Understanding NTP Software ODA™

An NTP Software® White Paper

Abstract

While conventional archiving has served its users well for years, the explosive growth of user (file or unstructured) data coupled with platform consolidation has created the need to look at archiving in new ways. ODA™ greatly increases the ROI on any archiving investment while improving the quality of the data in the archive.

Doc # 6076EF

The information contained in this document is believed to be accurate as of the date of publication. Because NTP Software must constantly respond to changing market conditions, what is here should not be interpreted as a commitment on the part of NTP Software, and NTP Software cannot guarantee the accuracy of any information presented after the date of publication.

This White Paper is for informational purposes only. NTP SOFTWARE MAKES NO WARRANTIES, EXPRESS OR IMPLIED, IN THIS DOCUMENT.

NTP Software products and technologies described in this document may be protected by United States and/or international patents.

NTP Software, 20A Northwest Boulevard, Unit 136, Nashua, NH 03063, USA
1-603-622-4400

www.ntpsoftware.com

 Doc # 6076EF

Introduction

Over the last few years, user data (file data or unstructured data) has grown explosively. On most networks user data accounts for 80% or more of the total storage requirements. Managing this data effectively is now one of the Industry's most important tasks.

For years archiving has been a part of many companies' data management strategies, both for reasons of data protection and as a way to control costs. However, with the explosive growth of file data, managing the cost of archiving itself has now become a necessity. NTP Software ODA™, a new technology in the form of a product, dramatically increases the ROI of any archiving investment.

The costs of conventional archiving methods are essentially proportional to the amount of data in primary storage. As the quantity of this data increases – which it continues to do at an average rate of about 50% a year (per Gartner Group) – the costs of conventional archiving methods increase commensurately. In fact, some end-user organizations have already reached the point where their archiving windows are too small to accommodate all of the processing that needs to be done on an on-going basis.

The solution is a paradigm shift to NTP Software ODA™. With ODA, the majority of the operating cost of conventional archiving is eliminated while all of its benefits are preserved. In addition, ODA improves the overall quality of the data that reaches the archive.

The Benefits of File System Archiving

The benefits of archiving are generally well known and well understood, to the point where most companies do it in some form or another. As we move from conventional archiving to ODA, the goal is to extend and enhance the existing benefits while greatly reducing overall cost and complexity. Therefore, we seek to maintain and enhance these key elements in the total solution:

- Automate the migration, storage and retention of unstructured data (file data) in accordance with IT policies
- Improve the performance of primary storage, control primary storage costs, and simplify management
- Reduce overall storage, backup and help desk costs, while controlling data loss and ensuring that all information remains protected and in compliance with corporate policies

The right archiving strategy also has the secondary benefit of reducing the costs associated with search and legal discovery by creating a centralized, indexed archive that can be searched on demand.

Enhanced by ODA™

Overall, the case for archiving is compelling. ODA™ preserves all the benefits as noted above while removing one of conventional archiving's largest cost components completely (file system scans). In addition, a best-practice implementation of ODA increases the business value of the data in the archive and further reduces storage and operating costs.

Overcoming the Challenges of Conventional Archiving

All archiving solutions have three cost elements:

1) The cost of determining what to move
2) The cost of actually moving the selected data
3) The cost of maintaining the archive

Every one of these costs is proportional to the amount of data involved in the process. For example, moving more data costs more than moving less; maintaining a bigger archive costs more than maintaining a smaller archive, etc. Therefore, if we wish to control the cost of archiving, we must manage the amount of data involved at each of the above steps in the process.

From analysis we know that items 1 and 3 dominate the cost equation. In conventional archiving methods, the actual cost of moving the data is relatively small compared to the cost of finding what to move and maintaining it in the archive. To manage cost-effectively, it becomes important to go where the money is and to attack the cost of finding what to move as well as trying to reduce the amount of data that ends up in the archive.

Conventional archiving methods typically use an expensive strategy to determine which files should be moved. Specifically, they regularly scan the entirety of primary storage to find files that meet one or more simple criteria. If the file meets the criteria, it gets moved. The problem with this is that the user organization is paying the same cost over and over again to repeatedly scan the same files and move very few of them. In addition, as primary storage grows over time, the cost of this repetitive scanning grows proportionately. Eventually, as larger organizations are beginning to discover, the archiving window becomes too small to accommodate all the processing that needs to be done to determine what to move. At this point conventional archiving is no longer practical or economically viable. A solution has turned into a problem.

Another challenge is that the simple rules used by conventional archiving move files without regard to their business value or future needs. This results in very few files ever being deleted and further accelerates the cost. In fact, users often come to see the archive as an alternative to deletion and keep more data online than they would have done if archiving were not an option. This means that not only does the cost of finding potential files to archive grow without bounds, but so does the archive. Most organizations will ultimately get to the point where the cost of the conventional approach cancels out all of its benefits.

ODA™ – A New Paradigm

The solution to the problems posed by conventional archiving methods is to think about the opportunity differently. Conventional archiving looks at archiving as though it were a project: "We need to find all the files to archive and get them moved." However, archiving is not a one-time event; it is really a continuous process. The economics of a continuous process are different from those of a one-time event. For continuous processes to be cost-effective, you must minimize their recurring costs. One time costs (the cost to move a particular file – which only happens once) are relatively unimportant. This is what the conventional archiving model misses.

Primary storage as it exists at any given moment is just fine as long as it remains in a steady state. Only when users add to the existing contents of primary storage does something need to change. If no new data were ever added, nothing would ever need to be archived. This is the critical insight.

It is user activity that drives the need for archiving, not the passage of time.

It is this re-conceptualization of "archiving as a continuous process that needs to occur as a result of day-to-day user activity" that gives us the opportunity to mitigate costs and leads us to the ODA solution.

ODA™ – What is it?

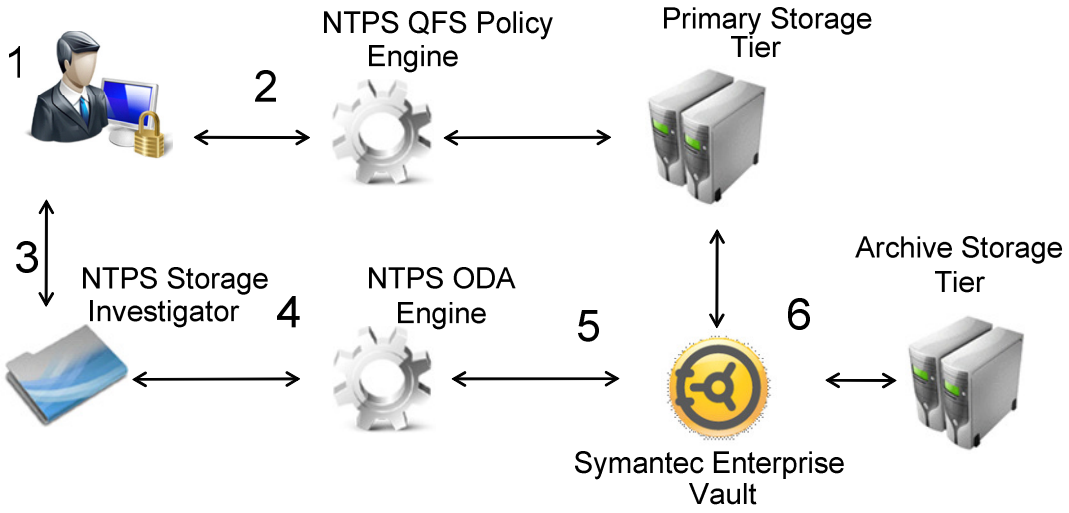
NTP Software ODA™ relies on the end-users' own activity to discover archiving opportunities. ODA uses what are often existing storage management policies that include elements such as quotas or filters to discover an archiving opportunity. Once ODA is triggered it either sends a message to selected users that their recent activity has resulted in the need to remove or archive some files from primary storage and asks them to take action, or it acts automatically.

In situations where it is appropriate to engage the user community, they are presented with a summary of the data for which they are responsible via a simple, friendly interface. From this interface they can view their data in a variety of ways that are appropriate to managing storage, for example by size, age, or type.

It is at this point that the end-users add real value to the overall process. They can use their unique knowledge of the business value of their files by selecting the ones most appropriate to archive (as opposed to the files they know they will need to use next week or next month). Archiving only the most appropriate files reduces operating costs. At the same time the users can take advantage of the opportunity to delete files – further reducing overall cost.

But one of the greatest benefits of ODA should not be overlooked: there is no longer any requirement to scan the entire file system. Scanning terabytes of data to find files to archive takes hours and generates millions of I/Os. In conventional archiving this has to be done over and over again – every week or every month. With ODA, these costs are completely eliminated – forever!

The specifics of what happens in ODA implemented in conjunction with NTP Software QFS® and Symantec Enterprise Vault™ are illustrated on the next page.



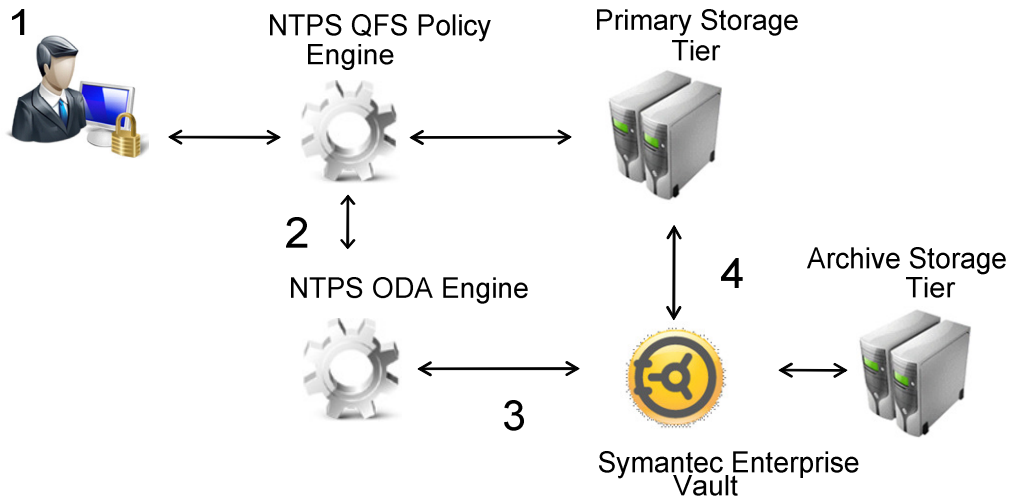
1. User attempts to store data, triggers policy
2. QFS® notifies User
3. User selects files to archive with NTP Software Storage Investigator™ (SI)
4. SI notifies ODA Engine
5. ODA Engine activates Enterprise Vault
6. Files moved by Enterprise Vault

Fully Automatic ODA™ – What is it?

Fully automatic ODA (which includes Quota-Driven Archiving™) uses the same mechanism and technology as ODA in general. However, in some cases it may not be appropriate to have users participate in making archiving decisions. For example, there are times when all the data has to be preserved for legal reasons – regardless of whether the user wants to or not. Deletion is not an option; the files simply have to be moved to the archive. In such cases decisions are automatic, so the process should be as well.

In these situations archiving is still driven by user activity in primary storage (and not by expensive file system scans) but the users are not engaged in the process of deciding what to archive. Instead the ODA Engine applies its logic / rules in a granular manner and makes decisions on behalf of the user community about what files are to be archived.

The specifics of what happens in Automatic ODA implemented in conjunction with NTP Software QFS® and Symantec Enterprise Vault™ are illustrated on the next page.



1. User attempts to store data, triggers policy
2. QFS® notifies ODA™ Engine
3. ODA™ Engine activates Enterprise Vault
4. Files moved by Enterprise Vault

Conclusion

Conventional archiving, while having provided real benefits over the years, is now something whose costs need to be managed – to the point where some end user organizations can no longer accommodate the resources required for the file system scans that drive it.

NTP Software ODA™ completely eliminates these scans and their cost while adding value in other areas as well. The net result is a dramatic boost in the ROI of any archiving solution and an increase in the business value of the data that ultimately comes to be archived.

NTP Software Professional Services

For further assistance in creating the best Storage Management Policies, please contact your NTP Software representative at 1-800-226-2755 or 1-603-622-4400.

NTP Software Professional Services offers training and consulting services in support of the deployment and configuration of your Storage Resource Management software.

http://www.ntpsoftware.com