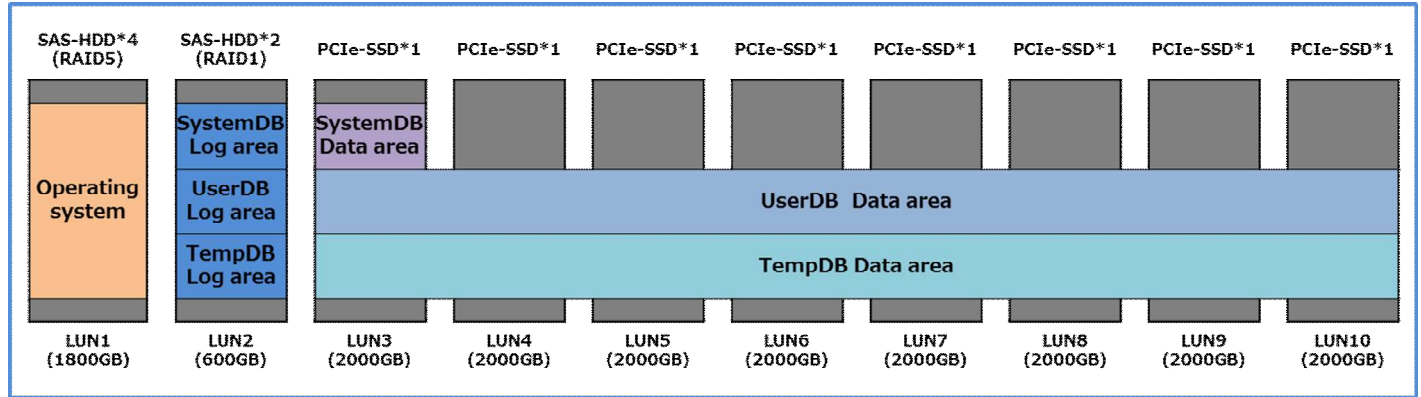


Table 3. Usage of storage areas

Use	Hardware	LUN	RAID Configuration	Capacity
Operating system	SAS-HDD	LUN 1	Hardware RAID 5	1800 GB
User DB area (transaction log) Temp DB area (transaction log) System DB area (transaction log)	SAS-HDD	LUN 2	Hardware RAID 1	600 GB
Hot Spare	SAS-HDD	-	-	600 GB
User DB area (data file) Temp DB area (data file) System DB area (data file)	PCIe-SSD	LUN 3	-	2000 GB
User DB area (data file) Temp DB area (data file)	PCIe-SSD	LUN 4	-	2000 GB
	PCIe-SSD	LUN 5	-	2000 GB
	PCIe-SSD	LUN 6	-	2000 GB
	PCIe-SSD	LUN 7	-	2000 GB
	PCIe-SSD	LUN 8	-	2000 GB
	PCIe-SSD	LUN 9	-	2000 GB
	PCIe-SSD	LUN 10	-	2000 GB

3-4. FlashMAX® III Settings

Table 4 defines the settings for the items and values in FlashMAX® III.

Table 4. FlashMAX® III Settings

Setting	Value
FlashMAX® III operations mode	Maximum Capacity

1) FlashMAX® III Operations Mode

The FlashMAX® III Operations Mode should be chosen depending on the characteristics of applications. For DWH, "Maximum Capacity" mode should be selected.

3-5. Software configuration

Windows Server®

Windows Server® 2012 R2 Standard was used in the Express5800/A2040b DWFT Reference Architecture configuration.

Table 5 defines the settings for the items and values in Windows Server®.

Table 5. Windows Settings

Setting	Value
Power option	High performance
Drive format	Format user DB area at 64K
Paging file size	768 GB
Lock pages in memory	Enabled for SQL Server® execution accounts
Performing volume maintenance tasks	Enabled for SQL Server® execution accounts

1) Power option

Configured to run in power-saving mode can have a harmful effect on SQL Server®. In some instances, a latency of 20 to 25% may occur, so use the voltage control setting that prioritizes performance.

2) Drive format

When formatting drives, the allocation unit size is set to 64KB, the maximum value. With sequential I/O, DWFT Reference Architecture is optimized and places data consecutively. As a result, a larger allocation unit size delivers higher performance. Table 6 shows the settings for formatting drives for each use.

Table 6. Drive format

Use	LUN	Size	Drive letter	File system	Allocation unit size
Operating system	LUN 1	Maximum	C:	NTFS	Default *1
User DB area (log) Temp DB area (log) System DB area (log)	LUN 2	Maximum	E:	NTFS	64K
User DB area (data) Temp DB area (data) System DB area (data)	LUN 3 - LUN 10	Maximum	Arbitrary	NTFS	64K

*1. Do not check the Quick format and Compress files and folders boxes when formatting.

3) Paging file size

Disk space assigned must be large enough to meet virtual RAM and paging requirements. Generally, either 250GB or 1.5 times system RAM, whichever is greater, must be available as free disk space. It was set to 768GB (1.5 times system RAM) for verification by NEC.

4) Lock pages in memory

Local security policy: memory page lock authority is given to the account executing the SQL Server® engine service. Processes executed by an account that has memory page lock authority will avoid paging of secured memory. This prevents the buffer pool from being paged out and improves performance.

5) Performing volume maintenance tasks

Local security policy: execute volume maintenance tasks authority is given to the account executing the SQL Server® engine service. Processes executed by an account that has execute volume maintenance tasks authority will skip 0 filing processes when initializing files. This will reduce the block period when adding, expanding and recovering data files, and improve throughput.

The SQL Server® version is SQL Server® 2014 Enterprise Edition. The system databases (master, MSDB, model) are moved to LUN 3, which is reserved for system databases. Move the master with the activating options (-d, -l), and move MSDB and model with the Alter Database phrases.

Table 7 indicates the SQL Server® settings and their values.

This is the setting used at the validation by NEC. This can be adjusted to the actual user environment.

Table 7. SQL Server® Settings

Setting	Value
Startup option	-E
Trace flag 1	-T1117
Trace flag 2	-T834
Max server memory	236GB
Max degree of parallelism	Row store verification:12 Column store verification:48
Resource Governor	Row store verification:12% Column store verification:25%

1) Startup option (-E)

When -E is designated, extents will be assigned continuously until 64 extents are assigned to one data file in the file group. This improves the data sequence, so the workload in the data warehouse (sequential I/O) is processed efficiently.

2) Trace flag 1 (-T1117)

Enabling -T1117 will batch process all files in the file group with extensions when automatic file extensions are generated. This will help balance data allocation and level off I/O.

3) Trace flag 2 (-T834)

When this trace flag (-T834) is enabled, SQL Server allocates Windows large-page onto

memory for the buffer pool. This improves performance as the efficiency of the translation look-aside buffer (TLB) in the CPU is increased.

4) Max server memory

When max server memory is set to the default, the SQL server will continually secure memory and maintain a larger cache. In addition, the use of physical memory is monitored regularly. If insufficient memory is detected, memory will be released where possible. However, memory cannot always be released immediately, such as when it is in use. In that case, paging will be used and performance will deteriorate. Because of this, it is recommended that max server memory is set to a fixed size. When determining the size of max server memory, the maximum memory size used by the operating system and other processes must be subtracted from all physical memory. It was set to 241664 MB (approximately equal to 236GB) for verification by NEC.

5) Max degree of parallelism

Max degree of parallelism is generally recommended, for a data warehouse workload, to be set to a value equal to the equipped logical cores.

With regard to the workloads used in the NEC verification, max degree of parallelism was set to "row store: 12 (logical processors / cpu socket)" and "column store: 48 (logical processors)".

6) Resource Governor

The Resource Governor controls the CPUs that can be assigned to queries and the volume of memory resources, which must be determined for each individual user. Request_Max_Memory_Grant_Percent is an important tuning item for improving query I/O throughput and responses. With regard to the workloads used in the NEC verification, maximum performance was achieved by setting Request_Max_Memory_Grant_Percent to "row store: 12 percent" and "column store: 25 percent".

4. Pre-verification using Express5800/A2040b DWFT Reference Architecture

Use Express5800/A2040b DWFT Reference Architecture to execute pre-verifications and to confirm that a hardware configuration with a core-balanced architecture has been created.

4-1. SQL Server® DWFT Benchmark

The SQL Server® DWFT Benchmark issues benchmark queries against the test tables at regular intervals. Measurements are taken for each session in which queries are issued. This reference document shows the results of 24 measurement sessions.

1) Rated User Data Capacity (TB) [Primary Metric]

Calculated value based on the Row Store Relative Throughput, the Column Store Relative Throughput, available storage and the physical system memory. The calculations assume a compression ratio of 5:1.

2) Maximum User Data Capacity (TB) [Primary Metric]

Calculated value based on the total disk capacity of all disks allocated to primary data storage and assumes a compression ratio of 5:1. This metric is not limited by the relative throughput metrics.

3) Row Store Relative Throughput [Primary Metric]

Calculated as a percent ratio of the Row Store throughput to the Row Store throughput of the DWFT reference configuration.

4) Column Store Relative Throughput [Primary Metric]

Calculated as a percent ratio of the Column Store throughput to the Column Store throughput of the DWFT reference configuration.

5) Row Store Measured Throughput (Queries/Hr/TB)

Calculated from the measured number of row store benchmark queries completed during the measurement interval, normalized to a 1TB database and expressed in Queries/Hr/TB.

6) Row Store Measured Scan Rate Physical (MB/Sec)

Measured physical I/O reads from disk during the measurement interval of the benchmark and expressed in MB/Sec.

7) Row Store Measured Scan Rate Logical (MB/Sec)

Measured user query throughput which includes reads from RAM/buffer cache and expressed in MB/Sec.

8) Row Store Measured I/O Throughput (MB/Sec)

Calculated midpoint of the measured Scan Rate Physical and the measured Scan Rate Logical and expressed in MB/Sec.

9) Row Store Measured CPU Utilization (%)

Average CPU utilization measured during the measurement interval of the benchmark and expressed in percentage of total CPU.

10) Column Store Measured Throughput (Queries/Hr/TB)

Calculated from the measured number of column store benchmark queries completed during the measurement interval, normalized to a 1TB database and expressed in Queries/Hr/TB.

11) Column Store Measured Scan Rate Physical (MB/Sec)

Measured physical I/O reads from disk during the measurement interval of the benchmark and expressed in MB/Sec.

12) Column Store Measured CPU Utilization (%)

Average CPU utilization measured during the measurement interval of the benchmark and expressed in percentage of total CPU.

4-2. Verification results

SQL Server® DWFT Benchmark verification

Table 8 shows the SQL Server® DWFT Benchmark verification results.

Table 8. SQL Server® DWFT Benchmark

Verified item	Verified result
Rated User Data Capacity (TB) [Primary Metric]	45 TB
Maximum User Data Capacity (TB) [Primary Metric]	55 TB
Row Store Relative Throughput [Primary Metric]	161
Column Store Relative Throughput [Primary Metric]	220
Row Store Measured Throughput (Queries/Hr/TB)	169 Queries/Hr/TB
Row Store Measured Scan Rate Physical (MB/Sec)	4,290 MB/Sec
Row Store Measured Scan Rate Logical (MB/Sec)	5,181 MB/Sec
Row Store Measured I/O Throughput (MB/Sec)	4,735 MB/Sec
Row Store Measured CPU Utilization (%)	91 %
Column Store Measured Throughput (Queries/Hr/TB)	1,429 Queries/Hr/TB
Column Store Measured Scan Rate Physical (MB/Sec)	1,394 MB/Sec
Column Store Measured CPU Utilization (%)	100 %

DWFT Certification #2014-031	NEC Express5800/A2040b - 45TB DWFT Reference Architecture			Report Date: 7/27/2015	
DWFT Rev. 5.4					
System Provider	System Name	Processor Type		Memory	
NEC	NEC Express5800/A2040b	Intel Xeon E7-8893v2 3.4 GHz (4/24/48)		512 GB	
Operating System		SQL Server Edition			
Windows Server 2012 R2 Standard		SQL Server 2014 Enterprise Edition Service Pack 1			
Storage Provider	Storage Information				
NEC	8x 2000 GB SSD for data and tempdb 7x 600 GB HDD for OS and log (RAID 5 + RAID 1 + HotSpare)				

Primary Metrics			
Rated User Data Capacity ¹ (TB)	Row Store Relative Throughput ²	Column Store Relative Throughput ³	Maximum User Data Capacity ¹ (TB)
45	161	220	55

Row Store					
Relative Throughput ²	Measured Throughput (Queries/Hr/TB)	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
161	169	4,290	5,181	4,735	91

Column Store					
Relative Throughput ²	Measured Throughput (Queries/Hr/TB)	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
220	1,429	1,394	N/A	N/A	100

The reference configuration is a 2 socket system rated for 25TB using 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.

5. Summary

The Express5800/A2040b DWFT Reference Architecture provides a balanced data warehouse solution. An optimized system can be created using a balanced data warehouse environment and implementing Microsoft's best practices.

The Express5800/A2040b DWFT Reference Architecture offers the following advantages:

- Optimum hardware configuration that complies with core-balanced architecture.
- Configuration with a verified hardware I/O balance and SQL Server[®] benchmarks.
- In case of the flash memory device failure, the data can be protected by hardware RAID mechanisms (RAID5).
- The cost and time to build the data warehouse are reduced.

Appendix 1. List of models used to configure DWFT Reference Architecture

The hardware and software used in this system is listed in Table 9 and Table 10.

Table 9. Hardware

Model*1	Qty	Product name
NE3400-040F	1	Express5800/A2040b
NE3351-100	1	Internal DVD-ROM Drive
NE3381-88	4	Power Supply Unit (1000W)
NE3343-H001 or -H002	1	Cable Arm (70mm or 110mm)
NE3400-008F	4	Intel® Xeon® E7-8893v2 Processor Kit
NE3302-H002	8	Memory Riser Card
NE3302-H010F	32	16GB MEM (2 x 8GB RDIMM)
NE3350-304	7	600GB 10Krpm SAS HDD
NE3303-168	1	RAID Controller(1GB, RAID 0/1/5/6/10/50/60)
NE3304-133	4	1000BASE-T Adapter(4 port)
NE3318-H003F	8	2.2TB PCIe-SSD

*1. Models shown are for use in Japan. They may vary outside Japan.

Table 10. Software

Model *2	Qty	Product name
----	2	Windows Server® 2012 R2 Standard Server (2 CPU License)
----	12	SQL Server® 2014 Enterprise Core (2 pack Core License)

*2. CAL license in accordance with the system will be required.

Appendix 2. Related information

- NEC Express5800 Series

<http://www.nec.com/en/global/prod/express/scalable/index.html> (For English)

<http://www.nec.co.jp/products/pcserver/rack/r140e4/index.shtml> (For Japanese)

Appendix 3. Inquiries

- NEC Express5800 Series

https://contact.nec.com/http-www.nec.com_tb_009express_en_882b3f/ (For English)

http://www.nec.co.jp/products/express/question/top_sv1.shtml (For Japanese)

©NEC Corporation 2015. All rights reserved.

This document may not be revised without the authorization of NEC Corporation.

This document is only meant to provide information, and may contain technical or editorial errors.

The contents of this document may be changed in the future without notification.

NEC Corporation does not guarantee the content of this document.

Microsoft®, Windows® and SQL Server® are registered trademarks or trademarks of Microsoft Corporation in the United States and other countries.

Intel® and Xeon® are registered trademarks or trademarks of Intel Corporation in the United States and other countries.

FlashMAX® is a registered trademark of HGST, Inc. and its affiliates in the United States and/or other countries.

Any other product names, company names, etc. are the trademarks or registered trademarks of their companies.

Sept. 2015