Dell Microsoft SQL Server 2008 Fast Track Data Warehouse Performance Characterization

A Dell Technical White Paper

Database Solutions Engineering
Dell Product Group
Anthony Fernandez
Jisha J
Executive Summary
Dell Microsoft Fast Track Configurations provide a CPU core-balanced approach in implementing a symmetric multiprocessor (SMP)-based Microsoft® SQL Server® data warehouse on Dell™ PowerEdge™ Servers and Dell| EMC Storage. Microsoft Fast Track provides proven performance and scalability expectations on data warehouse workloads. A Fast Track Data Warehouse reference configuration achieves a cost-performance balance between SQL Server data processing capability and throughput of available hardware components. The objective of this white paper is to complement the DELL Reference Configuration Microsoft SQL Server 2008 Fast Track Data Warehouse configuration guide by analyzing the performance using queries similar to standard TPC-H queries. This white paper also explores the performance effects of using different multipathing options provided by EMC® PowerPath®. The results show how the reference configuration building block was balanced and optimized by matching the Storage throughput capacity with the Database Server processing capacity.

THIS WHITE PAPER 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.

© 2010 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, and the DELL badge, and PowerEdge are trademarks of Dell Inc. EMC, PowerPath, and Clariion are registered trademarks of EMC Corporation. Microsoft, SQL Server, and Windows are registered trademarks of Microsoft Corporation in the United States and/or other countries. Intel and Xeon are registered trademarks of Intel Corporation in the U.S. 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.

July 2010
Contents

Executive Summary ................................................................. 2
Introduction .................................................................................. 4
  Audience and Scope ............................................................... 4
Dell Microsoft Fast Track Data Warehouse Reference Architecture .................. 4
Dell Microsoft Fast Track Performance Guidelines ........................................ 6
Evaluation Metrics: MCR and BCR .................................................... 7
Fast Track Performance and Different Multipathing Options ............................... 8
  Single Query Performance .................................................................. 8
  Workload Performance ...................................................................... 10
Conclusion ...................................................................................... 12
References ...................................................................................... 13

Tables
Table 1. Dell Microsoft Reference Architecture Components ............................... 5
Table 2. Sample Query Variants ..................................................................... 6
Table 3. Baseline MCR and BCR Performance .................................................. 8

Figures
Figure 1. Dell Microsoft Fast Track Reference Architecture ............................... 5
Figure 2. Effect of Using Different PowerPath Multipathing Policies With Single Threaded Queries (MAXDOP=1) ................................................................. 9
Figure 3. Effect of Using Different PowerPath Multipathing Policies With Multithreaded Queries (MAXDOP=0) ................................................................. 10
Figure 4. Effect of Using Different PowerPath Multipathing Policies With Single Threaded and Multithreaded Workload (MAXDOP=1 and MAXDOP=0) .................................................. 11
**Introduction**

Microsoft SQL Server Fast Track Initiative provides a concrete methodology and guidelines for implementing balanced data warehouse architecture for predefined customer workloads. All the key components of SQL Server architecture are evaluated and configured to optimize any data warehouse workload throughput. The goal of the configuration is to achieve a high performing data warehouse configuration by balancing the SQL Server processing capability and the overall hardware throughput.

Dell, partnered with Microsoft, is committed to deliver the high performance and cost-effective solutions for customers. The Dell Microsoft Fast Track Reference Architecture Configuration outlines the best practices at both the hardware and software layers to build a high performing data warehouse configuration based on the predefined customer workloads. For more information, please refer to the *DELL Reference Configuration Microsoft SQL Server 2008 Fast Track Data Warehouse* guide. A link to this configuration guide is provided in the “References” section of this white paper.

**Audience and Scope**

This white paper analyzes the performance of the Dell Microsoft Reference Architecture using queries similar to standard TPC-H queries. In addition the various multipathing options supported by PowerPath are also explored. This white paper may be useful for IT planners, architects, database administrators and Business Intelligence (BI) users who are interested in the data warehouse performance and configuration aspects.

**Dell Microsoft Fast Track Data Warehouse Reference Architecture**

Dell Microsoft Fast Track Reference Configuration bundles the highly efficient and cost effective Server and Storage along with Microsoft SQL Server 2008 R2. The reference configuration components are listed in Table 1 as follows.
Table 1. Dell Microsoft Reference Architecture Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>Dell™ PowerEdge™ R710</td>
</tr>
<tr>
<td>CPU</td>
<td>(2) Quad Core Intel Xeon® X5550 2.66GHz 8MB Cache, 6.4GT/s Intel QPI, 95W</td>
</tr>
<tr>
<td>Number of Cores</td>
<td>8</td>
</tr>
<tr>
<td>Memory</td>
<td>48GB RAM (6x8GB DIMMS @1333MHz)</td>
</tr>
<tr>
<td>PCI-E Slots</td>
<td>PCI-E Riser with two x4 Gen 2 slots (Slots 1 and 2)</td>
</tr>
<tr>
<td>Internal Storage Controller</td>
<td>PERC6/I or H700 512MB Cache</td>
</tr>
<tr>
<td>Internal Drives</td>
<td>(2) 73GB 6Gpbs SAS 15K HDD (6) 300GB 6Gbps SAS 15K HDD*</td>
</tr>
<tr>
<td></td>
<td>* Various sizes can be configured.</td>
</tr>
<tr>
<td>Network Adapters</td>
<td>Two Embedded Broadcom 5709C dual-port Gigabit Ethernet (four ports total)</td>
</tr>
<tr>
<td>Fibre Channel (FC) HBA</td>
<td>(2) Emulex LPe-12002-E (8Gb Dual Port FC HBA) FRWM 1.11 A5/Driver 2.20.006 Or (2) QLogic QLE2562 (8Gb Dual Port FC HBA)</td>
</tr>
<tr>
<td>FC Switches</td>
<td>(2) Brocade SW200E (8 Ports enabled each)*</td>
</tr>
<tr>
<td></td>
<td>* High Availability Configuration requires two FC switches.</td>
</tr>
<tr>
<td>FC Storage</td>
<td>(2) Dell</td>
</tr>
<tr>
<td>Operating System</td>
<td>Microsoft Windows® 2008 R2</td>
</tr>
<tr>
<td>Database Software</td>
<td>Microsoft SQL Server 2008 R2</td>
</tr>
<tr>
<td>Multipathing Software</td>
<td>EMC PowerPath 5.3 SP1 64bit</td>
</tr>
</tbody>
</table>

The complete Dell Microsoft Reference Architecture may be depicted as in Figure 1. It shows the Fast Track configuration with dual Fibre Channel HBA’s and switches for high availability.

Figure 1. Dell Microsoft Fast Track Reference Architecture
The EMC PowerPath multipathing solution provides flexibility and efficiency for the Dell EMC storage arrays, functioning as the backend for the SQL Server Fast Track Database Configuration. In addition, Dell recommends providing high availability to the database server using Microsoft Clustering Services (MSCS). For more information, please refer to the DELL Reference Configuration Microsoft SQL Server 2008 Fast Track Data Warehouse guide. A link to this configuration guide is provided in the “References” section of this white paper.

As part of the activities for the white paper, tests were conducted on the preceding configuration to analyze the overall data warehouse performance.

**Dell Microsoft Fast Track Performance Guidelines**

Experiments were conducted on several configurations in Dell Labs to optimize the performance achieved by the Dell Fast Track Reference Architecture. The experiments were carried out using different query variants (simple, average and complex) and also using a workload mix. The query complexity is decided depending on the number and complexity of JOIN statements.

All the queries were derived from queries similar to the standard TPC-H benchmark. Tests also evaluated the single threaded and multithreaded query performance.

The queries selected for the exercise are given in Table 2.

**Table 2. Sample Query Variants**

<table>
<thead>
<tr>
<th>Query Type</th>
<th>Selected Query</th>
<th>Query Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>Forecasting Revenue Change Query (Q6)</td>
<td>SELECT SUM(L_EXTENDEDPRICE*L_DISCOUNT) AS REVENUE FROM LINEITEM WHERE L_SHIPDATE &gt;= '1994-01-01' AND L_SHIPDATE &lt; dateadd (yy, 1, '1994-01-01') AND L_DISCOUNT BETWEEN .06 - 0.01 AND .06 + 0.01 AND L_QUANTITY &lt; 24</td>
</tr>
<tr>
<td>Average</td>
<td>Order Priority Checking Query (Q4)</td>
<td>SELECT o_orderpriority, COUNT(*) AS order_count FROM dbo.orders WHERE o_orderdate &gt;= '1993-04-01' AND o_orderdate &lt; dateadd(month, 3,'1993-04-01') AND EXISTS ( SELECT * FROM dbo.lineitem WHERE l_orderkey = o_orderkey AND l_commitdate &lt; l_receiptdate AND l_shipdate &gt; '1993-01-01' AND l_shipdate &lt; '1993-12-31' ) GROUP BY o_orderpriority ORDER BY o_orderpriority</td>
</tr>
<tr>
<td>Complex</td>
<td>Pricing Summary Report Query (Q1)</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>select l_returnflag,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>l_linestatus, sum(l_quantity) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sum_qty, sum(l_extendedprice) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sum_base_price, sum(l_extendedprice* (1-l_discount)) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sum_disc_price, sum(l_extendedprice* (1-l_discount) *(1+l_tax)) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sum_charge, avg(l_quantity) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>avg_qty, avg(l_extendedprice) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>avg_price, avg(l_discount) as</td>
<td></td>
</tr>
<tr>
<td></td>
<td>avg_disc, count_big(*) as count_order</td>
<td></td>
</tr>
<tr>
<td></td>
<td>from lineitem where l_shipdate &lt;= dateadd(dd, -90, '1998-12-01')</td>
<td></td>
</tr>
<tr>
<td></td>
<td>group by l_returnflag, l_linestatus</td>
<td></td>
</tr>
<tr>
<td></td>
<td>order by l_returnflag, l_linestatus</td>
<td></td>
</tr>
</tbody>
</table>

Note that the preceding queries were selected as a representation of the customer query complexities. Customers would have to select a standard SQL query (or a set of queries) that are specific to their data warehouse workload for the Fast Track performance evaluation.

**Evaluation Metrics: MCR and BCR**

Dell Microsoft Fast Track performance is mainly evaluated based on two parameters: Maximum Core Consumption Rate (MCR) and Benchmark Consumption Rate (BCR). MCR refers to the maximum throughput that can be achieved using the specified CPU configuration. BCR is the maximum throughput that may be achieved using the defined disk layout. Dell Microsoft Fast Track configurations ensure the optimal database performance by providing predefined configurations that have been tested to be balanced. The goal of Microsoft Fast Track configurations is to design configurations that are capable of consuming the available storage bandwidth represented by a balance between MCR and BCR results.

MCR evaluation is performed completely on the cached data, independent of the disk configuration. On the other hand, BCR is measured based on the disk I/O during the query load. The following formula is used to calculate the MCR and BCR in MB/s: \(( \text{Logical reads} / \text{[CPU time]} \times 8\text{KB} / 1024 \)\)

Detailed guidelines for measuring the MCR and BCR values may be found in the “Workload Testing” section of the Microsoft *Fast Track Data Warehouse 2.0 Architecture*. A link to this document is provided in the “References” section of this white paper.
For the proposed Dell reference architecture, the MCR and BCR values were measured to be as given in Table 3 as follows.

**Table 3. Baseline MCR and BCR Performance**

<table>
<thead>
<tr>
<th>MCR (Maximum Core Consumption Rate) – CPU Based</th>
<th>293 MBps</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR (Benchmark Consumption Rate) – Disk Based</td>
<td>296 MBps</td>
</tr>
</tbody>
</table>

The results show similar results for MCR and BCR in terms of throughput achieved. This proves that the maximum CPU consumption rate is capable of processing the available disk bandwidth thus creating a balanced configuration. This configuration maximizes the overall database throughput while minimizing unrealized resources.

The following sections of this white paper analyze the performance of Dell Microsoft Fast Track Data Warehouse for executing query variants and a workload mix similar to standard TPC-H queries utilizing BCR to measure disk I/O throughput.

**Fast Track Performance and Different Multipathing Options**

As explained in the **DELL Reference Configuration Microsoft SQL Server 2008 Fast Track Data Warehouse** configuration guide, customers have the option to utilize Microsoft Multipathing I/O (MPIO) or EMC PowerPath.

Microsoft MPIO has been tested by Microsoft extensively on all Fast Track configurations. It is recommended to use with the Round Robin with Subset multipathing policy.

PowerPath provides an efficient Multipathing solution for the Dell | EMC AX4-5 arrays. EMC PowerPath provides the flexibility of using different multipathing policies for accessing the external storage hardware. This section explores the performance of the Fast Track reference architecture on using some of the different multipathing policies. By default, PowerPath utilizes the Clarion® Optimized (ClarOpt) setting. The different multipathing policies explored are Basic failover (bf), ClarOpt (co), Least Blocks (lb), Least I/Os (li) and Round Robin (rr).

The MAXDOP query option controls the query parallelism. MAXDOP option 1 enables a single thread of execution; whereas, MAXDOP of zero utilizes multiple threads. The test queries were executed using both the MAXDOP 1 and 0 query options to analyze the single threaded and multithreaded query performance, respectively. For more information on MAXDOP options, refer to the “References” section of this white paper.

**Single Query Performance**

Simple, Average and Complex queries were executed individually on the database. On each of the test iterations, all the queries were executed separately with the different PowerPath policies mentioned earlier. Tests were conducted with the MAXDOP query options 1 and 0. Using the MAXDOP option 1 will ensure that the query is utilizing only a single core of CPU at any point of execution. The results of the exercise may be depicted as in Figure 2 as follows.
The different PowerPath policies produced comparable database throughput for all the query types with MAXDOP option 1. The result shows that with a single thread of execution, having a single core accessing a single database volume, the database may be configured with any of the PowerPath policies for optimum output.

Similarly, all the test queries were run with MAXDOP option 0 to simulate a multithreaded query operation. The results are as given in Figure 3 as follows.
We can see that even in the case of multithreading, all the different multipathing policies are delivering almost the same query performance. To summarize, when queries of similar complexity are executed against the reference configuration, all the PowerPath multipathing policies delivered similar results.

**Workload Performance**

A sample workload based on simple/average/complex queries with a 70/20/10 percentage mix, respectively, was executed on the Dell Microsoft Fast Track Reference Configuration. The workload tests were executed with both MAXDOP options 1 and 0.

Figure 4 represents the BCR performance on using the MAXDOP options 1 and 0 as follows.
As per the results observed, three of the PowerPath policies, Least blocks, Least IOs and Round Robin, give enhanced performance for a 70/20/10 workload mix similar to a TPC-H workload mix using single thread execution (MAXDOP=1). For a multithread execution model, Least IOs seems to outperform all the other PowerPath policies.
Conclusion

Dell Microsoft Fast Track Data Warehouse provides specific guidelines for deploying a low cost and high performing data warehousing solution. All the key database configuration components, including all the hardware and software elements, are optimized to cater to large sequential workloads.

Dell Microsoft Fast Track Reference Configuration achieves an efficient balance between the SQL Server processing capability and the aggregate hardware throughput for specific database size requirements.

Dell EMC PowerPath provides the flexibility and ease of using different multipathing options for database connectivity. As per the single threaded and multithreaded experimental results, the Least I/Os (li) policy is seen to provide consistent performance for all query variants (simple, average and complex) and for a workload mix of 70/20/10 (simple/average/complex, respectively). Based on the nature of the queries expected, customers may choose the multipathing policy accordingly.

To conclude, Dell Microsoft Fast Track Reference Configuration provides an excellent end-to-end solution with definite configuration and performance guidelines for implementing an out-of-the-box, balanced data warehousing solution.
References

DELL Reference Configuration Microsoft SQL Server 2008 Fast Track Data Warehouse

Dell SQL Server Solutions
www.dell.com\sql

Dell Services
www.dell.com\services

Dell Support
www.dell.com\support

Microsoft Fast Track Solutions
www.microsoft.com\fasttrack

An Introduction to SQL Server 2008 Fast Track Data Warehouse Architectures

Fast Track Data Warehouse 2.0 Architecture

Microsoft Dell Data Sheet
http://download.microsoft.com/download/D/F/A/DFAAD98F-0F1B-4F8B-988F-22C3F94B08E0/Dell%20Fast%20Track%202.0%20Datasheet.pdf

Introduction to New Data Warehouse Scalability Features in SQL Server 2008

Best Practices for Data Warehousing with SQL Server 2008
http://msdn.microsoft.com/library/cc719165.aspx

Microsoft MAXDOP