



realtimepublishers.comtm

The Definitive Guidetm To

Windows 2003 Storage Resource Management

VERITAS[™]

Evan Morris

Chapter 3: Analyzing and Planning Storage.....	60
Phase 2: Planning SRM	60
Outcome of Storage Analysis and Planning	61
Trend Analysis and Capacity Planning.....	62
Storage Management Decision Points	65
SRM Product and Process.....	67
SRM Product Evaluation	67
StorageCentral Installation.....	67
Auditing with StorageCentral	68
Tips for Selecting Objects.....	73
Working with Report Data.....	76
Customizing Reports.....	80
Preparing for the Next Phase	81
Eliminating Duplicate Files	82
Eliminating Unused Files.....	83
Eliminating Wasted Space	84
Reducing Consumption.....	85
Enforcement Policies	85
Support of Storage Applications.....	87
Summary.....	89

Copyright Statement

© 2004 Realtimepublishers.com, Inc. All rights reserved. This site contains materials that have been created, developed, or commissioned by, and published with the permission of, Realtimepublishers.com, Inc. (the "Materials") and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of Realtimepublishers.com, Inc or its web site sponsors. In no event shall Realtimepublishers.com, Inc. or its web site sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, indirect, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

The Materials (including but not limited to the text, images, audio, and/or video) may not be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, in whole or in part, except that one copy may be downloaded for your personal, non-commercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Realtimepublishers.com and the Realtimepublishers logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners.

If you have any questions about these terms, or if you would like information about licensing materials from Realtimepublishers.com, please contact us via e-mail at info@realtimepublishers.com.

Chapter 3: Analyzing and Planning Storage

In the previous chapter, we explored the built-in features, resource kit, and support tools of WS2K3. I defined the requirements for the storage analysis phase, then we took a look at how much of this work could be done using the available tools. Although there are many built-in tools at our disposal, there are still some tasks—such as detailed analysis about who is using storage—that will benefit from the features offered by third-party SRM products. We'll explore such tools in this chapter.

In addition, we'll begin to explore the next phase—planning an SRM deployment. We'll use the information that we captured about both storage capacity and storage performance to determine a course of action (or *courses* of action, if you're used to contingency planning). I'll present you with a flowchart that will help to lay out the storage-management decisions that you must make; most notably, whether you decide to live within your existing capacity or to add capacity. During the course of this chapter, we'll look at using SRM tools for performing trend analysis and capacity planning.

Finally, we'll look at how SRM solutions can help eliminate duplicate files, unused files, and wasted space as well as reduce consumption. I'll show you how to perform an analysis of your current environment—highlighting SRM techniques and reporting—to improve your storage usage efficiency. I'll also explore the tools you can use to perform a detailed storage analysis, gathering information about who is using storage and how much.

Phase 2: Planning SRM

Table 3.1 shows the planning phase in the overall SRM-deployment methodology.

Phase	Process	SRM
Planning	What are the problems and the priorities to solve them? Can we find a better way?	Lay out the organizational policies in these areas: —Security Policies —Supported Storage Systems —Support of Storage Applications —Storage Quotas or Allocations —Restricted File Types —Enforcement Policies —Support and Escalation Plan —Backup and Disaster Recovery Procedures

Table 3.1: Planning phase of the SRM-deployment methodology.

In the previous chapter, we used several tools to take a snapshot view of storage usage, including looking at several angles of usage either by user, folder, or even the type of application being used. But a snapshot view is merely a point-in-time capture of information. To be of value for storage-resource planning, the snapshot views must include more than one point in time. We need to capture the information discussed in the previous chapter several times to evaluate growth patterns and plan for anticipated growth. In the next section of this chapter, we'll look at growth analysis, then we'll develop a flowchart for planning a course of action.

Outcome of Storage Analysis and Planning

The expected outcome of storage analysis and planning is twofold:

- Create a plan of action for storage design
- Develop an organizational storage policy

Although the first item is certainly important, it must be married with an organizational storage policy. As we develop technological solutions for managing storage, the solutions will be used to support the organizational storage policy that we'll develop in the next phase of the project. In fact, the two must complement each other, as a conflict between storage design and organizational storage policy is disruptive to business process. For example, if user space is limited by policy on a business-critical storage system, but the tools aren't in place to enforce the policy, then additional administrative burden is created to monitor and maintain the system.

The end result of this twofold strategy is to increase the availability of storage systems while reducing the cost of administrative maintenance. These goals are difficult to attain; as we increase the importance and usage of storage systems, we must also increase their fault tolerance and ability to recover from any disruption to service.

In the area of TCO, entropy is the enemy—if a system is left alone over time, it will begin to decay. For storage management, this decay means that if we don't put the tools in place to monitor and maintain the system in an efficient manner, the state of storage will move towards a natural decay in order and chaos. Perhaps you have seen the results of this corrosion for a particular file server: an ever-increasing number of files being stored that have diminishing access patterns but use ever more storage, which increases the backup and recovery window. In the next section, we'll look at the impact of increasing storage capacity.

One more point about the backup and recovery window, as the concept of Service Level Agreements (SLAs) can have a powerful impact in storage-capacity planning. When users are storing more and more information, throwing storage space at the problem can actually work—for a short time. But before long, business-critical or even mission-critical data can be compromised because it is mixed with files that are no longer essential, but all the data must be backed up or recovered during the same time period. This scenario can put you in a situation from which you cannot possibly recover in time to meet your SLA. Even if you don't have a formal SLA, you may have an informal Service Level Objective (SLO) that states the expectation that any downtime is unacceptable and you should do everything within your power and budget to reduce or eliminate it!

This situation is the driving force behind HSM—moving nonessential files away from business-critical and mission-critical information. The SRM plan can include how HSM will be incorporated, and the benefit to the company (which drives the cost justification) is the improved response and protection of business-critical and mission-critical information. Recently, the storage industry has begun calling this overall process information lifecycle management (ILM).

The next sections will illustrate how to use an SRM product to address each of these problems. In the previous chapter, I showed you how to use tools such as the Windows Server resource kit tools and Performance Monitor to perform a storage analysis; however, these tools are quickly left behind when compared with the functionality of SRM products.

Trend Analysis and Capacity Planning

In the previous chapter, I showed you how to use Windows Server's Performance Monitor for gathering disk-usage information. Unfortunately, Performance Monitor doesn't make the grade for performing trend analysis and capacity planning. You can use it to capture information at periodic intervals, showing increasing disk usage over a 30-day interval, but its real weakness is the lack of control over the level of detail. Performance Monitor is best suited for reporting on the level of logical disk drives. The information we'll want is disk usage on the file share or directory level. For example, Figure 3.1 shows a third-party tool's chart of Disk Space Used for a 30-day trend against a specific file share. This trend is fairly typical, showing disk cleanup followed by ever-increasing storage usage, with a recent spike. Also, the ability to tie this information back to the responsible user is paramount, and we'll run reports to do so.

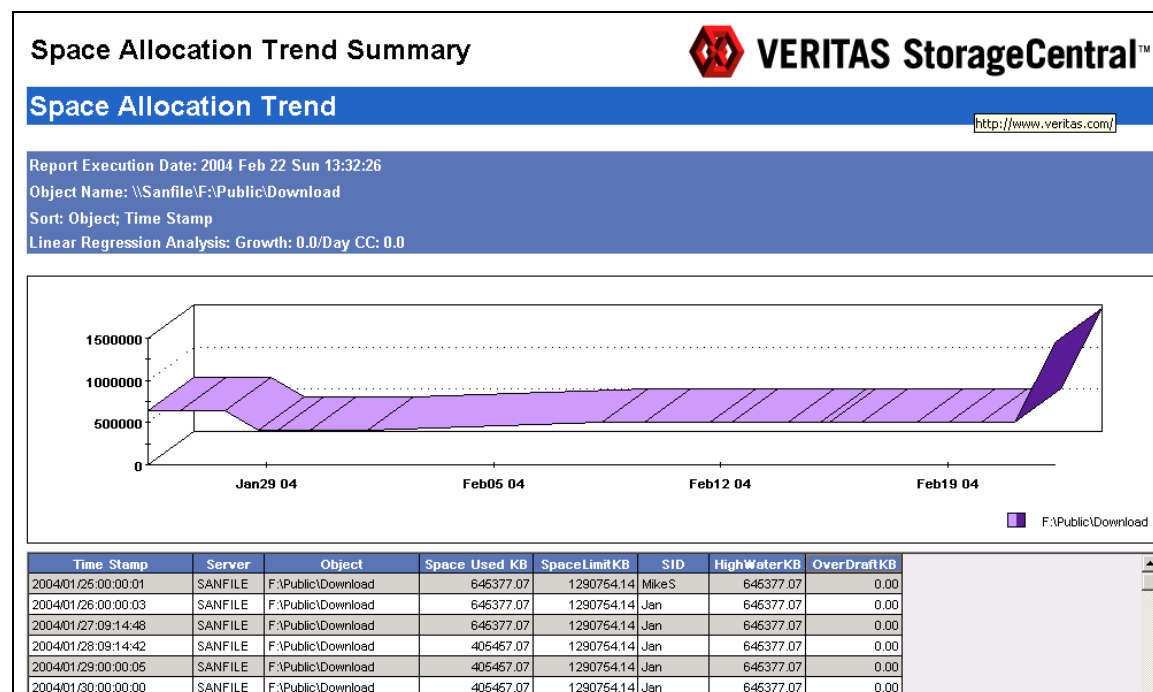


Figure 3.1: VERITAS' StorageCentral SRM Space Allocation Trend Summary.

To access the trend reports, you simply configure this tool to collect the data in advance. Right-click either the server or AD root (depending on whether you are using the Standard or Active Directory edition, respectively), and click Properties. On the Trending tab, select trending and any parameters you want to change, such as the time interval to collect data.

Figures 3.2 through 3.4 illustrate the serious consequences of the storage growth battle that you might be currently fighting. I've included this series of figures to illustrate—to you or perhaps even to convince your management—that the issues are more serious than at first glance. The series shows IT spending mapped against storage spending ranging from 20 percent to 60 percent per year. I've deliberately left the storage growth lower than some market predictions, which show storage demand increasing by more than 100 percent per year (doubling) for many companies. Granted, that number might be hard to sustain with recent economic events, but over a 10-year period, the bursts in growth may make a 20 percent to 60 percent rate easy to achieve.

Figure 3.2 shows IT resources increasing annually at 5 percent mapped against storage demand increasing annually at 20 percent. The column labeled Effective Ratio illustrates the impact on IT resources, given constrained resources to manage the growing storage. If storage demand is growing at a rate of only 20 percent, then IT resources relative to the amount of storage that must be managed reaches a ratio of 3.8 after a 10-year cycle. This ratio can be interpreted as meaning that for every dollar of IT spending, the company must be capable over time of increasing its efficiency to managing nearly four times as much storage resource allocation.

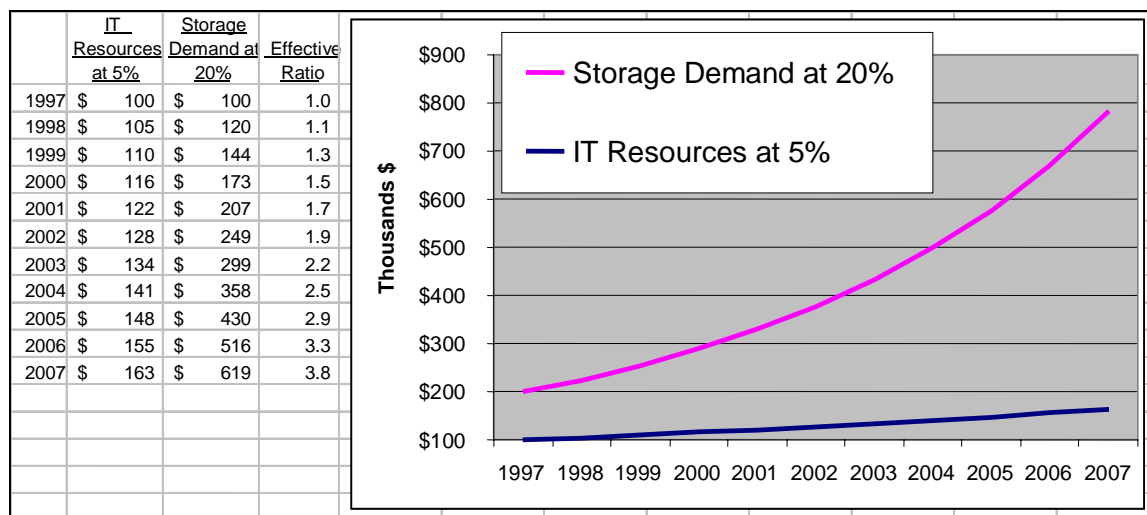


Figure 3.2: IT resources mapped against storage demand increasing at 20 percent.

Figure 3.3, similar to Figure 3.2, shows IT resources increasing annually at 5 percent mapped against storage demand; except in this case, storage demand is increasing annually at 40 percent. The column labeled Effective Ratio illustrates the impact on IT resources, given constrained resources to manage the growing storage. If storage demand is growing at a rate of only 40 percent, IT resources relative to the amount of storage that must be managed reaches a ratio of 17.8 after a 10-year cycle. This ratio can be interpreted as meaning that for every dollar of IT spending, the company must be capable over time of increasing its efficiency to managing nearly 18 times as much storage resource allocation.

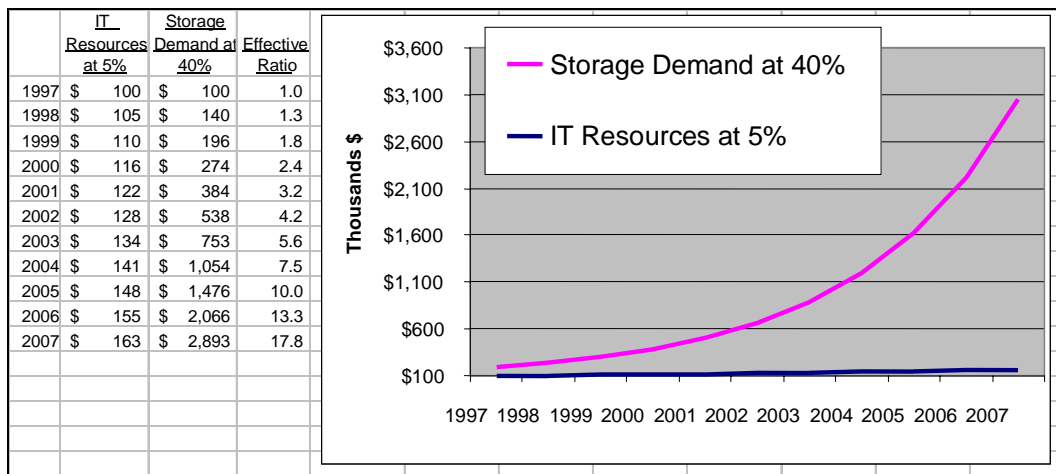


Figure 3.3: IT resources mapped against storage demand increasing at 40 percent.

Figure 3.4 shows IT resources increasing annually at 5 percent mapped against storage demand increasing annually at 60 percent. The column labeled Effective Ratio illustrates the impact on IT resources, given constrained resources to manage the growing storage demand. If storage demand is growing at a rate of 60 percent, IT resources relative to the amount of storage that must be managed reaches a ratio of 67.5 after a 10-year cycle. This ratio can be interpreted as meaning that for every dollar of IT spending, the company must be capable over time of increasing its efficiency to managing nearly 68 times as much storage resource allocation!

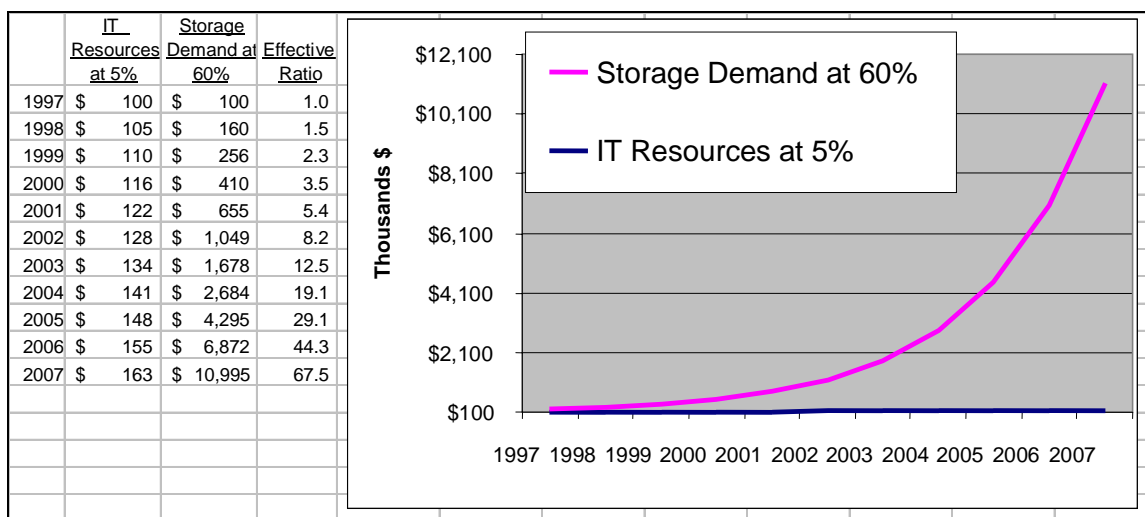


Figure 3.4: IT resources mapped against storage demand increasing at 60 percent.

Here is where the going gets even tougher—technological improvements in storage mean that for the same dollar amount, the amount of storage purchased increases at a constant rate. So you can use the information in the figures in two ways: either as a dollar-per-dollar comparison of administrative spending versus storage spending, and as a comparison of administrative spending versus the storage capacity that must be managed.

Storage Management Decision Points

You've analyzed existing storage and now you must decide the appropriate course of action. Based on either the capacity and performance analysis, are you going to upgrade your existing hardware to achieve more performance and more storage, or are you going to live within the existing capacity? If the latter is your main choice of action, SRM tools will be indispensable in achieving your aims. Even if your solution requires that you add more storage, SRM tools can be your best bet for managing the new storage.

Figure 3.5 illustrates the storage-management decision points in a flowchart and is designed to give you an overview of the SRM process. I'll refer back to this flowchart throughout the chapter, as it covers the primary concepts of SRM planning. The flowchart shows several options, or plans of action, depending on whether you're storage constrained or performance constrained (and you might be constrained different ways for different systems or applications—we'll look at the different storage applications later). First, let's look at the area covered in the flowchart in the SRM Product and Process section.



Some of the options can accomplish both goals: by adding newer, faster, larger systems, you get more storage space with higher performance. But this option assumes that you have the necessary budget. For those of you lucky few, I'll look at performance design and new technology later in this chapter.

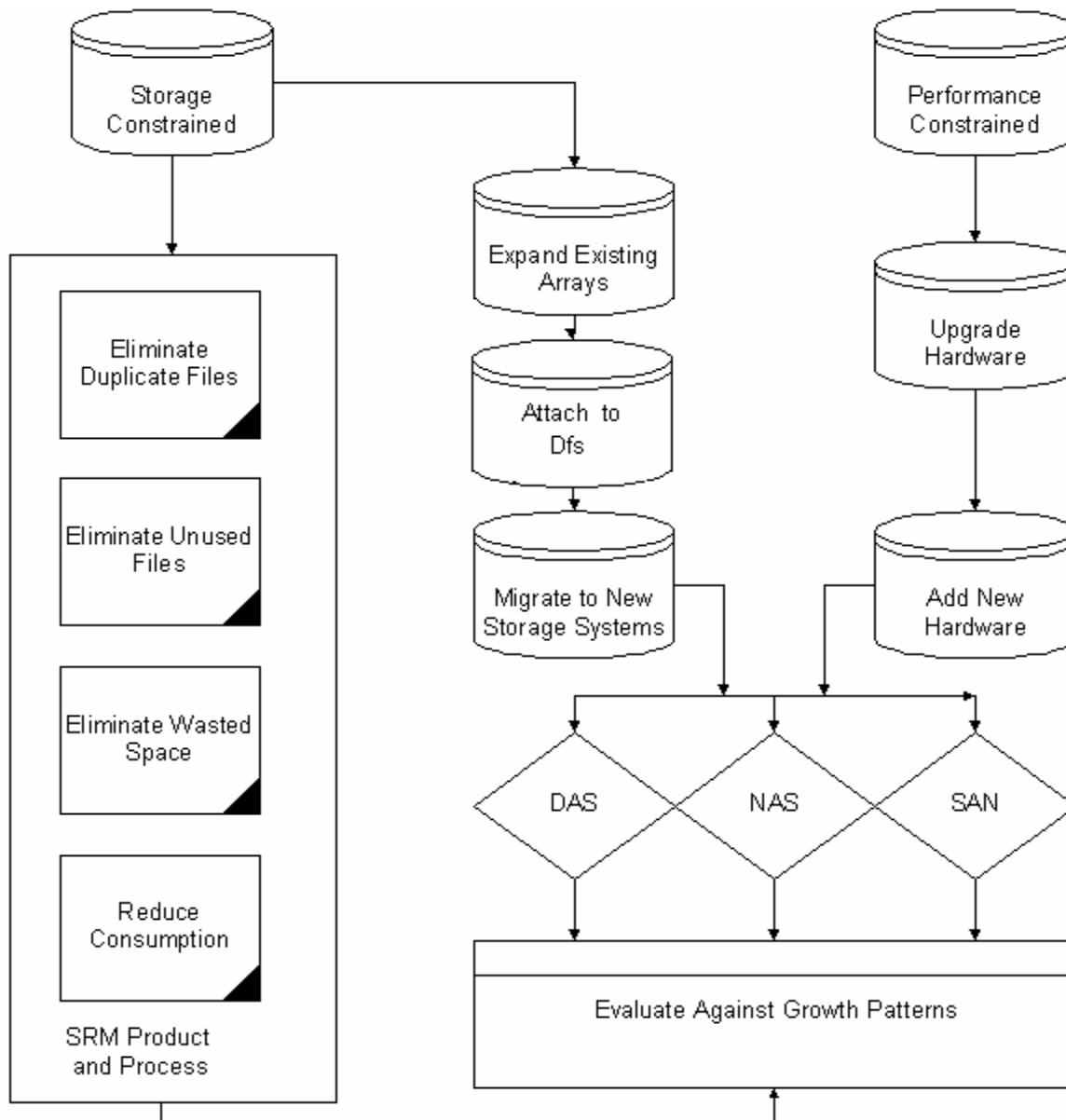


Figure 3.5: Storage-management decision points.

SRM Product and Process

In this section, we'll look at how SRM tools can assist in the capacity-planning process. The first step in the planning process is to identify what you will use the SRM tools to accomplish. The primary functions of SRM solutions that we'll include in the plan are to

- Eliminate duplicate files
- Eliminate unused files
- Eliminate wasted space
- Reduce consumption

Controlling future disk space usage has several benefits that might not be immediately apparent. The most obvious is that the disk space is available for business usage. But controlling disk usage through SRM also helps to improve the availability of servers by shortening backup windows and reducing the amount of data that may need to be recovered. Controlling disk usage through SRM can also reduce corporate liability when file blocking is used to eliminate unwanted file types (such as MP3s that might involve copyright violations).


SRM Product Evaluation

For the rest of this chapter, I'll show you how to use VERITAS StorageCentral SRM as an example of an SRM product, although you can certainly use any competing product. Ideally, you should be able to perform the same functions and generate the same reports. Extensive information that compares SRM products is available from industry trade magazines. If you already have another product, you can still use the SRM deployment methodology detailed here.

 You can access an evaluation copy of the StorageCentral SRM suite from <http://www.veritas.com>. You will need to provide an email address and fill out a brief survey for download registration.

StorageCentral Installation

VERITAS provides ample documentation with StorageCentral SRM. I'll just point out a few items to consider before you install this product. The first consideration is the service account that will be used to run the StorageCentral services. As the installation wizard explains as it guides you through the product installation, if you don't specify a service account, StorageCentral will use the LocalSystem account. As with the Performance Monitor service account, you should change the StorageCentral account from LocalSystem to a domain account. This modification is necessary for the service to be able to connect to other computers. Thus, if you want to be able to report against network drives or use email notification, you'll want to use a domain account for the service startup account.

 If you get an InstallShield error 1607 "unable to install," try running the setup with the source file copied to a local hard drive. Such a simple error can throw you; I attempted to run the install with the downloaded setup file on my network share and received this error.

Another issue that you might encounter, is that StorageCentral depends on the Messenger service for receiving notifications from agents running on other servers (see Figure 3.6). On WS2K3, the Messenger service is disabled by default; during installation, you will receive a pop-up message that notification messages sent to this server by agents will not be received. The default property for the Messenger service was changed in WS2K3 and it is no longer started automatically (as a security precaution). You can change the service startup, but this issue will only affect you if you are running StorageCentral in a multi-server environment in which you will be pooling the information to a central server.

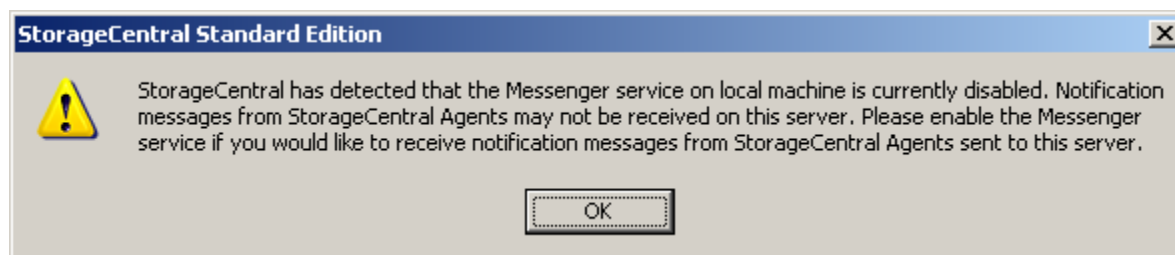


Figure 3.6: The Messenger service is required for receiving notifications.

Figure 3.7 shows another prompt that you might encounter during installation. You will receive this prompt if you do not specify email server properties during setup. As it states, you can configure these settings later, so no need to be concerned about it or let it delay your installation.



Figure 3.7: Prompt during installation if no email server properties entered.

Auditing with StorageCentral

Auditing with StorageCentral is a simple three-step process:

1. Select the report(s)
2. Select the object(s)
3. Run the report(s)

Figure 3.8 shows a report running; in this case the Sample Storage Audit, which is a good place to start. The figure also shows a list of the Report Sets that are available.

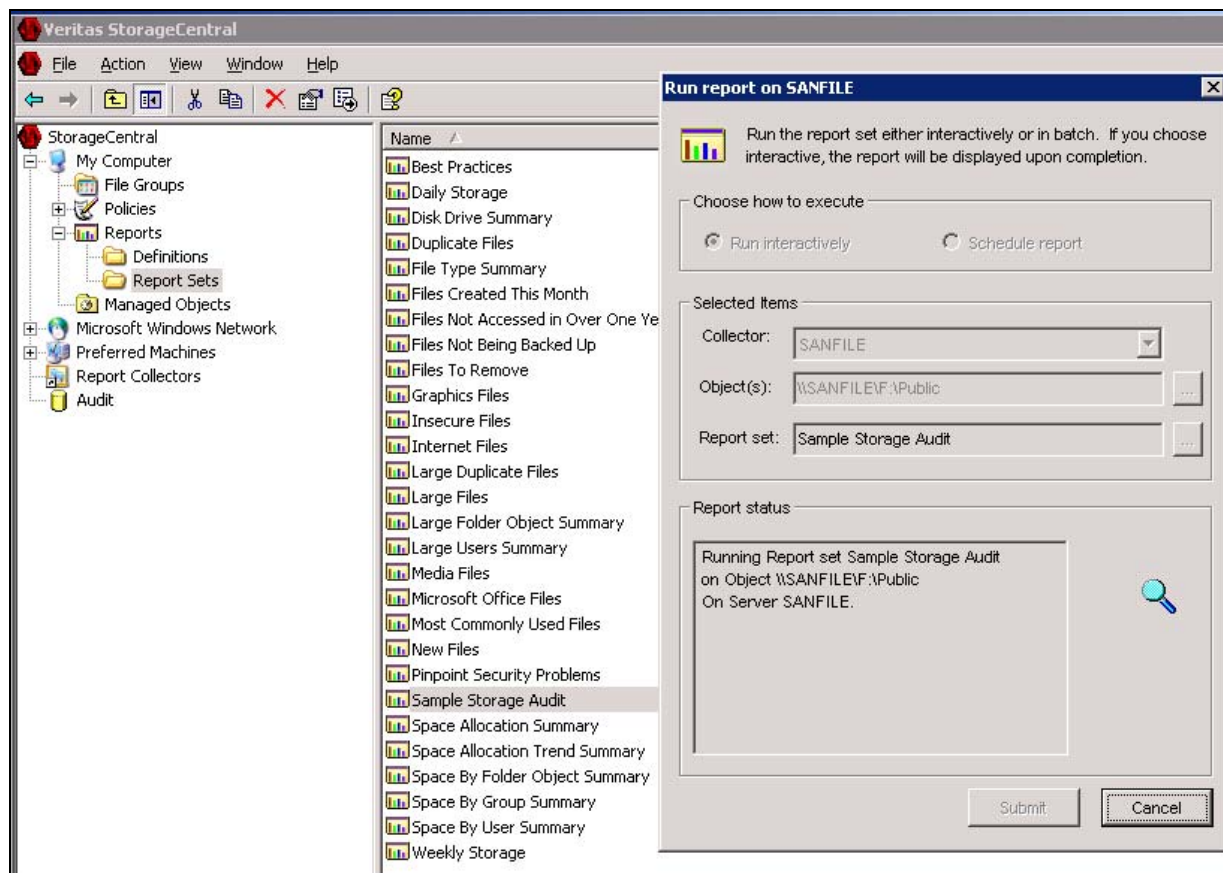


Figure 3.8: Running the Sample Storage Audit.

To run this report, launch the StorageCentral console, expand My Computer if you are running the report on this same machine, expand Reports, and select Report Sets. Alternatively, you can right-click Preferred Machines, and select the option to create a new Preferred Group. You can then select the reports on one of those servers. Next, right-click on the Sample Storage Audit, and select Run Report.

☞ Double-clicking on the report set is the same as right-clicking and selecting Run Report. To get to the report properties, you will need to right-click and select Properties.

At this point, you will see a dialog box similar to that shown in Figure 3.9, except that you will not have any objects selected (as I have already selected `\\sanfile\f:\public`). To select the objects, which are the folders on disk that you want included in the report, click the ellipses button next to the Object(s). At this point, you can also select whether to run the report just once or—even better—set up a schedule to run the report every Monday morning to start your week off with some good information.

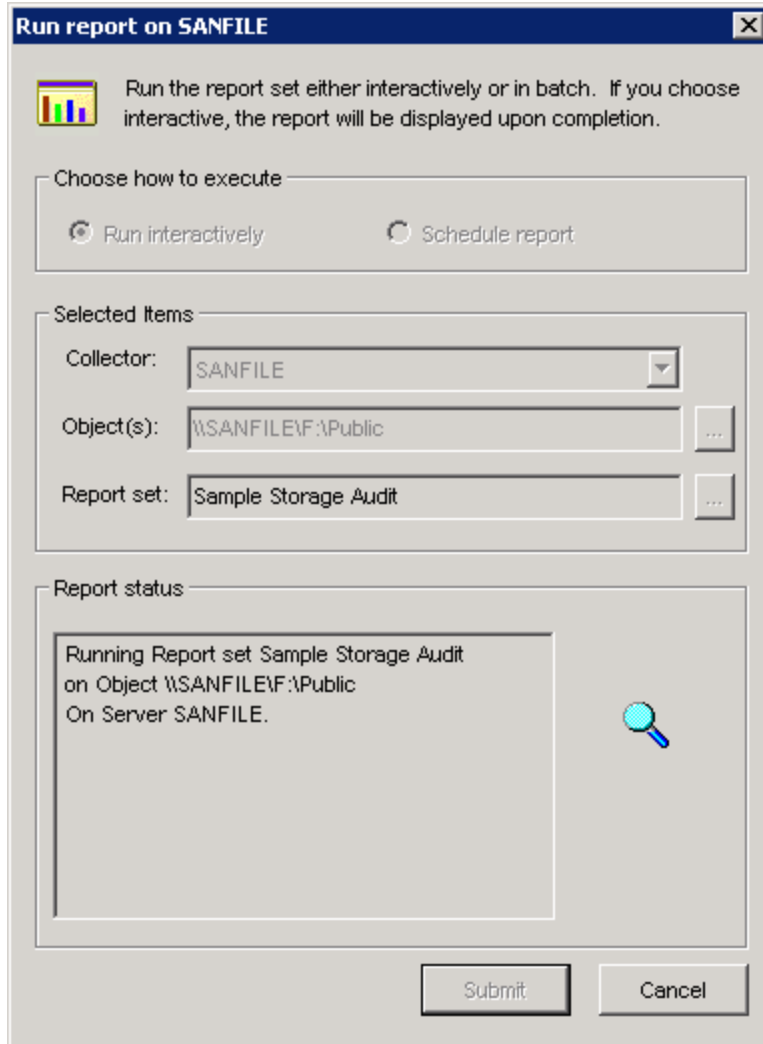


Figure 3.9: Running the report interactively against the selected object(s).

Figure 3.10 shows the dialog box in which you select the folder objects for the report. Subfolders are automatically selected, and you can select multiple disks and folders.

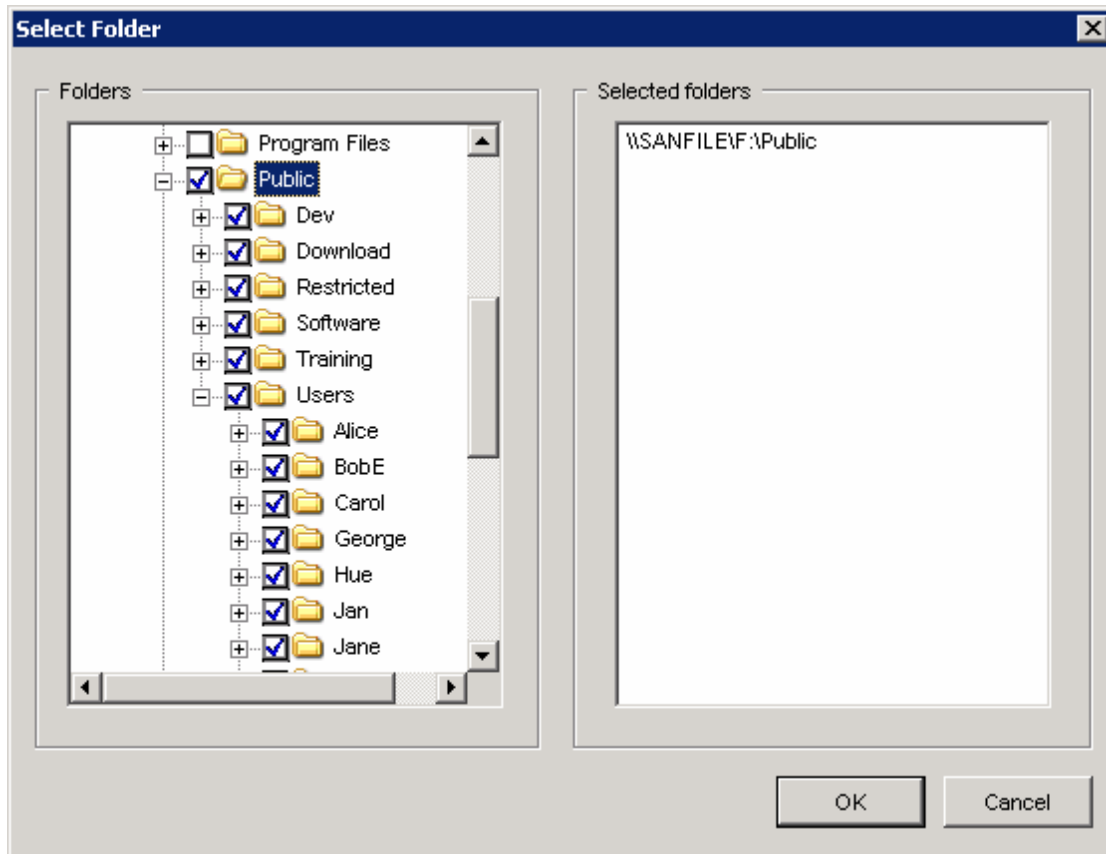


Figure 3.10: Selecting the objects for reporting information.

Figure 3.11 shows the HTML Report generated automatically for the Sample Storage Audit. If you click on Sample Storage Audit, and choose Properties, you will find where you can change the output properties and location. We'll look at how you can manipulate CSV files in particular in a moment. Notice that on the Content tab, you can select the Report Definitions that compose this report. In this case, the Sample Storage Audit is mainly Space Allocated by User. So, at this point, you should get a feel for the relation between definitions and reports.

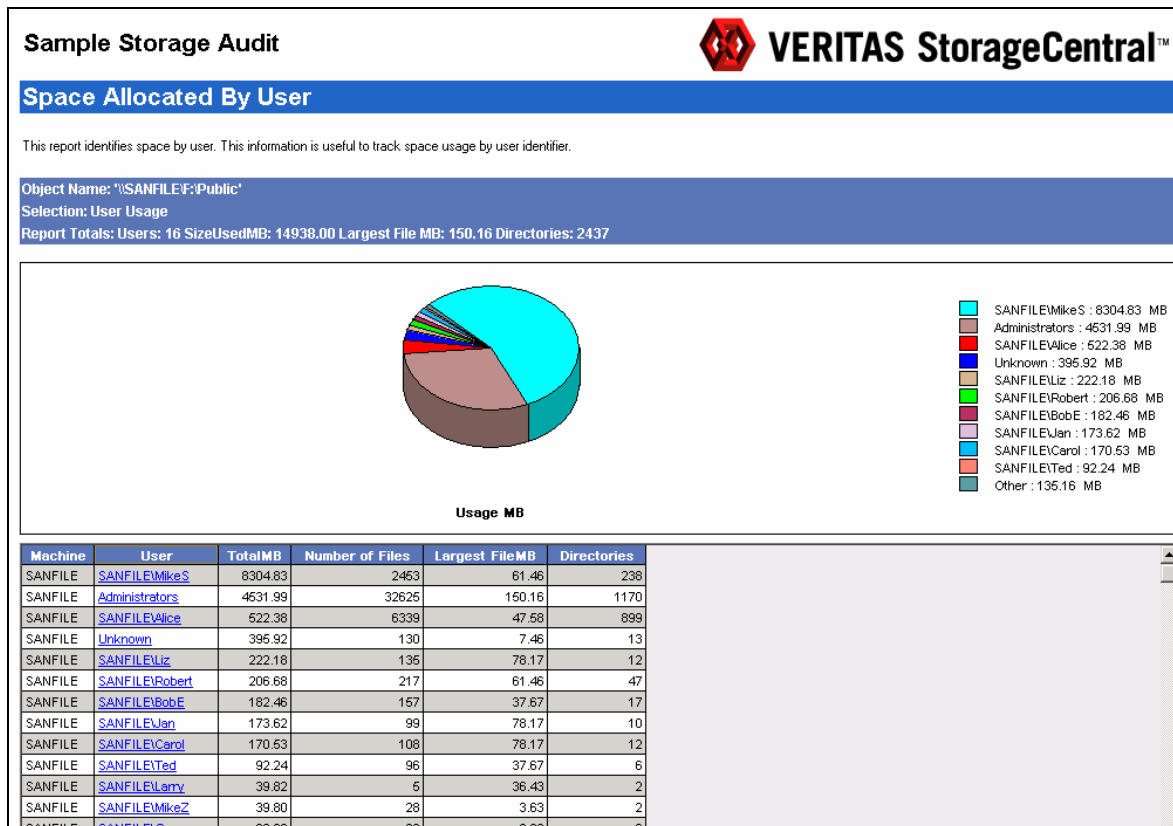



Figure 3.11: The HTML Report generated automatically for the Sample Storage Audit.

From the report THAT Figure 3.11 shows, you get a quick graphical view of the top users of your public share. So now let's take a look at the types of files that are eating up that space and whether they are files you should be managing.

Another useful report to run to get you started is the Duplicate Files report. This report provides information that helps you immediately recoup the cost of StorageCentral. Through this report, you can instantly see the largest files and how many copies there are. Is it time to set up a \\Public\MovieTrailers share just so that you can maintain only one copy of those 50MB movie trailers as opposed to every other user keeping a copy in his or her home share? Whether you take drastic steps depends on your corporate policy, but performing this audit will help you to develop that corporate policy

 We'll explore developing SRM corporate policy in the next chapters.

Tips for Selecting Objects

When selecting objects to run your reports against, you can benefit from a few tricks. For example, you can add a server with network shares to the list of preferred machines, but you can't simply retrieve from a mapping or connection from a server with StorageCentral installed. The remote server also needs StorageCentral installed, otherwise, you'll receive the report that Figure 3.12 shows.

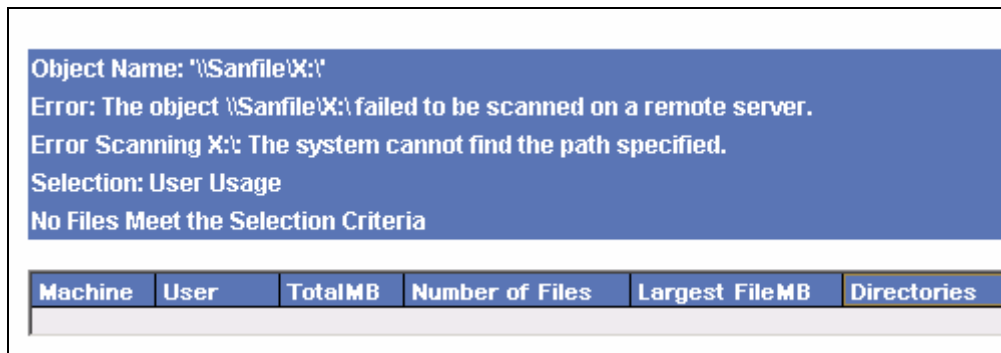


Figure 3.12: Unable to retrieve data directly unless StorageCentral is installed.

StorageCentral reports are fairly processor intensive (even canceling a report takes a while to shut down). In addition to scheduling reports for later, the product gives you the ability to select and run multiple reports at one time, as Figure 3.13 shows. You can use this feature to group a bunch of reports together because they can take a few minutes to run against the thousands and thousands of files that a company keeps. As the following figure shows, when you're selecting multiple reports, you can configure the settings for each report (by clicking Properties). As the reports finish running, they'll generate the Active HTML, which allows you to deal directly with the files from within the report, with the right-click enabled command menus. Alternatively, if you need to continue working with StorageCentral, you can launch another instance of the console while the first is still running.

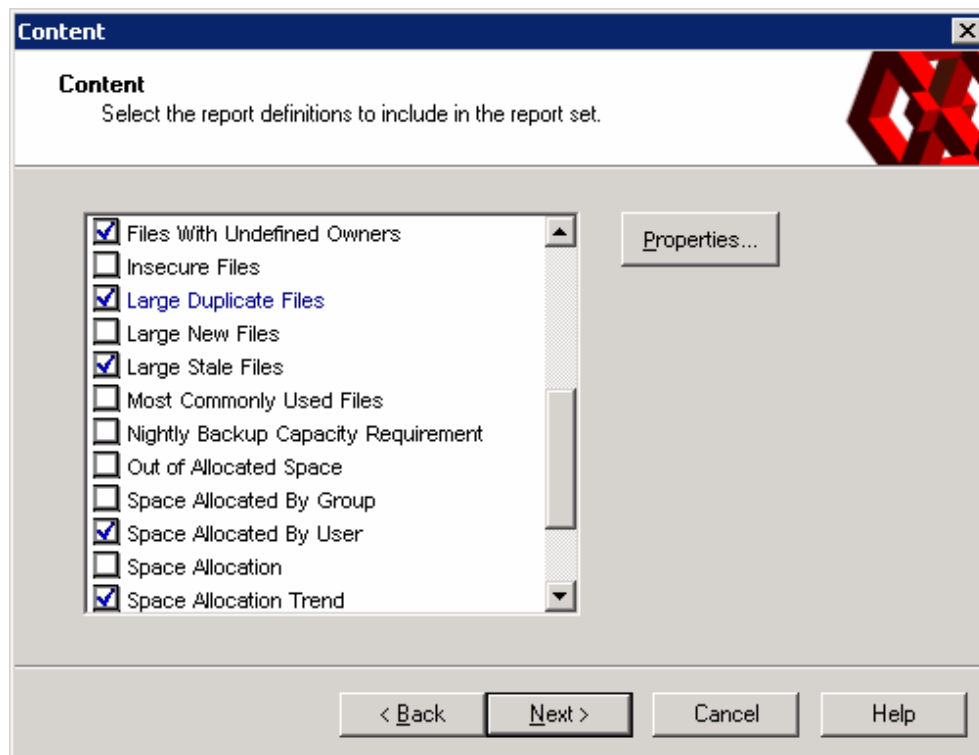


Figure 3.13: Selecting multiple reports.

Rather than hunt through the reports list, selecting reports individually, I prefer to use the Daily Storage, which generates the following reports by default:

- Files by Type
- Large New Files
- Large Stale Files
- Most Commonly Used Files
- Nightly Backup Capacity Requirement
- Space by Disk Drive

Rather than grab additional reports individually from the reports list, I prefer to select additional reports in the existing reports properties. This functionality lets you select one report and let it run while you get other work done. You then have multiple reports ready and waiting in the console for when you need them.

The following list highlights additional reports and their benefits:

- **Space by User**—The heart and soul of file server administration and SRM is knowing who is using what. Figure 3.11 shows the Space by User report in graph view. This report is useful for identifying who is using what.
- **Space by Directory**—This report shows the space used at a root point by all folders and files below that point. I'll use the Space by Directory report later as an example of working with data imported into Microsoft Excel.



The following circumstances provide an example situation in which the Space by Directory report came in handy. I was working on a project in which I would burn a CD-R disc and create a new project directory when the total size of files in the project directories hit 650MB. If I failed to keep track and the files exceeded 700MB, I would need to create a second folder to move about 650MB into and create the project CD-R disc. The Space by Directory report let me, at a glance, see which folders were ready to archive. Later, we'll look at quotas, which you can use to automatically notify you when the threshold is hit.

- **Duplicate Files**—As I mentioned earlier, this report is a high-return report, instantly allowing you to free up a good chunk of space. Give this report a little bit more time to run than the others, as it has a lot of comparing and sorting to do before presenting the results.

The reports contained in StorageCentral aren't static output. Right-click one of the files listed in the Duplicate Files report, and notice that you're given several menu options including Find, Delete, Copy to, and Move to. These functions let you perform file management directly within the reports. However, because the reports are geared for direct file management, they lack the ability to manipulate the onscreen view of the information; there is no way to sort or shift the columns or even copy the information to the clipboard to paste into another program. Figure 3.14 shows an example of the Disk Drive Summary report.

To change the view, such as the sorting of a report, you have two options. First, you can modify the report's properties and run it again. Second, you can output the report data and work with it externally.



Before you get too extreme deleting files in the Duplicate Files reports, remember that the deleted files don't go to the Recycle Bin even when you run the reports and delete files on the file server. Thus, you might want to offload the files to tape or cheap disk storage before deleting.

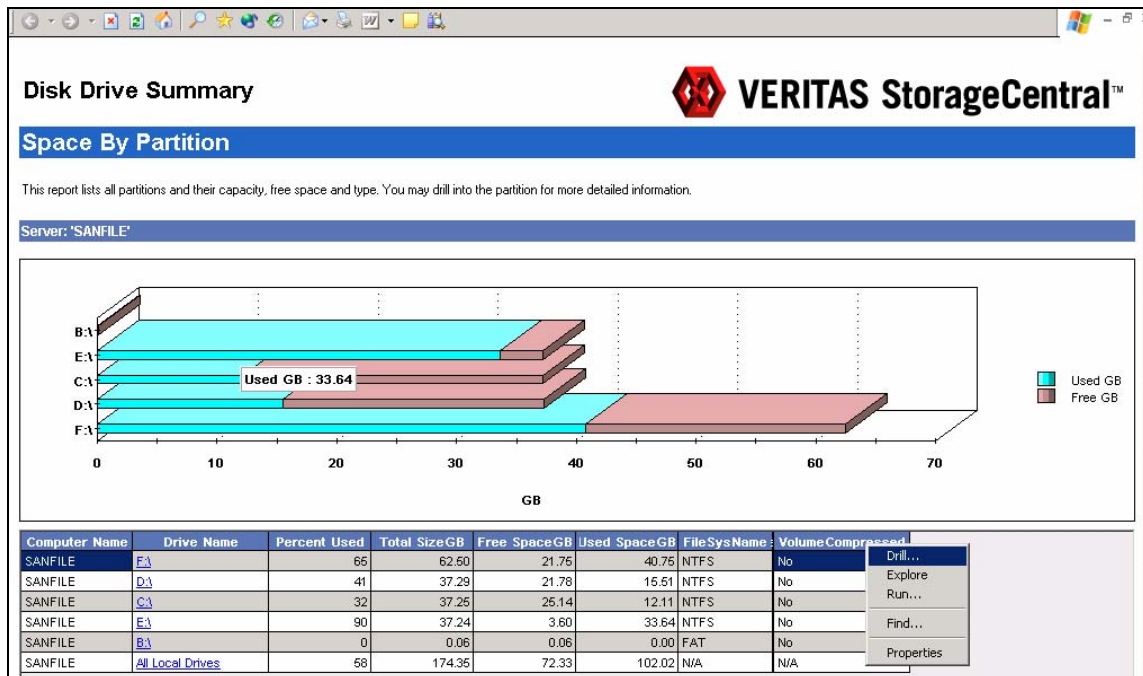


Figure 3.14: Disk Drive Summary report.

Working with Report Data

Let's take a look at exporting one of the reports and working with the data externally. Figure 3.15 shows the Disk Drive Summary Report in the .txt file format viewed in Notepad. The columns are separated by a pipe symbol (|). An alternative to using the Save As option, is to open the report properties, and on the Format tab, specify that you will output to an HTML CSV or ASCII file. Also on this report tab, you can choose whether to display the file sizes in bytes, kilobytes, megabytes, or gigabytes. By outputting to one of these formats, you can save the raw data for custom presentation or mathematical manipulation.

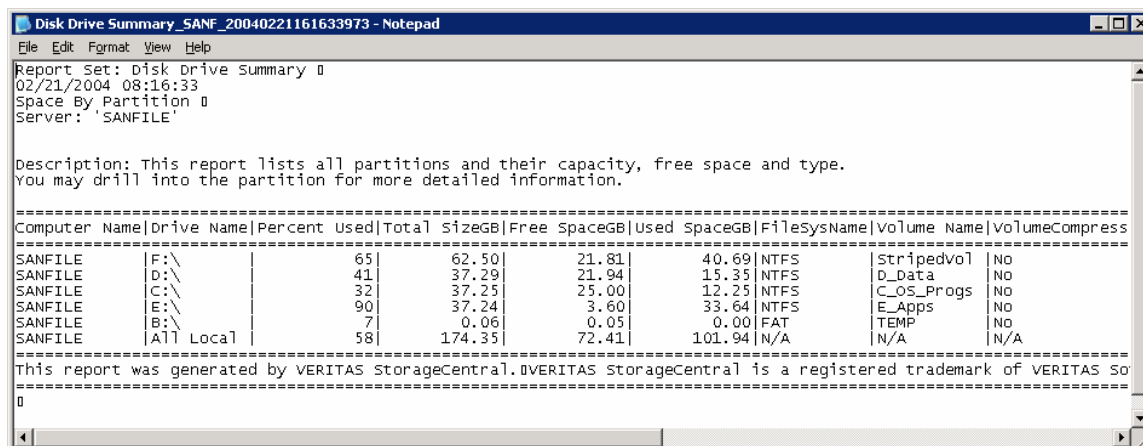


Figure 3.15: Disk Drive Summary Report in the .txt file format in Notepad.

As you can see in Figure 3.16, you don't gain much clarity from simply bringing the file into Microsoft Excel. The next step will be to break apart the data into cells and apply formatting.

Microsoft Excel - Disk Drive Summary_SANF_20040221161633973

File Edit View Insert Format Tools Data Window Help

Type a question for help

Report Set: Disk Drive Summary

Computer Name	Drive Name	Percent Used	Total SizeGB	Free SpaceGB	Used SpaceGB	FileSysName	Volume Name	VolumeCompress
SANFILE	F:\	65	62.50	21.81	40.69	NTFS	StripedVol	No
SANFILE	D:\	41	37.29	21.94	15.35	NTFS	D_Data	No
SANFILE	C:\	32	37.25	25.00	12.25	NTFS	C_OS_Progs	No
SANFILE	E:\	90	37.24	3.60	33.64	NTFS	E_Emusic	No
SANFILE	B:\	7	0.06	0.05	0.00	FAT	RAMDISK	No
SANFILE	All Local	58	174.35	72.41	101.94	N/A	N/A	N/A

This report was generated by VERITAS StorageCentral.
 VERITAS StorageCentral is a registered trademark of VERITAS Software Corporation.
 Other trademarks are the property of their respective owners.

Figure 3.16: Disk Drive Summary report in Excel before parsing and formatting.

After you apply formatting in Excel, as Figure 3.17 shows, the report is much easier to read, especially because you can apply formatting and eliminate unnecessary columns.

See the sidebar “Using Excel to Manipulate CSV Files” for more information about Excel.

Report Set: Disk Drive Summary

2/21/2004

8:16

Space By Partition

Server: 'SANFILE'

Description: This report lists all partitions and their capacity, free space and type.

Drive Name	Percent Used	Total SizeGB	Free SpaceGB	Used SpaceGB	FileSys Name	Volume Name	Volume Compressed
F:\	65	62.50	21.81	40.69	NTFS	StripedVol	No
D:\	41	37.29	21.94	15.35	NTFS	D_Data	No
C:\	32	37.25	25.00	12.25	NTFS	C_OS_Progs	No
E:\	90	37.24	3.60	33.64	NTFS	E_Apps	No
B:\	7	0.06	0.05	-	FAT	TEMP	No
All Local	58	174.35	72.41	101.94	N/A	N/A	N/A

This report was generated by VERITAS StorageCentral.

Figure 3.17: Disk Drive Summary report in the .txt file format in Excel after formatting.

Using Excel to Manipulate CSV Files

The following list highlights a few quick tips for using Microsoft Excel to manipulate CSV files.

- Set your cursor just below the field headings in column A (such as Machine Name, Directory; usually cell A9 in my worksheets), and select Alt + Window, Freeze Panes. This setting will let you scroll about the report and still be able to see the column headings. For maximum viewing, set your scroll lock button before freezing the panes, and scroll the A9 up to the top position.
- With your cursor still in what is now the anchor cell, A9, select all cells by pressing Ctrl + Shift + End and you're ready to sort data however you want. Just select Alt + Data, Sort and notice that the Sort by options now show the field headers such as Machine, Directory, File Owner, Dir Size Used KB, and so on.
- Do you find that you're constantly adjusting the print settings of CSV reports? Open a new workbook and select Alt + File, Page Setup. Enter your print setup options; I usually select the fit-to-one page option (whenever possible), zero out the margins, center horizontally and vertically, and enter the page headers and footers, such as illustrated in Figure 3.18.

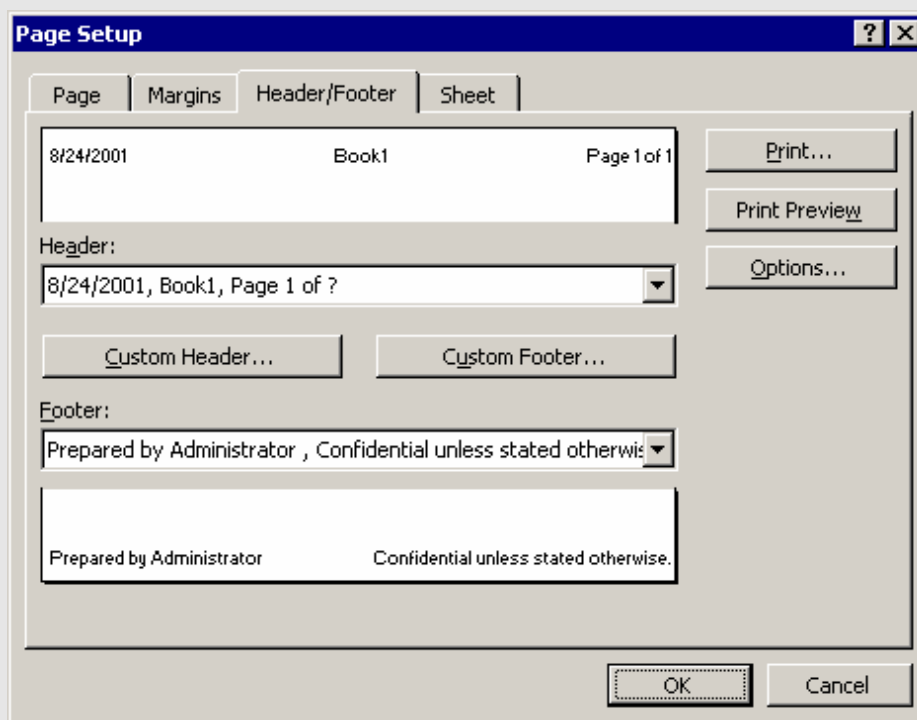


Figure 3.18: Sample page setup options for printing CSV reports using Excel.

Save the file in %ProgramFiles%\Microsoft Office\Office\XLStart where %ProgramFiles% is usually C:\Program Files. The file will be available when you start Excel. (You can even change directory by using

```
cd %ProgramFiles%
```

at a command prompt.) If instead you save the file as a template in the Templates folder (%APPDATA%\Microsoft\Templates), it will be available when you select Alt + File, New. The difference is that instead of opening the CSV file, you'll select Alt + Data, Get External Data, Import Text File, enter your CSV file (or *.csv to see all), choose Delimited, and select the comma as the delimiter in the Text Import Wizard.

- Do you find that you get an error when trying to save a CSV file because you the file is still open in Excel and you still need the file open in Excel? By using the above method for bringing the CSV data in (Get External Data from the Data menu), it creates a new worksheet and doesn't lock the CSV file. (I still have to delete the first row and the first column.)

Figure 3.19 shows the output of the Space by Folder report imported into Excel. Not a pretty picture, which is exactly why I included it here. There is always the risk of including too much information in a report and losing the focus on what you set out to audit. In this example, I wanted to know the size of all files and folders below the root user folder, which is known as the Branch Size Used. The Branch Size Used is the sum of all files and folders below a particular folder, which can be handy information. The other option is to shell out to a command prompt and run the `Dir /S` command, but even that can take a very long time to generate the last line of information, which is really all that you want to know.

Directory	File Owner	Used KB	Compre	y Dirs	y Files	Branch Size	Compressed	Dirs	Files
FileSRV32	Unknown	-	-	13	13	246,212	246,212	2,331	3,612
FileSRV32\WFileSRV32\FFileShare\Alice	FileSRV32\Alice	-	-	4	6	221,746	221,746	447	1,596
FileSRV32\WFileSRV32\FFileShare\Alice\Download	FileSRV32\Alice	6,617	6,617	5	10	52,671	52,671	54	964
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560	FileSRV32\Alice	222	222	2	4	22,859	22,859	23	321
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\Simulations	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images	FileSRV32\Alice	628	628	2	26	22,637	22,637	21	317
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages	FileSRV32\Alice	68	68	-	54	68	68	-	54
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m0	FileSRV32\Alice	221	221	19	41	21,940	21,940	19	837
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m00	FileSRV32\Alice	24	24	-	7	24	24	-	7
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m01	FileSRV32\Alice	549	549	-	70	549	549	-	70
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m02	FileSRV32\Alice	184	184	-	27	184	184	-	27
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m03	FileSRV32\Alice	15,805	15,805	-	81	15,805	15,805	-	81
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m04	FileSRV32\Alice	425	425	-	44	425	425	-	44
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m05	FileSRV32\Alice	353	353	-	65	353	353	-	65
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m06	FileSRV32\Alice	287	287	-	35	287	287	-	35
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m07	FileSRV32\Alice	617	617	-	54	617	617	-	54
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m08	FileSRV32\Alice	381	381	-	52	381	381	-	52
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m09	FileSRV32\Alice	626	626	-	77	626	626	-	77
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m10	FileSRV32\Alice	190	190	-	34	190	190	-	34
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m11	FileSRV32\Alice	1,016	1,016	-	69	1,016	1,016	-	69
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m12	FileSRV32\Alice	435	435	-	67	435	435	-	67
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m13	FileSRV32\Alice	229	229	-	23	229	229	-	23
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m14	FileSRV32\Alice	224	224	-	36	224	224	-	36
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m15	FileSRV32\Alice	280	280	-	38	280	280	-	38
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info	FileSRV32\Alice	2,141	2,141	-	3	2,141	2,141	-	3
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin	FileSRV32\Alice	-	-	5	5	-	-	26	26
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM	FileSRV32\Alice	-	-	1	1	-	-	1	1
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras	FileSRV32\Alice	-	-	1	1	-	-	1	1
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML	FileSRV32\Alice	-	-	14	14	-	-	19	19
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help	FileSRV32\Alice	-	-	3	3	-	-	3	3
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files\lib	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files\lib\MenuLib	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files\lib\MenuLib\PPC	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files\lib\MenuLib\PPC\SideBars	FileSRV32\Alice	-	-	-	-	-	-	-	-
FileSRV32\WFileSRV32\FFileShare\Alice\Download\1560\WcbFile\Images\courseimages\m16\media\Info\Wcbadmin\COM\ADsType\Docs\Extras\DNS\WMI\HTML\Help\help_th_tbar_files\settings_files\find_files\lib\MenuLib\PPC\SideBars\Support	FileSRV32\Alice	-	-	-	-	-	-	-	-

Figure 3.19: A difficult-to-read report shown in Excel.

When working with Space By Folder or other reports on directories, one trick is to change the Number of Levels from the default of 1, as Figure 3.20 shows. If you run this report at the root of a share or drive, you might not get enough detail on those user directories and other subfolders. Note that this is done on the Report Set that you are running and not in the Report Definition. Simply open the report properties, select the Content tab, and select the report properties button. On the Settings tab, select the Number of Levels input field.

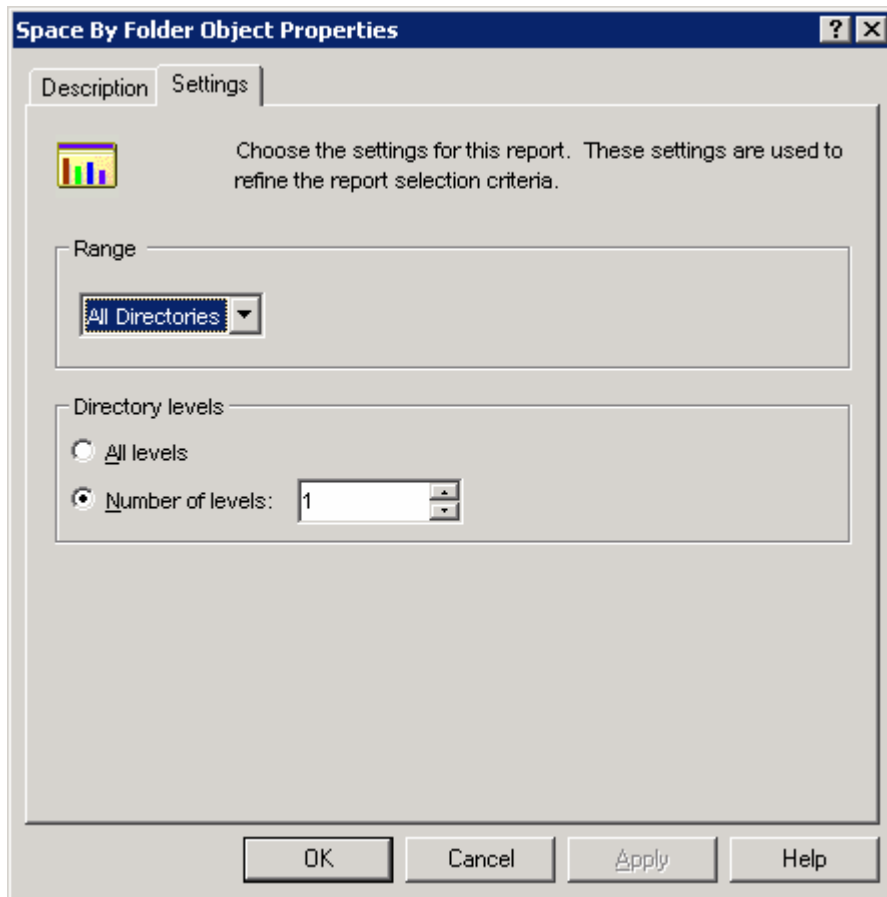


Figure 3.20: Number of Levels for Folder Object Reporting.

Customizing Reports

When you look at Report Definitions you'll see two icons, which are highlighted in Figure 3.21. The two icon types represent standard report definitions (top icon, with the hammer) and custom report definitions.



Figure 3.21: Report definition icons.

The first type has a limited set of properties (description, settings, and detail), and the second type offers the full set of properties that Figure 3.22 shows. The engine behind many of the reports in StorageCentral is the query. For the most part, you should be able to use the many default reports supplied with the product. However, you might want to either modify an existing query structure or create your own from scratch. Figure 3.22 shows a sample query expression.

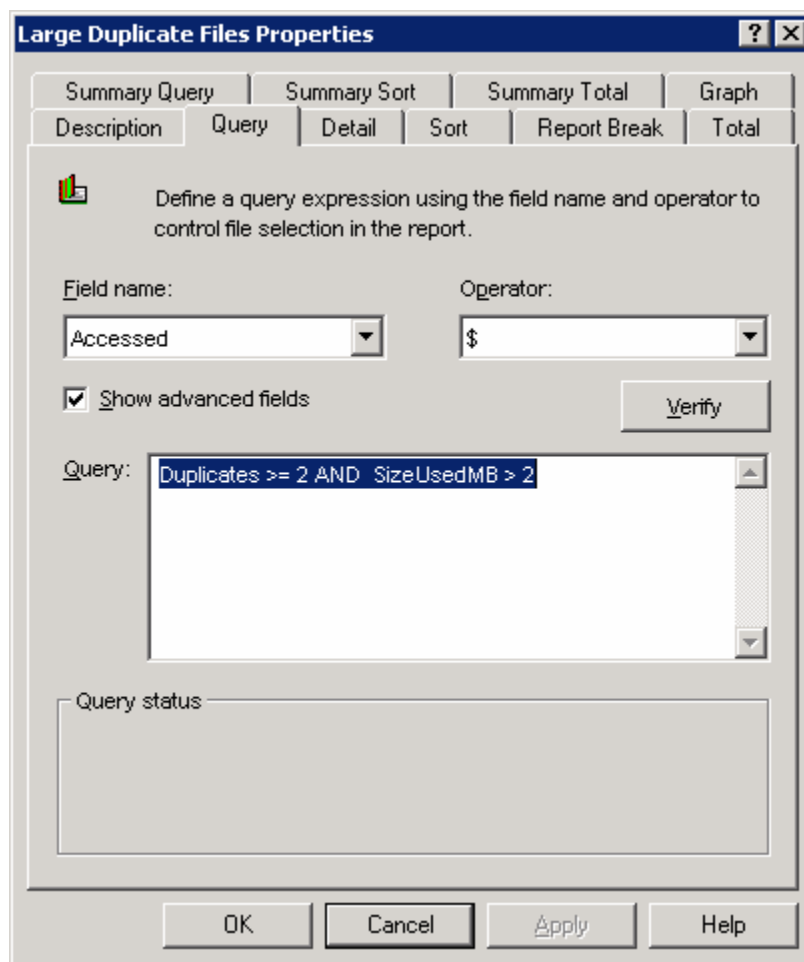



Figure 3.22: Sample query expression in Custom Report Properties.

Preparing for the Next Phase

At this point, you have the capability to generate reports that will give you the following information:

- Storage utilization (space used and free space)
- Storage users
- Types of files stored and where
- Duplicate files
- Storage hierarchy (where storage is being used and administered)

From this information, you can begin to generate a gap analysis, which compares where you are now with where you want to be, both in terms of your existing storage resources and taking advantage of the new WS2K3 features. Let's start by exploring how to eliminate duplicate files.

 We'll continue this analysis in the next chapter as well as begin the next phase, planning an SRM deployment.

Eliminating Duplicate Files

Figure 2.23 shows a sample SRM report for duplicate files. In this case, the amount of wasted space from duplicate files is significant, 833MB of the gigabyte or so of storage.

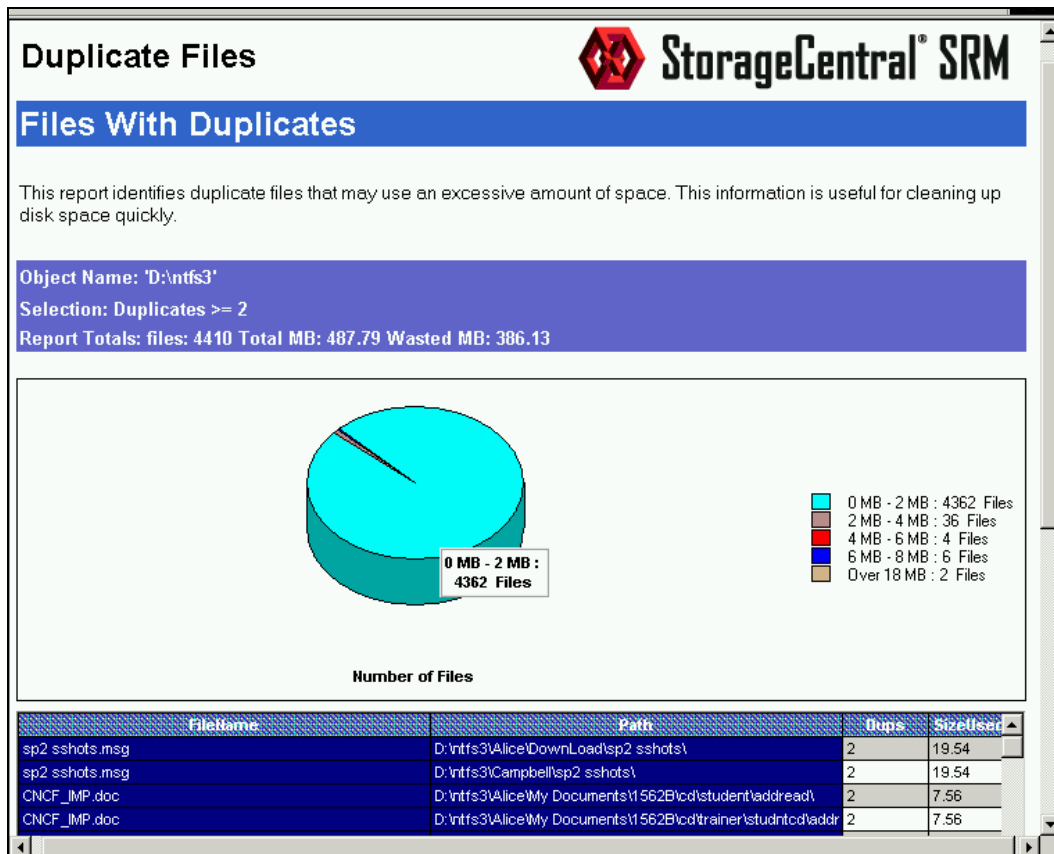


Figure 3.23: Report showing wasted space from duplicate files.


By looking at the report, you can understand the reasons that the duplicate files are being created:

- People are saving the same files; for example, email attachments that are sent to many people
- People are working on the same projects and files; for example, video and white paper files in the sample report
- People are saving setup and application installation files; for example, the downloaded setup files in the sample report
- There are files that share a common filename even though they're not actually duplicate files. (The example in this report being the Outlook.OST, which is an offline synchronization folder.) This occurrence isn't so much a problem that needs a solution as a caution to the administrator that some files might not be duplicates but have the same name and in rare cases are the same size.

Thus, your SRM plan must deal with the source of each of these reasons for creating duplicate files. First, resolve the situation in which people are using the file share as a dumping ground for files received by email. This behavior creates duplicate versions of the files because users are receiving many of the same files and instead of maintaining the single-instance storage that an email system (in this case Exchange Server) would maintain, they're creating duplicate files. The solution to this problem is twofold:

- The original file author could merely send out a link to the original file (I'll show a good method for doing so as we develop a storage-management solution).
- If the users have sufficient space in their mailboxes, they may be less inclined to dump the attachments.

Another reason for the duplicate files is that many of the users are working on a joint project, and they've copied the files to an individual share. The solution to this problem is more complex, as collaboration and workflow applications are designed to address this problem. The document author might say "I have created this document, and you are enabled to add to it," but the end user might still feel the need to keep a personal copy.

 In the report showing duplicate files, there is a column that provides Revised Days. From this report, you can immediately see that some of the files are quite old and might fall into the next category, unused files.

Eliminating Unused Files

Unused files are most often older files that no one is currently accessing, as illustrated in the wasted files report shown in the next section. A more difficult-to-manage category of unused files are those that have been orphaned, usually by a user who is no longer working for the company. For those of you who provide working space to contractors and vendors, you know how difficult a task cleaning up files can be. Who wants to wait for the files to age to clean them up? SRM products can generate a report for files with undefined access control entries (ACEs), which displays all the files that have an owner security identifier (SID) in an ACE that has been deleted or is invalid.

Eliminating Wasted Space

Wasted space can come in several forms: from files that are no longer being used, duplicate files, prohibited files, or all of the above. The wasted space report, which Figure 3.24 shows, contains a column labeled Wasted, which uses a string to indicate why the file was flagged for this report. The string *a* is for aged, *e* is for expired, *l* is for large, and *d* is for duplicate. Not shown is *o* for over-allocated, as I have not set any quotas on the users or directories. The wasted space report is helpful for allowing you to see why space is being used and getting to the root cause.

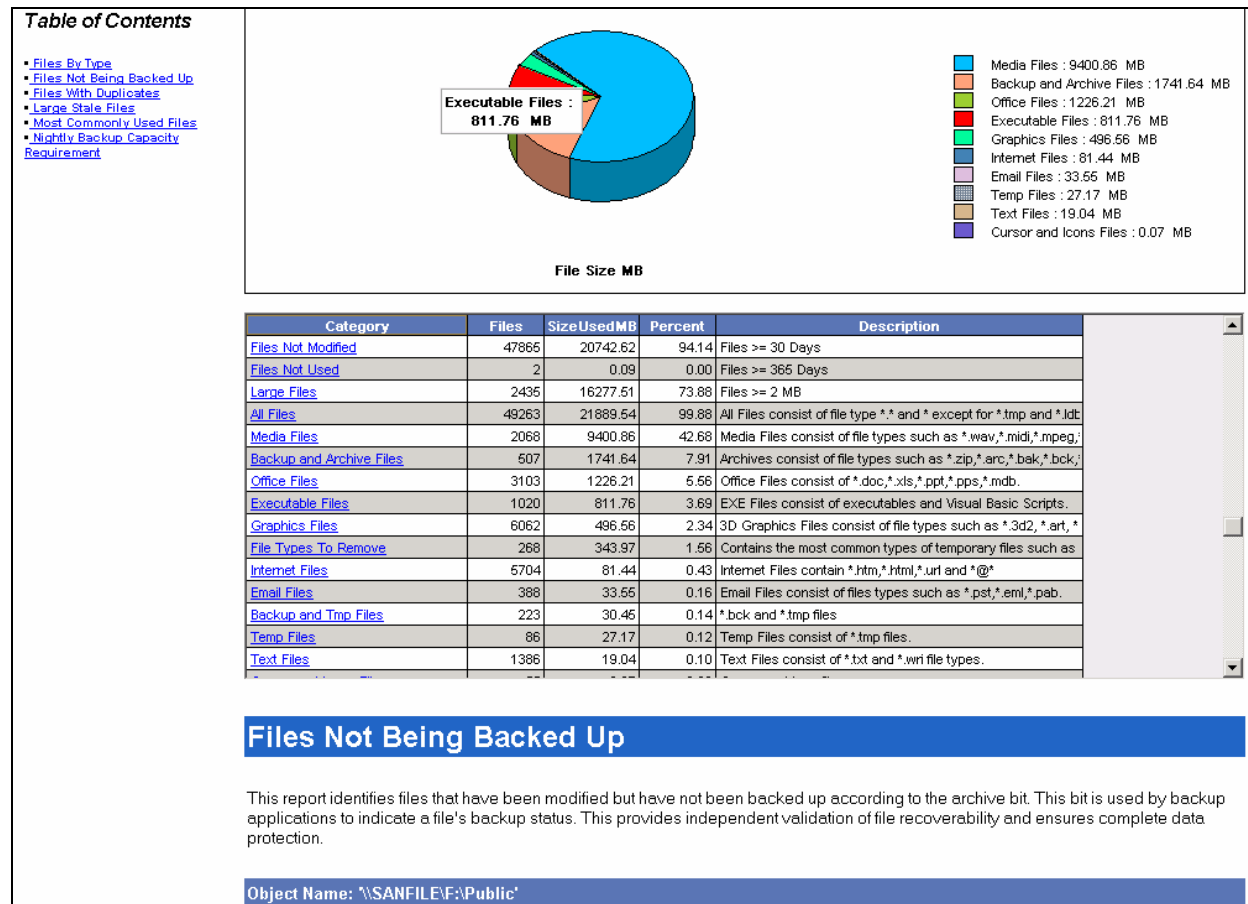


Figure 3.24: Best practices report in StorageCentral SRM.

Reducing Consumption

So you've run these reports and done all that you can to remove older, unneeded files and developed a solution to reduce duplicate files. There are several additional methods for reducing the level of storage consumption. The most severe is to set immediate storage quotas or allocations on the storage used without any contingency for alternative storage locations. The least severe is to provide a storage location alternative and assist users in migrating from one storage system to another.

One method of reducing consumption is to remove files that fall under the category of restricted file types, the most infamous of late being .MP3 files. You can remove these files by scavenging existing shares and removing them, or you can prevent future storage of the files by using an SRM product.

Enforcement Policies

The process of implementing storage quotas can be initiated by a soft touch (for example, by providing information messages and warnings) or by immediately imposing a hard limit and preventing further file storage until the user cleans up storage to stay under the quota threshold. Informational messages can be used to let the users know that a hard limit is coming.

As an administrator, you might also be faced with making the initial decision of where to define the quota limits. A third-party SRM product can be useful for this decision. For example, the StorageCentral SRM product offers a template called Baseline directory utilization that allows you to set the future quota at 200 percent of the current usage. Figure 3.25 illustrates using the predefined template as a soft quota enforcement policy. This setting gives users a little bit of breathing room before they hit their quota, but can also reward the space hog at the time the SRM product is first deployed, so you might want to look at overall usage reports before setting this quota.

Notice also the footnote in the dialog box, stating "Not applicable with Network Appliance." This note is new to StorageCentral SRM and applies to NAS using Network Appliance Filers. If you are using or planning to use one of these storage solutions, you should read the StorageCentral SRM Help file information about why the exceptions shown below apply. The Filer uses a different file screening or blocking policy that does not match StorageCentral, so you'll want to learn how to best deal with the differences.

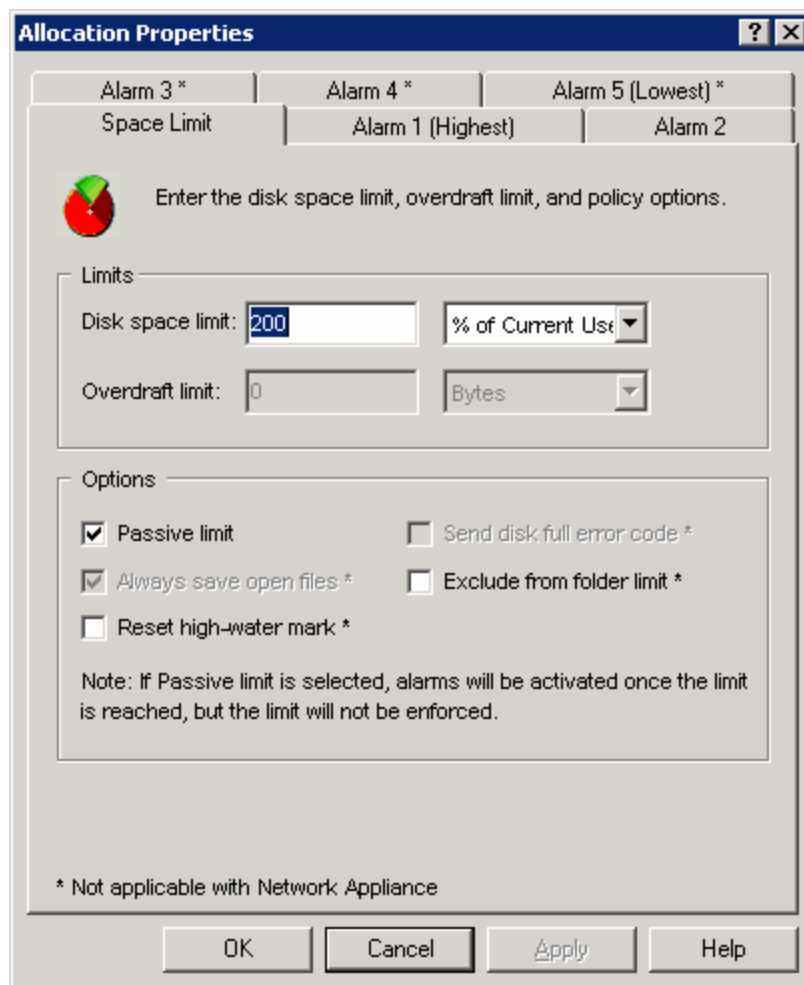


Figure 3.25: Allocation Policy for 200% of Used Space as a soft quota.

How you implement storage policy is entirely up to you and your organizational policy; if you want to increase acceptance of your quota-management system, people must benefit from the SRM product and policy. You must make clear the fact that quota management is in the end users' and the organization's best interests and not just something that is being done so that you, the administrator, can leave work on time. In the next chapter, we'll look at defining administrative boundaries and assigning roles as part of the SRM process.

The next step in the planning process is to look at different storage applications and the various storage location alternatives. The plan will include which type of storage you are providing and how you will manage it—which tools are available and what policies will be set. Perhaps some of these storage alternatives will be new to your organization, for example, applications such as Windows SharePoint Services (which was formerly known as SharePoint Team Services—STS—which explains why the download file is named STSV2.exe).

Support of Storage Applications

In the first chapter, I introduced the concept of storage location alternatives and gave a simple scenario in which I send you an email with an attachment that you want to save and share with others. Let's look deeper at each of the storage location alternatives, both from the benefits to the end user (the main reasons that they would choose to use one over the other) and the impact that each has on the storage administrator.

Table 3.2 shows the advantages, disadvantages, administrative considerations, and deployment recommendations for these types of storage: mailboxes, personal mail folders (PSTs), public mail folders, local file folders, network shares, Microsoft collaborative applications, intranet or Internet Web servers, and database applications.

Storage Location	Advantages	Disadvantages	Administrative Considerations	Recommendations
Mailbox	Saved in context of message; usually some simple search functions	Mailbox sizes are usually very limited; may be needed for business-critical communication	Must usually back up the entire mail store, so you must consider limits placed on mailbox sizes; depends on deployment of messaging application, such as Exchange Server	Keep the limits on mailbox sizes and use file shares; see the section about HSM solutions for multi-class storage tiers
Personal folder (PST)	Saved in context of message; usually some simple search functions	Depending on location, may not be RAID protected storage (local computer); more prone to corruption, very difficult to share information	Difficult to back up if the file is open, loss of single-instance-storage (one file or instance in the messaging database shared among many users); no quota on file or disk usage	Do not use for shared or critical information
Public Folder (email)	Saved in context of message; better search functions through content indexing; open for sharing among multiple users	Often difficult to navigate tree (separate from mailbox folders)	Requires well-designed public folder hierarchy and some level of end-user training to be useful; depends on deployment of messaging application, such as Exchange Server	Create structure and encourage usage instead of personal folders for shared information
Local computer folders	Personal/private information; ease of use; roaming information (notebooks)	Lack of fault tolerance (RAID, regular backups) and usually lesser-quality hard drives; notebooks can be lost or stolen	Distributed management (which can be a good thing or not); difficult to share or locate information	Use network file sharing with offline folders to synchronize information

Storage Location	Advantages	Disadvantages	Administrative Considerations	Recommendations
Network file share	Easiest structure (no Web pages or database to create); direct to the end user with very little administrative intervention. Storage appliance can add features such as file screening and data replication	Used to be a problem for roaming users until offline folders allowed complete synchronization	Requires a third-party utility to really manage the storage and provide end-user quotas; can use replication to place copies of information closer to end users	Use network file sharing with offline folders to synchronize information
Web Storage System (WSS)	Content indexing and permissions on stored files (similar to NTFS file shares); can be used for collaborative workflow (document versioning, and controlled check-in and out); maintains single-instance storage (one copy of a file for many users)	Requires learning the new application (for collaboration)	Depends on deployment of application, such as SharePoint or Exchange WS2K3; can use built-in replication model to place copies of information closer to end users	Consider for collaborative workflow and publishing environments; benefits over Web server (see below)
Intranet or Internet Web server	Broad variety of client browser support (for reading and linking or indexing information)	Must go through Web master to publish; difficult for end user to act as document editor or publisher, which can lead to stale information	Depending on the amount of information to be published and the intended audience; not for use as a collaborative storage system	Useful for read-only publications and distributing software over the Internet
Database system	Rapid retrieval of information based on indexes and relational structures	Takes a lot of effort to set up and use, requiring database specialists who must act on behalf of the end user or information consumer	The only choice for bits of information that must be gathered into tables and structured for it to be useful (for example, human resources personnel information)	Will continue to be used for information that must be fairly well structured to be shared

Table 3.2: Storage location alternatives.

There are at least seven storage location alternatives for you to provide to your end users. Whether you offer one or all, you must develop the ability to include them in your SRM plan and process.

Summary

In this chapter, we started the next phase, planning the SRM deployment. We used the information captured on both storage capacity as well as storage performance to determine a course of action, which I laid out in a flowchart. The flowchart identified the storage management decisions that you must make, most notably whether to live within your existing capacity or to add capacity, even if for performance reasons.

We then looked at how SRM solutions can help eliminate duplicate files, unused files, and wasted space and reduce consumption. We also covered the tools and utilities you can use for storage management. Phase 1 of your SRM deployment is analysis, so I presented some analysis tools, showing what you can do natively in Windows Server with the Performance Monitor as well as with the Windows Server resource kit tools. Later I showed why you may need a third-party SRM suite to improve the audit process and prepare you for the next phase, planning your SRM deployment.

In the next chapter, we'll cover structuring the SRM project and using storage-management tools to make better use of either your existing storage or your newly deployed storage. We'll also explore the options creating additional storage capacity using the Windows Server RSS, a basic HSM system. We'll do a storage performance design example for Windows Server, using Exchange Server 2003 and file servers as examples to illustrate storage performance applications. Next we'll look at expanding storage arrays through Dfs and the hardware selection process involved in migrating to new storage systems. This discussion will include exhausting the core Windows Server functionality and SRM features (including quota management) before turning to a more comprehensive SRM solution. We'll cover product functionality that assists toward your SRM goals—eliminating duplicate files, eliminating unused files (aged and orphans), eliminating wasted space, and reducing excess consumption—through setting disk quotas.