Technical Report

# Microsoft Office SharePoint Server 2007 with EBS Best Practices Guide

Microsoft Business Unit, NetApp
June 2011 | TR-3819

**TABLE OF CONTENTS**

**LIST OF TABLES**

**LIST OF FIGURES**

# 1 INTRODUCTION

## 1.1 PURPOSE AND SCOPE

This document describes the architecture of a solution for deploying Microsoft® SharePoint® Server 2007 on NetApp® storage. It discusses how to size the SharePoint 2007 environment and provides detailed information about SharePoint performance.

## 1.2 INTENDED AUDIENCE

The report is intended for SharePoint and storage professionals who design, test, deploy, and manage corporate infrastructures. It is designed to serve as a reference architecture for customers, consultants, and others who are considering a SharePoint deployment with NetApp products. To use methods and procedures mentioned in this technical report, NetApp assumes the reader has working knowledge of the following:

- Microsoft Office SharePoint Server 2007
- Microsoft Office SharePoint Server 2007 architecture and administration
- Microsoft SQL Server® 2005/2008
- NetApp SnapDrive® 6
- NetApp SnapManager® for Microsoft SharePoint (SMOSS)
- NetApp SnapManager for SQL Server (SMSQL)
- NetApp Data ONTAP®

# 2 MICROSOFT OFFICE SHAREPOINT SERVER 2007 BEST PRACTICES

## 2.1 SELECTING A TOPOLOGY TYPE

This section describes how to design the information architecture, including logical structure, approximate sizing, and reliability and performance requirements for the environment in which to run the SharePoint product and technologies and SQL Server.

**SMALL SERVER FARM**

**Figure 1) Small server farm.**
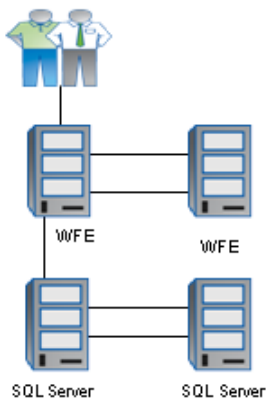
The small server farm topology consists of two servers and aims to serve small office environments. Adding a second Web front-end server and a database server can provide high-availability options.

This computer can optionally run the components for backward compatibility with Microsoft with one or more computers running Microsoft SQL Server 2000/2005/2008.

**MEDIUM SERVER FARM**

**Figure 2) Medium server farm.**
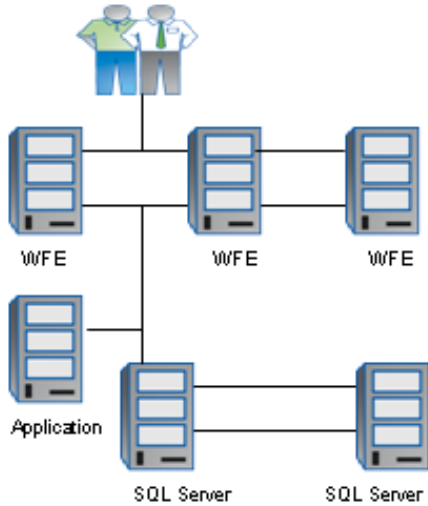


This solution is intended for larger offices utilizing a variety of the MOSS functional capabilities: intranet sites, collaboration, workflow, and indexing/search. Four servers build up this farm topology. Adding another Web front-end server and a database server can provide high-availability options.

**LARGE SERVER FARM**

**Figure 3) Large server farm.**

This solution is for large enterprise SharePoint deployments leveraging all of the functionality of MOSS: My Sites, business intelligence, Forms Services, intranet, content management, collaboration, workflow, and Indexing/Search. Eight initial servers make up this farm topology. The addition of more Web front-end servers and clustered database servers can provide further scalability and high-availability options.

## 2.2    DETERMINING THE CAPACITY OF SHAREPOINT SERVER MEMORY

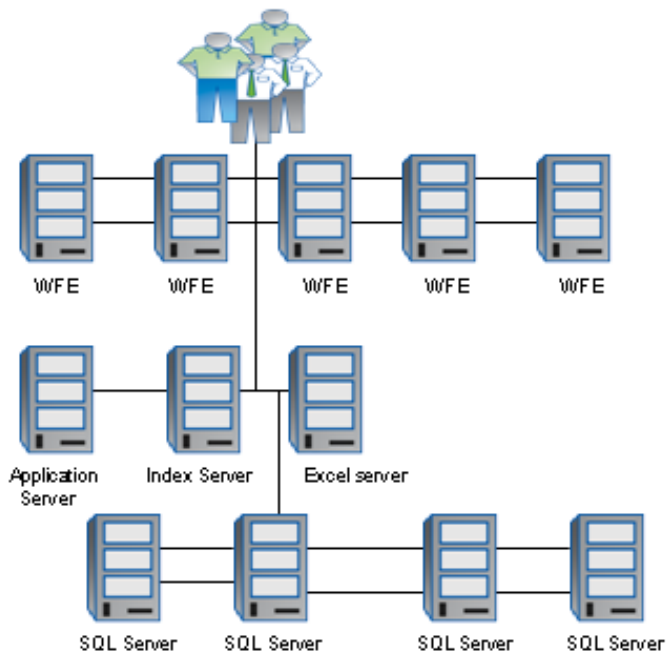Install Microsoft Office SharePoint Server 2007 with SP2 on a computer that has a new installation of Windows® Server 2008 and all critical updates. For the best performance of SQL Server hardware, 4GB is the minimum required memory for small-size deployment, 8GB is recommended for medium-size deployments, and 16GB and above are recommended for large deployments.

## 2.3    DEDICATING A SERVER TO RUN SQL SERVER

Assign a dedicated server to run SQL Server 2005/2008. Do not run other farms on this server unless you deploy a single server.

For SQL Server databases hosting SharePoint products and technologies, NetApp recommends installing the SQL Server 64-bit version on a 64-bit operating system unless you have a significant business reason not to. This is because we anticipate that this is the last version of the SharePoint product and technologies that will run on 32-bit operating systems and databases for optimal performance and that will host the SharePoint product and technologies on the latest version of SQL Server with the latest service pack.

## 2.4    SHAREPOINT 2007 CONFIGURATION ORDER

Microsoft recommends configuring Office SharePoint Server 2007 in the order listed below. This order makes configuration easier and enables the services and applications to be in place before they are required by the server types.

Install the central administration Web application on the following servers:

Application server, such as a query server that runs Excel calculation services, but not an index server (for performance reasons); if the farm will have an application server, install Office SharePoint Server 2007 on the server first, which also installs the central administration site.

- All the front-end server Web servers
- The index server
- Query servers, if separate from the index server
- Avoid configuring the index server as a query server
- Other application server (optional)

## 2.5    ADDING A SERVER TO THE SHAREPOINT FARM

To achieve a minimal server farm configuration initially, install SQL on a single server followed by the installation and configuration of Office SharePoint Server 2007. This server will operate as the Web server and application server (choose Complete on the Central Administration page). You can add more servers at a later time if you need to expand the farm.

## 2.6 DISABLING THE CENTRAL ADMINISTRATION SERVICE ON ADDITIONAL SHAREPOINT SERVERS

Central administration service is not required on index servers. Disable administration services on all index servers in farms with multiple index servers. However, do not stop this service on the index servers that host the central administration Web site.

**Figure 4) Disabling central administration.**
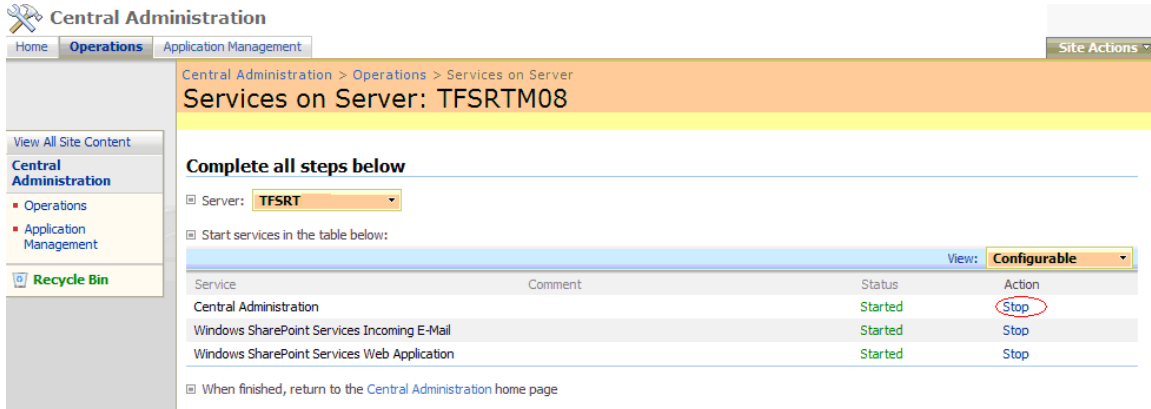


## 2.7 DISABLING THE WINDOWS SHAREPOINT SERVICE WEB APPLICATION SERVICES

**Figure 5) Disabling Windows SharePoint Service Web application services.**
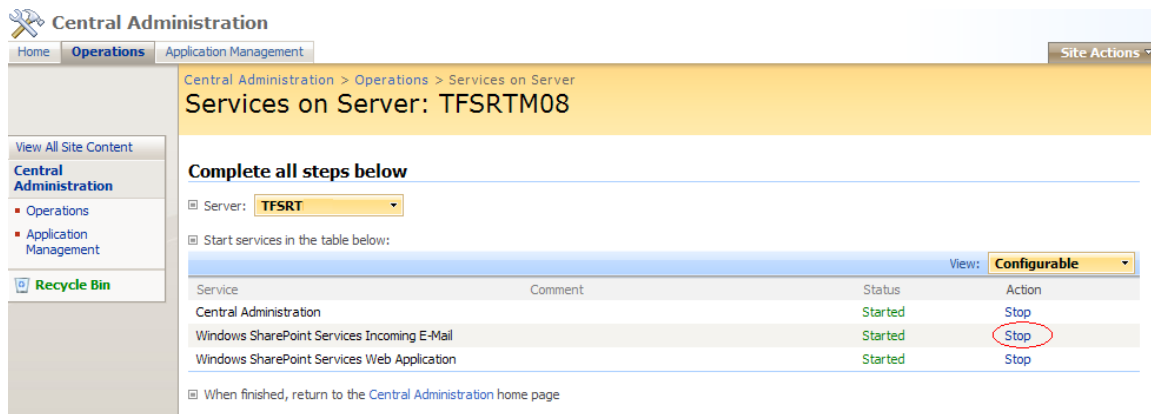


## 2.8 LIMITING CONTENT DATABASE SIZE

To enhance the manageability and performance of the SharePoint product and technologies, the content database should be smaller then 100GB. If the design requires a database larger than 100GB, split the content from a site collection that is approaching 100GB to a new site collection in a separate content database to avoid performance or manageability issues. This helps to enable the database to be backed up in reasonable windows.

## 2.9 USING THE MICROSOFT BEST PRACTICES ANALYZER FOR SHAREPOINT

Run the Microsoft Best Practices Analyzer tool called SharePointbpa.exe. This tool provides a detailed report listing the recommendations you can make to the environment to achieve performance, scalability, and uptime.
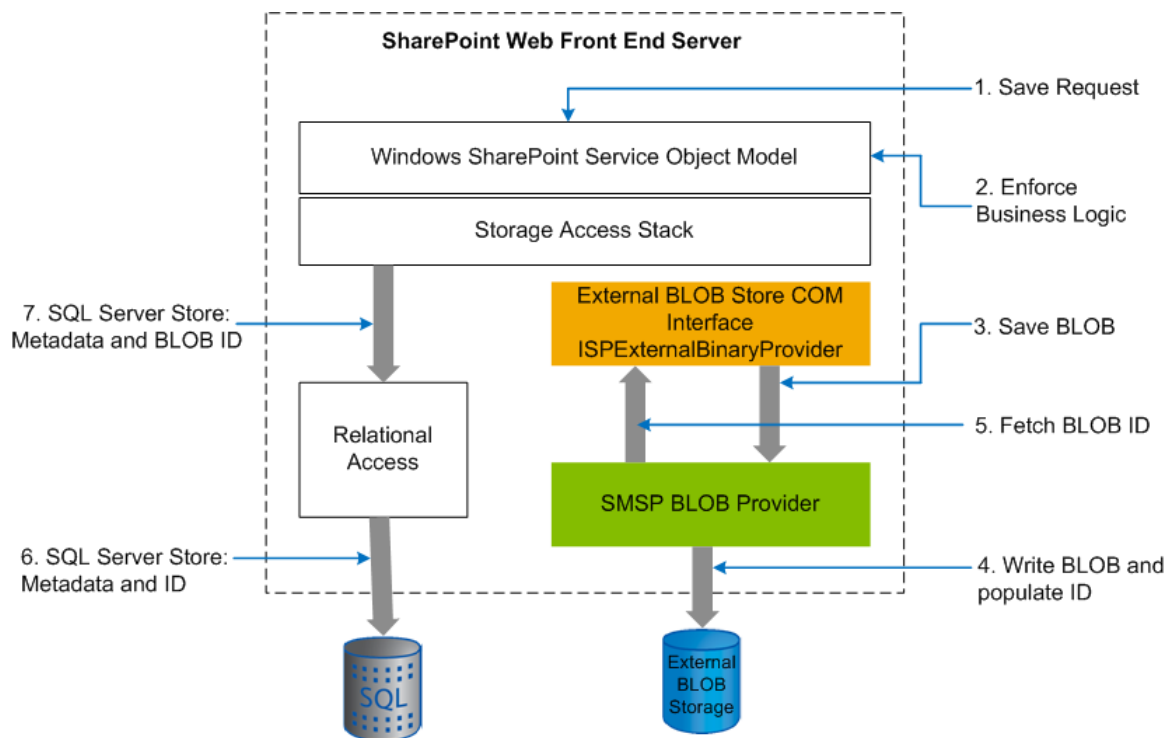
# 3  EXTERNAL BLOB STORAGE

## 3.1   OVERVIEW OF EXTERNAL BLOB STORAGE

The industry understands that Microsoft SQL Server is highly efficient in managing structured data. Because large deployments of SharePoint take place and the major part of the content represented as unstructured data is text files, images, office files, videos, and so on are represented as binary large (BLOB) data files.

SQL Server, with its fast query engine, can process large structured data and is not the best option for storing unstructured binary datastreams because unstructured binary data does not expose itself to the rich SQL query environment. Since BLOB data is not able to leverage one of SQL Server's most useful features, its query engine and value as a BLOB store are limited.

The solution, then, is to use a simple, cost-effective storage medium to function as an external BLOB store (EBS). The EBS runs parallel to the site's SQL Server content database that stores the site's structured data. To coordinate the two datastores, you must implement a COM interface (ISPExternalBinaryProvider), which uses simple semantics to recognize file Save and Open commands and invokes redirection calls to the BLOB store when it recognizes BLOB datastreams.

**Figure 6) External BLOB storage architecture.**



When you implement this interface, you can use the object known as the External BLOB Store Provider—the EBS Provider—which coordinates store and retrieve operations for the BLOB store. The EBS Provider also makes sure that the SQL Server content database contains metadata references to the associated BLOB streams in the external BLOB store.

## 3.2 ARCHITECTURE OF EXTERNAL BLOB STORAGE

When you save a document in SharePoint without the implementation of the External BLOB Store Provider, the entire file along with metadata and the BLOB data is saved to the SharePoint content database. SnapManager for Microsoft Office SharePoint Server provides the option to configure some business rules based on which BLOB data associated with these files can be extracted from the content databases and saved in the external BLOB storage.

Figure 7 shows the different components involved in external BLOB storage. The middle-tier storage stack routes BLOB datastreams and uses the EBS Provider to store BLOB data in the external BLOB store and returns the metadata for the BLOB, which helps in the retrieval. The SharePoint object model is completely unaware of the schematics of the EBS Provider.

**Figure 7) External BLOB storage architecture.**
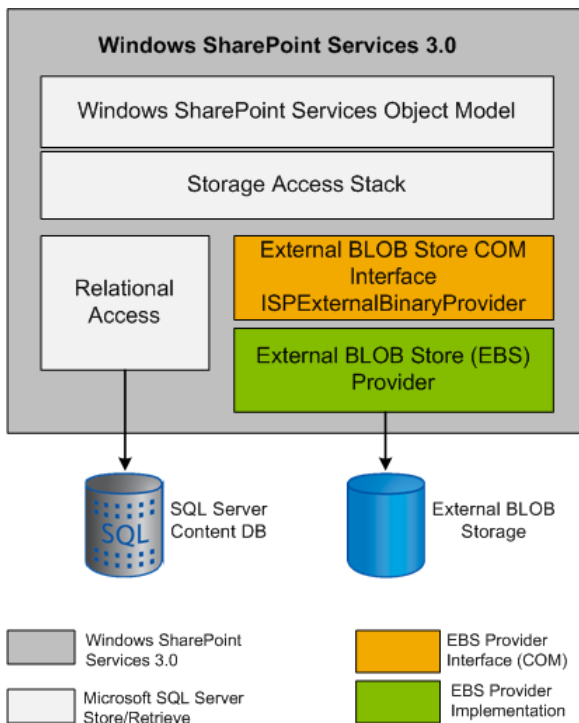


The deployment of external BLOB storage is a custom implementation using the solution from various independent software vendors (ISVs). SnapManager for Microsoft Office SharePoint Server capitalizes on the custom implementation of external BLOB storage.

## 3.3 BENEFITS OF EXTERNAL BLOB STORAGE

By using external BLOB storage, we can free important SQL resources by moving the BLOB data out of it and allowing the SQL database engine to perform with efficiency with the structured content. The content that stays in the content database is the metadata for BLOBs that have been moved out of the SQL database into the BLOB storage. For best performance, Microsoft recommends limiting the database to 200GB to allow a considerably large store for metadata. You can store unstructured content from the SQL database to inexpensive disks, which also helps reduce costs.

## 3.4 STORAGE PLAN FOR EXTERNAL BLOB STORAGE

External BLOB storage is typically on an external SMB file system as a network share. In the case of SnapManager for SharePoint Server, the BLOB data can be saved on an SMB/CIFS share or a LUN over SAN protocols in a NetApp storage system. NetApp recommends using the EBS store on an SMB/CIFS share on the NetApp NAS systems using inexpensive SATA disks drives.

- SAN configuration—Mount the LUNs to the Windows hosts using FCP/iSCSI protocol. This is where the SQL databases would be stored.
- NAS configuration—SMB/CIFS shares should be used for the EBS data (recommended).

The rule of thumb is to keep the SMB/CIFS share for the EBS store on a separate dedicated volume that is not shared by any other SMB/CIFS share or any other files.

Sizing storage guidelines for a BLOB store are:

- Average number of documents that are archived each time = N
- Average size of the documents = M
- Size of CIFS volume = (active BLOBs + orphan BLOB within delayed deletion window) + Snapshot$^{®}$ copies x snapshot average size

You can use 7200 RPM SATA disks for the EBS store and keep the 15K RPM FC/SAS disks for the SQL Server or, for that matter, any high-speed disk for a SQL workload.

## 3.5 PERFORMANCE RESULTS OF USING EXTERNAL BLOB STORAGE

Here are some of the conclusions from the tests performed on the external BLOB storage deployment.

- **EBS relieves SQL Server lock pressure**: Because Microsoft Office SharePoint Server 2007 writes all data for a given content database to a single table, there will be considerable lag in the SQL Server wait time due to concurrent data uploads. In this case data can be documents, spreadsheets, presentations, images, and so on.

**Figure 8) SQL Server locks average wait time with native SQL Server.**
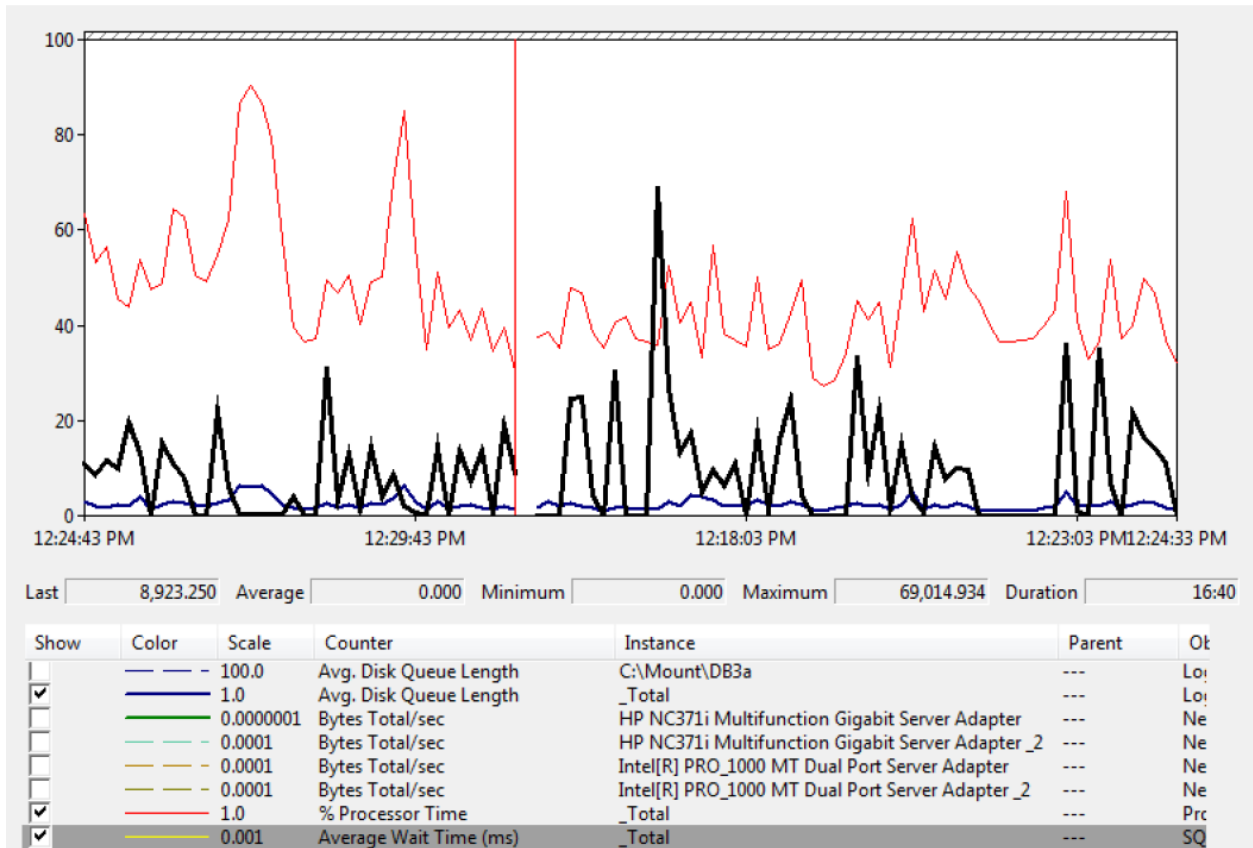
- **EBS reduces SQL Server CPU usage**: As EBS removes some I/O operations from SQL Server, the overall CPU load that SQL Server generates with BLOBs going directly to EBS is significantly less than for the same workload running against a native SQL Server BLOB store.

**Figure 9) SQL Server CPU utilization and lock times.**
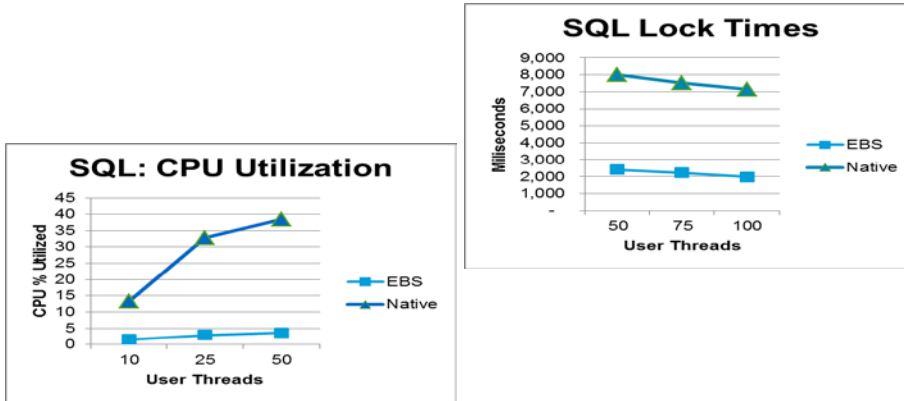


As shown in **Error! Reference source not found.**, SQL Server CPU utilization is much higher for native than for EBS even under a relatively light workload. Similarly there is a significant decrease in the SQL Server wait time across various workloads on an EBS-enabled document library compared to that of a native SQL Server.

- **EBS reduces disk I/O on the SQL Server**: Since EBS redirects BLOB read/write activity from SQL Server to the SMB share on the NetApp storage, there is a significant reduction in SQL Server disk I/O. SQL Servers that are disk bound will see a significant performance improvement after EBS is enabled.

**Figure 10) SQL Server average disk queue length with native SQL Server.**
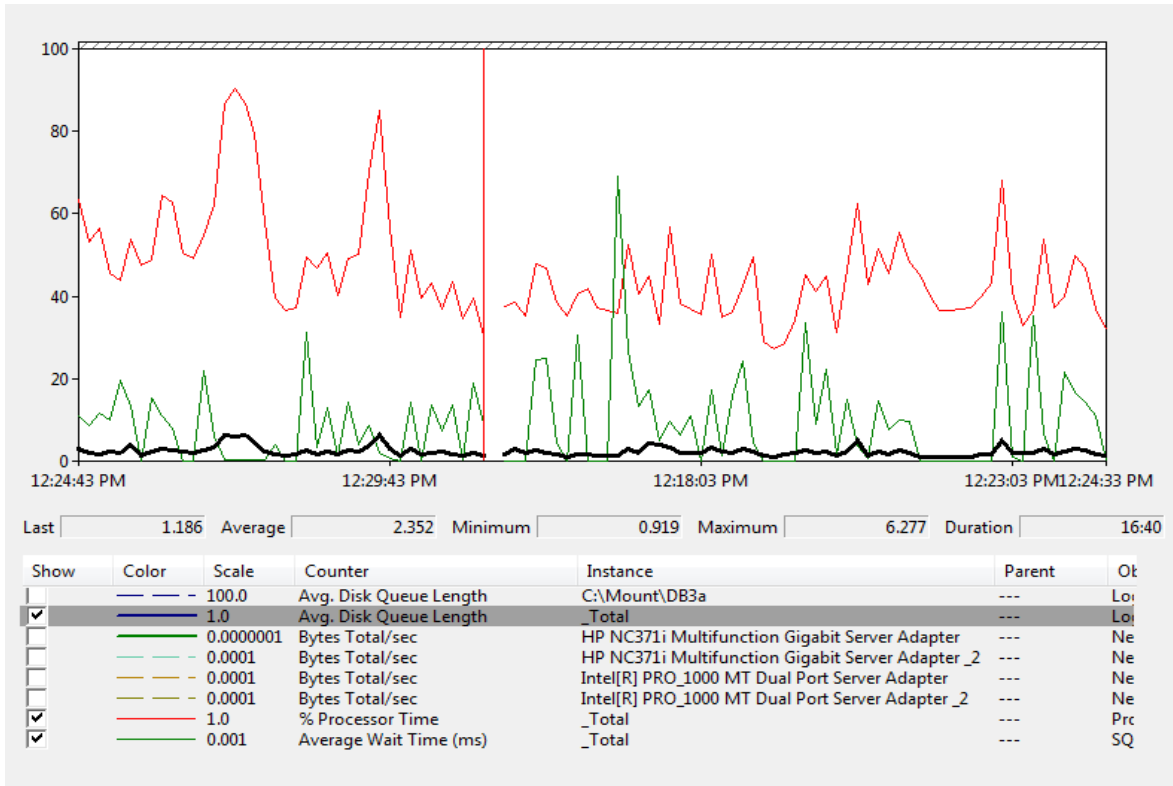
- **EBS increases SharePoint scalability**: The EBS-enabled site was able to serve more than twice the number of requests per second (RPSs) compared to native SQL Server databases for the same set of data.

For details on the EBS performance results and SQL performance improvement from using EBS, see TR-3874: Performance Report of SharePoint 2007 EBS Provider for NetApp Storage System.

# 4 SQL SERVER BEST PRACTICES

## 4.1 PLANNING FOR STORAGE PERFORMANCE

It is important to plan SQL Server storage for storage capacity, not performance or IOPs. With advances in disk technology, the disk drive's increase in storage capacity has outpaced the increase in IOPs by almost 1,000:1. Hence, the IOPs capacity is the standard to use while planning the SQL Server storage configuration. This is discussed in detail in the layout section of this report.

## 4.2 USE CLUSTER OR HIGH AVAILABILITY

A cluster is a collection of servers known as nodes that together provide a single, highly available system for hosting applications such as SQL Server 2008.

Microsoft clusters provide a highly available environment that can protect against SQL Server failures of hardware, operating systems, device drivers, or applications. If one node in a cluster is unavailable because of a failure, the service fails over to another node, which begins providing service, replacing the failed node.

## 4.3 SQL SERVER PART OF AD DOMAIN

Security and account management in SQL Server is through AD user accounts using integrated security. This method allows greater security at multiple levels and makes user management easier.

## 4.4 ENABLE SQL SERVER TO KEEP PAGES IN MEMORY

SQL Server dynamically allocates and deallocates memory based on the current state of the server. In an attempt to grab a substantial amount of memory, SQL Server may not be able to react quickly enough and the operating system may swap some of SQL Server's memory disk.

You can prevent the SQL Server memory from being swapped, known as Lock Pages in Memory. To do this the account that the Microsoft SQL Server runs under must be given the Lock Pages in Memory user rights.

DATABASE FILE SIZES AND AUTOGROW INCREMENTS

| Best Practice |
| --- |
| Set database file sizes and autogrow increment appropriately. |

SQL Server supports the ability to grow both data and log files automatically as they fill. NetApp recommends setting the file sizes appropriately and growing them manually at times of minimal system use on a planned basis. Autogrow should only be used as a safety net to prevent the files from becoming full and making the database read-only at times when unpredicted substantial growth occurs.

When database files are expanded there is an impact on performance. This impact is minimized but not eliminated through fast file initialization.

## DATABASE FILE GROUPS

| Best Practice |
|---|
| Plan DB file groups based on workload. |

SQL Server provides the ability to choose how and where to lay out your database tables and other structures on disk. The primary structure to control this behavior is a file group. Database structures are assigned to file groups, which contain files on disk where that data can be stored.

The placement of these data files is critical to the I/O performance of your database. NetApp recommends setting up file groups based in the database workloads.

## PROCESSORS AND MEMORY

| Best Practice |
|---|
| Retain the default values for processors and memory. |

When a SQL Server is first installed, most of its tunable parameters are set to default values. NetApp recommends leaving these parameters at their default levels on a server dedicated to SQL Server use. These values can be changed only in cases in which other workloads run in the same server or when issues arise from the use of the defaults.

By default, SQL Server runs at a standard priority and makes all processors in the system available for use. Also by default, SQL Server uses as much memory as it needs until it notices that memory pressure is starting to build. If other processes start consuming memory, SQL Server begins decreasing its memory footprint appropriately to decrease the possibility of swapping.

## TEMPDB

Tempdb is the oft-ignored component of SQL Server. The tempdb database is used by SQL Server to store intermediate results during query and sort operations. The tempdb can have a significant impact on any highly scaled client/server application.

Proper planning is required when tempdb disk I/O is limited. It is difficult to support a large SharePoint environment.

| Best Practices |
|---|
| • Size your tempdb. Since autogrowth causes this database to grow significantly, this is essentially a free option. When you consider that every time SQL Server is restarted the tempdb resets to its original size, you can see how SQL Server can spend a lot of time extending the tempdb. The tempdb total size should be set to 25% of the predicted size of your largest database. Autogrow should remain enabled. If the tempdb data file(s) autogrow beyond your predicted size, then they should be permanently resized, taking into account maximum perceived growth. <br>• Add tempdb data files such that the number of data files equals the number of CPU cores present in the SQL Server machine. Each data file, .mdf, and .ndf file should be equal in size. <br>• Place the tempdb database file on a LUN. <br>• Do not allow other data files from other databases to be placed on any of the tempdb logical units. |

For more information on creating SharePoint databases, see http://technet.microsoft.com/en-us/library/cc262869.aspx.

## STORAGE LAYOUT

This section provides examples of SharePoint and SQL Server designs for NetApp storage and takes into consideration environments that use SnapManager for SQL Server.

**Design Example 1: Small Server Farm**

To attain the optimal layout for SQL Server databases with NetApp storage systems, a basic set of best-practice recommendations needs to be followed. One best practice is to have a separate aggregate for each SQL Server instance. The content database data and log files should reside in separate volumes.

**Table 1) Small server farm.**

| Topology | # DB | SQL Instances | Avg. DB Size | DB Information | # Aggregate | # Volume |
|----------|------|---------------|--------------|---------------|-------------|----------|
| Small Farm Server | 35 | 1 | 10GB | Content DB | 1 | |
| | | | | Content DB TLogs | | 1 |
| | | | | SQL System DB and SharePoint | | 1 |
| | | | | Temp DB | | 1 |
| | | | | SnapInfo | | 1 |

The SQL Server and SharePoint system databases can share volumes but need to reside on separate LUNs. This means you need separate LUNs for the master database, model database, and the Microsoft database, or MSDB. Because the tempdb has a tendency to grow at a very fast rate, it needs to be placed on a separate volume. Placing the tempdb on a separate volume avoids the creation of a Snapshot copy of the tempdb.

For smaller production hosts, the SnapInfo directory can share a volume with the transaction logs.

**Figure 10) Small server farm design.**
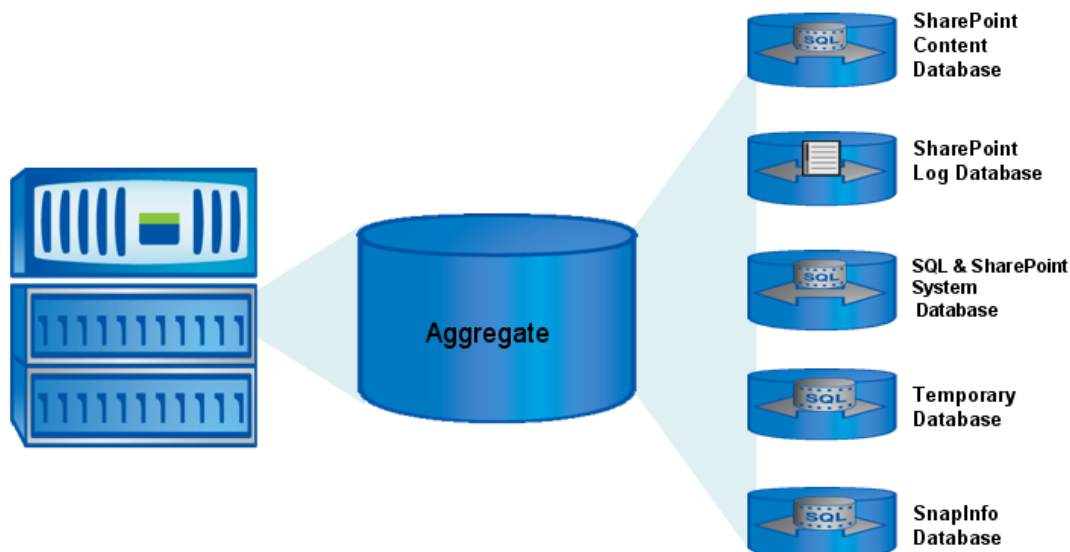


**Design Example 2: Medium to Large Server Farm**

In case of medium to large production hosts, it is best to have a dedicated volume for the content databases, Tlogs, SQL db, SharePoint Config db, Temp, and SnapInfo.

A SharePoint content database contains the collection of all content in the site collections that comprise the portal instance. SharePoint databases grow rapidly; with this in mind it is important to consider growth and overhead factors when determining how much content will eventually be stored in a given content database.

In large production environments, the size of the content databases grows rapidly. Crossing the Microsoft soft limit of 100GB, larger databases can cause performance to depreciate. Traditional backups/restores consume a lot of time and have manageability issues. SharePoint is query intensive, so the greater the number of databases, the more the distribution of the query load.

NetApp suggests a single volume for larger databases to isolate the incremental changes to a database such that they can be captured faster by the Snapshot process.

| Best Practices |
| --- |
| • Construct and size your content databases. Once content database sizes have been predicted, NetApp recommends creating the content databases using a script that generates the empty databases based on other recommendations in this document. Autogrow should be left on to prevent issues down the road. The content databases should be monitored so that the databases don't exceed predetermined limits. |
| • Separate each data file to exist on a separate LUN. |
| • Do not allow data files from different content databases to share LUN space. |

For more information on creating SharePoint databases, see http://technet.microsoft.com/en-us/library/cc262869.aspx.

SharePoint configuration databases are not very read or write intensive. SharePoint configuration databases and SharePoint admin content databases can be placed on a single volume and resources can be shared.

| Best Practices |
| --- |
| • Do not have multiple data files per database. |
| • All database data files for these databases can share a logical unit on shared physical disk spindle resources. |
| • The log files for these databases can share a logical unit on a shared physical disk spindle. |

**Search databases**

Search databases are a significant component of the SharePoint environment. Search databases store metadata value for every document that is served by a shared service provider. The other components that get stored in the search database are the history log, search log, crawl statistics tables, links table, and statistics tables.

Careful planning is required when deploying SharePoint, because SSP servers have many site collections consisting of many content databases. Microsoft recommends a limit of 50 million content items per index server. Placing search databases in a dedicated volume will yield good performance results.

**Temp databases**

Temp databases often ignore a component of the SQL Server—the tempdb. The tempdb database is used by SQL Server to store intermediate results during query and sort operations. So the tempdb can have a significant impact on any highly scaled client/server application.

Proper planning is required when tempdb disk I/O is limited, because then it is difficult to support large SharePoint environments.

The SnapInfo directory must be large enough to hold the maximum number of log backups that will be retained.

**Table 2) Medium to large server farm layout.**

| Topology | # DB | SQL Instances | Avg. DB Size | DB Information | # Aggregate | # Volume |
| --- | --- | --- | --- | --- | --- | --- |
| Large Farm Server | 100 | 1 | 25GB | SQL Instance 1 | 1 | |
| | | | | Content DB | | 3 |
| | | | | Content DB TLogs | | 3 |
| | | | | SQL System DB | | 1 |
| | | | | Temp DB | | 1 |
| | | | | SharePoint Config DB | | 1 |
| | | | | | | |
| | 50 | 1 | 50GB | SQL Instance 2 | 1 | |
| | | | | Content DB | | 2 |
| | | | | Content DB TLogs | | 2 |
| | | | | SQL System DB | | 1 |
| | | | | Temp DB | | 1 |
| | | | | SnapInfo | | 1 |

| Topology | # DB | SQL Instances | Avg. DB Size | DB Information | # Aggregate | # Volume |
|---|---|---|---|---|---|---|
| | | | | | | |
| | 35 | ` | 200GB | SQL Instance 3 | 1 | |
| | | | | Content DB | | 35 |
| | | | | Content DB TLogs | | 35 |
| | | | | SQL System DB | | 1 |
| | | | | Temp DB | | 1 |
| | | | | Search DB | | 1 |
| | | | | | | |
| | 5 | 1 | 150GB | SQL Instance 4 | 1 | |
| | | | | Content DB | | 5 |
| | | | | Content DB TLogs | | 5 |
| | | | | SQL System DB | | 1 |
| | | | | Temp DB | | 1 |

**Figure 11) Medium to large server farm design.**
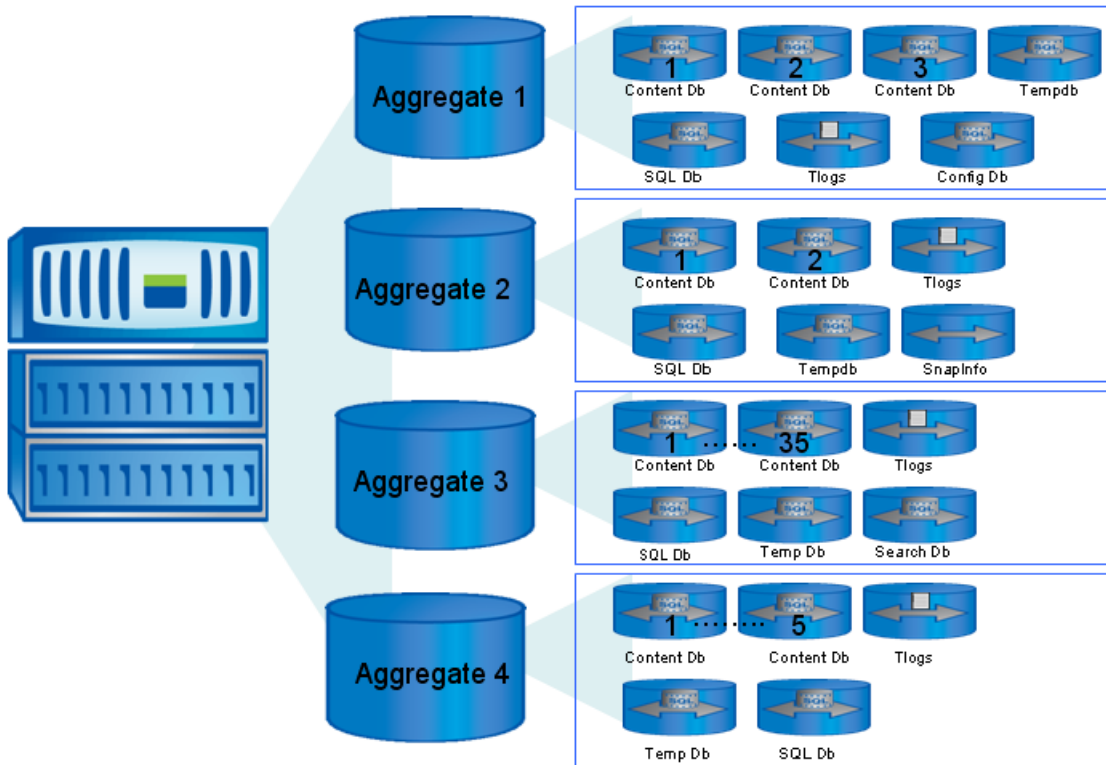


**Note:**

SnapManager recommendations are:

1. Databases per storage system volume: 35

2. Databases per LUN: 35

3. LUNs per SQL Server computer: 120

4. SQL Server instances per SQL Server computer or Windows cluster: 16

SnapManager does not prevent you from creating configurations that exceed these limits. Depending on the databases, you might encounter performance issues with less than the recommended maximums, or you might have no problem with greater than the recommended maximums.

When you map databases to storage system LUNs, you need to follow certain guidelines for the placement of the databases. When a database spans multiple LUNs, these LUNs should not be used by any other database. System and user databases should not be combined into a single LUN. Doing so will impose a streaming-based backup on the user databases as well as restrain Snapshot copy creation. Always distribute databases larger than 2TB across multiple LUNs after buffering them with several secondary data files. For sizing a MOSS environment on NetApp storage, use the NetApp tool available.

# 5  SUMMARY

This paper discusses general best practices and storage layout recommendations for Microsoft Office SharePoint Server 2007 environments along with EBS on NetApp storage. The combination of NetApp storage solutions and Microsoft Office SharePoint Server 2007, along with EBS, enables the creation of enterprise-level database storage designs that can meet today's most demanding application requirements.

# REFERENCES

Capacity Planning for tempdb:
http://msdn2.microsoft.com/en-us/library/ms345368.aspx

Estimate Performance and Capacity Requirements for Search Environments:
http://technet.microsoft.com/en-us/library/cc262574.aspx

Performance Recommendations for Storage Planning and Monitoring:
http://go.microsoft.com/fwlink/?LinkID=105623&clcid=0x409

Creating SharePoint Databases:
http://technet.microsoft.com/en-us/library/cc262869.aspx

SharePoint Disaster Recovery:
http://media.netapp.com/documents/tr-3714.pdf

NetApp Deduplication with SharePoint:
https://fieldportal.netapp.com/ci_getfile.asp?method=1&uid=8112&docid=12005

**SQL Database Layout Best Practices on NetApp Storage Systems**

TR 3693:  Microsoft SQL Server Relational Engine: Storage Fundamentals for NetApp Storage
http://media.netapp.com/documents/tr-3696.pdf

TR 3431: Snapmanager 5.0 for SQL Server: Best Practices Guide
http://media.netapp.com/documents/tr-3431.pdf

**SharePoint Sizing:**

Sizing Spreadsheet:
https://fieldportal.netapp.com/ci_getfile.asp?method=1&uid=8112&docid=11920

TR 3615i: SharePoint 2007 Sizing Guide
https://fieldportal.netapp.com/ci_getfile.asp?method=1&uid=8112&docid=9608

# ACKNOWLEDGEMENTS

www.netapp.com