

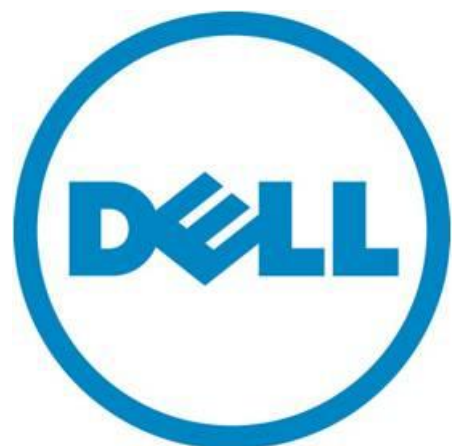
Dell SMB Reference Configuration for Microsoft® SQL Server® 2008 R2 Fast Track Data Warehouse with the Dell BOOMi Integration Capability

A Dell Technical White Paper

Dell | Database Solutions Engineering

Dell Product Group

August 2011



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.

© 2011 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*, 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.

October 2011

Contents

- Introduction 3
 - Audience and Scope 3
- Microsoft SQL Server Fast Track Data Warehouse 3
- Dell Fast Track Data Warehouse Reference Architecture 4
 - Hardware Component Architecture 4
 - Dell PowerEdge R510 Server 5
 - Processors 5
 - Memory 5
 - Internal Storage Controller (PERC H700) Settings 5
 - Stripe element size 6
 - Read policy..... 6
 - RAID configuration 7
- Application Configuration 7
 - Windows Server 2008 R2 SP1..... 7
 - SQL Server Configuration 8
 - Internal Storage System 8
- Performance Benchmarking 9
 - Baseline Hardware Characterization using Synthetic I/O 9
 - Fast Track Workload Evaluation 11
 - Calculating MCR 11
 - Calculating BCR 12
- Populating Your Data Warehouse 12
- Building an Integration Process 13
 - Configuring Integration Automation 16
- Deploy an Integration Process..... 17
 - Addressing Enterprise-Class Integration 18
 - Reviewing Overall Integration Performance 18
 - Reviewing Individual Process Executions 18
 - Exploring Atoms 19
 - Subscribing to Process Alerts 19
- Conclusion 20
- References..... 21
- Appendix 22

Summary of Microsoft SQL Server Fast Track results 22

Tables

Table 1. Tested Dell Fast Track Reference Architecture Component Details 4
Table 2. Dell Fast Track Reference Architecture Solution Details 5
Table 3. Mount Point Naming and the Storage Enclosure Mapping 8

Figures

Figure 1. Proposed Dell Fast Track Reference Architecture..... 4
Figure 2. Virtual Disk Settings 6
Figure 3. Internal Storage Controller Settings 7
Figure 4. RAID Configuration 7
Figure 5. Storage System Components 9
Figure 6. SQLIO Line Rate Test from Cache (Small File) 10
Figure 7. SQLIO Real Rate Test from Disk (Large File) 11
Figure 8. Dell Boomi 13
Figure 9. Dell Boomi Build Environment Real-Time Integration Testing 16
Figure 10. Dell Boomi - Process Execution View 18

Introduction

Data Warehousing is used for integrating, storing and analyzing data in order to perform trend analysis, business intelligence reporting and various types of predictive analysis. With today's never ending data growth and complexity, it's becoming a tedious job for customers to balance capacity and performance within the data warehouse system. Growing data volumes and loading challenges, OLAP query complexity, and number of users are causing response times to increase. IT executives are looking for solutions that offer lower cost, easier management, and better performance.

There are many challenges in designing a database configuration for OLAP workloads. One is ensuring an optimal balance of I/O, storage, memory and processing power.

Dell™ and Microsoft® jointly developed guidelines and design principles to assist customers in designing and implementing a balanced configuration specifically for Microsoft SQL Server® data warehouse workloads to achieve “out-of-box” scalable performance.

Another challenge is data integration and cleansing, especially where data is coming from multiple sources. Based on direct input from customers, Dell has incorporated Dell Boomi, an efficient cloud-based on-demand application and data integration service as part of the balanced infrastructure.

This whitepaper describes the architecture design principles needed to achieve a balanced configuration for the Dell PowerEdge™ R510 server using the Microsoft Fast Track Data Warehouse 3.0 guidelines.

Audience and Scope

This whitepaper is intended for customers, partners, solution architects, database administrators, storage administrators, and business intelligence users who are evaluating, planning, and deploying an optimally balanced data warehouse solution. The scope of this whitepaper is limited to the data warehouse. Analytics tools and reporting services that use the data warehouse are outside the scope of this whitepaper.

Microsoft SQL Server Fast Track Data Warehouse

In order to overcome the limitations of traditional data warehouse systems, Microsoft has come up with a cost effective solution that optimally balances the hardware and software capabilities of the system. It provides an easy to deploy data warehouse infrastructure by mainly focusing on storage tuning and database layout. Fast Track Data Warehouse (FTDW) its implementing data warehouse solutions differently. As most data warehousing queries scan large volumes of data, FTDW designs are optimized for sequential scans/reads. These proven methodologies yield performance much better than that of traditional data warehousing systems. Based on this fact, DELL has made a deep study on the FTDW architecture and come up with a reference guide that helps customers implement FTDW on DELL hardware.

Dell Fast Track Data Warehouse Reference Architecture

In order to optimize data warehouse stack component performance, we must properly tune each layer. The following sections explain the tuning of selected hardware and software.

Hardware Component Architecture

Redundant and robust tests have been conducted on DELL’s PowerEdge servers to determine best practices and guidelines for building a balanced FTDW system.

Figure 1. Proposed Dell Fast Track Reference Architecture



PowerEdge R510
1 * Intel Xeon® X5675 3.07Ghz Six Core CPU
Windows2008 R2 SP1
SQL Server2008 R2

Configuration availability may be further enhanced by configuring database clustering using multiple servers.

Table 1. Tested Dell Fast Track Reference Architecture Component Details

Component	Details
Server	PowerEdge R510 (BIOS: 1.6.3)
CPU	(1) Intel® Xeon® CPU X5675 @3.07GHz (HT Enabled)
Number of cores per socket	6
Total Number of CPU Cores	6
Memory	48GB RAM (3 * 16 DDR3 DIMMs @1066MHz)
Internal Hard Drives	12 x 600GB 15K RPM Serial-Attach SCSI 6Gbps 3.5in Hotplug Hard Drive
Operating System	Microsoft Windows® 2008 R2 SP1 Enterprise Edition
Database Software	Microsoft SQL Server 2008 R2 Enterprise Edition

Table 2. Dell Fast Track Reference Architecture Solution Details

Solution Description	Configuration ID
Dell Fast Track 3.0 Configuration PowerEdge R510	7068273
Dell Fast Track 3.0 Configuration PowerEdge R510 with Dell Boomi	7068280

Dell PowerEdge R510 Server

The Dell PowerEdge R510 server is a 2-socket, 2U high-capacity, multi-purpose rack server offering an excellent balance of internal storage, redundancy, and value in a compact chassis.

The PowerEdge R510 server was developed with a purposeful design, energy-optimized options, the performance of the Intel Xeon processor 5500 and 5600 series, DDR3 memory, and enterprise-class manageability. For more technical specifications of R510 Server, please refer to the *Power Edge R510 Technical Guide*, a link to which is provided in the References section of this document.

Processors

The Microsoft Fast Track 3.0 reference guide describes how to achieve a balance between components such as storage, memory, and processors. In order to balance the available internal storage and memory for the Dell PowerEdge R510, a single Intel Xeon X5675 Six core processor operating at 3.07GHz speed was used.

Note: For environments that require high processing capabilities, a second socket can be populated.

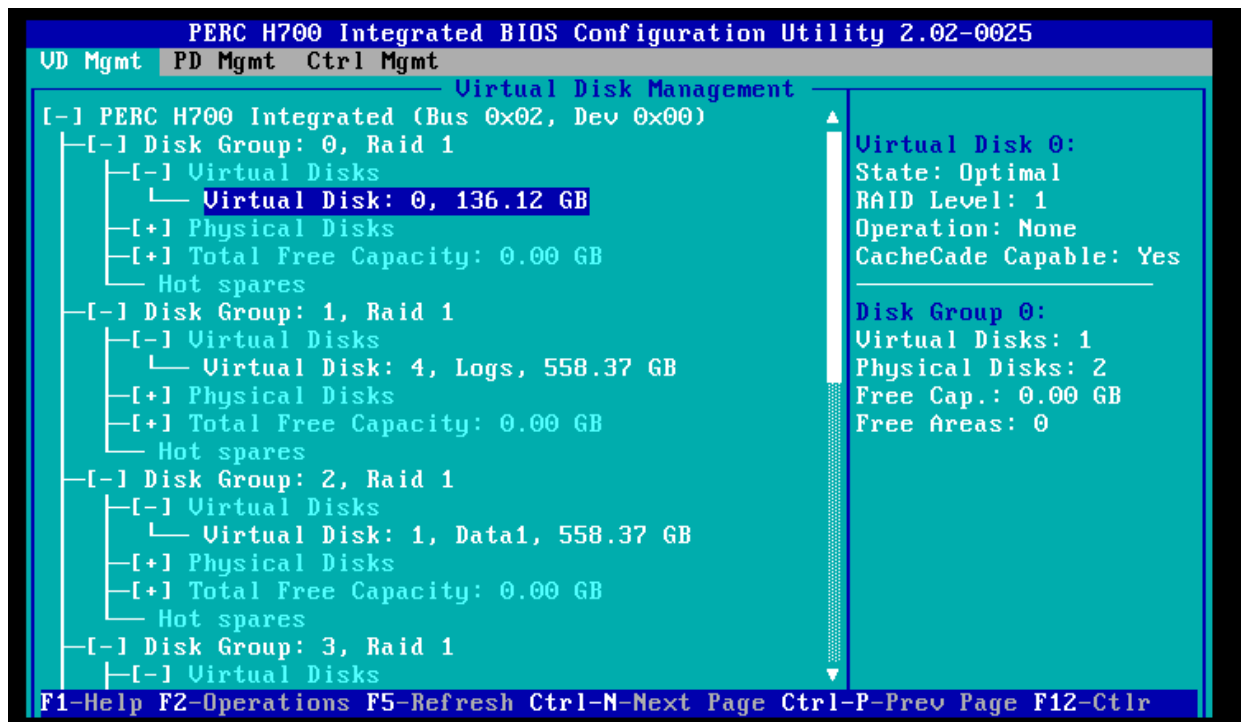
Memory

For FTDW architecture, Microsoft recommends using 8GB of memory per processor core. With enough memory installed on the system, the large-scale queries involving Hash joins and sorting operations will benefit from SQL Server offloading operations from the Tempdb to Memory. Selection of Memory DIMMS will also play a critical role in the performance of the entire stack. In our test configuration, we have configured the database server with 48GB of RAM running at 1066 MHz speed. Refer to the *Microsoft Fast track 3.0 Reference Guide* for detailed recommendations on system memory configuration.

Internal Storage Controller (PERC H700) Settings

The Dell PERC H700 is an enterprise level SAS 2.0 RAID controller that provides disk management capabilities, high available, 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.

Figure 2. Virtual Disk Settings



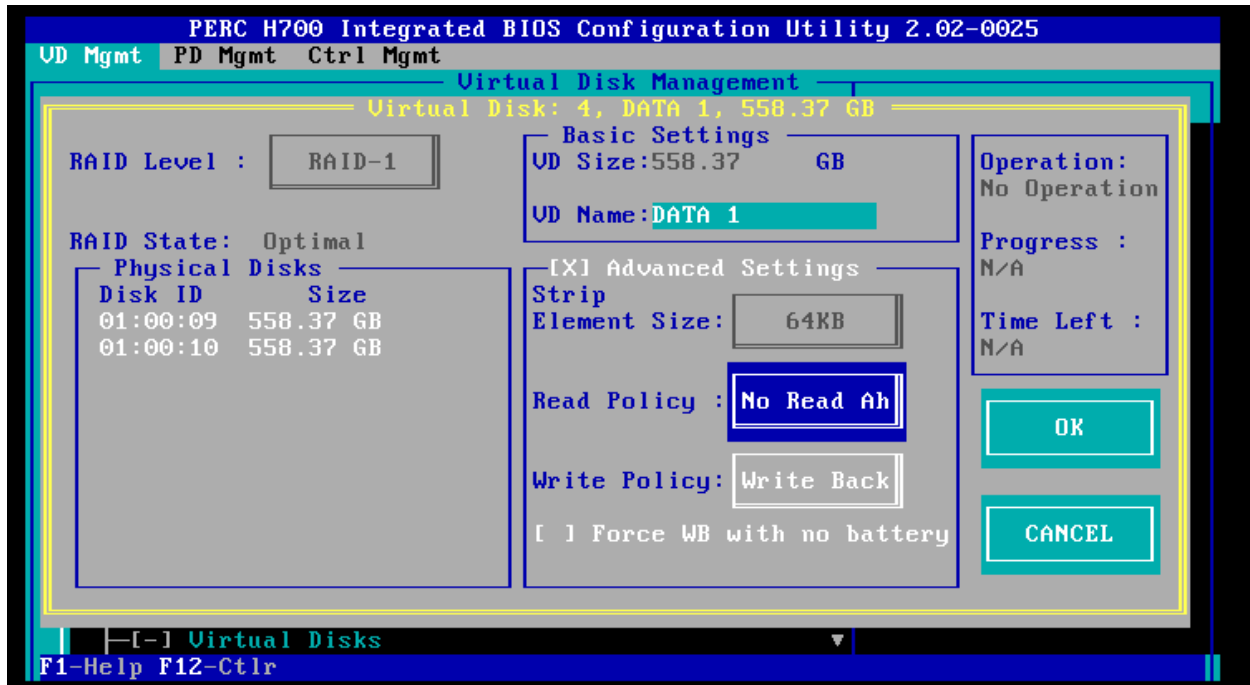
Stripe element size

By default, the PERC H700 creates virtual disks with a segment size of 64KB. During the Fast Track validation testing we utilized stripe element sizes of 64KB and 1MB to compare and contrast any performance improvements. For most workloads, 64KB default size will provide an adequate stripe element size. We recommend testing various sizes depending on the workload characteristics of your configuration.

Read policy

The default setting for read policy on the PERC H700 is “adaptive read ahead”. During testing, however, we observed that changing the setting to “No read ahead” improved the overall performance by 4%; we attribute the improvement to unnecessary read ahead during large sequential I/O requests. Adaptive read ahead typically improves performance in small random workloads.

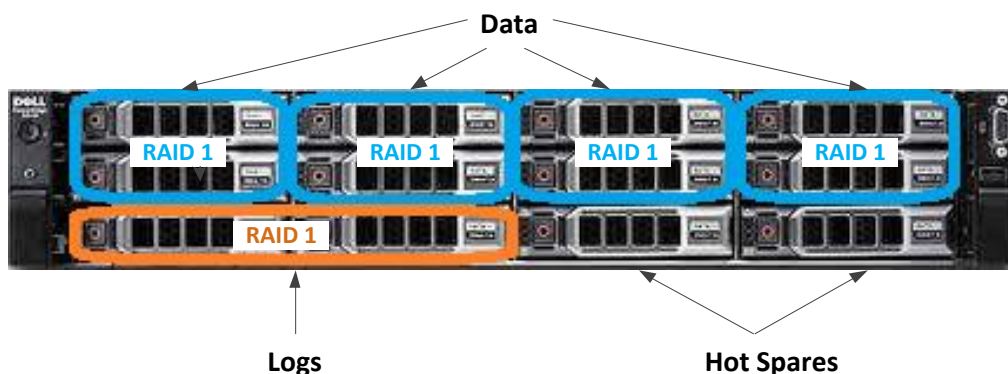
Figure 3. Internal Storage Controller Settings



RAID configuration

One of the most critical decisions that we have to make when deploying a new storage solution is which RAID type(s) to use as that choice heavily impacts the performance of the application. We have configured the proposed Fast Track configuration using RAID 1 disk groups for database data files and database log files. Four RAID 1 data disk groups and one RAID 1 log disk group were created. Figure 3 shows the proposed RAID configuration.

Figure 4. RAID Configuration



Application Configuration

The sections below explain the settings applied for the operating system and database layers.

Windows Server 2008 R2 SP1

Default settings were used for the Windows 2008 R2 SP1 operating system.

SQL Server Configuration

These startup options were added to the 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. This improves 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 that you pre-allocate the data file space rather than allow auto grow.
- Enable **Lock Pages In Memory**. Refer to Appendix for more information.
- -T834: Large Page Allocations.

Note: This flag should be tested thoroughly on a case-by-case basis. The validation for this configuration does not include this trace flags. For more information on trace flags refer to the Appendix.

Internal Storage System

The Fast Track reference architecture 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:

- Four RAID 1 Disk Groups were created, each consisting of 2 disks. These RAID groups were dedicated for the primary user data.
- One RAID 1 disk group was created of 2 disks. This RAID group was dedicated to host the database transaction log files.
- The remaining 2 disks were assigned as the storage hot spares.

For the entire internal storage setup, there were eight disks dedicated to hold the primary user data and two disks to hold the database log files.

For Fast Track architectures, we recommend that you use mount point rather than drive letters for storage access. It is also very important to assign the appropriate virtual disk and mount point names to the configuration in order to simplify troubleshooting and performance analysis. Mount point names should be assigned in such a way that the logical file system reflects the underlying physical storage enclosure mapping. Table 2 shows the virtual disk and mount point names used for the specific reference configuration and the appropriate storage layer mapping. All the logical volumes were mounted to the “C:\FT” folder.

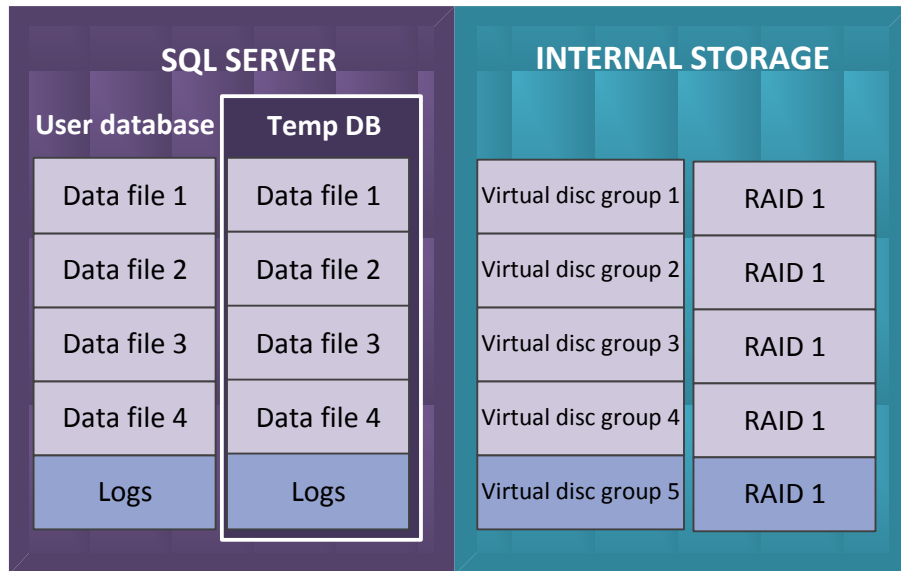
Table 3. Mount Point Naming and the Storage Enclosure Mapping

Disk Group	Virtual Disk	Virtual Disk Label	Logical Label	Full Volume Path
1	1	se1- v1	Data1	C:\ft\se1-v1
2	2	se1- v2	Data2	C:\ft\se1-v2
3	3	se1- v3	Data3	C:\ft\se1-v3

4	4	se1- v4	Data4	C:\ft\se1-v4
5	5	log1-v5	log	C:\ft\log\SE1-log

Figure 5 represents the storage system configuration for the proposed Fast Track reference.

Figure 5. Storage System Components



The production, staging, and system temp databases were deployed per the recommendations provided in the *Microsoft Fast Track Data Warehouse 3.0 Reference Guide*.

Performance Benchmarking

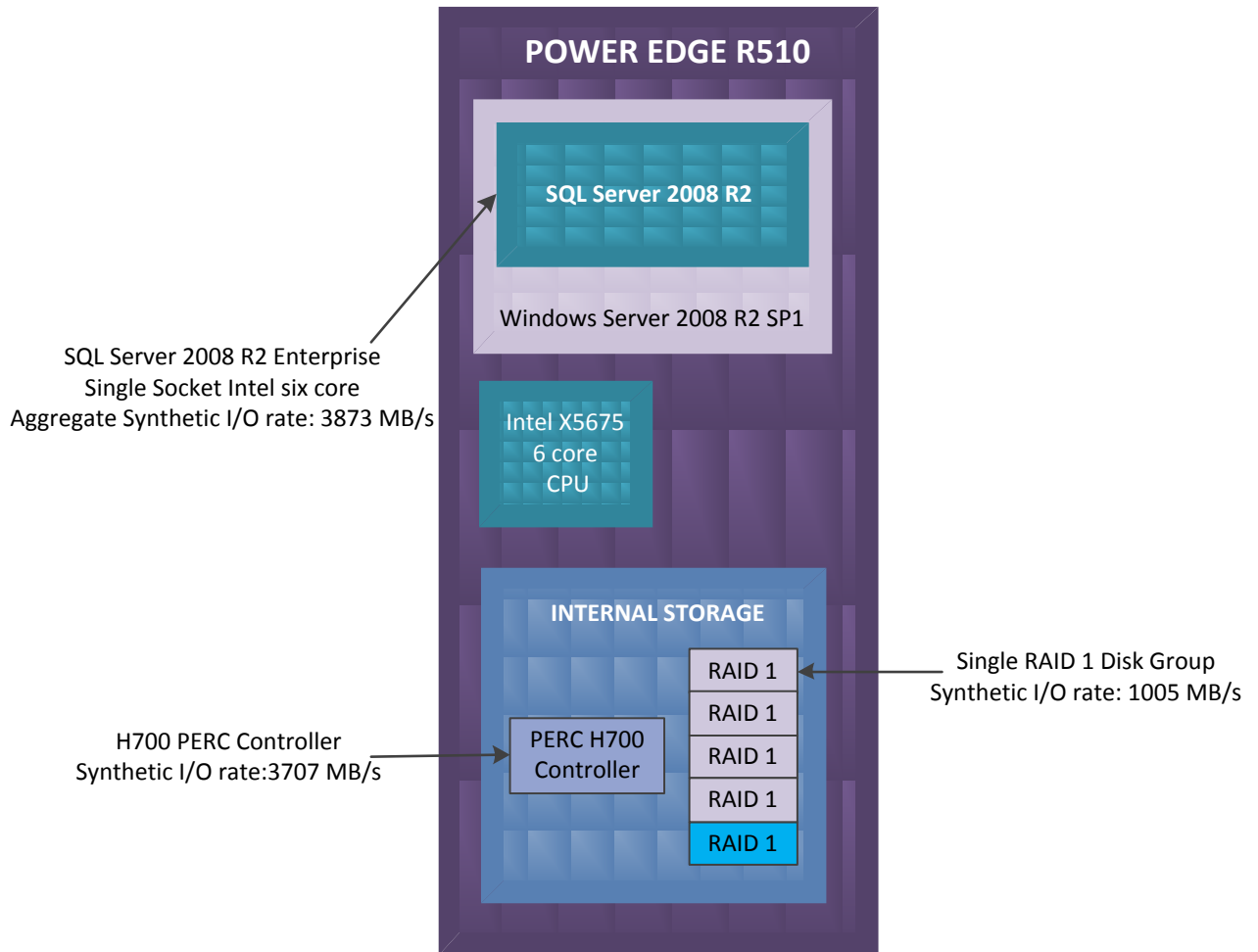
Microsoft Fast Track guidelines help to achieve optimized database architecture with balanced CPU and storage bandwidth. The following sections describe the performance characterization activities carried out for the validated Dell Microsoft Fast Track reference architecture.

Baseline Hardware Characterization using Synthetic I/O

You must thoroughly analyze the storage hardware to make sure that the storage backend is capable of delivering the maximum possible throughput. This will ensure that the performance of the system is not bottlenecked in any of the intermediate layers.

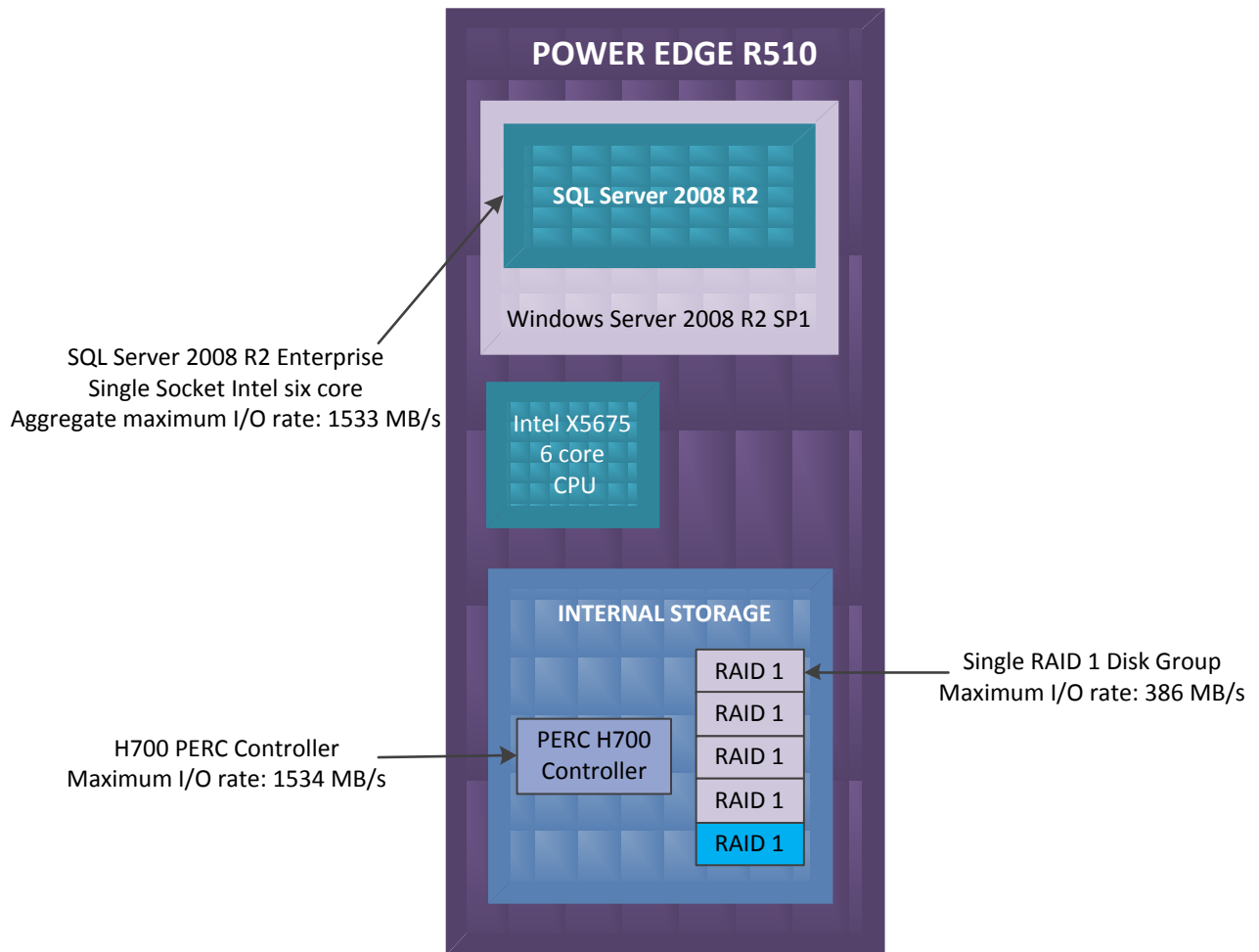
The disk characterization tool, SQLIO, was used to validate the configuration. Please refer to the *Fast Track Reference Guide* (link provided in the reference section) for detailed guidelines. Figure 6 shows the baseline performance numbers achieved for the validated reference architecture. The results show the maximum baseline that the system can achieve from a cache called Line Rate. A small file is placed on the storage, and large sequential reads are issued against it with SQLIO. This test verifies the maximum bandwidth available in the system to ensure no bottlenecks are within the data path.

Figure 6. SQLIO Line Rate Test from Cache (Small File)



The second synthetic IO test with SQLIO was performed with a large file to ensure reads are serviced from the storage system hard drives instead of from cache. This shows the maximum real rate that the system is able to provide with sequential reads.

Figure 7. SQLIO Real Rate Test from Disk (Large File)



Fast Track Workload Evaluation

The performance of a Fast Track database configuration is measured using two core metrics: Maximum CPU Consumption Rate (MCR) and Benchmark Consumption Rate (BCR).

Calculating MCR

MCR indicates the per core I/O throughput in MB or GB per second. This is measured by executing a pre-defined query against the data in the buffer cache, and measuring the time taken to execute the query against the amount of data processed in MB or GB. MCR value provides a baseline peak rate for performance comparison and design purposes.

For the validated configuration with one Intel X5675 six core processors, the system aggregate MCR was 1722 MB/s. The realized MCR value per core was 287 MB/s.

Calculating BCR

BCR is calculated in terms of total read bandwidth from the storage hard drives and not from the buffered cache as in the MCR calculation. This is measured by running a set of standard queries specific to the data warehouse workload. The queries range from I/O intensive to CPU and memory intensive, and provide a reference to compare various configurations.

For the validated Fast Track configuration, the aggregate BCR was 1021 MB/s.

During the evaluation cycle, the system configuration was analyzed for multiple query variants (simple, average and complex) with multiple sessions and different degrees of parallelism (MAXDOP) options to arrive at the optimal configuration. The evaluation results at each step were validated and verified jointly by Dell and Microsoft. A summary of the Fast Track results are offered in the Appendix.

Populating Your Data Warehouse

Integrating multiple data sources efficiently, easily, and accurately is essential to any enterprise data warehouse. Dell helps make this possible with the Dell Boomi Integration Cloud. As the recipient of the CODiE award for best cloud integration platform for the past three years, Dell Boomi is widely recognized as the industry's leading application and data integration provider.

By building our data and application integration platform in the cloud, we've created a fundamentally different approach. Dell Boomi helps to reduce TCO as compared to traditional integration tools or custom coding. Because Dell Boomi is delivered in the cloud, there is never any software to maintain, and minimal IT resources are required for ongoing management. In addition, our library of pre-built connectors eliminates the need to build and maintain custom code. Dell Boomi also strengthens data security because data never passes through the Boomi platform; data is only processed through the Atom runtime engine located in the cloud or in the customer's own infrastructure, so the customer can select where the data flows.

Dell Boomi's intuitive visual design interface makes it easy to begin creating integration processes immediately. Using familiar point-and-click drag-and-drop techniques, users can build very simple to very sophisticated integrations with exceptional speed. As the Dell Boomi community expands, so does the number of connectors and process maps, and many are already built and ready for use, easing the job of connecting new applications. Training costs are low due to the platform's simplicity and ease of use.

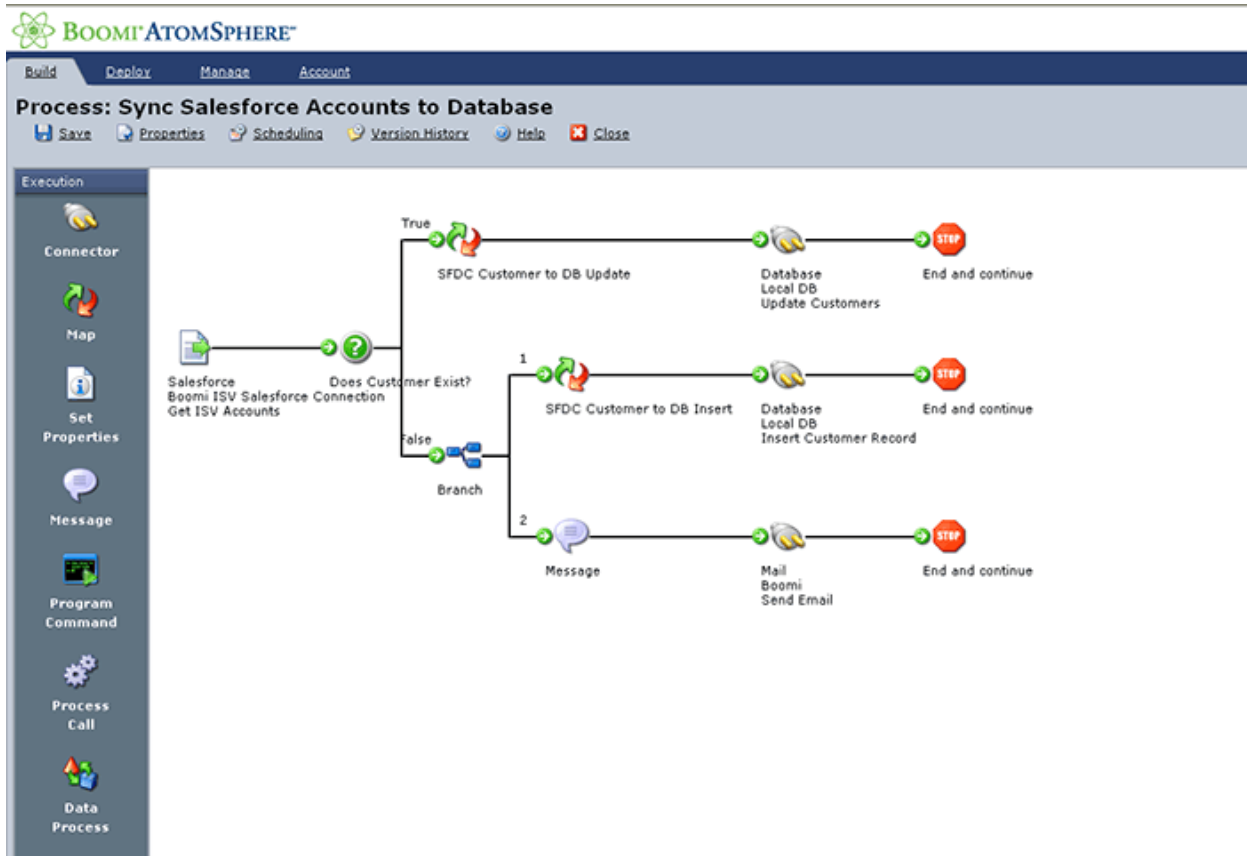
Dell Boomi

Dell Boomi is the market-leading provider of on-demand application and data integration services and the creator of AtomSphere®, the industry's #1 Integration Platform as a Service (iPaaS). AtomSphere connects providers and customers of SaaS, cloud and on-premise applications via a pure SaaS integration platform that does not require software or appliances. ISVs and businesses alike benefit by connecting to the industry's largest network of SaaS, PaaS, on-premise and cloud computing environments in a seamless and fully self-service model. Dell Boomi allows customers to create real-time (SOA), B2B and ETL-based integration processes.

Building an Integration Process

Since its inception, Boomi has focused on simplifying the creation of integration processes for application integration, data integration (ETL), and B2B integration. By identifying the common steps needed to automate complex integration scenarios, a series of common integration components have been created and are available to all Boomi users. When developing an integration process, these components are connected to create an end-to-end integration workflow.

Figure 8. Dell Boomi



Standard Boomi integration components include:



Connector

connect to any application or data source

Always the first and last steps of an integration workflow, the Connector enables access to another application or data source. The connector sends/receives data and converts it into a normalized XML format. A Connector's primary role is to "operationalize an API" by abstracting the technical details of an API and providing a wizard-based approach to configuring access to the associated application. Connectors are also configurable to capture only new or changed data, based on the last successful run date of an integration process.



Data Transformation

transform data from one format to another

While core to any integration, the data stored in various applications is rarely, if ever, semantically consistent. For example, a customer record represented in one application will have different fields and formats from that of another application. Using Boomi's Data Transformation components, users can map data from one format to another.

Any structured data format is supported, included XML, EDI, flat file, and database formats. While transforming data, the user can also invoke a variety of field-level transformations to transform, augment, or compute data fields. Over 50 standard functions are provided. Users can also create their own functions and re-use them in subsequent projects.



Decision

execute business logic and data integrity checks

Boomi's Decision components provide true/false data validation that enables users to explicitly handle a result based on the programmed logic. For example, an order can be checked against the target system to see if it has already been processed. Based upon the outcome of the data check, the request will be routed down either the 'true' or 'false' path. Another example includes checking products referenced in an invoice to ensure they exist before processing the invoice.



Cleanse

data cleansing and validation

Integrations are only as successful as the quality of data that gets exchanged. Boomi's Cleanse components allow users to validate and "clean" data on a field-by-field, row-by-row basis to ensure that fields are the right data type, the right length, and the right numeric format (e.g. currency). Users

have the option of specifying whether they wish to attempt to auto-repair bad data, or simply reject rows that are “dirty”. All validation results are routed through a “clean” or “rejected” path which allows users to explicitly handle either scenario.



Message

user-defined dynamic notifications

For any step of the integration workflow, Message components can be used to create dynamic notifications that make use of content from the actual data being integrated. This allows the creation of messages like "Invoice 1234 was successfully processed for customer ABC Inc." Connectors are then used to deliver the message to the appropriate end point.



Route

dynamic, content based routing

Route components examine any of the content in the actual data being processed, or use numerous other properties available to the user (such as directory, file name, etc.) and route the extracted data down specific paths of execution.



Split

intelligent data splitting and aggregation

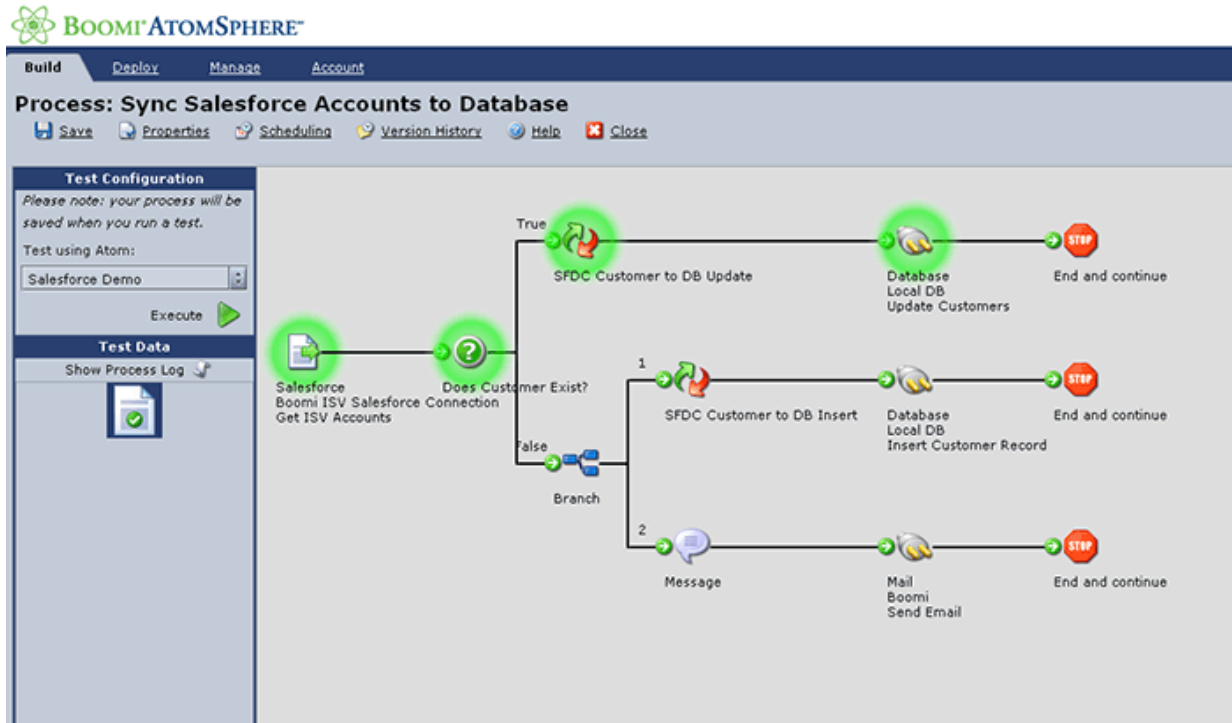
Split components re-organize data into logical representations such as business transactions. For example, you can take incoming invoice headers and details, and create logical invoice documents to ensure the applications being integrated process or reject the entire invoice vs. discrete pieces.

Testing Integrations

real-time

To further simplify and shorten the integration development cycle, users can test their integrations directly in the build environment. Integration testing includes watching as the data moves through the integration process, viewing the actual data, monitoring exactly where the integration fails, and click-thru on the failed step to examine error messages.

Figure 9. Dell Boomi Build Environment Real-Time Integration Testing



Configuring Integration Automation

Integration processes can be configured for “hands off” execution. AtomSphere provides both real-time as well as batch-style automation options depending on the requirements of the integration, as defined below:

- Event-based Invocation - Direct Atom Invocation

Included in every Boomi Atom is a lightweight HTTP Server. Data can be HTTP-posted to a specific Atom, and that data will be processed in real time.

- Event-based Invocation - Remote Atom Invocation

Boomi also provides a Trigger API, which allows you to securely invoke an integration process that is running inside an Atom, regardless of where that Atom may be running, without opening any holes in the firewall. This is a very powerful option when you wish to provide external access to integration processes or trigger an integration based on some event in another application.

- Schedule-based Invocation

Included in every Boomi Atom is a schedule manager, capable of invoking integration processes based on a schedule configured by the user. Invocations can be scheduled to run as frequently as every minute. The schedule-based invocation option requires no changes to external applications.

Deploy an Integration Process

Once your integration processes are built, they are deployed into lightweight runtime engines known as Boomi Atoms. The Boomi Atom is the "secret sauce" to Boomi's technology, and allows your integration processes to run wherever needed and as many times as needed, enabling nearly infinite scalability. The Boomi Atom contains within it one or more complete end-to-end integration processes, and executes those integration processes with absolutely no dependency on Boomi's platform. This allows for several deployment options to accommodate any integration requirement.

- SaaS to On-premise

On-premise applications are typically kept behind a firewall, with no direct access via the Internet, and no access even via a DMZ. To handle this requirement, the Boomi Atom can be deployed on-premise to directly connect the on-premise applications with one or more SaaS/Cloud applications. No changes to the firewall are required, and the Atom supports a full bi-directional movement of data between the applications being integrated. In this deployment style, your data does not enter Boomi's data center at any point.

- SaaS to SaaS

When all the applications being integrated can be accessed via a secure Internet connection, the Boomi Atom can run in Boomi's cloud for a zero-footprint deployment. Boomi manages the uptime of the Atom in this configuration. Your data is completely isolated from any other tenants in Boomi's platform.

- Cloud to SaaS/Cloud

If you are hosting your own applications either as an ISV or for internal use, you can deploy the Boomi Atom into any cloud infrastructure that supports Java, such as Amazon, Rackspace, OpSource, etc. This offers direct connectivity between your application and the applications you wish to connect. In this deployment style, your data does not enter Boomi's data center.

Managing Changes to Integration Processes

Because applications are inherently dynamic in nature, so are the integrations that connect them. AtomSphere provides significant capabilities for managing integration changes in a very simple, non-disruptive, and auditable manner. Using a "Test Mode", integration changes can be tested and verified without impacting any integrations currently running in production. Once tested, the new integration process can be deployed to production to replace the outdated process.

For more advanced testing requirements, AtomSphere provides Environments functionality. Environments are distinct and persistent collections of Atoms, enabling you to separate integration elements such as test versus production login credentials. Environments are the recommended approach when the applications being integrated have dedicated test/QA environments. Once defined and tested, processes can easily be promoted from one environment to another, with a full audit trail of what was promoted, when it was changed, and by whom.

Addressing Enterprise-Class Integration

Business critical integration processes demand a different level of reliability and scalability. To ensure that key integration processes are reliable and highly available, the Enterprise Edition of AtomSphere features Molecules, an enterprise-grade version of Boomi's patent-pending Atom™ technology. When deployed across multiple physical servers, Molecules enhance load balancing and ensure the reliability of mission critical integration processes.

Manage Your Integration Processes

Regardless of where Boomi Atoms are deployed or how many processes are deployed, the Atom's unique architecture enables the centralized management of ALL integrations from the Boomi AtomSphere platform. The Manage functionality of the Boomi platform enables users to monitor the health and activity of all Atoms, review detailed logs of what processes ran and when, how long they took to run, the result, how many objects were processed, etc.

Figure 10. Dell Boomi - Process Execution View

The screenshot shows the Boomi AtomSphere 'Process Execution View'. On the left, the 'Atom Explorer' panel lists several integration processes: Salesforce Demo (green checkmark), Salesforce Demo 2 (red X), Salesforce Demo 3 (red X), Salesforce Demo 4 (green checkmark), and Salesforce_Peachtree Atom (green checkmark). The main 'Executions' panel shows a table with the following data:

Time	Process	Atom	In	Out	Details
01/12/2009 3:26:15 PM	Synch Database Accounts & Address to Salesf	Salesforce Demo	1	4	Log

Below the execution table, there are sections for 'Inbound Data' and 'Outbound Data'. The 'Inbound Data' section shows a table with columns for 'Actions' and 'Size (kb)'. The first row shows an action with a size of 0.91 kb.

Reviewing Overall Integration Performance

The Process Execution view allows users to filter by date every integration process that ran in their account. Users can track individual integrations by status, which Atom ran the integration, how much data was processed, and how long the integration took to run. On demand, users can click-thru to a detailed log of the execution and the Atom will temporarily serve this log up to the browser session.

Reviewing Individual Process Executions

For each execution of an individual integration process, users can view the details of that execution, including each individual data set that was processed, both inbound and outbound. After viewing the execution details, there are several options:

- View the actual data that was sent or received with an external application (dependent on security permissions.) This data will be retrieved dynamically because data sets are never stored in Boomi's data center.
- Retry the data, which will request the Atom to retry the data sets you have selected, and pass them to a new process execution.
- View the link between inbound and outbound data sets. When processing large amounts of data, users can click on an individual data set to view the resulting data that was created from it.

Exploring Atoms

Regardless of where Atoms are deployed, the Atom Explorer is the central location to view the status of the Atoms, see where they are running, and set rules on how much history the Atom should retain. An Atom will appear offline when it has not communicated with the platform within a 5-minute window. You can use the Atom Explorer to access and manually execute the integration process in a specific Atom.

Subscribing to Process Alerts

For proactive notification of failures, users can subscribe to Alerts that broadcast via RSS any failed integration process, or any unresponsive Atom. For a simple view of all integration activity in an account, users can also subscribe to a monitor feed, and embed this into any another application that supports RSS. Alerts provide the process name, the Atom, the status (success or failure), and the date/time of the execution. By using this alert, users can quickly pinpoint the execution history for detailed diagnostic activity. Atom Alerts will also send notifications when an Atom goes offline, or comes back online, specifying the Atom name, the status, and the date/time of the status change.

Conclusion

The Dell Microsoft Fast Track Data Warehouse architecture provides a uniquely well-balanced data warehouse implementation solution. By following the best practices at all the layers of the stack, a balanced data warehouse environment can be achieved with a greater performance benefit than the traditional data warehouse systems.

The Dell Microsoft Fast Track Architecture provides the following benefits to customers.

- Tested and validated configuration with proven methodology and performance behavior.
- A balanced and optimized system at all levels of the stack by following the best practices of hardware and software components.
- Avoidance of over-provisioning of hardware resources.
- High availability at all the levels of setup (host, Switches and storage).
- Help to avoid the pitfalls of improperly designed and configured system.
- Reduced future support costs by limiting solution re-architect efforts because of scalability challenges.
- Dell BOOMi integrated on-demand data and application integration service to maximize data source flexibility while minimizing cost.

References

Dell SQL Server Solutions

www.dell.com/sql

Dell Services

www.dell.com/services

Dell Support

www.dell.com/support

OLTP and OLAP

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

Microsoft Fast Track Data Warehouse and Configuration Guide Information

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>

How to: Enable the Lock Pages in Memory Option

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

SQL Server Performance Tuning & Trace Flags

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

Power Edge R510 Technical Guide

<http://www.support.dell.com/support/edocs/systems/per510/en/index.htm>

Dell Boomi

<http://www.boomi.com>

Appendix

Summary of Microsoft SQL Server Fast Track results

The performance of the system configuration was analyzed with RAID1 for simple, average & complex query variants with 5, 10, 20, 40 sessions. Various maximum degrees of parallelism (MAXDOP) options were tested to arrive at the optimal configuration for this workload with MAXDOP=6.

Following table shows the results for simple work load with MAXDOP = 6 (4 disk groups, RAID 1)

Session	Physical I/O throughput total (MB/s)	Logical I/O throughput (MB/s)
5 Session	915	1093
10 Session	939	1119
20 Session	919	1224
40 Session	872	1378

The results in the following table are for medium work load with MAXDOP = 6 (4 disk groups, RAID 1)

Session	Physical I/O throughput total (MB/s)	Logical I/O throughput (MB/s)
5 Session	704	1009
10 Session	767	1145
20 Session	756	1183
40 Session	650	1038

The following results show the benchmarked rates from memory, physical disks and peak for the validated solution. These results can be used to compare solutions and choose an appropriate configuration that meets the throughput requirements.

Benchmark Scan Rate Logical MB/s	Benchmark Scan Rate Physical MB/s	FTDW Rated I/O MB/s	FTDW Peak I/O MB/s
1204	838	1021	1378