White Paper

Dell Reference Configuration

Deploying Microsoft® SQL Server™ 2005 Standard Edition with SP2 using the Dell™ PowerVault™ MD3000i iSCSI Storage Array

Abstract

This white paper provides an architectural overview and configuration guidelines for deploying Microsoft SQL Server 2005 on Dell PowerEdge servers with the Dell PowerVault MD3000i iSCSI storage array. Using the knowledge gained through joint development, testing and support with Microsoft, this white paper documents "best practices" that can help speed SQL Server 2005 solution implementation and help simplify operations, improve performance and availability. The white paper also evaluates the performance of the MD3000i and demonstrates that it is well-suited as a storage system for OLTP-type workloads.

October, 2007

October 2007 Rev. A01

Table of Contents

ABSTRACT **1**

INTRODUCTION **4**

DELL SOLUTIONS FOR SQL SERVER 2005 4
OVERVIEW OF THIS WHITE PAPER 4

ARCHITECTURE OVERVIEW **5**

HARDWARE CONFIGURATION **6**

STORAGE CONFIGURATION 6
Configuring Storage Connections *6*
Configuring Disk Groups and Virtual Disks *8*
SERVER CONFIGURATION 10
Configuring NIC Teaming *10*
Configuring NICs for iSCSI Connections *10*

SOFTWARE CONFIGURATION **11**

OPERATING SYSTEM CONFIGURATION 11
Establishing iSCSI Sessions *11*
Enabling Disk Partition Alignment *12*
Configuring the File System *13*
Optimizing the Server for Network Applications *14*
SQL SERVER 2005 CONFIGURATION 14
Configuring TempDB Files *14*

PERFORMANCE CHARACTERIZATION AND ANALYSIS **16**

SOLUTION DELIVERABLES LIST **17**

CONCLUSIONS **19**

TABLES AND FIGURES INDEX **20**

REFERENCES **20**

Introduction

Dell™ PowerEdge™ servers and Dell PowerVault™ storage systems are ideal choices for deploying highly reliable and sustainable Microsoft® SQL Server™ 2005 databases. This white paper is intended to help IT professionals design and configure SQL Server 2005 database solutions – using Dell servers and storage systems – that apply "best practices" derived from laboratory and real-world experiences. This white paper documents Dell's recommended approach for implementing a tested and validated solution for SQL Server 2005 SP2 database on Dell PowerEdge 9th generation servers, Dell PowerVault MD3000i iSCSI storage arrays, and Microsoft Windows Server 2003 R2 Standard Edition with SP2.

Dell Solutions for SQL Server 2005

Dell Solutions for SQL Server 2005 are designed to simplify operations, improve utilization and cost-effectively scale as your needs grow over time. In addition to providing price/performance leading server and storage hardware, Dell solutions for SQL Server 2005 include:

- **Dell Tested and Validated Configurations for SQL Server 2005** – in-depth testing of SQL Server 2005 configurations with Dell servers and storage; documentation and tools that help simplify deployment
- **Integrated Solution Management** – standards-based management of Dell Solutions for SQL Server 2005 that lower operational costs through integrated hardware and software deployment, monitoring and update
- **SQL Server Licensing** multiple licensing options that can simplify customer purchase
- **Dell Enterprise Support and Professional Services for SQL Server** – offerings for the planning, deployment and maintenance of Dell Solutions for SQL Server 2005

Dell PowerEdge servers and Dell PowerVault Storage help to minimize operating costs with price/performance leadership - Dell currently holds price/performance leadership for TPC-E and seven of the top ten TPC-C price/performance leadership positions with SQL Server 2005.[1]

For more information concerning Dell Solutions for SQL Server 2005, please visit www.dell.com/sql .

Overview of this White Paper

The balance of this white paper will provide the reader with a detailed view of deploying Microsoft SQL Server 2005 with the Dell PowerVault MD3000i iSCSI storage array. It will provide an overview of the tested reference configuration, best practices for configuring the hardware and software components, and pointers for obtaining more information.

[1] Source: TPC-E by Price/Performance Version 1 and TPC-C by Price/Performance Version 5. Results as of October 2007. See www.tpc.org for current results.

Architecture Overview

A sample architectural overview of the Dell Solution for SQL Server 2005 is shown in Figure 1. In this figure, the Dell PowerVault MD3000i and a Dell PowerEdge server are interconnected using a direct-attached topology. The iSCSI connections can also be made via an IP SAN, as illustrated in Figure 2. The architecture consists of the following components:

- Client systems that will access data stored within the SQL Server database (**Note:** in an *n*-tier architecture, this also includes other server systems that access the database)
- A client-server network made up of network controllers, cables and switches
- Dell PowerEdge servers running Windows Server 2003 and SQL Server 2005
- Server-storage interconnect using Internet SCSI (iSCSI), including the following:
  - Gigabit Ethernet adapters
  - Microsoft iSCSI Software Initiator
  - Gigabit Ethernet switches (for IP SAN configurations)
- Dell PowerVault MD3000i storage system with up to two MD1000 expansion enclosures, supporting up to 45 hard-disk drives

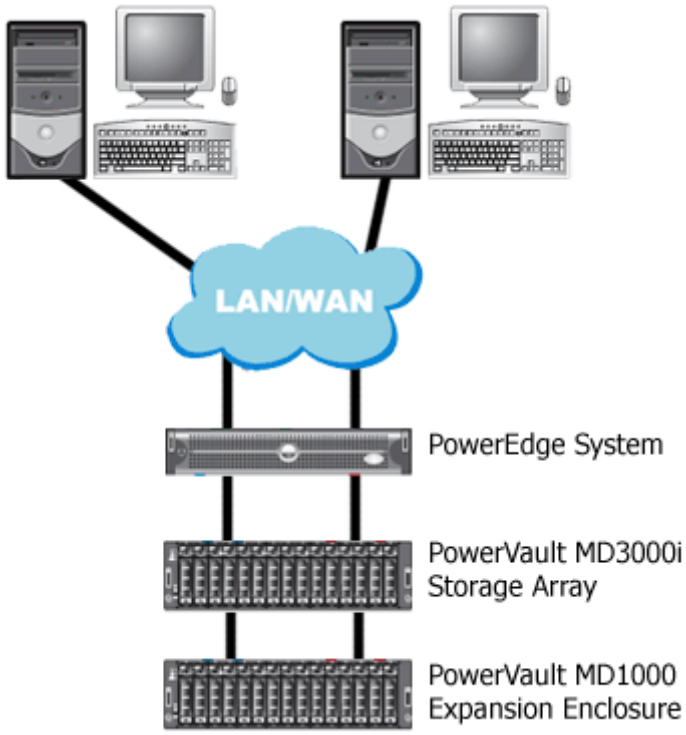

Figure 1 - Architectural Overview of Dell SQL Server 2005 configuration

Dell PowerEdge servers are designed to deliver the highest performance for critical enterprise applications including database, messaging, web services, and infrastructure applications. As proprietary systems are increasingly replaced by industry-standard systems, applications such as databases, high performance computing clusters, and messaging systems can take advantage of the performance and scalability of the PowerEdge servers. Combined with Dell storage systems, customers can easily deploy these PowerEdge servers as building blocks of a scalable enterprise, deploying solutions that combine computing resources, storage resources, and software.

The PowerVault MD3000i is an iSCSI storage array built with the high-availability features that make it suitable for deploying mission-critical applications hosted on Dell PowerEdge servers. In a duplex – or redundant – configuration, the MD3000i includes two RAID controllers and employs a mirrored write

cache. The storage system also provides multi-path I/O management which helps ensure that storage processing continues without disruption. Other high-availability features include hot-pluggable, redundant power supplies, cooling modules and disk drives, active disk scrubbing and non-disruptive firmware upgrades. These features are consistent with the requirements for the Microsoft SQL Server Always On program[2], and Dell has prepared a statement of compliance with SQL Server Always On Technologies for the MD3000i which is available at www.dell.com/sqlalwayson.

Because each MD3000i RAID controller provides two iSCSI ports for host connections, it is possible to directly-attach either one or two host servers using redundant data paths. When deployed in an IP SAN, with gigabit Ethernet (GbE) switches, up to 16 host servers are supported. The Dell Modular Disk Storage Manager (MDSM), which is used to manage the MD3000i, makes it possible to assign virtual disks to a specified server or set of clustered servers.

The MD3000i can be configured with up to two MD1000 expansion enclosures to increase the overall capacity of the storage array or to fulfill the hard-disk drive count requirements of certain applications, including databases. When using an IP SAN, a Dell PowerEdge server can connect to, and use virtual disks from, up to four MD3000i storage arrays. This provides scalability up to 180 total hard-disk drives with an aggregate raw capacity in excess of 70 TB[3].

Hardware Configuration

Storage Configuration

In order to successfully deploy the MD3000i storage system into a SQL Server 2005 database solution, it is important that the enclosure be connected and configured appropriately.

Configuring Storage Connections

Figure 2 illustrates a reference IP SAN implementation of a PowerEdge server and an MD3000i storage array where the data and log files for the SQL Server 2005 database reside. Using standard GbE NICs (optionally with TOE enabled) and GbE switches in a redundant configuration, commands and data flow over multiple iSCSI links between the server and the storage array. The multi-path I/O driver provides failover for connections to alternate RAID controllers and allows load balancing across connections to the same RAID controller. This means that the I/O load for a given MD3000i virtual disk can be distributed between the GbE NICs in the server when the solution is configured in this manner. In the figure, two separate and isolated networks are used for iSCSI traffic. Such a configuration isolates iSCSI traffic from all other network traffic, ensures redundancy, and simplifies deployment. However, it is also possible to employ VLANs or use other techniques to segment network traffic appropriately.

If the MD3000i and the server are connected directly, the host should have at least one connection to each RAID controller to ensure that failure of a GbE NIC port, a cable, or an MD3000i storage controller does not disrupt communication between the server and the storage array. A simplified rendition of this configuration appears in Figure 1.

[2] For information about the SQL Server Always On program, see http://www.microsoft.com/sql/alwayson.
[3] RAW capacity does not account for space required for RAID redundancy or hot spare disks. MB, GB, and TB are defined as multiples of 1000, and this capacity requires the use of 400 GB 10k rpm disks.

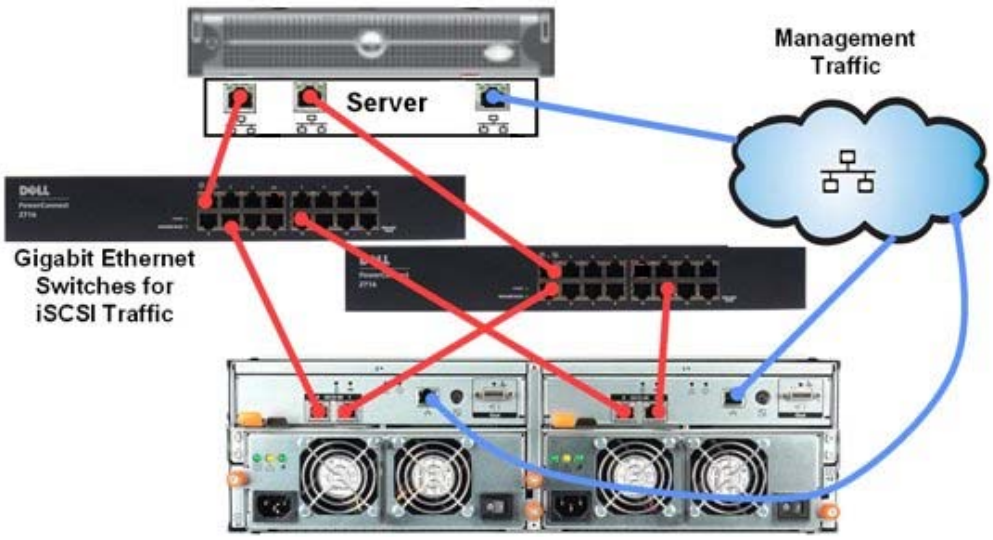

Figure 2 - Cabling an MD3000i Storage Array in an IP SAN

As many as two MD1000 enclosures can be cascaded from the MD3000i (see Figure 3), provided that the MDSM software, the EMM BIOS and firmware, and the cabling methodology are appropriate. For more information, consult the documentation for the PowerVault MD3000i array.

WARNING: If an MD1000 enclosure that is being used for expansion was previously connected to a server using a Dell PERC 5/E host-based RAID controller, the data on that enclosure will *not* be preserved when it is attached to the MD3000i. Instead, data should be backed-up so that it can later be restored on the expanded MD3000i system. All existing data on the MD1000 physical disks must be cleared before attaching the expansion enclosure to the MD3000i. The disks can be cleared either by clearing the MD1000 configuration, or by deleting all logical drives and hot spares from the MD1000; either of these tasks should be performed while the MD1000 is still connected to the PERC 5/E. New virtual disks that make use of the hard-disk drives in the MD1000 will need to be created from the MD3000i; this task is performed by using the MDSM management software.

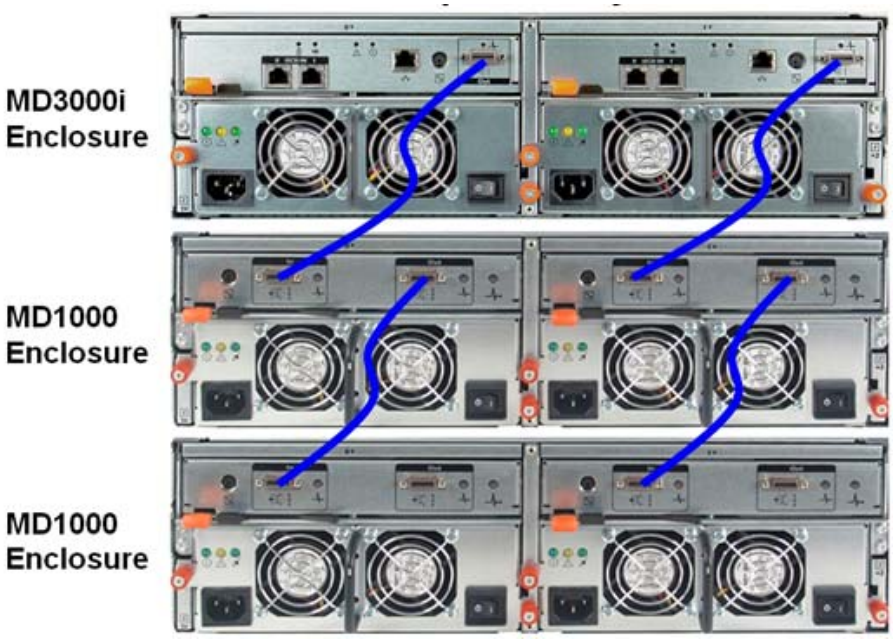

Figure 3 - MD3000i System with two MD1000 Enclosures Cascaded

Configuring Disk Groups and Virtual Disks

The physical disks in the MD3000i storage array provide the physical storage capacity for the SQL Server database. Before data can be stored, the MD3000i physical storage capacity must be configured into components, known as disk groups and virtual disks. A disk group is a set of physical disks that are logically grouped and assigned a RAID level. Each disk group created provides the overall capacity needed to create one or more virtual disks, which are logical entities that the server uses to store data.

SQL Server 2005 uses three specific storage areas; TempDB, Transaction Log File, and Data files for user defined databases. TempDB is a system database which is automatically created during the installation of SQL Server 2005. TempDB is a shared working area for all databases on the server, for various activities, including temporary tables, sorting, processing subqueries, building aggregates to support GROUP BY or ORDER BY clauses, queries using DISTINCT, cursors, and hash joins. In addition to TempDB, SQL Server 2005 also maintains a Transaction Log File for each database. The Transaction Log File stores the details of all the modifications made to a SQL Server database and the details of the transactions that performed each modification. This information is critical for maintaining database consistency and for aiding database recovery. In SQL Server 2005, data for user defined databases are stored in a separate data storage area. This data file contains user defined database objects such as tables, indexes, and stored procedures. All databases have at least one data file and at least one transaction log file. Data from the database tables is physically stored within the data file(s).

It is a best practice to separate the TempDB and Transaction Log files onto their own virtual disks on separate disk groups. The separation of TempDB and log files can enable better I/O performance by ensuring that these files do not share the same physical disks. Figure 4 illustrates a sample disk group and virtual disk configuration which separates TempDB, transaction logs, and data on distinct physical disks.

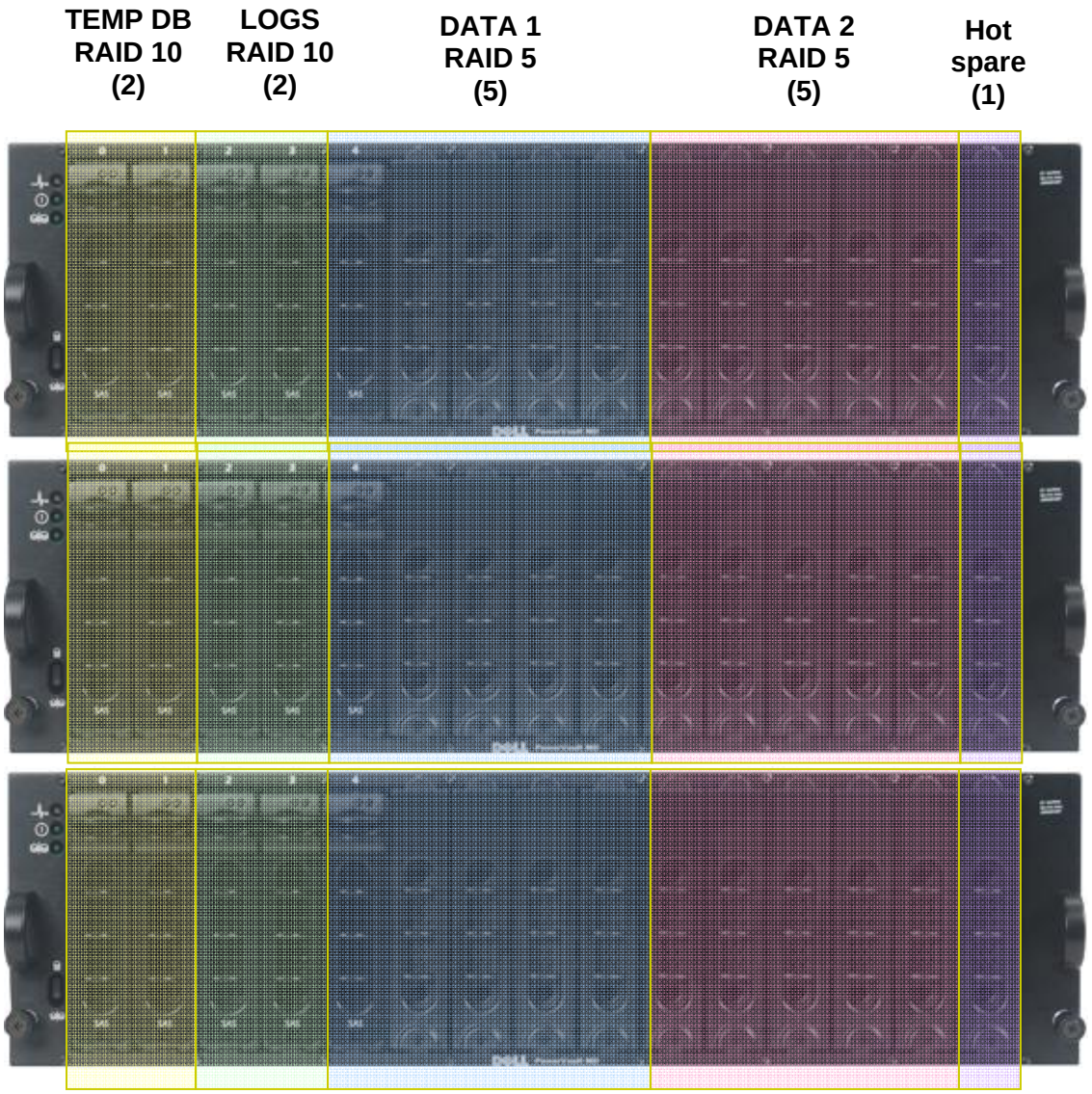

Figure 4 - Separation of Disk Groups and Virtual Disks within a PowerVault MD3000i and two MD1000 Storage Arrays for TempDB, Transaction Logs, and Data optimized for Data Warehousing

The recommended number of virtual disks for data should equal the number of physical CPU sockets[4].

Table 1 illustrates recommended virtual disk configuration and sizes for single-socket host servers.
Table 2 illustrates recommended virtual disk configuration and sizes for dual-socket host servers.

Virtual Disks	Minimum Size	Disk Group RAID	Used for
Data1	Database size	10, or 5 for read-only	data
TempDB	Depending on required TempDB size	10, or 0	TempDB
Logs	10% of database size	10, or 1	logs

Table 1 - Single-Socket Server - Storage Configuration

[4] "Storage Top 10 best Practices", Microsoft TechNet, October 17, 2006; http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/storage-top-10.mspx

Virtual Disks	Minimum Size	Disk Group RAID	Used for
Data1	Database size /2	10, or 5 for read-only	data
Data2	Database size /2	10, or 5 for read-only	data
TempDB	Depending on required TempDB size	10, or 0	TempDB
Logs	10% of database size	10, or 1	logs

Table 2 - Dual-Socket Server - Storage Configuration

RAID 10 is considered the optimal choice for virtual disks used in a SQL Server 2005 implementation because it offers a good mix of performance and fault tolerance by combining attributes of mirroring and striping[5]. When possible, the disk groups on which the transaction logs, TempDB and data virtual disks reside should be configured with RAID 10.

Because additional drives are required to implement RAID 10, it may not be the most practical choice for certain database deployments. In these cases, the following recommendations apply. RAID 0 may be used for the disk group containing the TempDB virtual disk. However, if RAID 0 is used, system availability will be impacted if a TempDB disk is lost. For disk groups that contain transaction log virtual disks, RAID 1 provides protection from drive hardware failure, and only requires 2 physical disks. For disk groups providing virtual disks for the data files, RAID 5 can provide a cost-effective alternative, especially for predominantly read-only workloads such as a data warehouse database. However, RAID 5 is not optimal for write-intensive workloads, such as in an OLTP database; RAID 5 tends to have significantly lower write performance than RAID 10, due to the reading and writing of parity blocks in addition to the reading and writing of database data.

Each virtual disk created on the MD3000i storage array should be mapped to a single partition and a single NTFS volume. For details on operating system partitions, please refer to the "Configuring File System" section below.

Server Configuration

An optimal database solution requires proper configuration of not only the storage system, but also of the server hardware and software.

Configuring NIC Teaming

To help guard against network access failures, PowerEdge servers running SQL Server 2005 can be configured to provide redundant links to the client network. Using NIC teaming software, two NIC interfaces of the PowerEdge server can be bonded together to operate under a single IP address. The NIC teaming software provides load balancing and failover functionality, balancing the workload and routing network traffic between the two NIC interfaces. If a failure occurs, affecting one of the NIC interfaces – examples include switch port failure, cable disconnection, or failure of the NIC itself – network traffic is routed to the remaining operable NIC interface. Failover occurs transparently to the SQL Server 2005 database with no network communication interruption or changes to the server IP address.

Configuring NICs for iSCSI Connections

As illustrated in Figure 2, it is recommended that at least two GbE ports on the PowerEdge server hosting SQL Server 2005 be used for iSCSI connections; such a configuration provides redundant links to the PowerVault MD3000i storage array by means of the MPIO driver, and NIC teams should not be used for the iSCSI interfaces. Dual-port GbE NICs are supported, but using ports on separate adapters can provide additional protection from potential data loss or corruption in case of an adapter failure. Single GbE port, cable, switch, or MD3000i storage controller failures are isolated when the recommended connection methodology is applied. If the GbE NICs used for the iSCSI connection feature a TCP Offload Engine (TOE), it can be enabled; tests have shown that TOE provides less than a 5% increase in iSCSI throughput.

[5] "Physical Database Storage Design", Kathy Lu, Lewis Bruck, Microsoft TechNet, February 12, 2007. http://www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx

Software Configuration

Operating System Configuration

Disks from the MD3000i storage array are made available to Windows Server 2003 so that they can be used to store SQL Server 2005 data, log, or TempDB files. These disks must be configured within the Windows Server 2003 operating system to provide an optimal environment for the SQL Server 2005 solution. Other operating system and SQL Server software parameters can also be configured to optimize the performance of the overall solution.

Establishing iSCSI Sessions

In order to use the virtual disks provided by the MD3000i storage array, the Microsoft iSCSI Initiator software must be installed on the server. The initiator makes it possible to use common GbE adapters to communicate with the iSCSI targets provided by the MD3000i. When establishing iSCSI sessions between the SQL 2005 database server and the storage system, it is recommended to enable the use of multi-path I/O (MPIO), and to specify the source and target ports in the Advanced Settings page of the initiator software. Figure 5 shows these settings in the Microsoft iSCSI Initiator software; the drop-down list shows that a total of four available target ports, because this system was configured as illustrated in Figure 2. Using MPIO and creating multiple iSCSI sessions provides failover between, and load balancing among, the iSCSI sessions that connect the database server and the MD3000i storage array.

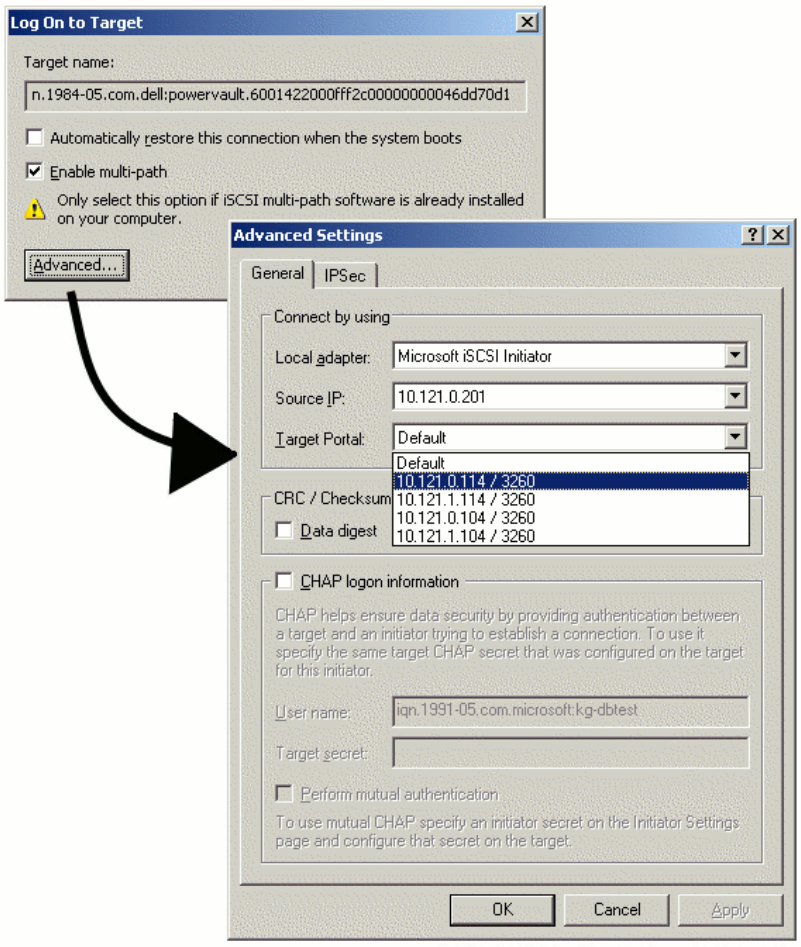

Figure 5 - Microsoft iSCSI Software Initiator Configuration

Enabling Disk Partition Alignment

Performance of the operating system and applications is affected by the placement of metadata on the virtual disk, which causes disk partitions to be misaligned. The fdisk utility and Windows Disk Manager place a Master Boot Record (MBR) on every Windows "Physical Disk" device. On a PowerVault MD3000i storage array, a track is equivalent to a segment or a stripe element with the default size of 128 KB (or 256 blocks). The MBR specifies 63 hidden sectors on the device. This causes subsequent data structures to become misaligned with respect to the track boundary. If a single I/O operation crosses a track boundary, it can consume added resources or cause additional work by the storage array leading to performance degradation (e.g. two separate writes may be required). Figure 6 illustrates a misaligned disk partition where the 64 KB I/O has spread to the first and the second MD3000i tracks.

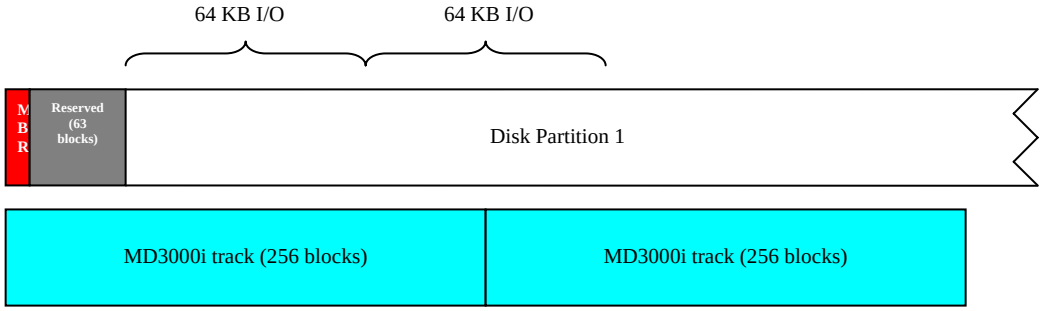

Figure 6 - Misaligned Disk Partition

In Windows Server 2003 SP1 or later, the diskpart utility is available for aligning partitions. Diskpart allows the creation of a primary partition at any desired block address rather than the default block address of 63. This enables partitions to be properly aligned to minimize performance loss due to track boundary crossing. Figure 7 illustrates an aligned disk partition where the partition starts at block address 128 or 64 KB, instead of the default partition location of block address 63.

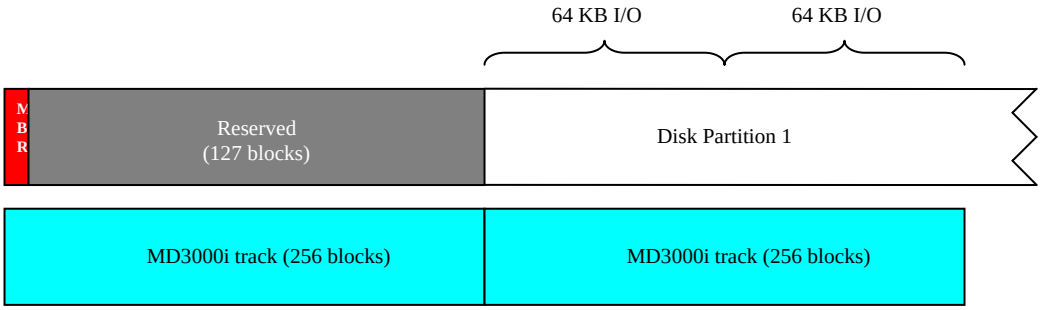

Figure 7 - Aligned Disk Partition

Follow the procedure described in the "Dell PowerEdge Systems Dell|Microsoft SQL Server 2005 SP2 Database on Microsoft Windows Server 2003 Standard Edition with SP2 Deployment Guide" to align the primary partition for Windows Server 2003 basic disks on MD3000 storage arrays. For more information on aligning partitions, see the article "Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks", EMC PowerLink, September, 2005 at the following web location - https://powerlink.emc.com/nsepn/webapps/btg548664833igtcuup4826/km/live1/en_US/Offering_Technical/White_Paper/H1445_Using_Diskpar_Align_Partition_ldv.pdf

Configuring the File System

In SQL Server 2005, each database consists of at least one data file and at least one transaction log file. As mentioned in the Section "Configuring Disk Groups and Virtual Disks" above, the recommended number of virtual disks for data should equal the number of physical CPU sockets, and each virtual disk created in the storage system is mapped to the operating system as a Windows "physical disk" with a single partition. To achieve optimal performance that scales with heavy workloads, Microsoft recommends that the number of data files configured for a SQL Server 2005 database equal the total number of CPU cores installed on the server[6]. Therefore, within each of the NTFS volumes used for data, the recommended number of data files should equal the number of cores per CPU socket.

Each partition on a Windows "physical disk" is assigned a unique drive letter. It is recommended that all partitions be formatted using the NTFS file systems, and employ 64 KB clusters.

For example, a Quad-Core, single-socket system should have four data files for each data partition as illustrated in Figure 8 and Table 3.

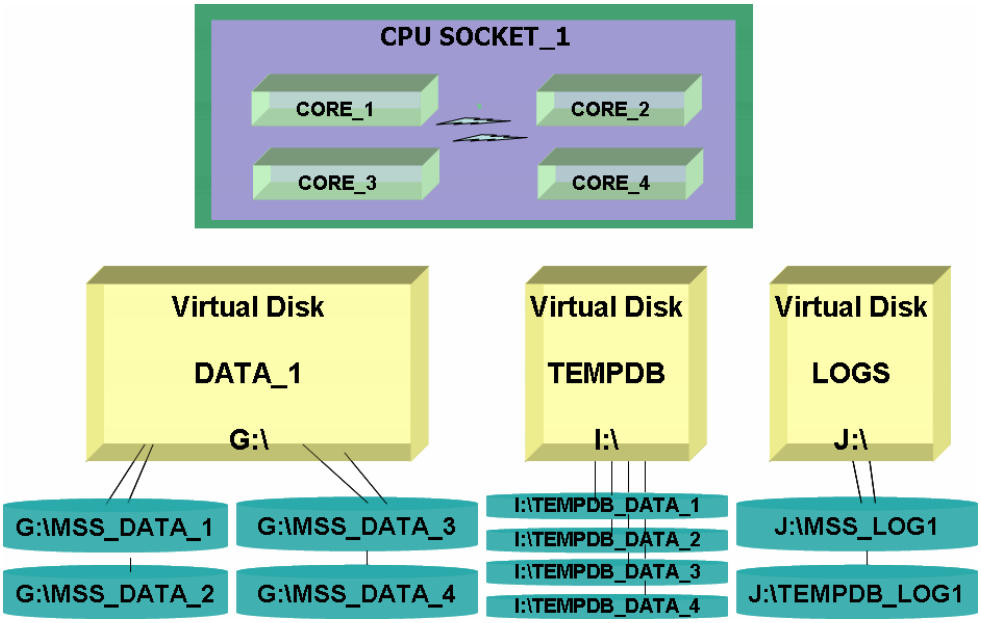

Figure 8 - Virtual Disk to Partition Mapping and Configuration for a Quad-core Single-socket PowerEdge server hosting SQL Server 2005

Drive Letter	Folder Name	Format	Used for
G:	MSS_DATA_1, MSS_DATA_2, MSS_DATA_3, MSS_DATA_4	NTFS 64 KB	data
I:	TEMPDB_DATA_1, TEMPDB_DATA_2, TEMPDB_DATA_3, TEMPDB_DATA_4	NTFS 64 KB	TempDB
J:	MSS_LOG1, TEMPDB_LOG1	NTFS 64 KB	logs

Table 3 - Quad-core single-socket Server Storage Configuration

Table 4 illustrates a sample file system configuration for a quad-core, dual-socket host server. Details for configuring the TempDB files are covered in the "SQL Server 2005 Configuration" section below.

[6] Lu and Bruck, op. cit.

Drive Letter	Folder Name	Format	Used for
G:	MSS_DATA_1, MSS_DATA_2, MSS_DATA_3, MSS_DATA_4	NTFS 64 KB	data
H:	MSS_DATA_5, MSS_DATA_6, MSS_DATA_7, MSS_DATA_8	NTFS 64 KB	data
I:	TEMPDB_DATA_1, TEMPDB_DATA_2, TEMPDB_DATA_3, TEMPDB_DATA_4, TEMPDB_DATA_5, TEMPDB_DATA_6, TEMPDB_DATA_7, TEMPDB_DATA_8	NTFS 64 KB	TempDB
J:	MSS_LOG1, TEMPDB_LOG1	NTFS 64 KB	logs

Table 4 - Quad-core Dual-socket Server Storage Configuration

The number of TempDB data files should also equal the number of CPU cores. Refer to section "Configuring TempDB files" below for more details.

Optimizing the Server for Network Applications

By default, the "Maximize data throughput for file sharing" option is selected in the Windows Server 2003 Network Connection configuration. This option may limit the memory available for SQL Server 2005 operations because this option gives priority to applications which perform buffered I/O operations by caching their I/O pages in the system cache. For a PowerEdge server hosting SQL Server 2005, it is a best practice to select the "Maximize data throughput for network application" option in the Network Connection section of the Windows Server Control Panel. Follow the procedure described in the "Dell PowerEdge Systems Dell|Microsoft SQL Server 2005 SP2 Database on Microsoft Windows Server 2003 Standard Edition with SP2 Deployment Guide" to optimize network performance for SQL Server.

SQL Server 2005 Configuration

The location of the data files and log files for each database provided by SQL Server 2005 are configured individually, so it is relatively easy to make use of external storage for these files by following the guidelines outlined in this document. However, the location of the TempDB files can also impact the performance of a database solution, particularly with respect to large or complex queries.

Configuring TempDB Files

When the SQL Server 2005 database is initially created, the TempDB files are generally placed on server internal drives (e.g. the C: volume). Follow the procedure described in the "Dell PowerEdge Systems Dell|Microsoft SQL Server 2005 SP2 Database on Microsoft Windows Server 2003 Standard Edition with SP2 Deployment Guide" to move the TempDB files to external disk drives dedicated for TempDB files.

By default, a single TempDB data file is created during installation of SQL Server 2005. Adding additional TempDB data files can help avoid latch contention on allocation pages and can mitigate I/O performance issues. The recommended number of TempDB data files should equal the number of CPU cores in the server. To better utilize the allocation mechanism, all TempDB data files should be of equal size.

By default, the AUTO GROW option is set to "on" for TempDB files. However, expanding TempDB too frequently can lead to performance degradation. To avoid this issue, it is recommended to pre-allocate the TempDB space with a size large enough to accommodate the expected workload and set the file growth increment large enough to minimize TempDB expansions. In addition, Microsoft recommends setting the TempDB files FILEGROWTH increment to 10%[7].

[7] "Working with TempDB in SQL Server 2005", Wei Xiao, Matt Hink, Mirek, Sunil Agarwal, Microsoft TechNet, June 26, 2006, http://www.microsoft.com/technet/prodtechnol/sql/2005/workingwithtempdb.msp

By default, when a new data file is created or a data file is expanded as a result of AUTO GROW, the pages on the file are initialized by writing zeros before the file can be used. The file zeroing process can be time consuming and can lead to application timeout, especially during the AUTO GROW activity. One of the major improvements of SQL Server 2005 is the instant file initialization feature. This feature skips the file zeroing process when a data file is created or a file size is increased. Instead, the pages on the data file get overwritten with new data when pages are actually allocated. Microsoft's in-house testing for creating and growing files shows a significant performance improvement when the instant data file initialization is used[8].

To enable instant file initialization, you must assign the Windows SE_MANAGE_VOLUME_NAME permission to the Windows account that is designated as the SQL Server 2005 service account. By default, this permission is assigned to the local "Administrators" group; a user with system administrator rights can assign the permission by adding the SQL Server service account to the "Perform Volume Maintenance Tasks" security policy. Follow the procedure described in the "Dell PowerEdge Systems Dell|Microsoft SQL Server 2005 SP2 Database on Microsoft Windows Server 2003 Standard Edition with SP2 Deployment Guide" to enable instant file initialization.

[8] Lu and Bruck, op. cit.

Performance Characterization and Analysis

The core architecture of the PowerVault MD3000i and the PowerVault MD3000 are very similar. Both storage systems feature dual RAID controllers, share a common RAID engine, and provide support for the same number, type, and speed of hard-disk drives. In fact, the primary difference between these two storage systems lies in the ports that are used to connect host servers to the array. While each MD3000i RAID controller features two GbE host ports, the MD3000 controllers feature two SAS x4 wide ports.

The theoretical maximum throughput of a SAS x4 wide port is significantly higher than the theoretical maximum throughput of a gigabit Ethernet port. However, the overall performance of a database system is affected by many other factors; the storage subsystem throughput, bandwidth, and latency are all of concern, so it is important to perform tests to determine the respective capabilities of each of these storage systems.

The Dell Database Solutions Engineering team has conducted studies to evaluate the performance of the MD3000i and determine its suitability for database workloads. The team prepared an MD3000i and an MD3000 with identical physical disk configurations, disk groups, and virtual disks. These virtual disks were presented to servers running benchmarking and I/O generation tools, including the Microsoft BenchCraft TPC-E kit[9], the open source IOMeter tool, and the Oracle I/O Numbers toolkit. Several studies were performed using these tools, and the results were analyzed. Table 5 summarizes the observations that resulted from these studies.

I/O Characteristics	Observations
Mixed large and small random reads	The MD3000i exhibited 50-100% higher IO/s and lower latency than the MD3000 with this type of workload, but due to constraints of the GbE interfaces attained approximately 30-35% lower MB/s.
Mixed large and small random writes	The MD3000i exhibited 20-50% higher IO/s and MB/s than the MD3000 with this type of workload, with a marginally higher latency.
Small random reads	With this type of workload, the MD3000i exhibited 20-30% higher IO/s with marginally higher latency when compared to the MD3000.
Small random writes	With this type of workload, the MD3000i again exhibited 20-30% higher IO/s with marginally higher latency when compared to the MD3000.
Large sequential reads	With this workload, the constraints imposed by the GbE interfaces were an important consideration, and the MD3000i exhibited 25-35% lower bandwidth (MB/s) than the MD3000.
Large sequential writes	While less of a consideration than with the large sequential reads, the GbE interfaces again caused the MD3000i to exhibit 15-25% lower bandwidth (MB/s) than the MD3000.

Table 5 - Relative Performance of MD3000i and MD3000 with Different I/O workloads

These results indicate that iSCSI is a suitable technology, and that the MD3000i is a suitable storage system for database solutions with Microsoft SQL Server 2005. The MD3000i-based solution is a good fit for transaction-processing (OLTP) workloads, which predominantly feature small random writes. This solution continues to perform well when confronted with mixed workloads that introduce up to 30% large reads and/or writes, but the MD3000i begins to be constrained by the throughput of the GbE ports used for its iSCSI host connections. However, when raw throughput is of the utmost concern, such as with databases that are primarily intended as business intelligence systems data warehouses, the constraints imposed by GbE ports become more apparent. Further, because a software iSCSI initiator is used, some CPU overhead is incurred when compared to solutions that feature dedicated storage host-bus adapters (HBAs).

[9] In this context, the benchmarking tools were utilized only to generate a database workload for storage system comparison; the tools were used without the intent of producing publishable benchmark results.

Solution Deliverables List

This section contains the Solution Deliverables List (SDL) for the Dell solution for SQL Server 2005 on Windows 2003 Server with the Dell PowerVault MD3000i. It contains detailed listing of server and storage hardware configurations, firmware, driver, OS and database versions that were tested during the preparation of this paper.

Recommended Hardware/Software Requirements (For details, see below)		
	Validated Component(s)	**Minimum Single Node DB Configuration**
PowerEdge™ Nodes	PowerEdge 2950	1 Only
Memory	All valid Dell Server memory configurations (up to 32GB)	1 Gig
Dell PowerVault Storage	MD3000i Storage Array + MD1000 Expansion Enclosure(s) [†]	1 MD3000i with up to 2 MD3000
iSCSI Initiator	Microsoft iSCSI Software Initiator Intel® or Broadcom® Gigabit NICs, TOE optional	Installed, MPIO recommended for redundancy. 1 port (For External Storage) 2 ports recommended for redundancy
Ethernet Ports	Intel® or Broadcom® Gigabit NICs	1 port (for application server or client connectivity)
RAID Controllers (Used for internal storage only)	PERC 5/i	1
Internal Drive	All valid Dell Server internal storage configurations	73 Gig
SQL Server Software & Licenses	Microsoft® SQL Server™ 2005 SP2	Single node with appropriate licensing
Operating System	Microsoft® Windows Server™ 2003 R2 Standard Edition SP2	With appropriate licensing
Recommended Support Contract	Gold or Platinum Plus Support	

†: The MD1000 has been validated as an expansion enclosure in this configuration, not as a stand-alone storage device.

Supported Servers					
Dell PowerEdge Servers	**Model**	**BIOS**	**ESM/BMC Firmware**	**Internal SCSI RAID Firmware / Driver**	**Notes**
	PE2950	1.5.1 or higher	v1.33 or higher	PERC 5i –5.1.1-0040/2.8.0.32	Intel XEON Dual-Core or Quad-Core Processors
PERC 5/i	Firmware package version =5.1.1-0040; Driver version = 2.8.0.32				
Network Interconnect					
Ethernet Switches	All Dell Gigabit Ethernet Switches				
Intel Pro1000 PCIe Gigabit Family Of Adapters	Driver version = 9.9.12.0				
Broadcom NetXtreme II™	NDIS Driver = 3.0.5; VBD Driver = 3.0.9				
Broadcom NetXtreme™	Driver version = 10.19.01				
NIC Teaming (Not available yet for TOE NICs)					
Windows NIC teaming for Intel Adapters	Driver version = 8.4.2.0				
Intel Proset for Windows Device Manager	Driver version = 12.2.40.10				
Broadcom Advanced Control Suite	Management Application version = 10.0.8 BASP Teaming driver = 6.2.24				
iSCSI Network Interconnect and Initiator					
Ethernet Switches	All Dell Gigabit Ethernet Switches				
Gigabit Ethernet Adapters	See “Network Interconnect” above				
iSCSI Initiator	Microsoft iSCSI Initiator version = 2.05 or higher				
Storage Array					
Storage Arrays Supported (with Software)					
Storage Subsystem Disk Drives: 15K RPM SAS Drives **iSCSI Data transport speed**: 1 Gbit/s					
Dell PowerVault MD3000i	iSCSI RAID Controller Firmware = 06.50.32.60 (A00) or higher [†]				
MD3000i Storage Array MPIO driver	02.50.36.04 (from Dell MD3000i Resource CD version A01) or higher [†]				
Dell PowerVault MD1000 Expansion Enclosure	EMM Firmware = A04 or higher [†]				
Software					
Database	Microsoft SQL Server 2005 SP2				
Operating system	Microsoft Windows 2003 R2 Server Standard Edition SP2				

†: When using versions higher than those listed in this SDL, all of the storage array components should be updated to a set of versions that have been validated to interoperate properly; certain combinations may not meet these criteria. Refer to the release notes for the MD3000i and MD1000 firmware updates and ensure that all intended updates have been validated together.

Conclusions

Dell Solutions for SQL Server 2005 are designed to simplify operations, improve utilization and cost-effectively scale as your needs grow over time. This white paper provides a blueprint for setting up a standalone SQL Server 2005 database on Dell PowerEdge servers and Dell PowerVault MD3000i storage arrays (with MD1000 expansion enclosures). Based on observations derived from testing and analysis by the Dell Database Solutions Engineering team, this iSCSI storage solution offers a cost-effective and easy to deploy foundation for SQL Server 2005 databases with OLTP-type workloads.

The best practices described here are intended to help achieve optimal performance of SQL Server 2005. To learn more about deploying SQL Server 2005 on PowerEdge server and Dell storage, please visit www.dell.com/sql or contact your Dell representative for up to date information on Dell servers, storage and services for SQL Server 2005 solutions.

Tables and Figures Index

Table 1 - Single-Socket Server - Storage Configuration 9
Table 2 - Dual-Socket Server - Storage Configuration..... 10
Table 3 - Quad-core single-socket Server Storage Configuration 13
Table 4 - Quad-core Dual-socket Server Storage Configuration..... 14
Table 5 - Relative Performance of MD3000i and MD3000 with Different I/O workloads 16

Figure 1 - Architectural Overview of Dell SQL Server 2005 configuration 5
Figure 2 - Cabling an MD3000i Storage Array in an IP SAN..... 7
Figure 3 - MD3000i System with two MD1000 Enclosures Cascaded..... 7
Figure 4 - Separation of Disk Groups and Virtual Disks within a PowerVault MD3000i and two MD1000 Storage Arrays for TempDB, Transaction Logs, and Data optimized for Data Warehousing 9

References

1. "Storage Top 10 Best Practices", Microsoft TechNet, October 17, 2006. http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/storage-top-10.mspx

2. "Physical Database Storage Design", Kathy Lu, Lewis Bruck, Microsoft TechNet, February 12, 2007. http://www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx

3. "Working with TempDB in SQL Server 2005", Wei Xiao, Matt Hink, Mirek, Sunil Agarwal, Microsoft TechNet, June 26, 2006. http://www.microsoft.com/technet/prodtechnol/sql/2005/workingwithtempdb.mspx

4. "Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks", EMC PowerLink September, 2005. https://powerlink.emc.com/nsepn/webapps/btg548664833igtcuup4826/km/live1/en_US/Offering_Technical/White_Paper/H1445_Using_Diskpar_Align_Partition_ldv.pdf