

Acme Corporation Enterprise Storage Assessment

Prepared by:



Rocket Software, Inc.
150 N. Mathilda Place, Suite 304
Sunnyvale, CA 94086 • USA
Tel# +1.650.237.6100 / Fax# +1.650.237.9183

Table of Contents

Executive Summary.....	1
Introduction.....	3
Data Analysis	4
ROI Analysis.....	10
Supporting Information	12
Summary	13
Additional Benefits	14
Additional Rocket Arkivio Autoexplor Reports	13
About Rocket.....	14

Executive Summary

Thank you for allowing Rocket Software to provide an assessment of Acme Corporation's storage and data environment. The following information has been collected using Rocket Arkivio Autoexplore. Findings in this report include disk space utilization, file type distribution, file size distribution, growth trend analysis, and future growth projections. We have also included recommendations on how a well implemented ILM strategy could help gain efficiencies in your environment, and then followed up with a ROI analysis for implementing ILM.

In brief, we discerned the following important points from the data in the environment:

- Overall average utilization on tier 1 storage is 36%. Industry average is < 40%.
- 75.85% of file server data has not been accessed in over 6 months.
- 65.05% of file server data has not been accessed in over 1 year.
- XXGB of data is being backed up each week.
- XXGB of data is being sent off site each week.

The current storage environment consists of XX GB of file system data residing on tier 1 storage. XX GB of this data has not been accessed in over 6 months. This data could potentially be migrated to lower cost storage, while still remaining online and available to the end user should it be required. As this data would now be residing on separate storage, it could be backed up and sent offsite on a monthly rather than daily or weekly basis reducing backup times and the amount of hardware required to complete backups within required time limits. This would also reduce the time required to restore the servers thus increasing service levels while reducing potential production times.

The initial ROI, based on a migrating data that has not been accessed in greater than 6 months is \$217,000 annually. Given Acme Corporation's current growth this would represent a 3 year savings of \$644,000. The savings would be seen in the following areas.

- Avoid the purchase of expensive tier 1 storage by migrating infrequently accessed data onto lower cost tier 2 storage. This will result in a three year savings of \$317,000.
- Avoid backing up XX GB of data weekly. This will result in a longer useful life of current backup and recovery infrastructure. Through reduction of tapes and offsite storage and transportation this will result in a three year savings of \$17,000. This will also result in operational savings for personnel moving and tracking tapes.

- Reduce backup and recovery times. This will allow Acme Corporation to avoid upgrading their tape backup environment. This will result in a 3 year savings of \$237,000.
- The ability to classify data by file type. This would allow for the identification and rapid clean up of unwanted or inappropriate files. As an example, these could be in the form of temp files, install files, or mpeg's.
- The ability to manage storage on several servers as a single volume can represent a reduction in required man hours. These soft costs should also be considered, however are not shown in the estimated ROI.

A more in detailed analysis of Acme Corporation's storage environment follows.

Introduction

Nearly all organizations are experiencing an explosion in the demand for storage, particularly within data-intensive industries such as manufacturing, government, healthcare, and financial services. Gartner Group estimates the average company is growing their data at over 35% per year. The common approach to handle this growth has been to add higher capacity faster disk resources. This has proven to be very expensive both in disk costs, as well as tape resources required to backup and recover the data. Even with faster tape drives, many enterprises are finding it impossible to successfully backup all the user data in the time required. Almost more important, recovery times in the event of a hardware failure or disaster have become unacceptable.

A properly implemented life cycle management strategy can address these problems and provide significant cost savings.

- High cost fiber channel disk drives can be replaced by lower cost SATA drives on second or third tier storage. This is accomplished by keeping data that is least likely to be accessed on second tier storage, while keeping more critical high access data on first tier storage.
- On average, 45% of data currently on primary disk is able to be migrated to archive disk. That data no longer needs to be backed on a daily basis. This can reduce the amount of tape hardware required to perform backup. This also greatly reduces recovery time in the event of a failure.
- Separation of active and inactive data can also reduce the amount of data being sent off site on a daily basis. Further, in the event of a disaster, active data can now be recovered first. The data that hasn't been accessed in months can be recovered as time and hardware allows. This can also greatly reduce the quantity of hardware required at the DR site. In the event of a disaster, secondary disk can be implemented after the fact.
- Disk utilization can increase as automated policies can archive data when file system capacities reach preset limits.
- High disk consumers can easily be identified allowing charge back.
- Inappropriate or unneeded data can be rapidly identified and rapidly deleted.

Rocket Arkivio Autostor allows the systems administrator to scan, classify, group, manage, and migrate data. These activities can be set to run as scheduled policy actions requiring little or no intervention from the systems administrator.



Data Analysis

The following data analysis was made from auditing a representative portion of Acme Corporation’s storage environment. These numbers were then extrapolated to show an overall view of the entire network storage environment.

Disk Utilization

- The overall average storage utilization across all servers is 35.67%. The highest utilization is at 72.83% while the lowest utilization is at 2.91%. Redistribution of storage among servers may be possible depending on hardware constraints.(if the numbers deem necessary)

Top N Servers Based on Percentage of Capacity Used

(Any activity within the last 20 minutes may not be included.)

Rank	Server Name	Used Space (GB)	Capacity (GB)	% Used Capacity
1	FILESERVER	83.05	114.04	72.83 %
2	CADSERVER2	21.90	37.21	58.86 %
3	CADSERVER1	21.44	37.21	57.60 %
4	SUN-DEMO-02	25.68	46.05	55.77 %
5	linuxserver	98.80	373.71	26.44 %
6	BIGNAS	3.32	114.04	2.91 %

Figure 1: Servers percentage capacity used

- The highest utilized volumes are at 73.00%. These volumes will soon need to have capacity added.

Top N Volume Groups Based on Percentage of Capacity Used

(Any activity within the last 20 minutes may not be included.)

Rank	Volume Group	Used Space (GB)	Capacity (GB)	% Used Capacity
1	NetApp - F840	83.05	114.04	73.00 %
2	fileservr.agcs.com	83.05	114.04	73.00 %
3	High Performance CAD	38.89	66.43	59.00 %
4	cadserver2.agcs.com	21.90	37.21	59.00 %
5	Dell - PowerVault	43.34	74.43	58.00 %
6	cadserver1.agcs.com	21.44	37.21	58.00 %
7	Sun - StorEdge	25.68	46.05	56.00 %
8	sun-demo-02.agcs.com	25.68	46.05	56.00 %
9	linuxserver.agcs.com	98.80	373.71	26.00 %
10	HDS - Thunder	44.65	186.86	24.00 %

Figure 2: Volume group % capacity used

File Age Distribution

- Over all, 75.85% of the files currently residing on Acme Corporation’s tier 1 storage have not been accessed in the last 6 months. Depending on the utilization patterns of this data, it may be possible to migrate the majority of this data onto a tier 2 storage platform.
- 65.05% of files have not been accessed in last 6 months – 1 year.
- 10.80% of files have not been accessed in the last 6 – 12 months.
- 7.12% of files have not been accessed in the last 3 – 6 months.

File Age Distribution showing Size of Files based on Last Access Time

CAD Servers

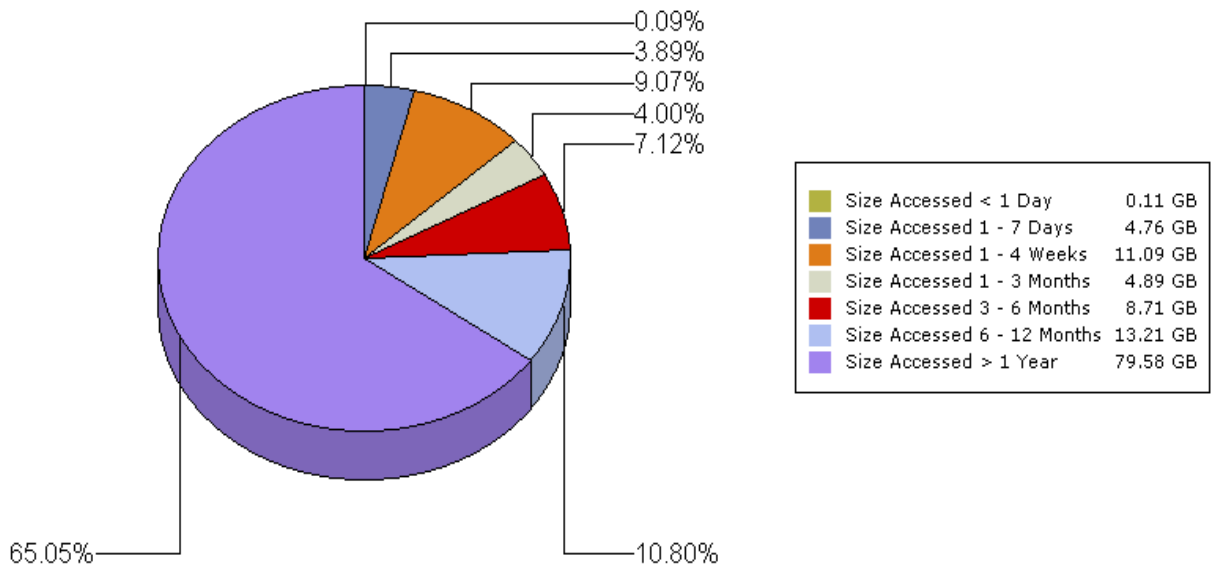


Figure 3: File age distribution

File creation date

- The following chart the age of files. This can be used for capacity planning as well as predicting growth for ROI purposes.
- % of files were created 3 years ago.
- % of files were created 2 years ago.
- % of files were created 1 year ago.

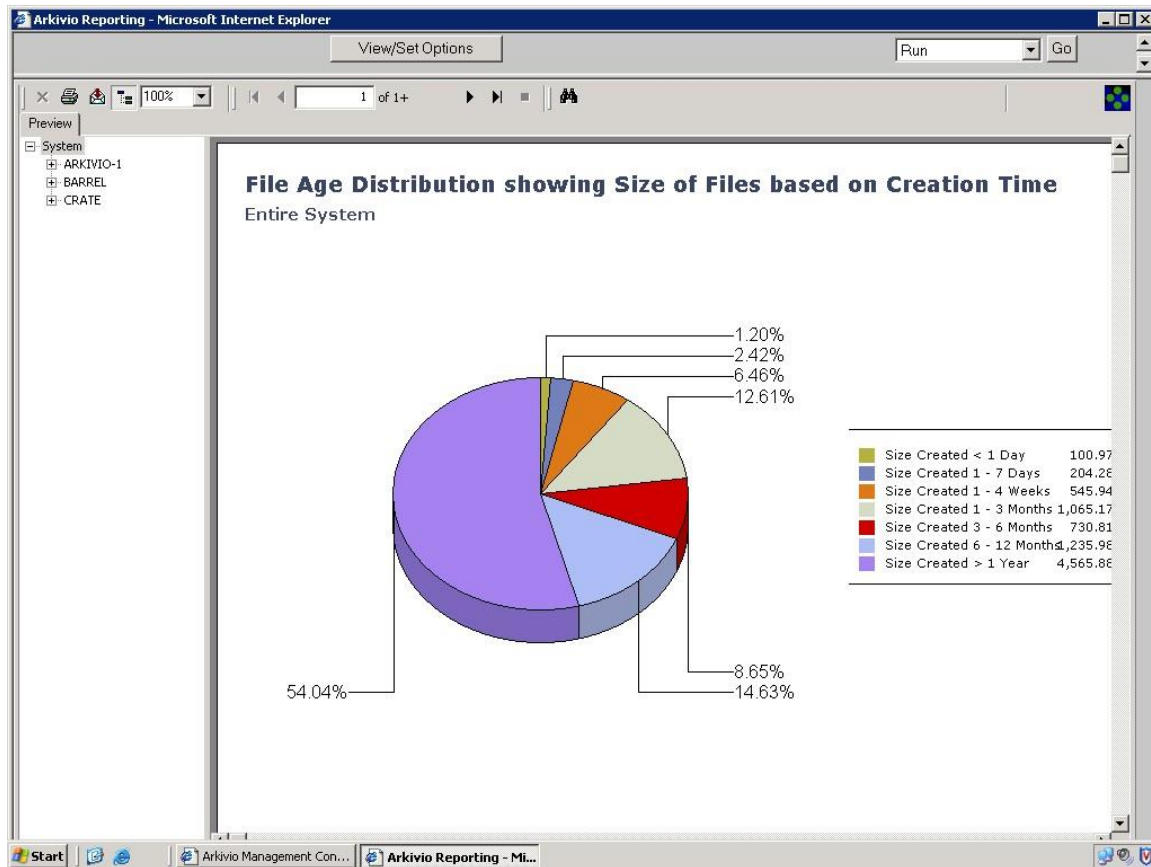


Figure 5: File creation date

Space consumption by file type

- The following chart shows the quantity of space consumed by various file types. This break down can be useful in ensuring the proper files are located on the appropriate storage. It can also be used for the identification and deletion of inappropriate files.
- 25.8% of the storage environment is comprised of CAD files.
- 24.7% of the storage environment is comprised of Office files.
- 24.7% of the storage environment is comprised of Excel files.

Space Consumption Grouped By File Type

Entire System

(File size shown in GB)

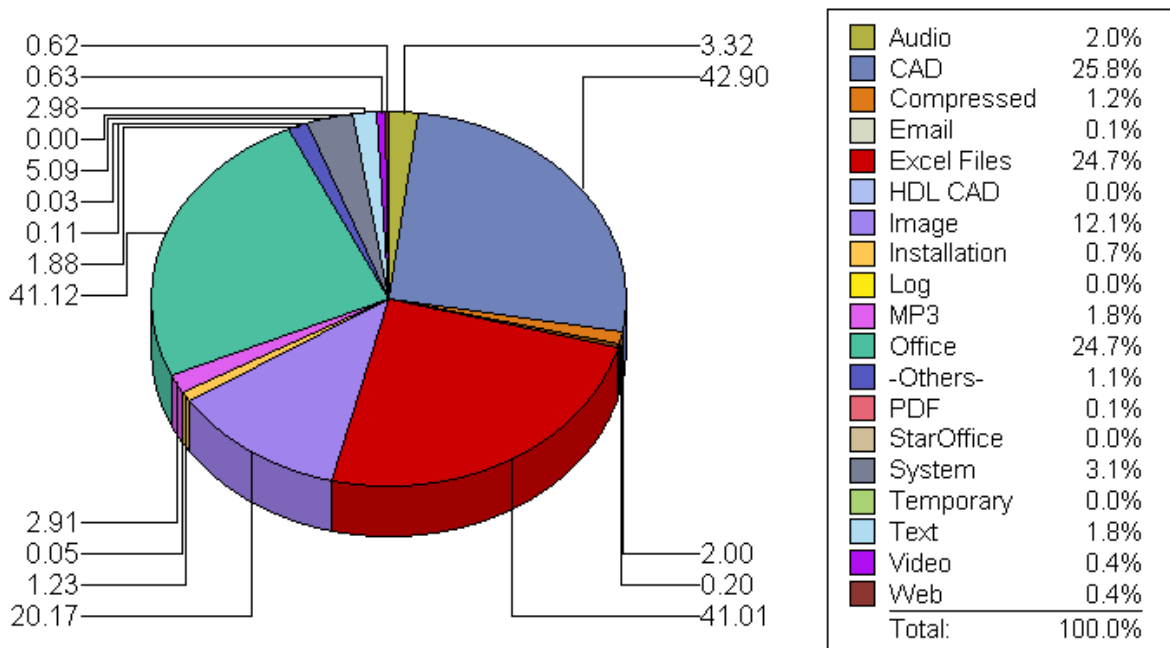


Figure 6: File type distribution



Duplicate files

- This chart shows the quantity of space being consumed by duplicate files.
- Approximately 122.7 GB of capacity is being consumed by duplicate files.
- Identification of this files can help administrators clean up wasted space.
- It is impossible to identify duplicate files without proper tools.
- Single instance storage can also be utilized to reclaim wasted space.

Total Wasted Space

122166937328

File Name	Last Modify	Time	duplicate files count	space wasted
mirage_4panel display(v3).eps	11/3/2004	20:59:13	5	1535593196
mirage_4panel display(v2).eps	11/16/2004	0:00:25	3	1353302318
Booth graphics v4.eps	12/6/2004	18:05:25	3	1264283086
Booth graphics v5.eps	12/6/2004	18:53:02	3	976284110
Booth graphics v6.eps	12/6/2004	19:00:35	3	976283694
Back_brochure.tif	5/31/2002	16:42:07	6	898315320
mirage_4panel display.eps	11/3/2004	20:48:52	3	829716038
front_brochure.tif	5/31/2002	19:34:52	5	721847408
new6th floor walk.avi	7/8/2004	16:08:19	2	636411392
Booth graphics3.psd	12/20/2004	17:55:25	3	628125160

Figure 7: Duplicate Files

Top file space consumption

- This chart shows the largest files consuming space in the environment. 138 GB of capacity is being consumed by a single backup file. This file could be sent to tape and deleted returning this capacity to the environment.
- In the Acme Corporation environment is if likely that 23.5% or 184.68GB of the environment can be recovered by simply deleting old backup files.
- The largest files may be nested 5 layer deep or even more in the file system. In file systems that contain hundreds of thousands or even millions of files, it is impossible for system administrators to locate and delete these sort of files without the proper tools to help them identify them. The attempt would consume large amounts of time and would still leave many files missed.

Top N Files by Size on each Volume

3DI-NETAPP

Volume: Data

Rank	Pathname	Size (MB)	Owner	Last Accessed Date
1	\\3di-netapp.3di.com\Data\IT\Exchange Backup\3di-mail2_3di-mail_backup.bkf	138,366.57		10/21/2005 9:07:57PM
2	\\3di-netapp.3di.com\Data\HOUA\ARCH\Airport_backup\Backup\Full Backup.bkf	41,469.12	chittenden	10/21/2005 1:33:04AM
3	\\3di-netapp.3di.com\Data\HOUA\ARCH\Airport_backup\Backup\Incremental Backups.bkf	2,696.19	chittenden	10/21/2005 1:38:07AM
4	\\3di-netapp.3di.com\Data\Q-MISC\PST Backups\fleishacker\archive backup	1,732.14	Fleishacker	11/18/2005 12:59:59PM
5	\\3di-netapp.3di.com\Data\Marketing\Individuals\Laura\grable\Laura's Data\Desktop\Personal\E-Mail\folders6.2.04.pst	1,517.08		10/22/2005 3:12:10AM
6	\\3di-netapp.3di.com\Data\Marketing\Individuals\Laura\grable\Laura's Data\Desktop\Personal\E-Mail\backup 3.10.2004.pst	1,406.64		10/22/2005 3:09:15AM
7	\\3di-netapp.3di.com\Data\Marketing\Individuals\Laura\grable\Laura's Data\Local Settings\Application Data\Microsoft\Outlook\outlook.ost	1,329.14		10/22/2005 3:17:58AM
8	\\3di-netapp.3di.com\Data\HOUR\Best Practices\graphics\models\backup\graphics\walkthru7-24.avi	1,177.77		10/21/2005 3:10:52PM

Figure 8: Top N Files by Size



ROI Analysis

The following data analysis was made from auditing a representative portion of Acme Corporation's storage environment. These numbers were then extrapolated to show an overall view of the entire network storage environment.

Current Approach

- Based on current growth trends, Acme Corporation will need to add XX GB of tier 1 disk over the next 3 years, at a cost of \$718,000 dollars.
- Acme Corporation will also need to add additional tape infrastructure at a three year cost of \$359,000, in order to backup and recover the additional data.
- Currently Acme Corporation is backing up XX GB of data in a weekly basis for on and offsite backups.
- Four copies of backups are kept resulting XX GB of data being written and saved on tape at any given time.

With ILM

- Acme Corporation would purchase the same amount of disk however this would be lower cost tier 2 disk at a cost of \$330,000.
- Acme Corporation would purchase Rocket Software's ILM software at a cost of \$100,000.
- Acme Corporation could migrate 45% of the data to second tier disk. A single backup of this data could be created. As the data would not change until the next archive cycle, no further backups would be required.
- Data could archived on a monthly or quarterly basis as required. This would reduce the amount of data on tape to XX GB.
- Acme Corporation would require far less tape infrastructure reducing the 3 year cost to \$103,500

Note that this ROI is based on the cost savings provided by freeing up space on already existing storage and includes MSRP pricing for typical second-tier storage devices. It does not include potential savings from backup/recovery optimization (including time and media costs), reduction in manual administration and reduction in potential downtime to

re-provision/reconfigure storage. The actual ROI may be much greater when these and other factors are considered.

	Year 1	Year 2	Year 3	3 Year Totals
Disk Cost (Current Approach)	\$180,000	\$234,000	\$304,200	\$718,200
Tape Cost (Current Approach)	\$90,000	\$117,000	\$152,100	\$359,100
Disk/software Cost (With ILM)	\$108,000	\$140,400	\$182,520	\$430,920
Tape Cost (With ILM)	\$0	\$45,000	\$58,500	\$103,500
Total Savings with ILM	\$87,000.00	\$150,000.00	\$195,000.00	\$432,000.00

Figure 9: Three year ROI worksheet

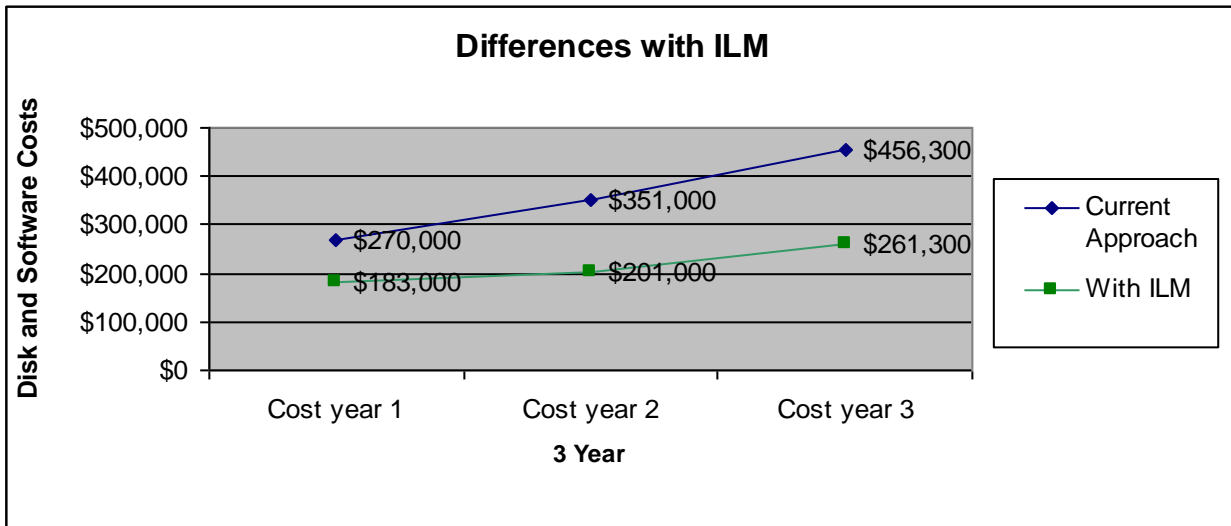


Figure 10: Three year savings with ILM

Other soft cost savings not included in ROI

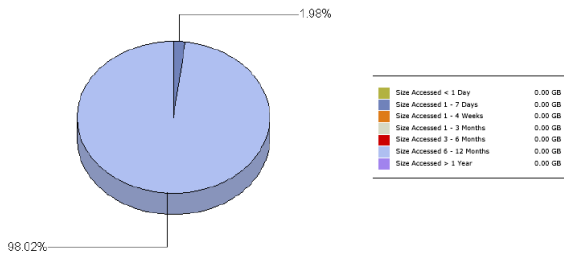
- Less time would be spent moving tapes.
- Automated movement of data should a file system start running out of space.
- Identification and automated deletion of inappropriate data.
- System administrators would be able to perform a single task on several servers at the same time rather than on a server by server basis.



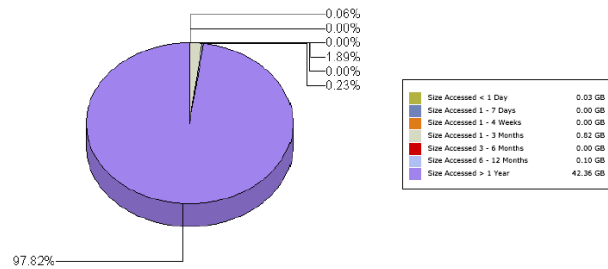
Supporting Information

Other useful information gather from Acme Corporation's environment. Add whatever other charts are meaningful to the customer.

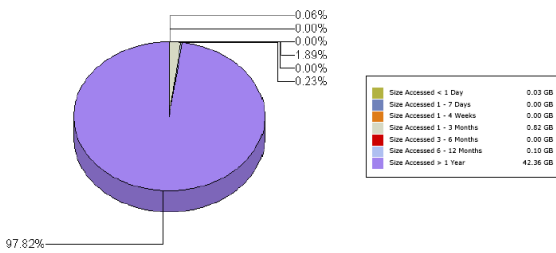
File Age Distribution showing Size of Files based on Last Access Time
BIGNAS Volume E



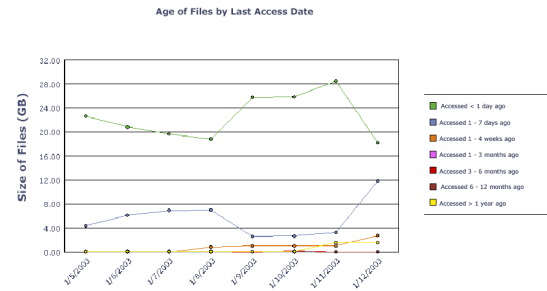
File Age Distribution showing Size of Files based on Last Access Time
FILESERVER



File Age Distribution showing Size of Files based on Last Access Time
FILESERVER



File Age Distribution Trends showing Size of Files based on Last Access Time
Entire System



Summary

- Acme Corporation has the opportunity to implement a comprehensive Information Lifecycle Management system that will reduce hard dollars expenditures on disk and tape by \$432,000 over a three year period.
- In addition to reducing costs, an ILM solution will position Acme Corporation to control future growth.
- Automated policies can be implemented to archive data when file systems are running out of space increasing systems availability thus improving service levels.
- Labor expenditures on tape movement can be reduced
- A true data archive can be implemented and policies can be created to delete expired data allowing data retention to be controlled.
- System administrator workload can be greatly reduced
- Inappropriate data can be rapidly identified and removed.
- Data can be viewed at a system, user, group, or enterprise level allowing for proper capacity planning.
- Disk consumption can be viewed at a user or group level allowing charge backs.
- ILM can assist in server and storage consolidation.

Additional Benefits

Fast payback – Rocket Arkivio Autostor improves the utilization of existing storage resources, thereby eliminating unused capacity and deferring incremental storage purchases.

Lower future storage acquisition costs – Rocket Arkivio Autostor allows an enterprise to maximize use of new low-cost storage devices, as well as prolong life of existing high-cost storage, which leads to a decrease in the average cost of future storage purchases.

Improve storage administration efficiency – Rocket Arkivio Autostor is very easy to use and automates many storage administration tasks lowering both labor and operating costs, typically the largest component of storage TCO.

Essential application of Online Archive and Tiered Storage (SATA)

Rocket Arkivio Autostor simplifies tiered storage management and automated data lifecycle management by utilizing a framework designed expressly for open systems environments. Freeing up primary storage capacity improves application performance and maintains service levels by allocating the correct type of storage to suit the requirements.

Essential application for Storage and Server Consolidation

Rocket Arkivio Autostor can quickly identify data and storage resources to help plan a consolidation effort. After identifying the data and planning the consolidation, Rocket Arkivio Autostor transparently and automatically migrates data away from the devices to be decommissioned to the appropriate new devices. Data can be placed directly into a multi-tiered environment, at which point policies can be configured to maintain the tiered infrastructure over time.

Essential application for Back-up Optimization

Rocket Arkivio Autostor quickly identifies and isolates the fixed content to optimize back-up and most importantly, restore functions. Not only will this lower cost and save time in the backup process, but the restoration process will now be optimized to only require restoration of the most business-critical data in the event of a loss.

Essential application for Compliance & Data Retention

Rocket Arkivio Autostor automates the management of regulated and fixed content data over its lifecycle to ensure regulatory compliance. This can be through regular snapshots of high-value data or through long-term archival of data with specific requirements. The grouping and classification mechanisms within Rocket Arkivio Autostor allow for the correct policies to be applied to the correct data.

RocketAutostor delivers a rapid return on investment that enterprises require in these difficult economic times. The remainder of this paper will discuss these financial benefits in more detail.

Additional Rocket Arkivio Autoexplore Reports

In addition to the analysis in the previous section, the full set of reports is provided upon which this analysis was based. The following is a description of these reports and how to best utilize them.

1. Capacity:
 - a. System-wide
 - b. By Volume / Server
2. File Age Distribution
3. By File Types
4. By Largest Consumers

There are 24 base reports, most of which have options to display data organized in specific, relevant fashion:

- Resource Usage:
 - Space Consumption
 - Volume Group Space Consumption
 - Space Consumption Trends
 - Volume Group Space Consumption Trends
 - Top N Space Consumption
 - Top N Capacity Consumption
- Data Usage
 - File Age Distribution
 - File Age Distribution Trends
 - Stale File Analysis
- Capacity Planning
 - Space Consumption Trends
 - File Age Distribution Trends
- Availability
 - Server Agent Availability Status
 - Server Agent Availability Status Trends
- Operations
 - Migration & Recall Trends
 - Operation Success/Failure Trends
- Policy Monitoring
 - Data Migration Distribution
 - Policy Results
 - Policy Results – Cost Savings
 - Copy Action Distribution
 - Move Action Distribution
 - Delete Action Distribution
 - Move Action History
 - Copy Action History
 - Delete Action History

These reports can be rerun with optional settings to make use of groupings of File Type Groups, File Groups and Volume Groups:

- Resource Usage:
 - Space Consumption / Trends
 - File Size
 - File Type, or
 - File Group
 - Volume Group Space Consumption / Trends
 - Total
 - File Type, or
 - File Group
 - Top N Space Consumption
 - Top N Capacity Consumption
 - Volume Groups
 - Servers, or
 - Volume/Vol. Group
- Data Usage
 - File Age Distribution (Access/Mod Time)
 - No Breakdown
 - File Group
 - File Type, or
 - User Group
 - File Age Distribution Trends (Access/Mod Time)
 - Filter for:
 - No Breakdown
 - File Group
 - File Type, or
 - User Group
 - Over Time Span of:
 - Last 7 Days
 - Last 30 Days
 - Last 90 Days
 - Last 180 Days
 - Last Year
 - Last 3 Years
 - Custom Date Range
- Capacity Planning
 - Stale File Analysis (Access / Mod. Time)
 - Number or File Size Over Server | Volume
 - Space Consumption Trends
 - File Age Distribution Trends
- Availability
 - Server Agent Availability Status
 - Server Agent Availability Status Trends
- Operations
 - Migration & Recall Trends
 - Operation Success/Failure Trends
- Policy Monitoring
 - Data Migration Distribution
 - Policy Results
 - Policy Results – Cost Savings
 - Copy Action Distribution
 - Move Action Distribution
 - Delete Action Distribution
 - Move Action History
 - Copy Action History
 - Delete Action History



About Rocket Software

Rocket Software was founded in 1990 with an initial focus on developing and delivering software solutions to IBM mainframe customers deploying IBM DB2. Within the first five years, Rocket licensed all of its intellectual property to IBM, and all of Rocket's products were rebranded and licensed to enterprises through IBM's worldwide sales channel.

Since those early years, Rocket has expanded its software OEM business to cover a wide spectrum of enterprise infrastructure products in the areas of Business Intelligence, Storage, Networks, Terminal Emulation, Integration, Security, and Database. And in addition to IBM, Rocket has licensed its products to a number of major OEMs that include Microsoft, RSA, EMC, Nortel, Motorola, and HP.

In parallel with a significant R&D investment in developing software products, Rocket Software began a series of strategic acquisitions in 2000. As a result, Rocket now has well over 100 software products that are licensed both through OEM partners and resellers, and via direct sales through Rocket's acquired branded sales channels, such as Aldon, Arkivio, BlueZone, CorVu, Mainstar, Servergraph, and Seagull.

Rocket values the technical skills and deep domain expertise of the employees at the companies we acquire. We endeavor to keep the teams intact and entrust them to manage and expand the product lines that they initially created. This philosophy has allowed Rocket to grow to more than 1000 employees around the world, including Eastern and Western Europe, Australia, Asia and North America.