

The Basics of the Windows Server 2008 Distributed File System (DFS)

Many File Servers and File Shares

It's a common scenario.

Over time, a company ends up with a large number of file servers that were created at different times and by different teams. The users are left to figure out where their files are stored, mapping many drives to file shares and learning complex UNC paths like \\server23\ffiles and \\finance.asia.company.com\spreadsheets and even something like \\10.1.23.45\scratch. To complicate things further, when an old file server is replaced, paths are broken and mapped drives stop working. It can be challenging. It turns out it doesn't have to be that way. Windows Server already provides a way to manage a large set of file shares as a consolidated list, under common namespaces.

Adding the DFS Services

DFS Services are available on all currently supported versions of Windows Server, but there are significant improvements in the Windows Server 2008 editions. Domain-based DFS namespaces require the use of Active Directory.

DFS Namespaces

A DFS namespace is really a place where you store all the links to all your file shares. From an administrator point of view, think of it as a folder where you keep a list of file shares. Users will see it as a single share with folders and they will have no idea they are actually navigating across a set of servers to get at files and subfolders.

When configuring DFS, you have a choice of using a domain-based or a stand-alone namespace. If you already have Active Directory deployed, you should consider using a domain-based namespace. If you're not using Active Directory, your only choice is a stand-alone namespace. The main advantage of a domain-based namespaces is the configuration will be stored in Active Directory. The path for users will refer to the name of the domain (not a single server) and it won't need to change over time.

With a stand-alone DFS, the server name becomes part of the path to the namespace and is subject to change if the namespace server is changed.

Creating a Namespace

After you install the service, your next step is to create a namespace.

Start "DFS Management" from the "Administrative Tools". On the tree on the left side, click on "Namespaces".

Click on "New Namespace..." Action on the panel on the right to start the "New Namespace Wizard".

Enter the name of the namespace server and click “Next”.

Select the type of Namespace and click “Next”. In my case, I’m using a domain-based namespace with 2008 mode enabled. On the review screen, click “Create” and then click “Close” when it’s done. You will end up with a brand new namespace with no folders:

Adding Folders to the Namespace

After you create the namespace, you will add folders to it, specifying the associated folder target. This means pointing to the actual file shares, making each one appear to users as a folder under the namespace. Before you do that, you want to think long and hard about the folder structure you’re creating. A basic goal of DFS is to create a stable infrastructure that will not constantly change on your users.

To start, click on the “New Folder...” action on the panel on the right.

Multiple Targets

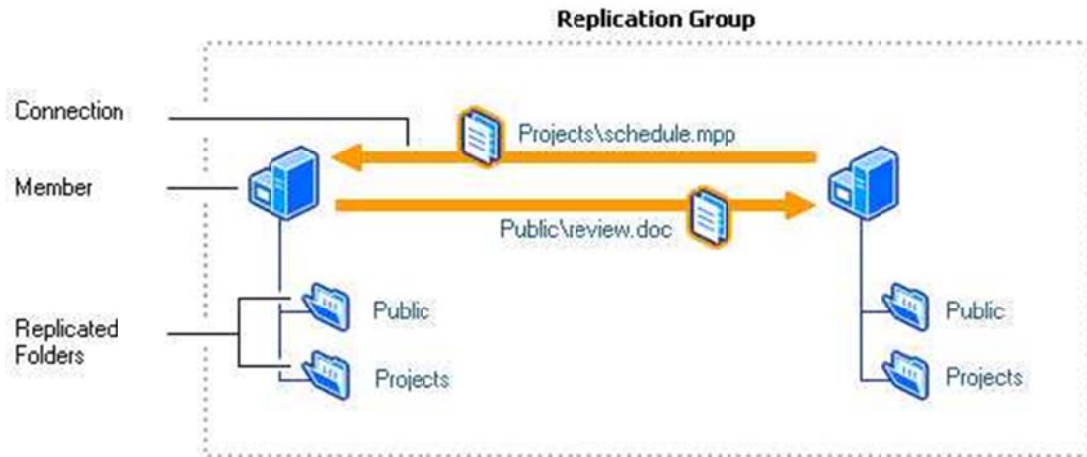
It’s useful to have multiple copies of the same data stored in different file servers. One reason for that is fault-tolerance (if one server is unavailable, you can still access the other one). The other reason is to choose the copy of the data that is closer to you. If you’re in a branch office and you want to access a very large file, you would rather get a copy from a server in that branch.

If you’re using domain-based DFS, the clients will be directed to the target that is closest to them. If there is no target nearby, the clients will be pointed to a remote one. This is similar to what happens when clients are looking for a domain controller. DFS uses the site information in Active Directory to determine which server works best.

DFS Replication

At this point, you’re probably thinking: How does the data get copied across multiple servers? Well, Windows Server includes a component to replicate data between file servers in case you need that. It’s called DFS-R (Distributed File System Replication). DFS-R was introduced with Windows Server 2003 R2 (replacing the old NTFRS with many advantages). DFS-R can be used for both domain-based and stand-alone DFS.

To replicate files between two (or more) shares, you need to create a replication group and specify a few things like which servers will replicate (members) and what they will replicate (replicated folders). DFS-R is aware of site topology. It also has options to control scheduling and the use of bandwidth (throttling). DFS-R uses Remote Differential Compression (RDC), meaning that only changes in the files are sent over the network, not the entire file.



Configuring Replication

When you create multiple folder targets for a folder, the DFS Management tool will offer you to automatically setup the replication group:



Assuming you have installed the DFS-R role service on both servers, it will go ahead and setup replication for you. It's done via the "Replicate Folder Wizard".

To start, it will ask you the replication group name and the replicated folder name: It will then verify if the server working as folder targets for that DFS folder can be a replication member. Our two servers are eligible:

Next, you need to select one of the servers as the primary member. This should be the server that contains the files you want to replicate and will be the authoritative server during the initial replication:

You will also need to select a topology. This is important if you have a larger number of servers and need to control how they replicate. For instance, if you have a central server in HQ and several branch servers, you can make sure the branch servers do not try to replicate with each other directly, but always replicate only with the HQ servers. In our case, with only two servers, "Full Mesh" will do: