

Windows 2003 Server - Distributed File System

July 24, 2008

Introduction

The Distributed File System is used to build a hierarchical view of multiple file servers and shares on the network. Instead of having to think of a specific machine name for each set of files, the user will only have to remember one name; which will be the 'key' to a list of shares found on multiple servers on the network. Think of it as the home of all file shares with links that point to one or more servers that actually host those shares. DFS has the capability of routing a client to the closest available file server by using Active Directory site metrics. It can also be installed on a cluster for even better performance and reliability. Medium to large sized organizations are most likely to benefit from the use of DFS - for smaller companies it is simply not worth setting up since an ordinary file server would be just fine.

Understanding the DFS Terminology

It is important to understand the new concepts that are part of DFS. Below is a definition of each of them.

Dfs root: You can think of this as a share that is visible on the network, and in this share you can have additional files and folders.

Dfs link: A link is another share somewhere on the network that goes under the root. When a user opens this link they will be redirected to a shared folder.

Dfs target (or replica): This can be referred to as either a root or a link. If you have two identical shares, normally stored on different servers, you can group them together as Dfs Targets under the same link.

The image below shows the actual folder structure of what the user sees when using DFS and load balancing.



Figure 1: The actual folder structure of DFS and load balancing

Windows 2003 offers a revamped version of the Distributed File System found in Windows 2000, which has been improved to better performance and add additional fault tolerance, load

balancing and reduced use of network bandwidth. It also comes with a powerful set of command-line scripting tools which can be used to make administrative backup and restoration tasks of the DFS namespaces easier. The client windows operating system consists of a DFS client which provides additional features as well as caching.

Setting Up and Configuring DFS

The Distributed File System console is installed by default with Windows 2003 and can be found in the administrative tools folder. To open, press Start > Programs > Administrative Tools > Distributed File System or in the Control Panel, open the Administrative Tools folder and click on the Distributed File System icon. This will open the management console where all the configuration takes place.

The first thing you need to do is create a root. To do this, right click the node and select New Root. Press next on the first window to be brought to the screen where you will have to make the choice of creating either a stand alone or domain root. A domain root will publish itself in Active Directory and supports replication, whereas a stand alone root does not. If you have an AD Domain Controller set up on your machine, I recommend choosing the domain root.

Note: The root would be the top level of the hierarchy. It is the main Active Directory container that holds Dfs links to shared folders in a domain. Windows 2003 allows your server to have more than one root - which wasn't the case in Windows 2000.

The next screen is the one where you have to select which trusted domains will be hosted. Since I only have one domain in my network, only domain.com is visible.

Once this is done you have to select a server on that domain - in my example it is netserv. The FQDN (Fully Qualified Domain Name) of this host server is netserv.domain.com.



Figure 2: inputting the host server name

The following screen allows you to specify the root name of your primary DFS root. You should give it something which will accurately define the contents of that share. In my example I have called this root "Company" - which would be a real name of an organization. You can change this to anything you want. You might wish to have a root called "Documents" - which would clearly state that one can expect to find anything related or specific to documents, and documentation in that root.

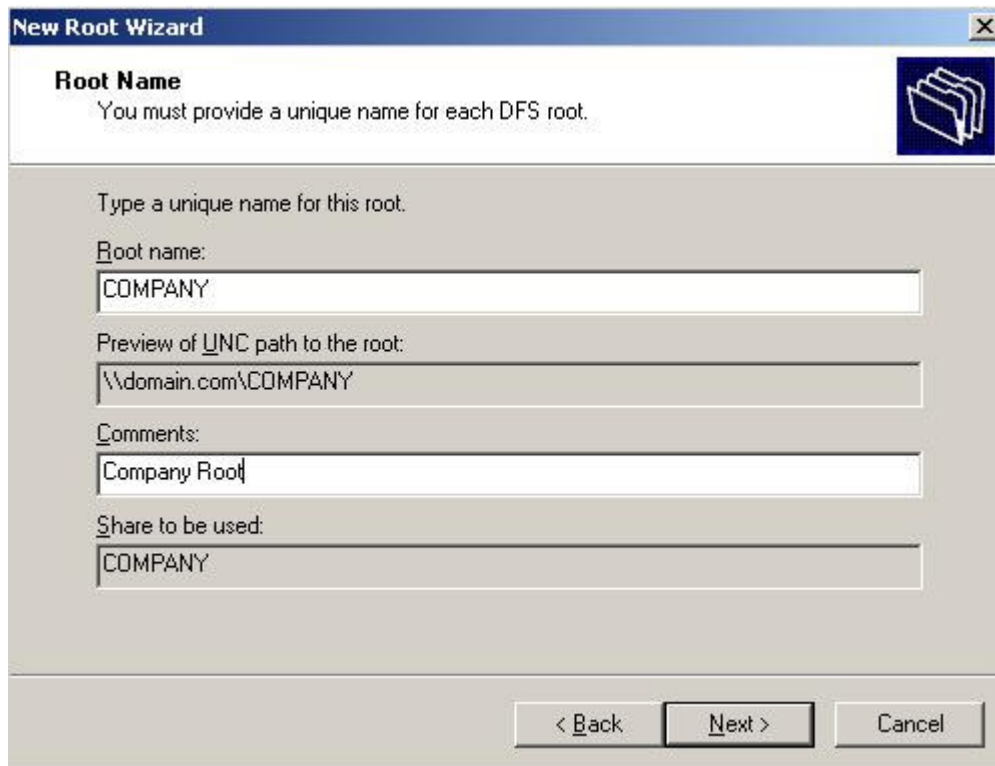


Figure 3: entering the dfs root name

You will now have to select the location of a folder in which all the files will be stored.



Figure 4: selecting the root share

Tip: for added security, when selecting a folder, try to choose one that is located on a partition other than that of the operating system.

Your DFS root is now configured and visible in the configuration console. Right click the root target and press Status to check if it is online or not.

A green check mark verifies that everything is working properly and that the node is online, whereas a red X means that there is a problem.

To add a new link, right click the root for which you want the link to be created, and select New Link. In the "New Link" screen, enter a name and path for the link and click OK. Repeat this for as many links as you need to create.

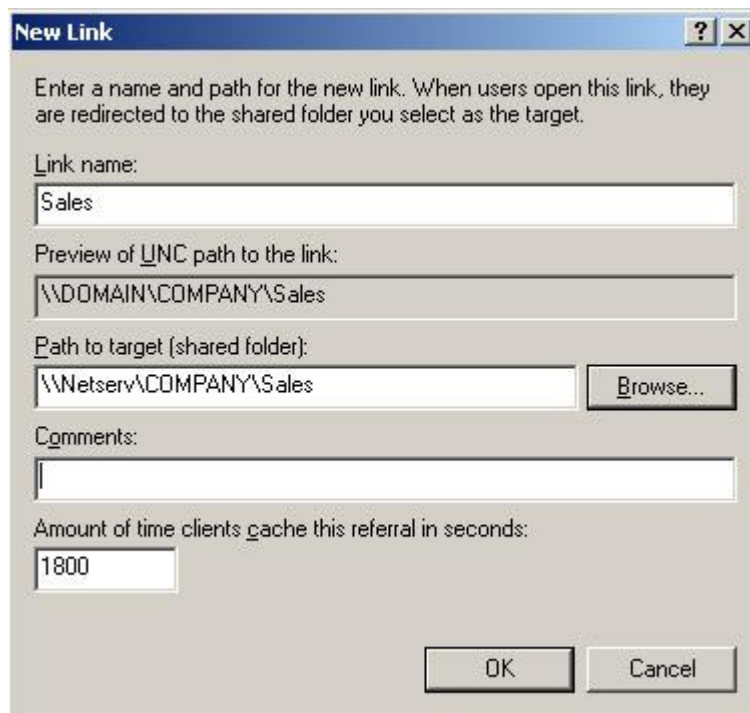


Figure 5: creating a new link

Links are visible right under the node. Below is a screenshot displaying the three links I have created for the COMPANY root.

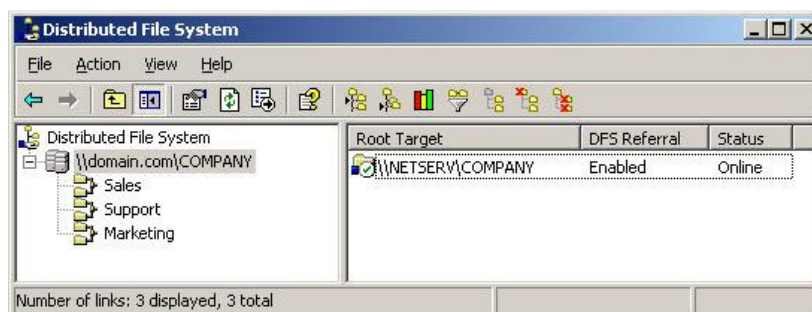


Figure 6: dfs root and three links in the DFS mmc console

Publishing the root in Active Directory

By publishing dfs roots in AD as volume objects, network users will be able to search for shares more easily and administration can be delegated.

To do this right click the desired dfs root, select Properties and go to the Publish tab. Enter the appropriate details in each box and press OK.

In the keywords section you can specify certain words that will help locate the dfs root when it is being searched for.

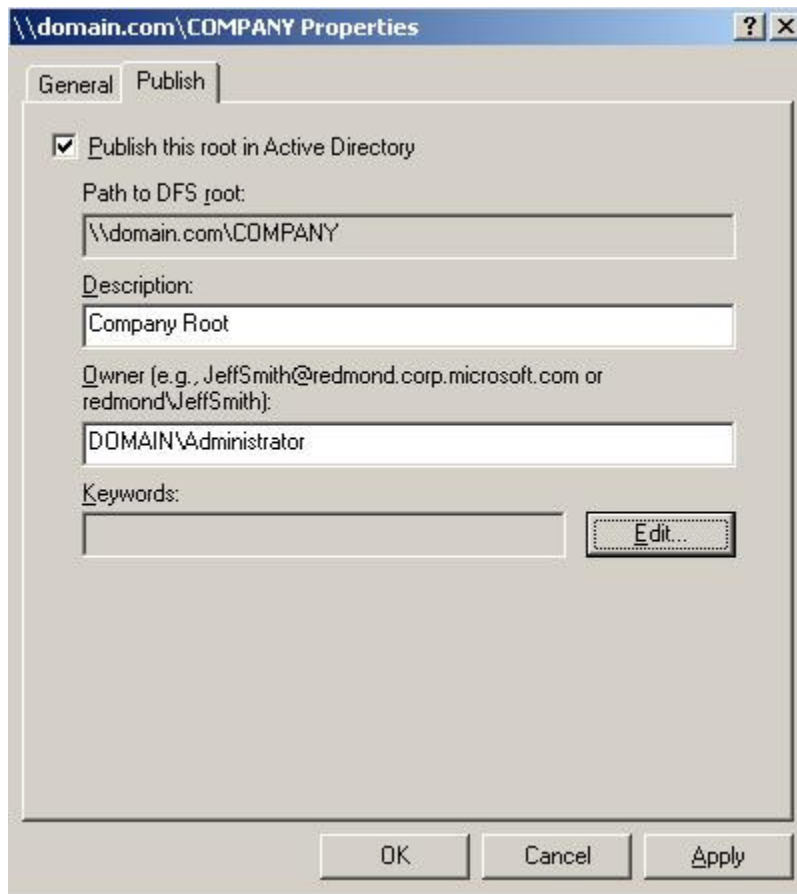


Figure 7: publish tab in the dfs properties window

The dfs root will now be published in Active Directory.

File Replication Services

There are two types of replication:

- * Automatic - which is only available for Domain DFS
- * Manual - which is available for standalone DFS and requires all files to be replicated manually.

The four ways in which replication can be achieved between two or more servers are:

- Ring
- Hub and Spoke
- Mesh
- Custom

The first three refer to network topologies and the last allows you to specify an advanced method of replication, which can be tuned to your needs.

The advantages and disadvantages of replication are as follows:

Advantages - client caching, integration with IIS, easy to administer and setup.

Disadvantages - limited configuration options, there is no method of programmatically initiating a replication session.