

Microsoft® SharePoint® Server 2010

Designing and Implementing a Large Farm

Dell™ Solutions Engineering

Ravikanth Chaganti and Kevin Guinn

May 2010



Executive Summary

A Microsoft® SharePoint® Server 2010 farm hosts service applications that provide many different functions for its users. These functions are surfaced through Web servers, and store their configuration information and data in SQL Server® databases. These interactions mean that a classical three-tier model is appropriate for SharePoint Server 2010. This paper provides a basic reference architecture that maps SharePoint roles to this model, and then extends this architecture to provide high-availability and redundancy at each tier.

The paper discusses key criteria and decision points for determining what server and storage hardware are appropriate for a large-sized SharePoint Server 2010 farm. It offers guidelines for determining how much storage space should be allocated, and when separate storage volumes are desirable. Finally, it demonstrates how services for multiple farms can be provided by a central farm. The intent of this paper is to enable administrators to understand important planning criteria, and to begin mapping their plan to a specific farm topology and well-suited hardware components. Dell offers more direct and specific guidance through consulting and other services.

A companion paper, *SharePoint Server 2010: An Introduction*, is available from www.dell.com/sharepoint. It offers an overview of SharePoint Server 2010, and provides common concepts and definitions that form a good basis for understanding the reference architectures presented in this paper.

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, *PowerEdge*, *PowerConnect*, and *PowerVault* are trademarks of Dell Inc. *Microsoft*, *Windows*, *Windows Server*, and *SharePoint*, and *SQL Server* are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. *AMD Operton* is a trademark of AMD. *Intel*, *Xeon* are registered trademarks of Intel Corporation in the US and other countries.

May 2010

Table of Contents

- Executive Summary ii
- Introduction 2
- Designing and Building a Large Farm..... 2
 - Information Architecture..... 2
 - Large Farm Topology 4
 - Database Considerations 6
 - System Architecture..... 6
 - Operating System and Database Editions 6
 - Database Storage 7
 - Optimizing I/O for SharePoint Databases 7
 - Estimating Storage Capacity 9
 - Considerations for WFE and Application Servers 11
 - System Architecture..... 11
 - Search Index Storage on Query Servers 11
- Building High Availability and Redundancy 12
 - High Availability for the Database Server 13
 - Windows Server Failover Clustering 13
 - SQL Database Mirroring 15
 - High Availability for Web Front-End Servers 15
 - Load-Balancing Web Front-End Servers 16
 - Redundancy for Service Application Components 17
 - Deploying the Search Service Application for Cross-Farm Sharing 18
- Summary 20
- Appendix A: Selecting a Dell Storage Array 21
- Appendix B: Selecting PowerEdge Servers 22

Introduction

Microsoft SharePoint Server 2010 builds on the capabilities that were offered in Microsoft Office SharePoint Server 2007 to provide a rich platform for collaboration, information sharing, and document management. SharePoint 2010 adds several new features, and introduces important architectural changes and product improvements. For a complete overview of SharePoint Server 2010, its features, and key dependencies, please see *SharePoint Server 2010: An Introduction*, a companion to this whitepaper that is available from www.dell.com/sharepoint.

The remainder of this paper defines reference architectures that illustrate how a large-sized SharePoint Server 2010 farm may be designed. There are many factors, such as information isolation policies, or the need to provide Enterprise Search capabilities for content outside of the farm, that can introduce the need for additional servers or more complex designs.

Designing and Building a Large Farm

A SharePoint server farm is a set of servers which collectively provides the services needed by a SharePoint deployment. Some of these services, or sets of services, comprise predefined roles and must be configured within the solution. Other services and components are optional, but they provide additional features and functionality that are often desirable. There are some constraints and best practices that help determine which components should be located on each server in the farm. Also, by considering how the components are distributed, the farm can be designed to more easily accommodate later growth.

NOTE: In SharePoint Server 2010, components generally provide functionality for a given service application. As a result, this paper may use the terms “role” and “component” interchangeably. In this context, SharePoint roles refer to one or more components that provide a farm service, and should not be confused with Windows Server roles, which generally include one or more Windows services to provide operating system functionality.

Information Architecture

A large farm will typically include multiple application pools, content databases, and site collections. The way that these entities are allocated and arranged helps determine the information architecture for the farm, which should be developed in conjunction with the farm’s hardware topology. Considerations for developing the information architecture are based on the intended uses of the farm and on the SLAs that govern these uses, including performance targets, data isolation, and backup windows. A farm may provide services to intranet, extranet, and internet environments; it may feature large document repositories, portals, enterprise search, and collaborative sites. The end-user experience will depend heavily on how hardware and software resources are used in the farm.

The information architecture for a large SharePoint farm should be designed to allow for flexibility and growth, and must account for this growth in terms of many overlapping factors. Some of these factors include physical storage, number of sites and sub-sites, number of list items, and number of users. The

scalability boundary¹ conditions listed in Table 1 serve as a starting point for discussing how information will be organized within the SharePoint 2010 farm.

Table 1: Scalability Limits in SharePoint Server 2010

Object	Advised upper bound	Discussion
Content Databases	300 per web application	With 300 content databases per Web application, end user operations such as navigating to the site or site collections are not affected. But admin operations such as creating a new site collection will experience performance degradation.
Site collections	500,000 per web application	Overall farm throughput has been seen to decrease as the number of site collections increases.
Sites & subsites	250,000 per site collection	This includes top level sites and the sub-sites underneath each top-level site
Documents	50,000,000 per library	The type and size of documents that are stored will impact this limit. Nested folders, views, and other organization techniques enable larger libraries.
Lists	2000 per site (or sub-site)	These guidelines are intended to maintain a good user experience when rendering list views. Some larger lists may be acceptable large list throttling is used
Items	5000 per list view	
Web parts	25 per page	This figure is an estimate based on simple Web Parts. The complexity of the Web parts dictates how many Web Parts can be used on a page before performance is affected.
Users in a Site Collection	2 million per site collection	You can add millions of people to your Web site by using Microsoft Windows® security groups to manage security instead of using individual users. This limit is based on manageability and ease of navigation in the user interface.

¹ For complete information on SharePoint 2010 software boundaries and limits, refer to [http://technet.microsoft.com/en-us/library/cc262787\(Office.14\).aspx](http://technet.microsoft.com/en-us/library/cc262787(Office.14).aspx)

The factors in Table 1 influence the scalability and performance considerations of the information architecture. For an optimal end-user experience, all basic farm navigation operations and page loads should be able to complete with sub-second response times. This is not always practical, but response times that exceed three seconds for high-frequency operations - such as loading a portal home page or a document library - are likely to generate user complaints. Predictions and modeling are an important part of the information architecture. However, when a farm is released to production, its use will evolve over time. In addition to planning for growth, administrators should monitor response times, indexing times, and other criteria as the farm is used. Based on the actual observations, assumptions may change and it may prove useful to reallocate roles among the servers in the farm.

Business and regulatory controls may also require various levels of data isolation. Strict data isolation requirements may not only necessitate additional content databases, web applications, and site collections, but may also stipulate that dedicated database servers be provisioned within a farm. If the farm needs to meet these types of requirements, then the information architecture and physical hardware topology may both be impacted.

Similarly, planning the information architecture can help improve search performance in the farm. A few factors that will influence the capabilities of the search infrastructure include the number of (internal and external) content sources that are indexed, how "Best Bets" and synonyms are assigned to keywords, and which properties and metadata are discovered during a crawl operation. Additionally, how crawl rules are defined, and if results removal is required for various audiences will influence the time required to complete an indexing operation. That time directly impacts the "freshness" of the index and, therefore, how rapidly new content can be found by end users.

Complete books can be written about many of the areas for which planning the information architecture of the farm is important, and an in-depth examination of these topics is beyond the scope of this paper. However, Dell Pro-Consult services offers detailed assessment, design, and implementation offerings that bring our experts on site to assist you with your specific SharePoint requirements and needs.

Large Farm Topology

The setup utility for SharePoint 2010 offers either a Stand-alone Installation or a server farm Installation. The Stand-alone installation is intended only for single-server deployments. Therefore, the server farm Installation option will be required for all application servers and web front-end servers in a large SharePoint farm. A typical large farm uses three-tier architecture: featuring database, application, and presentation tiers. SQL Server is hosted by at least one server to form the database tier, systems hosting the SharePoint application server components form the middle tier, and the Web front-end and search query components are distributed across several servers to form the presentation tier. An example of this architecture is illustrated in Figure 1.

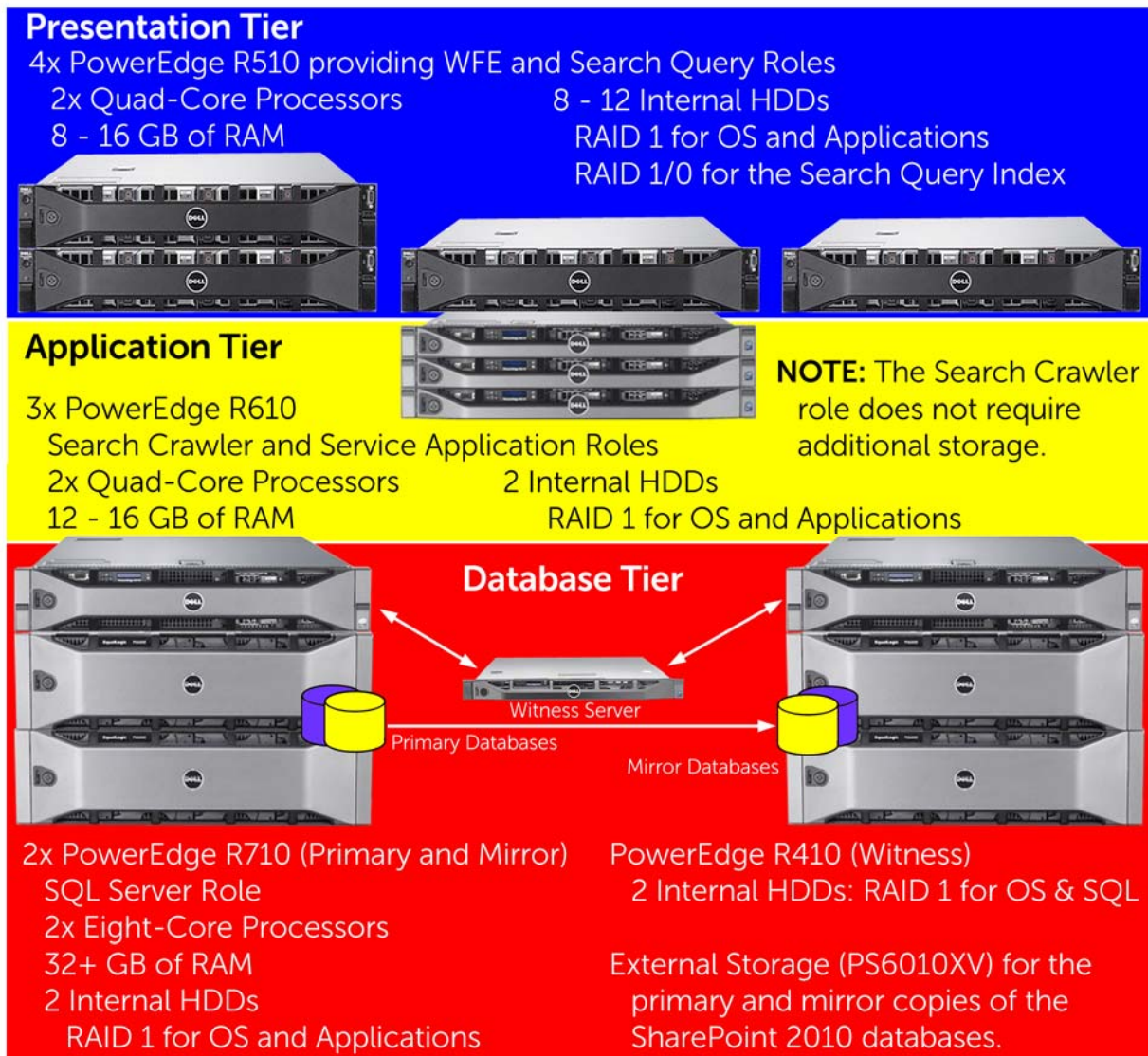


Figure 1: Example of a Large SharePoint 2010 Farm

In general, this type of large farm is expected to host between several hundred of gigabytes (GB) and a few terabytes (TB) of content and to provide services for up to 10,000 users. However, the way a SharePoint deployment is used can vary widely. The number and type of user requests ultimately determines whether a particular topology is suitable for the intended use of a SharePoint farm. For example, if much of the content is static or archival, then the data capacity of the farm can grow considerably without placing additional stress on the web servers. However, the additional content will cause a full search indexing operation to take longer to complete. Similarly, if enterprise search represents a significant proportion of the user activities, it can be beneficial to separate the query server role to dedicated servers. Because there are many such factors at play within a farm of this scale, it is extremely important to consider the information architecture as well as the hardware components when determining the farm topology.

Database Considerations

The first step toward deploying a SharePoint farm is to prepare the database server. All of the information that is made available from SharePoint sites is stored in SQL Server databases. SQL Server 2008 or SQL Server 2008 R2 is supported as the database for SharePoint farms. Several features of SQL Server 2008 or SQL Server 2008 R2 can be used to enhance the functionality or performance of a SharePoint farm. For example, database backup compression can reduce both the size of backup sets and the time required to complete a backup operation. Similarly, Transparent Data Encryption can play a role in providing additional security for data stored within a SharePoint farm.

System Architecture

SharePoint 2010 is available only for 64-bit extended (x64) architecture. This allows the SQL Server host to use RAM beyond 4 GB efficiently, which can help increase the number of connections and transactions that the database server can process. To fulfill this requirement, the host server must employ x64 hardware and run an x64 edition of a supported Windows Server operating system. Currently-shipping Dell PowerEdge servers with Intel Xeon® or AMD Opteron™ processors, such as those featured in this paper, employ x64 hardware and may be successfully used in a SharePoint farm.

For a typical large SharePoint farm, sixteen processing cores and 32 GB of memory are recommended as a starting point for the database server. If the corpus of data in the farm exceeds 2 TB, consider increasing the SQL Server RAM to 64 GB. Because the user load is likely to require good storage system performance in terms of IOPS, an external storage array is recommended. These arrays allow additional physical disk spindles to be allocated to the SharePoint databases, thereby increasing the number of IOPS that the database can process. The Database Storage section of this paper provides more information about the storage capacity and I/O requirements for important SharePoint 2010 databases.

Operating System and Database Editions

Choosing among the various editions of Windows Server 2008 or Windows Server 2008 R2, and SQL Server 2008² or SQL Server 2008 R2 is primarily a function of which system specifications and software features are important in the intended environment. The most important considerations for the operating system are listed in Table 2, and those for the database software are listed in Table 3. The high-availability features listed in these tables—such as failover clustering and database mirroring— are examined later in this paper.

Table 2: Critical Factors for Selecting an Operating System Edition

	Windows Server 2008 SP2 or 2008 R2 x64 Standard Edition	Windows Server 2008 SP2 or 2008 R2 x64 Enterprise Edition
Support for Failover Clustering	No	Yes, up to 16 nodes
Maximum Physical RAM	32GB	2TB

² SQL Server 2008 with Service Pack 1 and Cumulative Update 2 is the minimum required version of SQL Server 2008. SQL Server 2005 with Service Pack 3 and Cumulative Update 3 is also supported.

Table 3: Key Factors for Selecting a SQL Server Edition

	SQL Server 2008 or 2008 R2 x64 Standard Edition	SQL Server 2008 or 2008 R2 x64 Enterprise Edition
Support for Failover Clustering	2 nodes	Max number of nodes supported by OS
Database Mirroring	High-safety mode only	All high performance and high safety modes
SQL Server Reporting Services	Available	Available
SQL Server Analysis Services	Available	Available
Database Compression	Not Available	Available
Transparent Data Encryption	Not Available	Available
Resource Governor	Not Available	Available
Remote BLOB Storage	Not Available	Available

Consider the following example: mirrored database servers with 32 GB of RAM are sufficient to meet the current availability and scalability goals for a company's SharePoint farm, and high-safety (synchronous) mirroring is desired. There is also a business need to deploy transparent data encryption to enhance data security in the SharePoint content databases that are associated with some site collections that the farm will provide. If no other decision points are involved, the mirrored database servers for this example farm could employ Windows Server 2008 x64 SP2 or Windows Server 2008 R2 Standard Edition and SQL Server 2008 x64 or SQL Server 2008 R2 Enterprise Edition. If, however, the corpus of data in the farm is expected to expand beyond 2 TB during the service life of the solution, deploying the Enterprise Edition of Windows Server would allow the RAM in the SQL Server host to be increased beyond 32 GB in the future.

Database Storage

The overall performance of a database server is often constrained by its disk I/O subsystem. Dell™ EqualLogic, PowerVault, and Dell / EMC storage arrays all provide a storage subsystem that offers battery-backed cache, redundant array power supplies, multiple RAID levels, and other core availability features that are commonly requested in service level agreements for a SharePoint farm. For more information about these arrays, see Appendix A.

Disk I/O latency and disk queue length are critical factors to monitor as a solution is deployed and adopted. If frequent SharePoint operations, like page loads, take more than a second or two to complete, the end-user experience may not be acceptable. Additional disks or arrays can be added if the demands of a farm are placing too much stress on the database storage subsystem. The expansion capability of EqualLogic, PowerVault, and Dell / EMC storage arrays also provides a path that can be used to accommodate growth or increased use of the SharePoint farm.

Optimizing I/O for SharePoint Databases

The overall performance of a SharePoint solution can be limited by an I/O bottleneck within the database. As a result, it is important to consider how the solution will be used when determining RAID levels and designing the layout of logical disks for the SharePoint databases. The information architecture for the farm will help determine the number and size of content databases that are required, and what the I/O priorities are for each database. For site collections that feature many collaboration sites, there will be many write operations to the database. In contrast, portals and document libraries tend to favor reads. Table 4 outlines some of the functions that should have dedicated spindles, identifies a performance priority, and recommends optimal RAID and file configurations. The performance priority offers guidance for determining the order in which I/O

optimizations - such as faster disks, optimal RAID levels, and multiple SQL Server data or log files - should be applied.

Additionally, there are some optimizations that can be applied when volumes are created within the operating system. Windows Server 2008 and Server 2008 R2 use a 1024 KB NTFS volume offset that eliminates the need to provide manual partition alignment using diskpart.exe. NTFS volumes used for SQL Server data, logs, and TempDB should be configured with a 64 KB allocation unit size.

Table 4: Database Storage Recommendations

Function	Preferred RAID Level	Notes
Transaction Logs	RAID 1/0	Transaction logs volumes should be optimized for writes, and should allow ample room to allow for log growth.
Content Databases	RAID 1/0 for content databases that host site collections focused on collaboration or that will feature extensive write activity. RAID 5 for content databases that host site collections focused on portals or document repositories that feature predominantly read activity.	Consider dedicated content Databases for any site collection that is expected to have an unusually high level of activity. Use separate databases for site collections whose access patterns differ greatly, and optimize volumes by access pattern. Provision multiple data files for each configuration database.
TempDB	RAID 1/0	TempDB use can be influenced by SQL Server factors outside of SharePoint (such as rebuilding indexes). It is best to treat TempDB as if it will be heavily utilized.
Logging Database	RAID1/0	This database should be considered to be heavily utilized.
Search Crawler database	RAID 5 or RAID 1/0	Read-heavy, but poor write performance could increase the time required to complete a crawl.
Search Property database	RAID 1/0	Optimize for good read and write performance. Although the typical access pattern favors writes, poor read performance can impact search query times.

For these heavily-used databases, it is also a good idea to provision multiple data files. This allows SQL Server to parallelize I/O operations among the available processing cores on the host server. Consider creating no fewer than two data files per processor socket. Dell lab research suggests that between eight and 16 data files are optimal for a server with sixteen or more cores; more than 16 files becomes difficult to manage, and produces diminishing returns in terms of performance.

Estimating Storage Capacity

Estimating the total data storage space required for a SharePoint 2010 farm can be challenging, and is highly dependent on how SharePoint 2010 will be used within the organization. A departmental or company-wide portal may demand significantly more storage be allocated versus a team site or an individual MySite; however, there are likely to be many more of these smaller sites hosted within the farm. In addition, document libraries will generally require more storage than site features such as blogs, wikis, and other types of lists. If the versioning, recycle bin, or audit features of SharePoint will be used, then additional space must also be allocated for this data.

The general recommendation is that each content database should be limited to storing 200 GB of content. However, if more than one site collection shares a content database, the maximum recommended size of each site collection is 100GB. A large farm is likely to feature several site collections and to require multiple content databases. The allocation of containers (e.g., application pools and site collections) that well-planned information architecture prescribes will also influence how many separate content databases are required. Each content database is associated with a Web application, which may encompass one or more site collections.

Table 5 outlines a guideline³ and provides some examples for estimating the required amount of storage for the content database. Specifying a lower fill factor provides additional space in the database that can be used to store document versions and site recycle bins, or can be used to fulfill audit requirements. This “unused space” also serves as a buffer that can accommodate growth of the content stored in the farm.

Table 5: Estimating Content Database Size Requirements

	Size of content to be stored	Database overhead	Fill factor (for growth and versioning)	Minimum disk space to allocate
Guideline	X	20% (0.2 * X)	50 - 70% recommended	~1.7 * X (70% fill factor) ~2.4 * X (50% fill factor)
350 GB content	350 GB	70 GB	70%	600 GB
1 TB content	1 TB	205 GB	60%	1.92 TB
2 TB content	2 TB	410 GB	50%	4.8 TB

NOTE: In Table 5, the examples all yielded recommended disk allocation sizes well in excess of 200 GB. In these situations, it is recommended that site collections in the farm be split among several content databases, and that each individual general-use content database is no larger than 200 GB.

NOTE: Employing a Remote BLOB Store (RBS) will shift a significant amount of document and file storage to the RBS provider, and therefore reduce the size of associated content databases. Other list items are largely character-based, and will continue to be stored directly in the content databases.

³ At the time this paper was written, SharePoint Server 2010 had not yet been released, and updated formulas for estimating content database size based on document and list parameters were not yet available. This information is expected to be released in the future and made available on [http://technet.microsoft.com/en-us/library/cc298801\(office.14\).aspx](http://technet.microsoft.com/en-us/library/cc298801(office.14).aspx)

In addition to the content databases, the SQL Server host will also require space for housing the configuration, central administration, search (administration, crawl, and property), and service application databases. The configuration, central administration, and search administration are generally small (on the order of 1-2 GB). The search crawl and property databases are based on the amount of content in the farm. Based on initial Microsoft testing and analysis⁴, assume the following sizes for these databases:

- Crawl: 4.6% of the total combined size of all content databases
- Property: 1.5% of the total combined size of all content databases

In SharePoint 2010, trace logs & usage information are stored in the logging database. The default retention policy for this data is 14 days, but it can be configured using PowerShell. Longer retention periods will require additional storage space. Depending on how Usage and Health Data Collection is configured, this database can store very granular information to enable web analytics and usage monitoring; this information is collected and aggregated from every server in the SharePoint 2010 farm. For these reasons, the logging database can grow to a large size, and may do so within a short time if many different facets of health data are collected. This database should generally be housed on a dedicated volume, and should be monitored carefully if a large amount of data will be collected for an extended period of time.

Employing quotas to limit the size of individual sites and establishing governance policies to manage content and control the number of sites in the farm can help control the total space required by a given farm. Regardless, it is important to plan for capacity to increase over time, and to build a flexible and scalable infrastructure that will enable the farm to grow.

The Dell / EMC CX4-series arrays are well suited for storing typical SharePoint databases. Multiple LUNs can be configured to support these databases. If the farm grows and more IOPS are required to meet the needs of the farm, additional disk enclosures can be added to the CX4-series array. These arrays also provide hardware-based snapshots, replication, and other advanced storage features that administrators can use as part of a data protection strategy for SharePoint. For BI environments that aggregate large amounts of data, the 8 gigabit Fibre Channel connectivity offered by these arrays provides good bandwidth (in terms of MB/s) that can prove beneficial.

⁴ See "Storage and SQL Server capacity planning and configuration (SharePoint Server 2010)" at [http://technet.microsoft.com/en-us/library/cc298801\(office.14\).aspx](http://technet.microsoft.com/en-us/library/cc298801(office.14).aspx)

Considerations for WFE and Application Servers

In a large farm topology, the database server is typically hosted on one server (or a set of clustered or mirrored servers) and the remaining application server and Web front-end roles are distributed across several other servers. This topology enables additional servers to be added and roles to be reallocated as the farm grows to meet the changing needs of the business. To configure the SharePoint farm, complete the following high-level steps:

1. Deploy the database server.
2. Install the SharePoint 2010 software and any relevant prerequisites and updates on all application and Web front-end servers.
NOTE: During the installation, select the server farm option (to avoid provisioning a single-server solution), and then select the Complete option, even for systems that will initially serve as Web front-end servers. This option allows a server to host different roles as the farm changes over time.
3. Configure hardware or software load balancing for the Web front-end servers.
4. Start to configure the farm by running the SharePoint Configuration Wizard on one of the servers that will host the Central Administration site.
5. Use the SharePoint Configuration Wizard to add additional servers to the farm.
6. From the Central Administration interface (or PowerShell cmdlets), configure the components that will be hosted by each server in the farm.

System Architecture

As with SQL Server, SharePoint 2010 Web front-end and application servers require 64bit hardware and Windows Server 2008 x64 SP2 or Windows Server 2008 R2 operating system editions. It is uncommon for a SharePoint application server or Web front-end server to be able to benefit from the enhancements in Windows Server 2008 Enterprise Edition that are not available in Standard Edition. At the application tier and presentation tier, failover clustering is not employed for availability; instead, roles within the farm are configured to run on multiple servers. This means that, in most cases, Windows Server 2008 Standard Edition will be sufficient for the application and Web front-end servers.

In this large farm architecture, each server that hosts the Web front-end role should be configured with at least four processing cores and at least 8 GB of RAM. To take advantage of all three memory channels on certain Nehalem-based servers (such as the PowerEdge M610, M710, R610, and R710), consider deploying 12 GB on servers that host both roles. If searches are expected to constitute more than 20% of user activity, consider increasing the RAM further.

Search Index Storage on Query Servers

The large farm illustrated in Figure X features application servers that host the crawler component, which builds an index that makes it possible to search the content stored within the farm. A content source specifies the locations, depth, and type of content that will be crawled. In SharePoint 2010, the crawler and query components can each be hosted on multiple servers to provide load balancing and redundancy.

To improve search performance, deploy query server roles on the Web front-end servers so that all SharePoint search requests are processed and returned to the Web front-end component efficiently. The Query component can be installed on multiple servers, as can the Web front-end role. In SharePoint 2010, the search index that is stored on the query servers can be divided into index partitions, and each index partition can be assigned to one or more query servers in the farm. An index partition holds a part of the overall search index. For example, if there are three query servers in the

farm, three index partitions could each be assigned to a single query server. This way, each query server holds 1/3rd of the overall search index and therefore requires less disk space. Such a configuration, however, would not provide high-availability and could cause partial results to be returned to users. To provide availability, each index partition can be assigned to more than one query server.

Therefore, the guideline for disk size for query servers is based on the size of the crawl database, and the number of index partitions that are assigned to each query server. If a query server will store the entire index, allocate sufficient space for approximately 2.5 times the size of the crawl database (roughly 12% of the size of the crawled content). If there are multiple index partitions, then this storage allocation can be reduced based on the ratio of assigned partitions versus total index partitions. While this guideline provides a baseline, the index size will vary based on the definitions provided by the content source. For best results, store the search index on its own RAID 1/0 volume on each query server. Read performance helps provide query results more quickly, and write performance helps reduce the time required to complete a crawl operation.

In SharePoint 2010, crawler is stateless. This means that the crawler component does not store a copy of the index as it did in MOSS 2007. Therefore, there are not any special storage requirements for the search crawler server. Crawling content and building index partitions can be a CPU- and memory-intensive operation; if the crawler servers host other service applications, administrators should generally schedule crawls to occur at off-peak times.

Building High Availability and Redundancy

A business critical application hosted on a SharePoint Server solution that is not designed for high availability can lead to unacceptable amounts of service downtime and potentially even revenue loss. To ensure that hardware failures do not affect the business, farm architects must consider availability as a part of the solution design. Overall capacity and performance requirements will affect how to implement availability within the SharePoint farm. To determine whether to implement high availability within a solution, look at the expected uptime, the number of users affected if service becomes unavailable, and the productivity and business losses that will accumulate until service is restored. If the solution requires high levels of uptime to prevent productivity loss when the service is offline, consider high availability options for each of the server roles in the SharePoint solution. Typically, high availability is among the first considerations in determining the number of server computers required to deploy a server farm.

Within the scope of this paper, availability refers to service or server role availability only, and does not include high availability of data through backup and restore or data protection mechanisms. To mitigate any unexpected data loss, farm designers must evaluate various disaster and data recovery options as a part of SharePoint solution design. SharePoint 2010 offers out of the box backup and restore capabilities. Native backup and restore can be performed using the SharePoint central administration user interface or using Windows PowerShell. SharePoint 2010 offers granular backup and restore options, including restore from an unattached SQL database. For more information, see Backup and Recovery (SharePoint Server 2010) at [http://technet.microsoft.com/en-us/library/ee662536\(offic.14\).aspx](http://technet.microsoft.com/en-us/library/ee662536(offic.14).aspx). Alternatively, use Microsoft SQL Server 2008 backup and restore methods, or other mechanisms such as Microsoft Data Protection Manager 2010.

High Availability for the Database Server

Database downtime can be extremely disruptive to the farm. Therefore, limiting the potential for database downtime should be considered a high priority. The two most common ways to enhance the availability of the database tier are to deploy a failover cluster that runs a SQL Server instance or to employ SQL Server database mirroring. When used within a SharePoint farm, each of these methods has its own advantages and disadvantages. These are summarized in Table 6, and discussed in the following sub-sections.

Table 6: Availability Techniques for SQL Server

	SQL Failover Cluster	SQL Database Mirroring
Protection Provided	Data is stored on a single shared storage array and employs RAID, hardware redundancy, and multipath I/O for protection. The clustered SQL Server instance can run on any node in the cluster, providing tolerance for a wide range of hardware and software faults on the host servers. The design goal for a failover cluster is to militate against a single point of failure that might cause the database to remain offline.	Both the primary SQL Server and its mirroring partner have their own data copy, and each node's storage subsystem features RAID and hardware redundancy. Many SharePoint 2010 databases and service applications allow specification of mirror-aware connection strings that help facilitate automated failover. The fact that there are independent database servers and separate copies of the data provides protection against faults that may adversely impact one of the database host servers or its storage array.
Recovery Method	When there is a failover event, there will be some downtime as SQL Server services are started on the alternate cluster node and the database steps through the redo logs to ensure consistency. Once this is complete, services will resume without intervention.	SharePoint 2010 provides native support to SQL Server mirroring for many of the SharePoint 2010 databases. While creating web applications and service applications, specify a failover database server; in case of primary SQL server failure, SharePoint automatically switches to a secondary. The default time for switching to a failover server is 15 seconds.
Performance Impact	The overhead associated with running a clustered instance versus a standalone instance of SQL Server is generally considered negligible.	Logs are compressed and transmitted to the database mirroring partner. This adds some additional CPU, network, and I/O load to the primary SQL Server system.
Operating System and SQL Server Editions	Operating System: Enterprise Edition is required for failover clusters SQL Server: Standard Edition supports just two cluster nodes. Enterprise Edition supports as many nodes as the operating system.	Operating System: Standard Edition is generally sufficient. SQL Server: Standard Edition provides high-safety (synchronous) mirroring. Enterprise Edition adds support for high performance (asynchronous) mirroring. To enable automated failover high-safety mode and a witness server are necessary.

Windows Server Failover Clustering

A failover cluster allows a SQL Server instance to run on any one node that is a cluster member. If the hosting node fails, the instance will be relocated and start to run on an alternate node; this process is known as failover. During the installation of SQL Server 2008, options are provided for configuring a clustered instance and for adding an additional node to a clustered instance. The clustered instance is configured within a dedicated cluster resource group that includes the necessary SQL Server services, a common network name and IP address for the instance, and the shared NTFS volumes that will be used

for SharePoint databases, transaction logs, and TempDB storage. This resource group, and the SQL instance that it contains, will run on one physical node at any given point in time. A failover cluster hosting the SQL Server database for a large farm is illustrated in Figure 2.

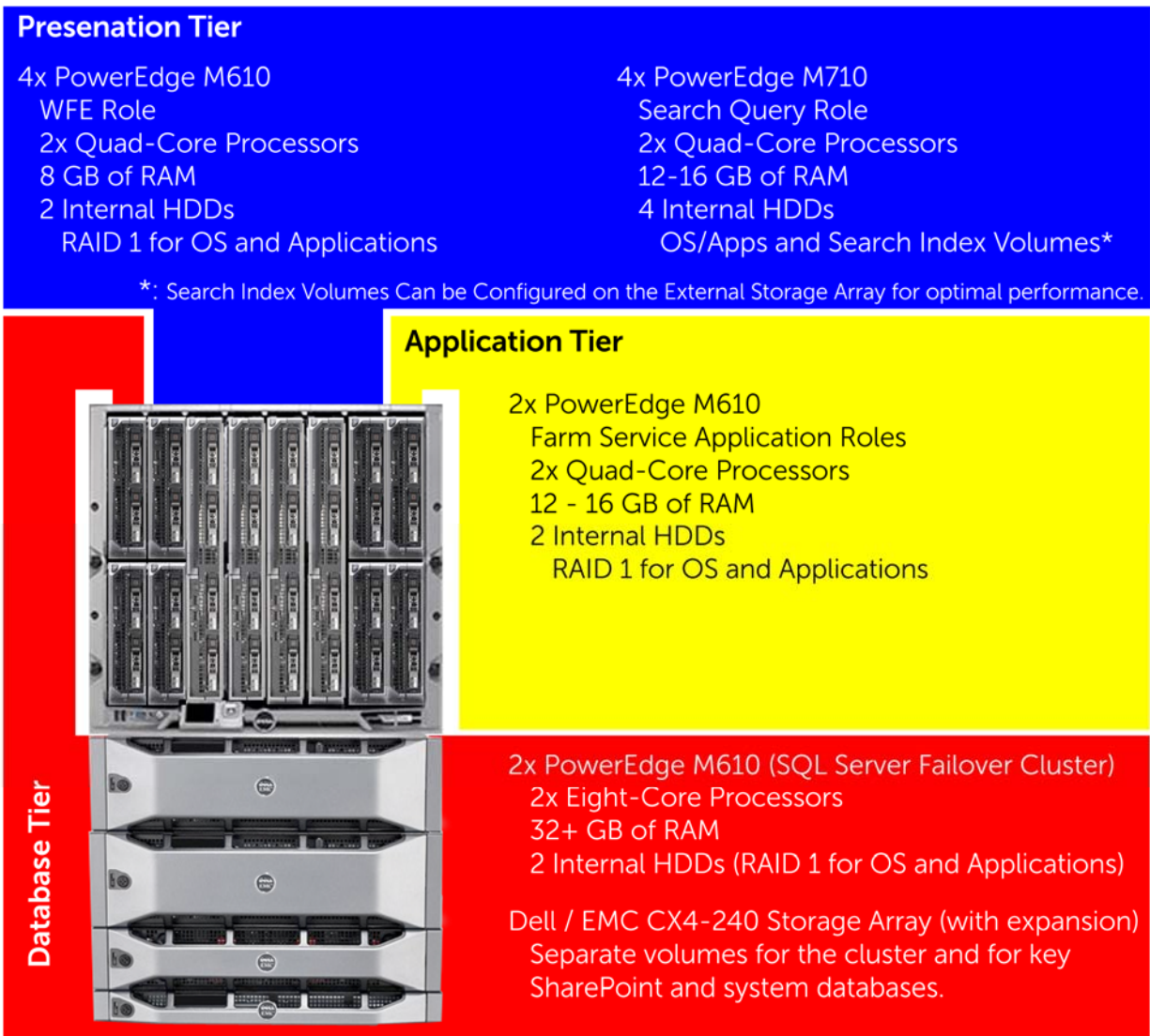


Figure 2: Large SharePoint Farm with a SQL Server Failover Cluster

NOTE: Instead of consolidating the roles, the farm in Figure 2 also features separate Web front-end and query servers. Such a configuration would be preferred if a significant percentage of the farm's activities involved search functionality; the crawler component could also be installed on the same host servers. The M710 was selected for query servers because it offers greater storage capacity than the M610. If either this capacity or the available IOPS are insufficient, then M610 servers could be used in conjunction with volumes configured on an external storage array.

The failover cluster employs a shared data storage array with integrated RAID controllers, such as the Dell / EMC CX4-series array in Figure X. The integrated RAID controllers in EqualLogic, PowerVault14, and Dell / EMC arrays help ensure that uncommitted writes - which reside in cache memory - are preserved in the event that a cluster node fails. These storage arrays also provide multipath I/O, which

protects the clustered SQL Server instance from failures related to the data paths between the host nodes and the storage array. In addition, expansion arrays can be connected to expand the capacity of these storage systems. Detailed installation and configuration details for failover clusters with Dell PowerEdge servers and the EqualLogic, PowerVault, and Dell / EMC arrays are available at www.dell.com/ha and support.dell.com.

SQL Database Mirroring

In SQL Server 2008, database availability can also be achieved using the database mirroring feature. This feature is implemented on a per-database basis. Database mirroring works by compressing transactions logs from the primary database, and shipping them to the mirror database on a separate server. The logs are applied to the mirror database, which allows both copies to reflect the same updates, inserts, and other changes to the database. SharePoint 2010 generates connection strings that help enable automatic failover to the mirrored database instance for many of its databases. For this to work best, configure mirrors in high-safety mode and employ a witness server; other types of mirror configurations do not allow for automatic failover. Such a database mirroring configuration was illustrated in Figure 1, which also features additional servers to provide availability for the Web front-end and application tiers of a large SharePoint 2010 farm.

Using Database mirroring, SharePoint 2010 can provide automatic failover⁵ of many SharePoint databases including configuration, content and service application databases. By default, all SharePoint 2010 databases are configured to use simple recovery model. To be able to use the automatic failover of the mirroring aware databases, you need to change the database recovery model to full recovery model. SharePoint 2010 does not support database mirroring for the web analytics staging database and the user profile synchronization database. Also, it is not recommended to mirror usage and logging database as this data can be quickly generated even in the case of a database failure.

High Availability for Web Front-End Servers

To increase availability and handle higher Web server workloads, Windows Server 2008 provides built-in software for Network Load Balancing (NLB). In an NLB cluster, multiple servers share a common IP address and distribute Web server workloads among nodes. If advanced load balancing heuristics, detailed monitoring, or statefulness for clients are required, consider a hardware load balancer.

Multiple Web front-end and query servers are shown in Figure 1 and Figure 2. In each of these sample farm architectures, one of these servers has been added for redundancy; provisioning an additional WFE server allows service levels to be maintained in the event of a failure. The query server role has also been either configured on all of the Web front-end servers, or on multiple dedicated servers to ensure availability.

The query component can be distributed across several servers. To ensure redundancy, each configured crawler index partition should be assigned to more than one query server. For example, consider a farm with four crawl index partitions (A, B, C, and D), and four servers hosting the query component (Q1, Q2, Q3, and Q4). To provide redundancy for the query role, each index partition should be mapped to at least two query servers. One such configuration is as follows:

- Query server Q1 is assigned index partitions A and B
- Query server Q2 is assigned index partitions B and C
- Query server Q3 is assigned index partitions C and D

⁵ [http://technet.microsoft.com/en-us/library/dd207314\(office.14\).aspx](http://technet.microsoft.com/en-us/library/dd207314(office.14).aspx)

- Query server Q4 is assigned index partitions D and A

In the above configuration, the redundant index partition assignments ensure that the loss of a single query server does not provide incomplete search results. If server Q1 were to fail, results from partition A would be handled by server Q4, and results from partition B would be returned by server Q2. The reference architectures illustrated in Figure X and Figure X can be configured in this manner to provide redundancy not only for the load-balanced Web front-end, but also for the query component.

Load-Balancing Web Front-End Servers

For the Web front-end server role, a load balancing solution can provide both performance and availability for the farm's services. Performance is derived from the fact that user requests to the SharePoint web sites are distributed among a pool of several servers that host the Web front-end role. Availability is derived by provisioning an additional server or servers to provide more capacity than is required to handle the volume of user requests to the Web sites.

Several load-balancing solutions are suitable for use with a SharePoint farm, including both hardware- and software-based options. Many of these solutions, such as Microsoft Internet Security and Acceleration (ISA) Server or hardware load balancers, can often provide better monitoring and more granular fault detection than is possible with the Microsoft Network Load Balancing (NLB) service. However, NLB is provided as a component of Windows Server 2008 and therefore offers a low-cost option for increasing the overall availability of a SharePoint farm.

When a load-balancing solution such as NLB is configured, a single network name and IP address is used to represent the pool of load-balanced servers. In addition to distributing the load among several nodes, NLB can detect a failed host and reroute the traffic to provide increased availability. However, NLB does not have the intelligence to detect many "soft faults," such as the IIS service not running on a node. Other solutions - such as hardware load balancers - are able to overcome this limitation, but were not tested during the development of this paper.

When a SharePoint Web application is configured in a farm with load-balanced front-end hosts, its Public URL should be configured to use the logical host name provided by the load balancing solution. The area outlined in red in Figure 3 shows where this is configured using the SharePoint Server 2010 central administration site. Similarly, the Public URL can be configured using the "URL" parameter of the `New-SPWebApplication` PowerShell cmdlet⁶.

NOTE: The Public URL field in SharePoint Server 2010 is analogous to the Load-Balanced URL field in MOSS 2007. This is the root name under which all sites in the Web application will be accessible to users. For example, consider the following scenario:

- The Web application in Figure 3 hosts a site collection for the Finance department that was configured to store its content under `"/finance"`.
- A site in that collection is named `"FY2010_Reports"`.
- Users would access this site by entering (or linking to) the following:
`http://SharePointNLB.example.local:25141/finance/FY2010_Reports/`

⁶ For a full description of this cmdlet, refer to its documentation on Microsoft Technet at [http://technet.microsoft.com/en-us/library/ff607931\(office.14\).aspx](http://technet.microsoft.com/en-us/library/ff607931(office.14).aspx)

NOTE: Unless default ports are used (e.g. port 80 for http and port 443 for https), users will have to specify both the host name and port name to access content. Advanced configuration options such as Alternate Access Mappings and Host Headers can be configured to make navigation easier. These settings may also be needed if the organization uses proxy servers or certain firewalls. Additional information about these options is available from the SharePoint Server 2010 documentation on Microsoft TechNet at [http://technet.microsoft.com/en-us/library/cc303422\(office.14\).aspx](http://technet.microsoft.com/en-us/library/cc303422(office.14).aspx).

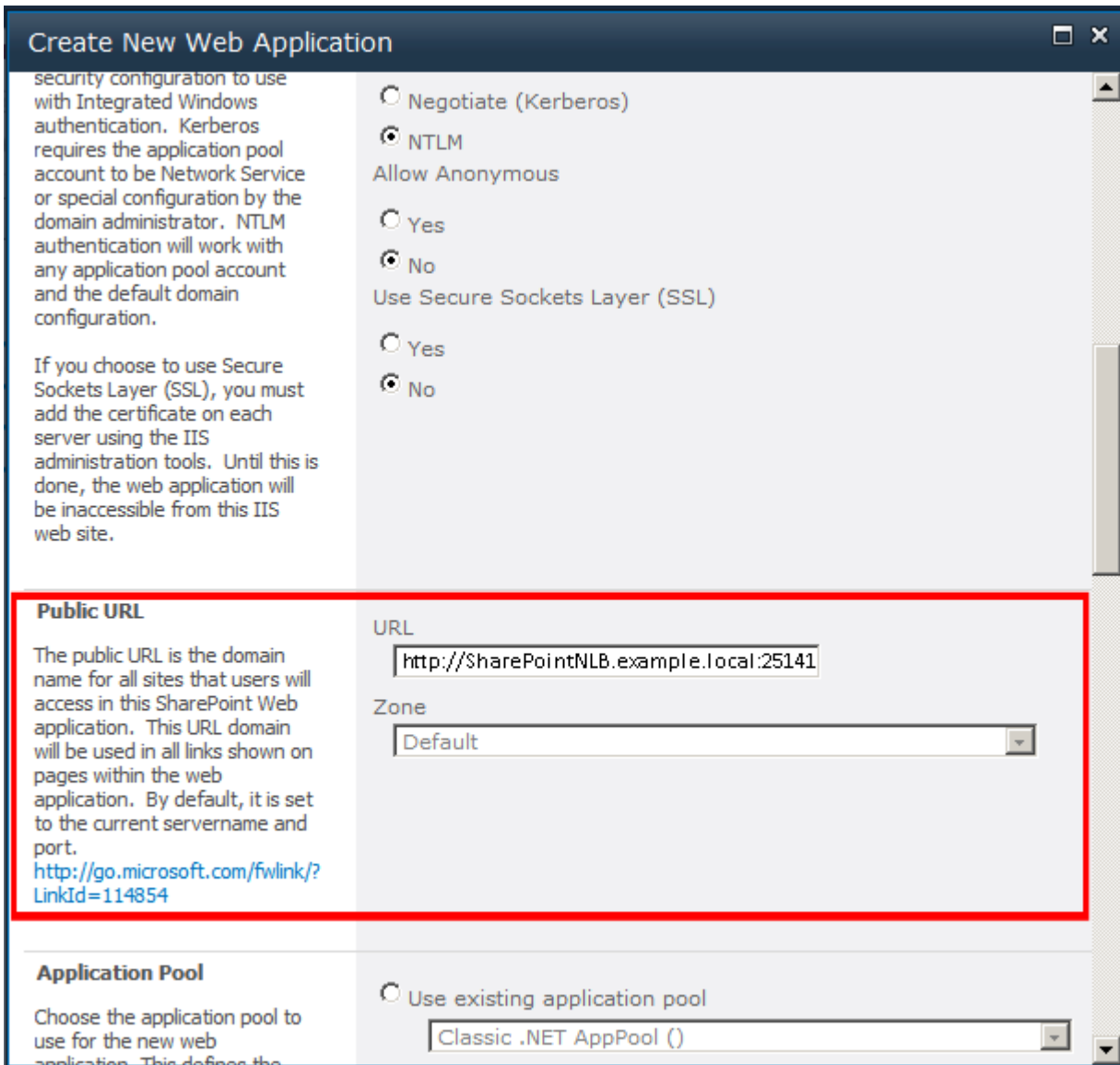


Figure 3: Web Application Public URL

Redundancy for Service Application Components

The application tier of a SharePoint farm hosts multiple components, such as the search crawler and other service applications (e.g. Excel Services, Word Automation, Managed Metadata, etc.). One major improvement is that the crawler component in SharePoint 2010 can be made highly available. This can be achieved by introducing more than one server to host this function in to the SharePoint farm, and

providing “N+1” hosts to create redundancy (i.e. if one crawler can complete a full crawl in the desired time, a second crawler can be added to provide availability for this function).

In the architectures illustrated in Figure 2 and Figure 3, high availability was provided for the crawler component by provisioning multiple servers at the application tier, and hosting the crawler component on more than one host. This same technique applies for most of the service applications as well.

NOTE: For farms that need to crawl a large corpus of content, it is also possible to provision more than one search crawler database. Each crawler database is associated with a unique search administration component, but can be shared among several crawler components. By protecting the crawl databases, provisioning multiple crawlers, generating index partitions, and assigning the index partitions to more than one query server, a SharePoint 2010 farm can provide end-to-end high availability for the search function.

Deploying the Search Service Application for Cross-Farm Sharing

In SharePoint 2010, many service applications can be shared across farms⁷. These service applications can be hosted in a central farm, and be consumed by other farms that require their services.

In a large SharePoint 2010 farm, it is recommended that service applications requiring high processing capabilities be run in a central farm. If the corpus of content requires that more than 40 million⁸ items be indexed, a separate search farm is highly desirable. Note that if a farm contains tens of millions of items, or indexes a significant amount of content outside the farm, it may also be preferable to deploy FAST Search Server 2010 for SharePoint. Creating a central search farm that provides cross-farm search services can reduce the administration overhead, and such a farm can be scaled-out relatively easily to accommodate further growth.

The following section outlines a central search service farm providing services to other farms in the organization. A similar configuration would apply if FAST Search Server 2010 for SharePoint was used, but the names and allocations of the search components would be different.

⁷ A listing of SharePoint Server 2010 service applications, including details about whether they can be shared among multiple farms, is provided in the “SharePoint Server 2010: An Introduction” whitepaper available from <http://www.dell.com/sharepoint>.

⁸ Refer to the scale-out decision points section of the Search architecture model <http://www.microsoft.com/downloads/details.aspx?FamilyID=22ffc029-2c08-457d-8311-ca457c6d160e&displaylang=en>

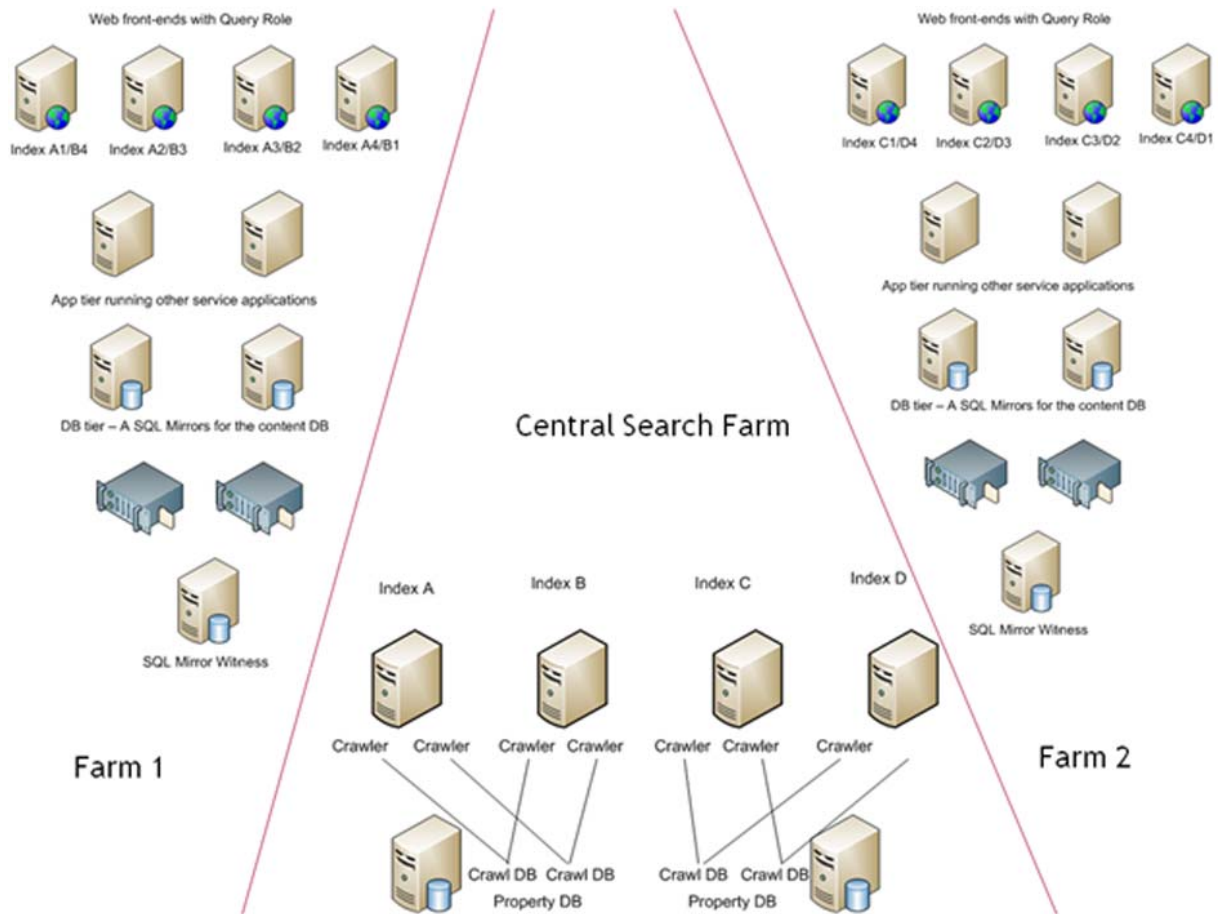


Figure 4: Centralized Search Farm

Figure 4 shows an implementation of shared search service application in a central farm. The central search farm publishes⁹ the search service application to other farms. Farm 1 and Farm 2 consume the enterprise search services provided by the central search farm, which implements a dedicated search application with four index servers. Each of the four index servers has two crawl components for higher levels of redundancy. The crawl components on index servers A and B are assigned to index farm 1, and crawl components on index servers C and D are assigned to farm 2. This can be achieved by using separate crawl databases and configuring host distribution rules that assign each content host to a dedicated crawl component.

In each of the farms, all four web front-end servers are configured host the query roles. Index partitions from each crawl component are propagated to the query servers. Since there are two index servers for each farm, the query components can be made redundant by assigning index partitions appropriately.

⁹ How-to publish a service application: [http://technet.microsoft.com/en-us/library/ee704545\(office.14\).aspx](http://technet.microsoft.com/en-us/library/ee704545(office.14).aspx)

Summary

A companion paper, *SharePoint Server 2010: An Introduction*, is available from www.dell.com/sharepoint. That paper provides a high-level overview of SharePoint Server 2010, and introduces key concepts and terminology that are helpful for understanding the reference architectures presented in this paper.

A large SharePoint 2010 farm consists of several servers, each of which is provisioned with different SharePoint components. Three-tier architectures are recommended - featuring a presentation tier that provides Web and search query functions, an application tier that provides search indexing and service application functions, and a database tier that hosts the SQL Server databases for the farm.

Because farms of this scale often contain important data or provide necessary business functions, the farm architecture should be designed with high availability and redundancy in mind. For the SQL Server database, the alternatives include failover clustering and database mirroring. SharePoint Server 2010 provides enhanced support for database mirroring, natively building SQL connection strings that are mirror-aware. At the application tier, SharePoint 2010 has enhanced its search architecture, and can now be more-easily configured to offer redundant search crawlers. Similarly, most service applications and the search query component can be hosted on more than one server. Finally, a load-balancing solution is recommended for the Web front-end servers.

When common services - such as search - are needed across farms, they can be provided by a central farm. This configuration allows the shared services to be designed, monitored, and managed independently of the farms that they serve.

Dell modular and rack-mounted servers are well-suited for all SharePoint 2010 farm roles. Some of these servers have been highlighted in the reference architectures presented in this paper, along with external storage arrays that can house the SharePoint databases. Additional details about Dell hardware are provided in the Appendices. Because SharePoint is both a content repository and a development platform, each farm is unique. Dell also offers services to help design, deploy, optimize and customize SharePoint Server 2010.

Appendix A: Selecting a Dell Storage Array

Dell offers several storage arrays that can be used to fulfill the capacity and performance need for the SQL Server databases and Index/Query Server volumes in a large SharePoint farm. The following table outlines some of the key considerations that impact the selection among the most popular of these arrays. Additional information about deploying these arrays with SQL Server is available at www.dell.com/sql.

Storage Array	Storage Technology	Capacity	Other Notes
EqualLogic PS Series	iSCSI	Sixteen 3.5" disks per array. Up to 12 arrays per group. One server can access volumes from multiple groups.	Supports thin provisioning that can be used to enable future growth. Arrays can be added non-disruptively. Volumes are designated based on capacity; administrators do not need to consider spindle count.
Dell / EMC CX4 Series	Fibre Channel or iSCSI	Fifteen 3.5" disks per array or expansion enclosure. Total number of disks varies based on the array model. CX4-120, CX4-240, CX4-480, and CX4-960 all offer up to the number of disks designated by the array's model number.	Supports thin provisioning that can be used to enable future growth.
PowerVault MD1220	SAS	Twenty-four 2.5" disks per array. Up to four arrays (96 disks) can be daisy-chained and connected to redundant x4 SAS busses. For optimal performance, limit daisy chaining and consider adding additional PERC H800 controllers.	Host-based RAID (PERC H800), so cannot be used with failover clusters. SAS, so cannot be attached to modular (blade) servers.

Appendix B: Selecting PowerEdge Servers

The table below introduces Dell PowerEdge servers that are well-suited for deploying large SharePoint farms, such as those illustrated in Figure 1 and Figure 2. The basic specifications that are most important for the SQL Server are listed first, followed by internal storage considerations that are important if the server will be used to host the Index Server or Query Server roles. The figures in this document showed a few representative samples of farms that can be built using different models of stand-alone and modular (blade) servers. It is reasonable to assume that systems with similar specifications can be readily substituted; regardless of their form-factor, these systems serve as strong building blocks for a large SharePoint farm.

Farm Role	Server	Notes
Database Server	R610	1U Rack form-factor dual-socket server. 12 total DIMM slots (6 per socket): for best memory bandwidth, populate all three channels. Four integrated GbE network ports. Two PCIe x8 expansion slots (e.g. for storage HBAs).
	R710	2U Rack form-factor dual-socket server. 18 total DIMM slots (9 per socket): for best memory bandwidth, populate all three channels. Four integrated GbE network ports. Up to four PCIe expansion slots (e.g. for storage HBAs or additional network ports).
	M610	Half-height modular form-factor dual-socket server (16 per M1000e chassis). 12 total DIMM slots (6 per socket): for best memory bandwidth, populate all three channels. Two integrated GbE network ports. Two dual-port expansion cards (e.g. for storage HBAs or additional network ports).
	M710	Full-height modular form-factor dual-socket server (8 per M1000e chassis). 18 total DIMM slots (9 per socket): for best memory bandwidth, populate all three channels. Four integrated GbE network ports. Four dual-port expansion cards (e.g. for storage HBAs or additional network ports).
	R905	4U Rack form-factor four-socket server. (AMD Opteron) 32 DIMM slots. Four integrated GbE network ports. Seven PCIe expansion slots (e.g. for storage HBAs or additional network ports).
	R910	4U Rack form-factor four-socket server. (Intel Xeon) 64 DIMM slots. Four integrated GbE network ports; or Two GbE and Two 10 GbE ports. Seven or Ten PCIe expansion slots (e.g. for storage HBAs or additional network ports).
	M910	Full-height modular form-factor dual-socket server (8 per M1000e chassis). 18 total DIMM slots (9 per socket): for best memory bandwidth, populate all three channels. Four integrated GbE network ports. Four dual-port expansion cards (e.g. for storage HBAs or additional network ports).

Farm Role	Server	Notes
Query Server, or Server Hosting both Web Front-End and Query Components	R410	1U Rack form factor dual-socket server. 8 total DIMM slots. Two integrated GbE network ports. One PCIe x16 expansion slot (e.g. for a storage HBA or additional NIC). Four 2.5" or 3.5" HDDs may provide adequate space for most index volumes. If additional IOPS are required, use external storage.
	R510	2U Rack form-factor dual-socket server. 8 total DIMM slots. Two integrated GbE network ports. Four PCIe expansion slots (e.g. for storage HBAs or additional NICs). Up to 12 3.5" HDDs, providing space and performance for most index volumes.
	R610	See basic specs in the "Database Server" section above. Six total 2.5" HDDs. External storage may be necessary to increase the capacity or performance of the index volume.
	R710	See basic specs in the "Database Server" section above. Six total 3.5" or eight total 2.5" HDDs. Four 450 GB 3.5" 15k rpm SAS drives in a RAID 1/0 configuration provides sufficient capacity for most index volumes, if additional IOPS are required, use external storage.
	M610	See basic specs in "Database Server" section above. Two total 2.5" HDDs. External storage (FC or iSCSI) should be used for the index volume.
	M710	See basic specs in the "Database Server" section above. Four total 2.5" HDDs. For farms with minimal search needs, it may be possible to house the index internally. However, external storage (FC or iSCSI) is still recommended.
Server Hosting Service Applications, or Web Front-End Server without a Query Component	R410	See basic specs in the sections above. For these roles, internal storage is not a consideration; the primary considerations are CPU cores and Memory capacity.
	R510	
	R610	
	R710	
	M610	
	M710	