

SHAREPOINT ARCHITECTURE FUNDAMENTALS

Whether you're a business user, manager, architect, developer, or IT pro, you'll want to understand the fundamental structure and core terminology of SharePoint because it will influence many choices you make, such as internal business ownership and total hardware costs. After reading this chapter, you'll understand the difference between sites and site collections, what comprises a portal in SharePoint, what shared services are, and why you could need as many as 20 servers or as few as one.

Functional Overview

Let's first review the key components of Microsoft Office SharePoint Server 2007, including the operating system, database services, Windows SharePoint Services, Office SharePoint Server applications, and shared services. It is important to understand which functionality is provided by which component and how these components relate to each other.

Operating System

MOSS 2007 is built on top of WSS 3.0, which, in turn, is built on top of the technologies and services provided by Microsoft Windows Server 2003. The core platform services use the Microsoft .NET 2.0 Framework. This combination of Windows and .NET provides SharePoint the following technologies:

- Internet Information Services 6.0 (for hosting Web applications)
- ASP.NET 2.x and above master pages, content pages (Web part pages), Web parts, personalization, membership, and navigation
- Windows Workflow Foundation (part of .NET 3.0)

Database Services

Microsoft SQL Server (either SQL 2000 SP4 or SQL 2005) is the relational database used for storing all content, data, and configuration information used by Office SharePoint Server 2007. Yes, that means that all content (including large documents) is stored in the database and not as files in the file system. Other relational databases, such as Oracle or MySQL, do not work and are not supported. If a separate database is not specified during installation, a specialized version of SQL Server 2005 Express is installed locally. The SharePoint-specific version of SQL Server 2005 Express has a database limit of 4GB and cannot be managed directly by SQL Enterprise Manager. For this reason, most organizations install MOSS in a farm configuration and specify a separate SQL Server machine for the database server.

Windows SharePoint Services

Windows SharePoint Services 3.0 builds on the operating system and database services to add additional features, such as team sites and collaboration features. Specifically, WSS 3.0 provides the following platform capabilities:

- **Storage.** Through content databases, which are literally SQL databases managed by SharePoint to accommodate the pages, data, and documents stored in the various portals, team sites, and workspaces
- **Management.** Administration pages with deep configuration options
- **Deployment.** Web farms, physical servers, and roles
- **Site Model.** Web application, site collection, and sites
- **Extensibility.** Features, Web parts, and templates

WSS provides more than just these core technology services. Microsoft decided to make WSS a powerful application out-of-the-box and thus provides the core collaboration features for MOSS:

- Document collaboration—check-in/out and versions
- Wikis and blogs
- RSS support
- Project task management (lightweight functionality, which should not be confused with Office Project Server 2007, also built on WSS 3.0)

- Contacts, calendars, and tasks
- E-mail integration
- Integration with Office client applications

Office SharePoint Server 2007: Applications and Services

Architecturally, Office SharePoint Server 2007 consists of a common set of shared services that support five server application components:

- **Portal.** Templates, people, audience targeting
- **Search.** Search center, cross-site search
- **Content management.** Authoring, publishing, records management
- **Business process.** Forms server, line of business (LOB) integration
- **Business intelligence.** Excel services, Key Performance Indicator (KPI) lists, Report Center (not to be confused with Business Scorecard Manager and Office PerformancePoint Server 2007, which both provide additional BI capabilities)

Each of these is built on the platform services and collaboration components of Windows SharePoint Services and the shared services components of Office SharePoint Server 2007.

Shared Services

Shared services provide the features that are used by multiple applications in MOSS 2007. What does that mean? Let's use an example—user profiles. You might want to use the user profile feature, which provides an out-of-the-box employee directory, including basic information (name and phone number, for example), along with some custom properties and a photograph. You might also want to create several different portals within your organization—for example, an Internet presence, an employee intranet site, and a collaboration portal for self-service team site use. You wouldn't want to create and manage three separate profile databases. In this case, the user profile service can be shared across the various portals—hence a *shared service*. Specifically, the following features are provided by shared services in MOSS 2007:

- User profile store
- Audiences
- Search services
- Usage reporting
- Excel services
- Business Data Catalog (BDC)
- Notification service for generating alerts
- Single sign-on services

So what exactly do shared services support? They support the fundamental element of SharePoint: sites. Some of a site's services are site-specific, while others are shared and provided by a Shared Services Provider (SSP), which we will discuss shortly. When a site is created, it is based on a template. You can think of a site as the cookie and a template as the cookie cutter. In the next section, we discuss sites, templates, and more about shared services.

Sites, Site Collections, Templates, and Shared Services Providers

Fundamental concepts in SharePoint create every portal, team site, workspace, Internet page, and extranet site:

- **Lists.** Lists are a data repository that can hold columns of data and/or documents. Visually, a list is represented by a Web part. It is analogous to a database table or Excel worksheet.
- **Sites.** A site consists of a data repository, visual elements, administration, and almost every other core element of the functionality and experience for the user. Visually, a site is represented as one or more Web pages, lists, and Web parts.
- **Web Application.** In IIS, a Web application is composed of an Internet Information Services (IIS) site with a unique application pool. While not a technically perfect definition, you can think of a Web application as a URL such as `http://my.intranet.com` or `http://sharepoint.intranet.com`.
- **Site Collection.** A site collection consists of a top-level site and its subsites. It is a logical unit for administration—there are settings that can only be configured at the site collection level (in other

words, at the top-level site). Each Web application can host many site collections.

- **Template.** A template defines what the site will look like, what lists comprise the site initially, how publishing will work on the site, and a number of other settings. It enables a site to be created via self-service using a precreated definition.
- **Shared Service Provider (SSP).** An SSP is a collection of shared services grouped and presented to one or more site collections. In WSS, no SSP is associated with a site. However, in MOSS, an SSP can provide a site with shared services (such as search, Excel services, and so on).

Sites and Site Collections

In WSS 3.0 and MOSS 2007, everything is a site. There is no longer the concept of an SPS *area*, as there was in SPS 2003. To explore the concepts of SharePoint, let's start with a simple example composed of a single Web server and its logical elements (see Figure 4.1). At the highest level, you have a physical server running Internet Information Server (IIS). Within IIS, you have a Web application, which maps to a URL (such as `http://myportal`), a port (such as 80), and an IP address (such as 192.168.1.4). Once a Web application is extended with SharePoint functionality, one or more top-level sites can be created. Each top-level site can contain one or more child sites. The collection of sites that includes a top-level site and all of its descendants down to the leaf site(s) is called a *site collection*. This is important given that much of SharePoint administration (quotas, backup and restore, permissions, Web part access, and so on) is based on the site collection concept.

After you determine which sites and portal sites your solution requires, the next step is to plan how these sites and portals are implemented across site collections. A site collection is a hierarchical set of sites that can be managed together. Sites within a site collection have common features, such as shared permissions, galleries for templates, content types, and Web parts, and they often share a common navigation. All sites in a site collection are stored together in the same SQL database. A portal site is often implemented as a site collection with the top-level Web site as the portal's homepage.

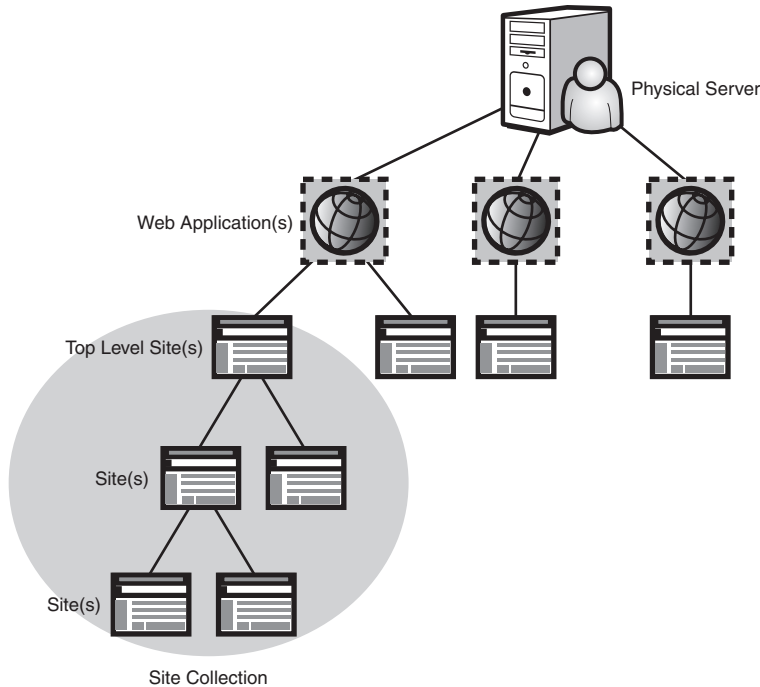


FIGURE 4.1 In SharePoint, a top-level site and its descendants are collectively referred to as a “site collection.”

In general, we recommend that you put each of the following types of sites in separate site collections right from the start. This will help you manage site collections and content databases better in the long run.

- Intranet portal sites
- Extranet sites
- Team sites related to a portal site or Internet site
- MySites (by default, each MySite is a site collection)
- Internet sites (staging)
- Internet sites (production)
- Lines of business within a conglomerate
- Document Center sites
- Records Center sites

For example, if you were to deploy a company intranet, a corporate Internet-facing site, and a records management repository, you’d want to create three site collections right at the start. This enables you to manage

the site collections individually, provide separate content databases, and more easily accommodate growth over time.

The downside of multiple site collections is that some features do not work *across* site collections. This is important because a large deployment of SharePoint will dictate multiple site collections. The following features do not work across site collections:

- **Content Types.** How common documents, forms, and other content are normalized in your organization.
- **Content by Query Web Part.** This Web part aggregates information from across sites but does not work across site collections.
- **Workflow.** When you deploy workflow, it is accessible only within the site collection it is deployed in.
- **Information Management and Retention Policies.** Records management policies are set at the site collection level, forcing organizations to deploy the same policy multiple times for large enterprises.
- **Search.** Certain features of search are configured at the site collection level.
- **Quotas.** You should absolutely define quotas so that users are used to limited storage from day one; also configured at the site collection level, which means that you will need to configure quotas separately at each top-level site.

Let's say you decided to create two site collections for project workspaces: one site collection for IT and one site collection for finance. Due to the site collection limitation described earlier, if you wanted consistent document metadata properties on a particular document type, you'd have to deploy the content type twice—once for each site collection.

So far, all of this is true for both WSS 3.0 and MOSS 2007 deployments. When you install MOSS 2007 over and above WSS 3.0, several additional capabilities are added to all sites—additional Web parts, additional templates, and more features, some enabled by shared services.

In summary:

- **WSS 3.0.** Site collections, sites, templates (collaboration only), and no shared services
- **MOSS 2007.** Site collections, sites, templates (collaboration, Internet presence, and portal), and shared services

NOTE “Web application” is the new terminology for an IIS Web site/IIS virtual server. IIS Web site = IIS virtual server = Web application.

Templates

Because a site is simply a Web part page with some administration and some lists to back it up, how do we get a different look and behavior for each site in SharePoint? Templates. Templates are collections of lists, Web part pages, and Web parts that are packaged together to define a starting point for your site. Given that everything is a WSS site in SharePoint, the template defines the look and behavior of the page(s). Table 4.1 lists the out-of-the-box templates in MOSS 2007. In addition to these templates, you can make an unlimited number of custom templates available to users. Templates are your building blocks, allowing you to quickly create complex solutions that include custom branding and functionality. Microsoft has provided an additional list of templates called the “fantastic 40,” which is a collection of additional templates for WSS 3.0, available at <http://www.microsoft.com/technet/windowsserver/sharepoint/wssapps/templates/default.aspx>.

Table 4.1 Out-of-the-Box Templates in WSS 3.0 and MOSS 2007

Category	Name	Best Suited For...	MOSS Only?
Collaboration	Team Site	Team Collaboration	
Collaboration	Blank Site	Custom Collaboration Applications	
Collaboration	Document Workspace	Actively working on documents	
Collaboration	Wiki Site	Adding, editing, and linking Web pages	
Collaboration	Blog	Posting information in chronological order, and others can comment	
Collaboration	Internet Presence Web Site	An Internet-facing corporate Web site with several various pages like publishing and search (site collection level)	Yes
Collaboration	MySite Host	A site used for hosting MySites (site collection level)	Yes
Collaboration	Records Repository	Storing documents that should not be modified after being added	Yes
Collaboration	News Home template	Creating and managing new articles	Yes

Category	Name	Best Suited For...	MOSS Only?
Collaboration	Publishing and Team Collaboration Site	Publishing Web content along with team collaboration	Yes
Collaboration	Publishing Site	Publishing Web content (like an intranet or Internet site)	Yes
Meetings	Basic Meeting Workspace	Tracking meeting data	
Meetings	Blank Meeting Workspace	Tracking meetings	
Meetings	Decision Meeting Workspace	Tracking meetings that lead to a decision	
Meetings	Social Meeting Workspace	Planning social events	
Meetings	Multipage Meeting Workspace	Tracking meeting data	
Enterprise	Document Center	Centrally managing documents (active, broadly published items)	Yes
Enterprise	Records Center	Centrally managing records (corporate "sealed" items)	Yes
Enterprise	Personalization Site	Hosting a MySite page for a user	Yes
Enterprise	Site Directory	Managing sites in a taxonomy	Yes
Enterprise	Report Center	Managing reports and BI information	Yes
Enterprise	Search Center	Hosting a central search page	Yes
Publishing	Publishing Site	A blank site for quickly publishing HTML Web pages	Yes
Publishing	Publishing Site with Workflow	A site for publishing Web pages on a schedule by using approval workflows	Yes
Publishing	Corporate Intranet Site	Creating an intranet site with news, sites directory, and search (much like the portal template in SPS 2003) (site collection level)	Yes

Shared Services Providers

As we mentioned before, a Shared Services Provider (SSP) is a set of services (for example, full-text search) housed under its own special site that provides functionality to one or more Web applications in MOSS 2007. (A WSS-only installation does not provide shared services, but a server farm with MOSS 2007 installed will enable shared services features to be used in WSS sites.) Each Web application in MOSS 2007 is associated with at

most one SSP, and more than one SSP can be configured per farm. In this context, a farm is a collection of SharePoint Web front-end servers that point to the same back-end data.

However, you might want to have more than one shared service provider, given that some of your sites might need some services configured one way, while other sites might need those same services configured another way. Typically, we recommend a single shared services provider unless you are using different authentication providers, considering multiple SSPs are very complex to set up and maintain.

Figure 4.2 shows how a single SSP can be applied to multiple sites. Going a step further, Figure 4.3 shows how a single farm can have two SSPs—one serving two portal sites and the other serving a portal site and a team site. So can an SSP service a team site? Yes, provided MOSS 2007 is installed in the farm. All this being said, however, most installations will probably use only one SSP for simplicity. The main reason for using more than one is to isolate services and service data from a security perspective. This is likely in hosted environments, in very private or restricted sites, in large organizations that have many independent business units, and in politically charged organizations that need to separate everything.

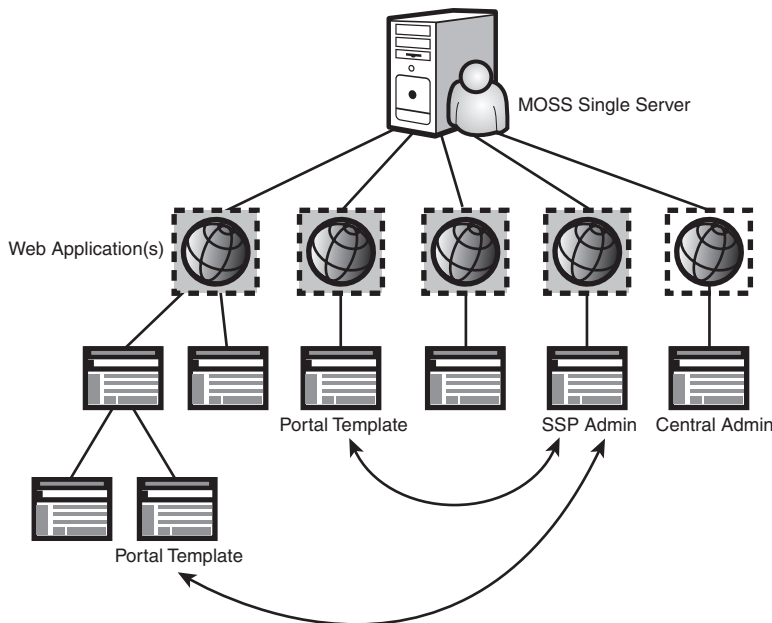


FIGURE 4.2 A single Shared Services Provider (SSP) can serve more than one site.

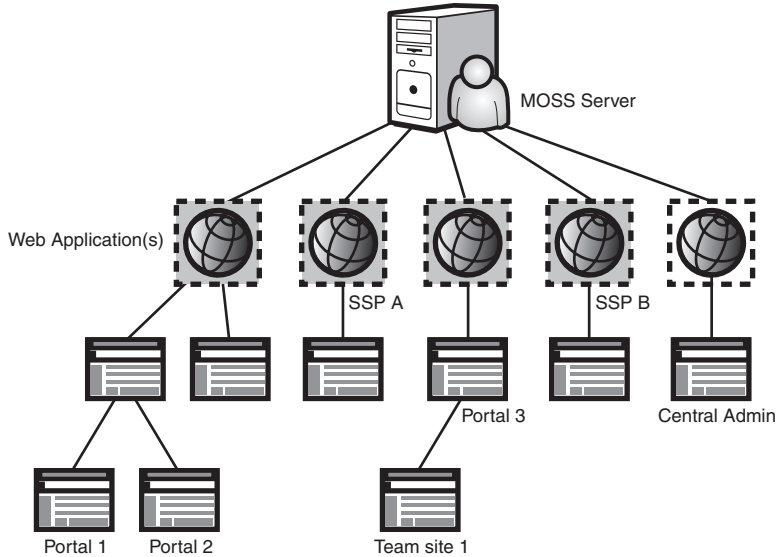


FIGURE 4.3 Typically, you'll need to configure only one SSP. However, you can create multiple SSPs if your environment needs them—for example, if you need to keep your business units separate.

The shared services in MOSS include search (scopes, content sources, indices, and so on), MySites, profiles, directory import, audiences, alerts, targeting, BDC, Excel calculation services, and usage reporting. MOSS 2007 does not support shared services over a WAN. This factor can impact design and deployment in large organizations.

Putting It All Together

To illustrate how sites, templates, and shared services work together, consider this: A “portal” (in SPS 2003 terminology) is now constructed simply by using a WSS site (after all, everything is now a site), plus a portal template (for example, the Corporate Intranet Site template), plus some shared services. This gets you a portal.

Sites, templates, and SSPs are important to understand, given that how you configure your portal and team sites largely depends on that understanding. Another consideration is who in your organization will administer various parts of your SharePoint environment. The next section covers how SharePoint administration is segmented and why it matters.

Understanding SharePoint Administration

Administration in SharePoint is a set of Web pages that allow both IT pros and business users to configure settings and add new content. In general, administration is broken out by role and grouped by type of task.

There are fundamentally three tiers to SharePoint administration:

- **Central Administration.** Where all global SharePoint settings are configured
- **Shared Services Administration.** Unique settings for each SSP
- **Site Level Administration.** Unique settings for each site

Central Administration

There is one Central Administration per farm; this enables settings such as topology, security, and enabled services.

Who? IT administrators

Where? Farm-level

How many? *One per farm*

Used for things such as adding a new physical server to the farm

The Central Administration site has three tabs: Home, Operations, and Application Management. Home gives you a link to tasks you'll need to complete to get your farm up and running (see Figure 4.4).

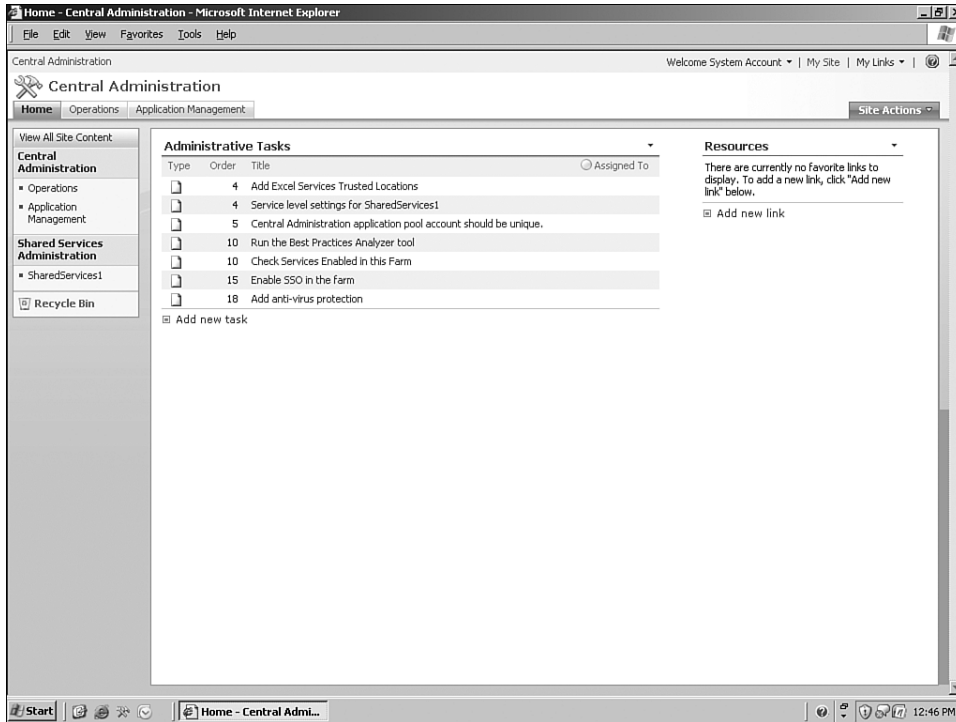


FIGURE 4.4 The homepage of SharePoint Central Administration provides you with the core tasks you'll need to undertake to get your farm working properly.

The Operations tab contains links to pages that help you manage your server or server farm, such as changing the server farm topology, specifying which services are running on each server, and changing settings that affect multiple servers or applications (see Figure 4.5).

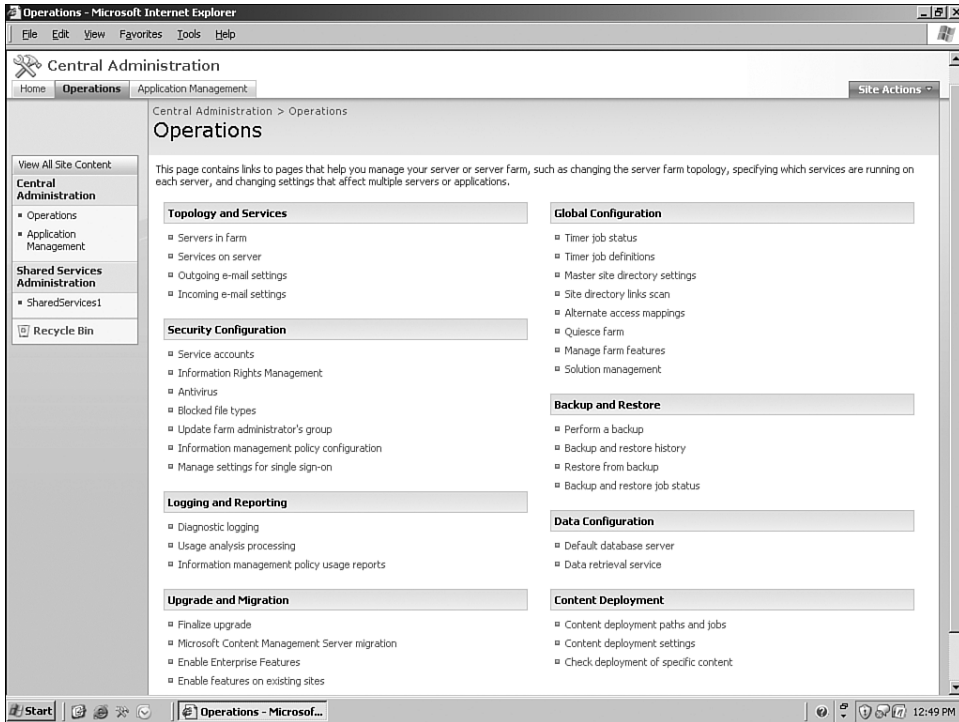


FIGURE 4.5 The Operations tab of SharePoint Central Administration provides physical and logical configuration settings for your farm.

Finally, the Application Management page contains links to pages that help you configure settings for applications and components that are installed on the server or server farm (see Figure 4.6). Central Administration also has a topology view, which provides a quick look at the farm's servers and what is running on them.

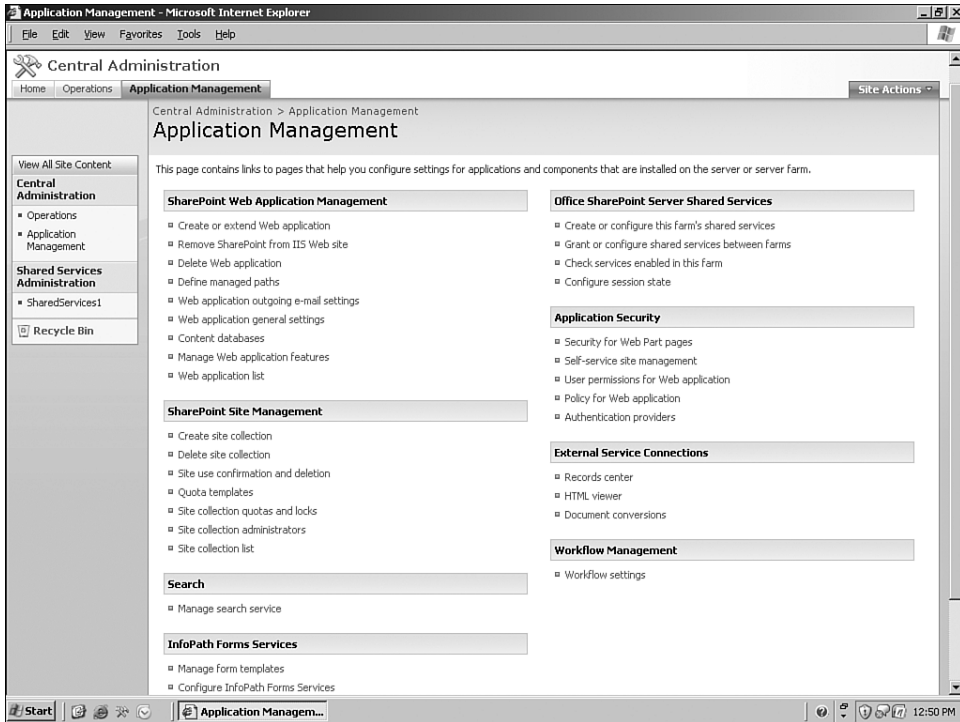


FIGURE 4.6 The Application Management tab of SharePoint Central Administration provides you with ways to configure your application components, such as Web application settings, workflow, and InfoPath forms.

Shared Services Administration

There is one Shared Services Administration per SSP. It provides services such as search, profiles, and audiences.

Who? IT Administrators (business unit)

Where? Can be used cross-farm—one SSP to many Web applications

How many? *One SSP per farm or organization*

Used for things such as creating a search content source

The primary usage of the SSP site is to provide an interface where IT pros within a business can manage their shared services (see Figure 4.7). This includes administration of user profiles, MySites, search, usage reporting, audiences, Excel services, and the BDC.

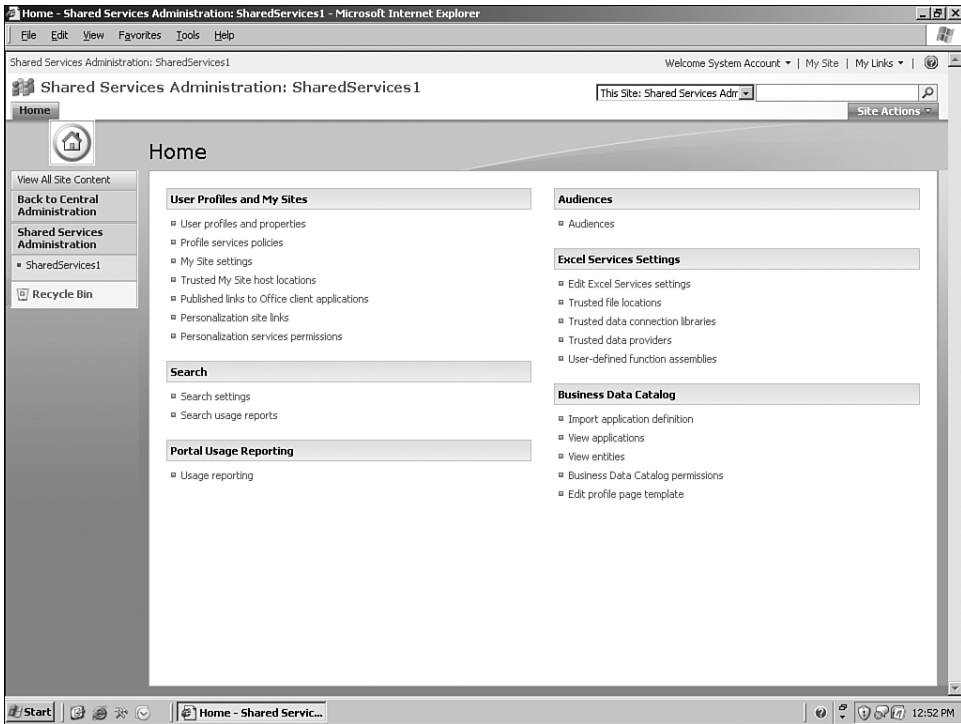


FIGURE 4.7 SharePoint Shared Services Administration also gets you to a place where you can administer shared services. There will be a link to each SSP you have configured. This provides methods to configure your shared application components, such as audiences, Excel services, and the BDC.

Site Settings and Site Collection Settings

Site Settings Administration for a specific site (with special options for a top-level site) enables the user to specify additional site collection settings.

Who? Business user or IT (site owner)

Where? Every site

How many? *One admin page per site, with an extra column for site collection settings for top-level sites*

Used for things such as site configuration, creating new lists, adding users to the site, storage, and site hierarchy

The primary usage of the site settings page(s) is to provide a UI where business users can manage their sites (see Figure 4.8). This includes the site-specific permissions, the look and feel of the site, and miscellaneous site settings. We recommend that business users who will be administering a site get adequate training on the site settings pages.

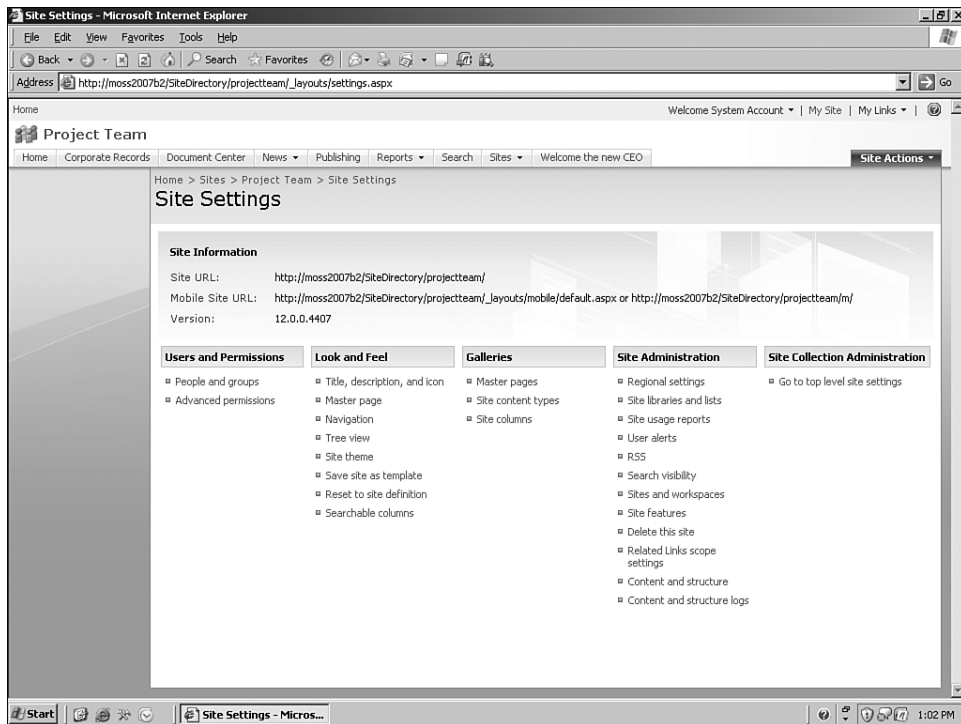


FIGURE 4.8 The administration page on a site lets a user (typically the site owner) configure site-specific items, such as site-level permissions, the lists and libraries stored within the site, and the look and feel of the site.

As you have seen, the various SharePoint configuration and administration settings require multiple administrators. You should carefully plan and designate which users should administer which pieces of the SharePoint administration puzzle.

In addition to these administration options, another major consideration for your SharePoint deployment is physical deployment. Specifically, this means, “How many servers do I need to deploy?” The next section helps you think this question through by helping you understand your options.

Physical Deployment Options

When considering deployment options for SharePoint, you are really considering your *topology*. In other words, you are determining how many servers you will deploy in your SharePoint farm and what roles they will play. In MOSS 2007, how you deploy SharePoint is very flexible.

There are no longer specific topologies that have to be adhered to as there were in SPS 2003 (small farm, medium farm, large farm), and there is no longer the notion of a “job server.” In MOSS 2007, servers have one of three roles:

- **Web Front End (WFE).** The SharePoint bits with just Web rendering enabled
- **Application Server.** Might include indexing, search, Excel calculations, project server, and other features
- **Database Server.** No SharePoint-specific software is installed (only SQL Server)

In light of this, we have an unlimited number of physical configurations to use when rolling out SharePoint. In this section, we’ll present several common deployment scenarios. Your environment will have special requirements around server roles, authentication, DMZ, and application services.

Here are some of the most common configurations you might want to consider—they are by no means the only choices.

Single-Server Deployment

A single server hosts all three roles (WFE, application server, and database) on a single machine. This is good for very small deployments, given that it's fast and easy. The major downsides include scalability issues (because there is no room to grow except for expanding things such as memory and processor) and availability issues (if the server goes down, SharePoint is down). From a logical perspective, all SharePoint objects are located on this server (content sites, SSPs, Central Administration, and databases).

NOTE Choose your deployment topology carefully. There is no direct upgrade from a stand-alone installation to a farm installation.

Two-Server Deployment

In a two-server scenario, one of the servers hosts the WFE and the application server, while the second server hosts the SQL Server database. This provides a way to manage the database separately but adds complexity without adding scalability or availability. This step adds a second tier to the deployment. In most organizations, this is the smallest deployment that is recommended for anything other than a demonstration environment or very small group.

Three-Server Deployment

By adding a server, which acts as an additional WFE/application server to the two-server deployment, we gain scalability (by being able to service more requests) and availability (by load-balancing requests so that if one server goes down, the system stays up and running on the other machine). The single point of failure is now the SQL machine.

Four-Server Deployment

By adding a machine to the database role and upgrading the SQL Server machine to a full cluster, we can achieve availability on both tiers of the environment. This is the smallest highly available environment, meaning that there are no single points of failure. Note that clustering is not a

simple upgrade but rather a re-install where you must move databases. (See <http://msdn2.microsoft.com/en-us/library/ms191295.aspx>.)

Five-Server Deployment

The next step you should consider is to start breaking out the application services for additional performance. For example, the indexing process is very CPU-intensive and should often be put onto its own server. In a five-server deployment, the fifth server would host just an application server, primarily serving as an indexing machine. This creates a three-tier environment, with a new application server tier being added.

N-Server Deployment

The beauty of the scale-out process is that we can continue adding servers at each of the tiers, depending on the needs of the business. Do we need to serve more Web pages per second? Add more WFEs? Do we need to dedicate processor time to calculating Excel sheets? Add some more application servers dedicated to Excel services? You get the idea. Let's say you decide that ten production machines make sense. You might then want to deploy a separate Internet farm in the DMZ and an intranet farm behind the firewall. In addition, you might want staging and testing machines so you can adequately test new features and Web parts. This could bring your server count to 20 servers or more. The next section provides three specific examples.

Deployment Examples

The most common examples of specific SharePoint usage models are the departmental usage model, the corporate intranet, and an Internet-facing deployment.

Departmental

A departmental solution is typically used for collaboration but might include a team portal and often consists of a single SharePoint server and a database (see Figure 4.9). However, savvy departments should deploy a second SharePoint server for availability.

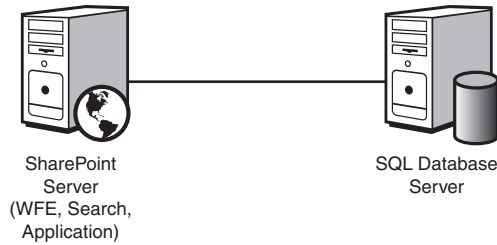


FIGURE 4.9 A departmental solution is often deployed as a single SharePoint server.

Corporate Intranet

A corporate intranet—serving anywhere from hundreds up to tens of thousands of employees—incorporates dedicated application servers (see Figure 4.10). All servers are deployed within the company firewall.

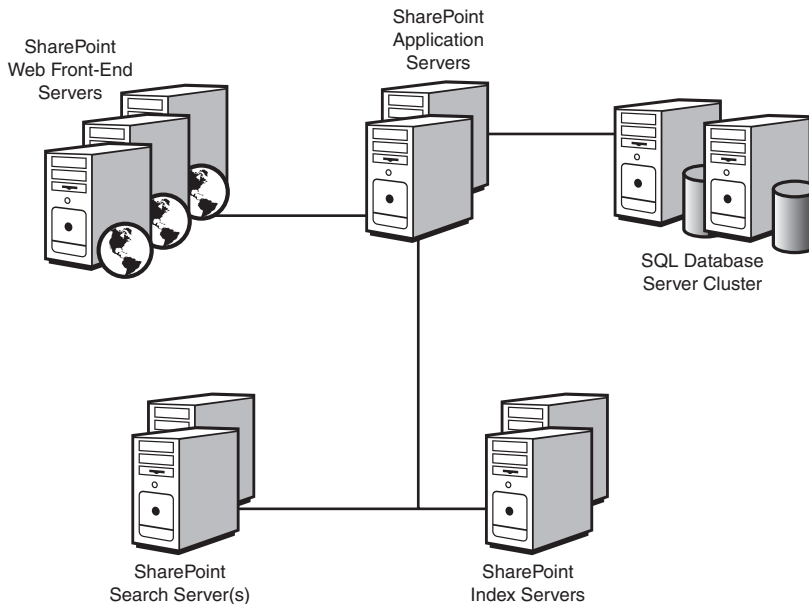


FIGURE 4.10 Corporate intranet farms typically consist of multiple WFEs, dedicated application servers, and dedicated search and index servers.

Internet (Web Content Management)

A corporate Internet presence gets a bit more complicated, given that you'll not only want to have enough Web servers to serve a large number of external users, but you'll also want an internal cluster for authoring purposes. SharePoint then deploys all content changes from the authoring cluster to your production cluster in a one-way manner. Figure 4.11 shows an example of a SharePoint deployment in a publishing environment.

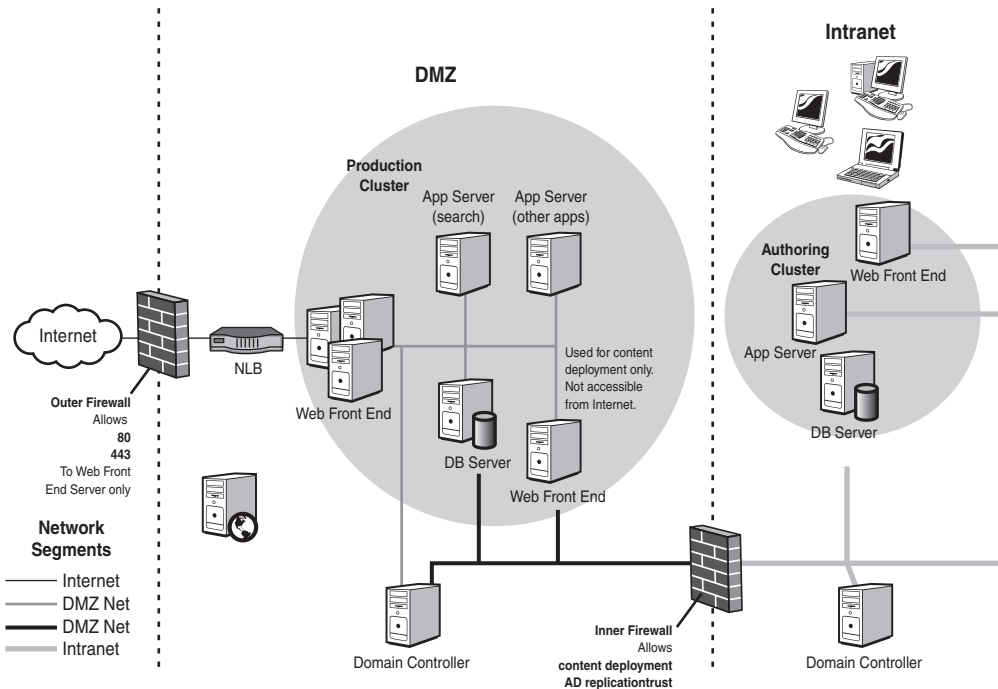


FIGURE 4.11 For your Internet presence, you'll want to include servers outside your corporate firewall (for Internet user access) as well as servers inside the firewall (for employee access).

When considering the question of how many servers you will need, the short answer for deployment is this: Carefully consider your usage plan (collaboration, portal, Web content management, and so on), uptime needs, number of users, application processing demands, geographic dispersion, and budget when determining your deployment architecture.

For a capacity planning tool, check out <http://technet2.microsoft.com/Office/en-us/library/eb2493e8-e498-462a-ab5d-1b779529dc471033.mspx>.

For a step-by-step guide to installing MOSS 2007 in a farm configuration, go to <http://technet2.microsoft.com/Office/en-us/library/601874ea-86c9-4611-bdaf-abe17bbb68161033.mspx>.

Key Points

- In WSS v3 and MOSS 2007, everything is a site. There is no longer the concept of an area.
- A site collection is a hierarchical set of sites that can be managed together. All sites in a site collection are stored together in the same SQL database. Some features do not work across site collections, so it's important to plan how you use them.
- WSS 3.0:
 - Site collections
 - Sites
 - Templates (collaboration only)
- MOSS 2007:
 - Site Collections
 - Sites
 - Templates (publishing, Internet, collaboration)
 - Shared services (provide profiles, search, BDC, and so on)
- “Web application” is the new terminology for an IIS Web site/IIS virtual server
- A Shared Services Provider (SSP) is a site that provides MOSS shared services to one or more Web applications:
 - Stand-alone WSS installation: no shared services
 - MOSS 2007: shared services
 - A “portal” is now a WSS site + a MOSS template + shared services
- There are no longer specific topologies that have to be adhered to (small farm, medium farm, large farm) and no longer a “job server.” In MOSS 2007, servers have one of three roles:
 - Web Front End (WFE)
 - Application server
 - Database server

