

Dell Quickstart Data Warehouse Appliance 2000

Reference Architecture



THIS DESIGN GUIDE IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2012 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the *DELL* logo, the *DELL* badge, and *PowerEdge* are trademarks of Dell Inc. *Microsoft*, *Windows Server*, and *SQL Server* are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. *Intel* and *Xeon* are either trademarks or registered trademarks of Intel Corporation in the United States and/or other countries. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

December 2012

Contents

Introduction	2
Audience	2
Reference Architecture	2
Hardware Component Architecture	2
Quickstart Software	4
Applications Settings	4
Windows Server 2008 R2 SP1	5
SQL Server 2012	5
Internal Storage Controller (PERC H710P Mini)	5
RAID Disk Group	7
Internal Storage System	7
Conclusion	9
References	10

Tables

Table 1: Dell Quickstart Reference Architecture Component Details	2
Table 2: Mount Point and Storage Enclosure Mapping	8
Table 3: Performance Metrics	9

Figures

Figure 1. Memory Slot Locations	3
Figure 2. Virtual Disk Settings	Error! Bookmark not defined.
Figure 3. Internal Storage Controller Settings	Error! Bookmark not defined.
Figure 4. RAID Disk Group Configuration	Error! Bookmark not defined.

Introduction

Data warehouses store large amounts of data accumulated from a wide range of sources. This data is generally used for trend analysis, business intelligence reporting, and various types of predictive analysis.

Because data continuously grows in both volume and complexity, organizations often struggle to balance data warehouse capacity and performance. Additionally, data integration and cleansing can be challenging, especially when data originates from multiple sources.

The Dell Quickstart Data Warehouse Appliance is an all-inclusive solution that combines a Dell PowerEdge™ R720xd server with Dell Boomi software and additional services and support. Its balanced configuration is validated on the Microsoft® Fast Track Data Warehouse for SQL Server® 2012 reference model, which is optimized specifically for SQL Server data warehouse workloads to achieve “out-of-box” scalable performance.

Audience

This document is intended for customers, partners, solution architects, database administrators, storage administrators, and business intelligence users who are evaluating, planning, and deploying the Dell Quickstart Data Warehouse Appliance.

Reference Architecture

The Dell Quickstart Data Warehouse Appliance is an integrated solution built on the Dell PowerEdge 12th generation server platform. The solution is optimized to deliver outstanding data warehouse performance. This hardware configuration was tested using Microsoft Fast Track Data Warehouse for SQL Server 2012 Guidelines. The software layer contains SQL Server 2012 Data Warehouse Appliance Edition running on Microsoft Windows Server® 2008 R2 SP1 and Dell Boomi data integration software.

Hardware Component Architecture

Dell conducted tests to determine best practices for building a balanced Fast Track Data Warehouse system on Dell hardware. Based on these results, the Quickstart Data Warehouse Appliance includes the hardware and software components listed in Table 1.

Table 1: Dell Quickstart Reference Architecture Component Details

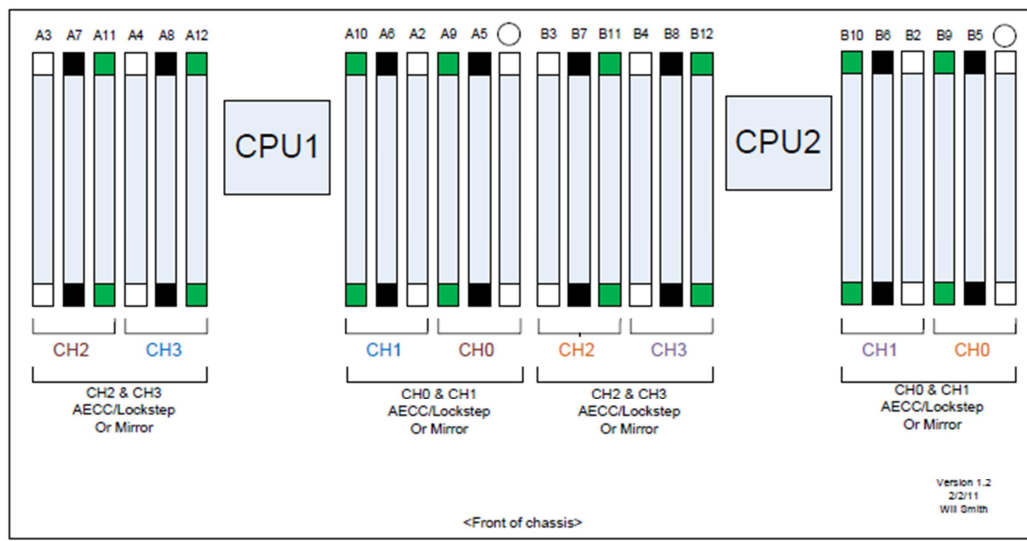
Component	Details
CPU	(2) Intel Xeon E5-2643 CPU @3.3GHz (Hyper-Threading Technology Enabled)
Sockets	2
CPU Cores	8
Memory	128GB RAM <ul style="list-style-type: none">(8) 16GB DDR3 DIMM @1600MHz

Internal Hard Drives	<ul style="list-style-type: none"> • (24) 900GB 10K 2.5" SAS (18 Data, 2 Logs, 2 Stage, 2 Hot Spares) • (2) 900GB 10K 2.5" SAS (2 drives for OS) Rear Bay
Internal Storage Controller	PERC H710P Mini
Operating System	Windows Server 2008 R2 SP1 Enterprise Edition
Database Software	SQL Server 2012 Data Warehouse Appliance Edition
Server Management	Dell Integrated Remote Access Controller (iDRAC7) Enterprise
Data Integration	<p>Dell Boomi (1-year subscription)</p> <ul style="list-style-type: none"> • Five standard connections. See more details at: http://www.boomi.com/products/evaluate/supported_applications • Five test connections. Test connections are equivalent to standard connections, but are used to integrate applications deployed in a sandbox environment. Test connections may be used to set up a redundant test, deployment, or QA environment. • Molecule. A group of multiple nodes that distribute processing across one or more physical machines. Used for integrations that receive or generate a high volume of documents. • Parallel processing. Allows part of a process execution to be executed in parallel to increase document throughput. Can help speed up slow or complex steps by spreading memory across multiple computing processes. • Advanced workflow. The Quickstart Data Warehouse is delivered with advanced features that are activated in AtomSphere for developing complex integration processes. Includes the following components: L Notify, Process Call, Find Changes, Add to Cache, Cleanse, Flow Control, Business Rules, and Catch Errors. • SOA Framework. SOA (Service Oriented Architecture) is a loosely-coupled architecture designed to meet the business needs of the organization. • Advanced User Security. The Advanced User Security feature provides additional AtomSphere roles, the ability to create custom roles, and the ability to use single sign-on (SSO) with SAML authentication.

For SQL Server 2012 reference architectures, Microsoft recommends using 64GB to 128GB of memory for one-socket configurations. This amount of memory improves the performance of large-scale queries involving hash joins and sorting operations by offloading operations from the Tempdb to memory.

Using RDIMM memory also plays a critical role in stack performance. The appliance utilizes 128GB of 1600MHz, DDR3 RDIMMs: four 16GB RDIMMs in slots A1-A4 (white connectors) and four 16GB RDIMMs placed in slots B1-B4 (white connectors) which is depicted in figure 1.

Figure 1: Memory Slot Locations



*White Connector = 1st DIMM population
 Black Connector = 2nd DIMM population
 Green Connector = 3rd DIMM population

See the *Fast Track Reference Guide* for detailed recommendations on system memory configurations.

Quickstart Software

In collaboration with Microsoft, Dell developed a wizard to simplify deploying the Quickstart appliance. The wizard configures the Quickstart appliance based on best practices of Microsoft Fast Track Data Warehouse for SQL Server 2012 Guidelines, and ensures the appliance is properly connected to the target domain and installed with the appropriate administrators.

Additionally, Dell has developed an internal disaster recovery system for the Quickstart Data Warehouse that allows the user to set the appliance back to the original factory state, regardless of the state of the operating system.

Applications Settings

The following section describes the operating system and database configuration settings.

Windows Server 2008 R2 SP1

- **Enable Lock Pages In Memory:** This policy setting is enabled to prevent the system from paging memory to disk. See [Enable the Lock Pages in Memory Option \(Windows\)](#) for more information.

SQL Server 2012

- **SQL Server Startup Options:**
 - **-E:** This parameter increases the number of contiguous extents that are allocated to a database table in each file as it grows to improve sequential access.
 - **-T1117:** This trace flag ensures the even growth of all files in a file group when auto growth is enabled. It should be noted that the Fast Track Reference Guidelines recommend pre-allocating data file space rather than allowing auto grow.
 - **-T834:** This optional setting can improve throughput rates for many data warehouse workloads by enabling large page allocations in memory for the SQL Server buffer pool. However, at this time SQL Server 2012 does not support the use of -T834 when ColumnStore Indexes (CSI) are used on the database. For this reason, you should evaluate your architecture before implementing this option. For more information about this and other trace flags, see [SQL Server Performance Tuning & Trace Flags](#).
- **SQL Server Maximum Memory:** For SQL Server 2012, Fast Track Reference Guidelines suggest allocating no more than 92% of total server RAM to SQL Server. If additional applications will share the server, the amount of RAM left available to the operating system should be adjusted accordingly. This appliance has the maximum server memory set at 119,808 MB.
- **Resource Governor:** The Resource Governor allocates a maximum of 25% of SQL Server memory resources to each session to reduce the maximum memory consumed per query. While it can be beneficial to limit the amount of system resources available to an individual session, this is best measured through analysis of concurrent query workloads. For this appliance, the Resource Governor memory grant was set at 25%. For more information, see [Using the Resource Governor](#).

Internal Storage Controller (PERC H710P Mini)

The Dell PERC H710P Mini is an enterprise-level RAID controller that provides disk management capabilities, high-availability, and security features, in addition to improved performance of up to 6Gb/s throughput. Figure 2 shows the management console accessible through the BIOS utility.

By default, the PERC H710P Mini creates virtual disks with a segment size of 64KB. Dell tested stripe element sizes of 64KB and 256KB to compare and contrast performance improvements, and found that for most workloads, the 64KB default provides an adequate stripe element size. It is recommended to test various stripe sizes depending on the workload characteristics of your configuration.

The default setting for the read policy is *adaptive read ahead*. Dell tested *adaptive read ahead*, *No read ahead*, and *Read Ahead* settings, and observed that the default setting provides the best performance.

Figure 2: Virtual Disk Settings

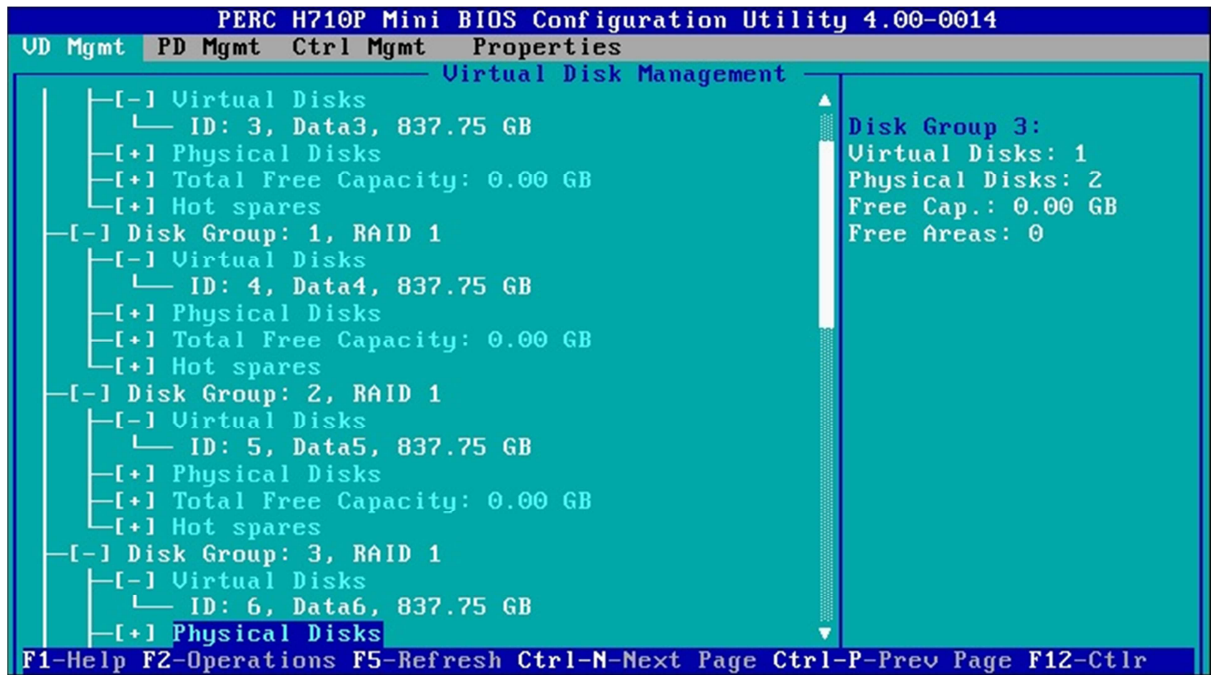
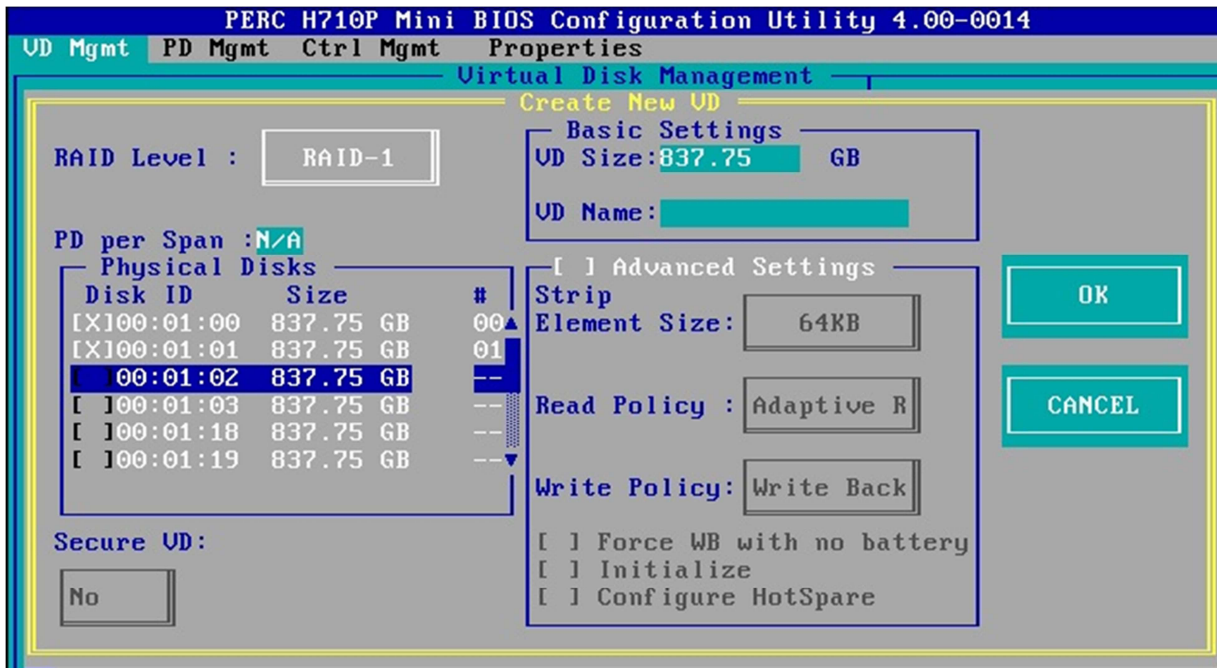


Figure 3: Internal Storage Controller Settings



RAID Disk Group

One of the most critical decisions when deploying a new storage solution is which RAID type(s) to use, as that choice heavily impacts application performance. Dell tested nine RAID 1 disk groups for database data files and one RAID 1 disk group for log files. This left four disks in the front of the chassis, two in RAID 0 to be used for staging data, and the remaining two for spare drives. The back two disks were used in a RAID 1 disk group for the operating system.

Figure 4: RAID Disk Group Configuration



Internal Storage System

Fast Track Reference Guidelines define three primary layers of storage configuration:

- Physical disk array (RAID groups for data and logs)
- Operating system volume assignment (virtual disk)

- Databases (User, System Temp, Log)

For each internal storage array:

- Nine RAID 1 disk groups were created, with each consisting of two disks. These RAID groups were dedicated for the primary user data.
- One RAID 1 disk group was created with two disks. This RAID group was dedicated to host database transaction log files.
- Two RAID 0 disks were created with one disk each. These RAID groups were created as a staging area to transfer data into the appliance.
- One RAID 1 disk group was created with two disks dedicated to the appliance operating system.
- Two disks were assigned as the global hot spares.

Dell used mount points rather than drive letters for storage access. Virtual disks and mount points were assigned using the same naming convention to simplify troubleshooting and performance analysis. All logical volumes were mounted to the “C:\FT” folder.

Table 2: Mount Point and Storage Enclosure Mapping

Disk Group	Virtual Disk	Virtual Disk Label	Logical Label	Full Volume Path
1	1	Cage1-Card1-vData1	Data1	C:\FT\PRI\Cage1-Card1-vData1
2	2	Cage1-Card1-vData2	Data2	C:\FT\PRI\Cage1-Card1-vData2
3	3	Cage1-Card1-vData3	Data3	C:\FT\PRI\Cage1-Card1-vData3
4	4	Cage1-Card1-vData4	Data4	C:\FT\PRI\Cage1-Card1-vData4
5	5	Cage1-Card1-vData5	Data5	C:\FT\PRI\Cage1-Card1-vData5
6	6	Cage1-Card1-vData6	Data6	C:\FT\PRI\Cage1-Card1-vData6
7	7	Cage1-Card1-vData7	Data7	C:\FT\PRI\Cage1-Card1-vData7
8	8	Cage1-Card1-vData8	Data8	C:\FT\PRI\Cage1-Card1-vData8
9	9	Cage1-Card1-vData9	Data9	C:\FT\PRI\Cage1-Card1-vData9
10	10	Cage1-Card1-vLog	log	C:\FT\LOG\Cage1-Card1-vLog
11	11	Cage1-Card1-stage	Stage	C:\FT\stage\Cage1-Card1-stage

The production, staging, and system temp databases were deployed per the recommendations provided in the Fast Track Reference Guidelines.

Conclusion

Microsoft Fast Track Reference Guidelines helped achieve an optimized database architecture on the Quickstart Appliance with balanced CPU and storage bandwidth. Table 3 shows performance metrics of the reference configuration.

Table 3: Performance Metrics

Category	Value	Description
FTDW Rated Data Warehouse Capacity	12 TB	This capacity rating is based on “up-to” capacity, but adjusted to account for Fast Track Rated I/O.
FTDW Rated I/O	1909 MB/s	Core performance metric for validation. Averages physical and logical I/O.
Benchmark Scan Rate - Logical	2164 MB/s	Reflects actual user query throughput which includes reads from RAM/Buffer cache
Benchmark Scan Rate - Physical	1654 MB/s	Reflects physical I/O read from disk during benchmark.
FTDW Peak I/O	3481 MB/s	Maximum observed I/O rate.
FTDW Rated CSI	4337.5 MB/s	Represents potential throughput using ColumnStore Index.

References

Dell | Microsoft SQL Server 2012 Solutions

<http://www.dell.com/sql>

Dell Quickstart Data Warehouse Appliance Solutions

<http://www.dell.com/quickstartdw>

Dell Services

<http://www.dell.com/services>

Dell Support

<http://www.dell.com/support>

OLTP and OLAP

<http://datawarehouse4u.info/OLTP-vs-OLAP.html>

Microsoft Fast Track Data Warehouse and Configuration Guide Information

<http://www.microsoft.com/fasttrack>

http://download.microsoft.com/download/B/E/1/BE1AABB3-6ED8-4C3C-AF91-448AB733B1AF/Fast_Track_Configuration_Guide.docx

An Introduction to Fast Track Data Warehouse Architectures

<http://msdn.microsoft.com/en-us/library/dd459146.aspx>

Enable the Lock Pages in Memory Option (Windows)

<http://go.microsoft.com/fwlink/?LinkId=141863>

SQL Server Performance Tuning & Trace Flags

<http://support.microsoft.com/kb/920093>

Power Edge R720 Technical Guide

<http://www.support.dell.com/support/edocs/systems/per720/en/index.htm>

Dell Boomi

<http://www.boomi.com>