Data Insight Self Paced Lab

Objective:

This lab is designed to introduce the high-level use cases within the Data Insight GUI. The Workspace provides an interactive view of the current environment. The reporting engine offers a variety of report packages designed to provide summary and detail data globally over various time periods.

There are no lab modules offered around the alerting section as this data required ongoing audit activity. This area offers event alerting and anomaly detection capabilities based on statistical deviance from calculated baselines.

About the environment:

VM Server Name
Matrix

Credentials
Username:
Administrator
Password:
password
Domain:
matrix

Once the desktop has loaded, the Data Insight console can be accessed through Internet Explorer, or Chrome. The credentials to log in are the administrator credentials used for VM login. To Log In, Click on the Data Insight Console icon on the desktop or open a browser and navigate to htts://localhost. If login is not working, ensure that cookies are enabled.

Lab Contents

Data Insight Use Case Test Scenarios

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

Data Ownership

Identify users, notify them, and clean up exposed data

Questions:

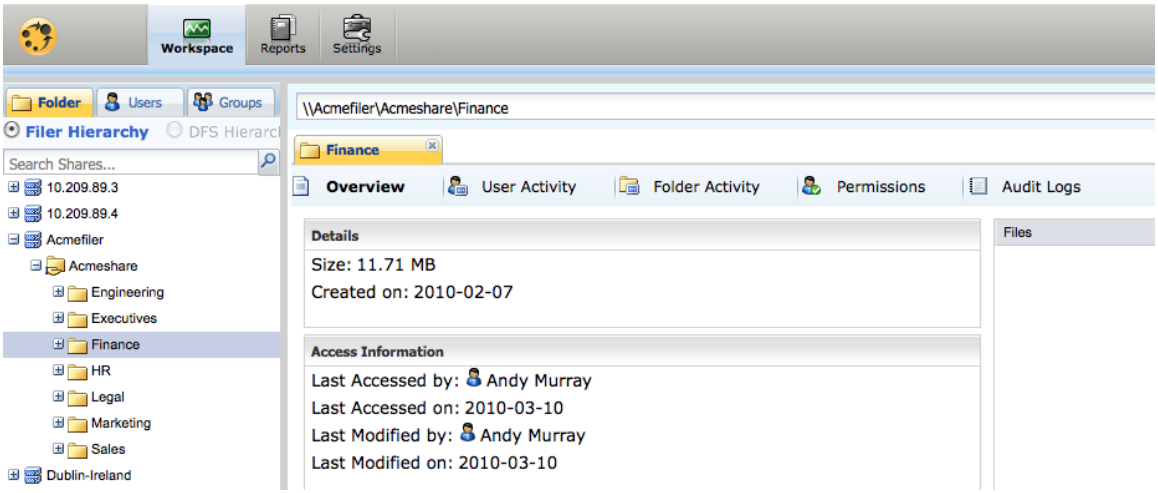
- Who is the inferred or calculated owner of a file or folder?
- Who else is using the same resources?
- What types of operations are being performed on this resource?
- How confident are we in the ownership calculation?
- Where are the most accesses occurring that impact this calculation?

The main idea of this use case is to be able to illustrate the person who is the most likely data owner based on usage and usage patterns.

Navigate to workspace / folders
Select a registered filer and drill down in the tree.
Pick a file or folder of interest. (Must have activity; pick something higher in the tree to increase odds of meaningful access history)

A good target folder for this lab is \\Singularity\DataNTAPCIFS. Click the + sign next to singularity, and the click on the DataNTAPCIFS folder. The overview should load in the main portion of the screen.

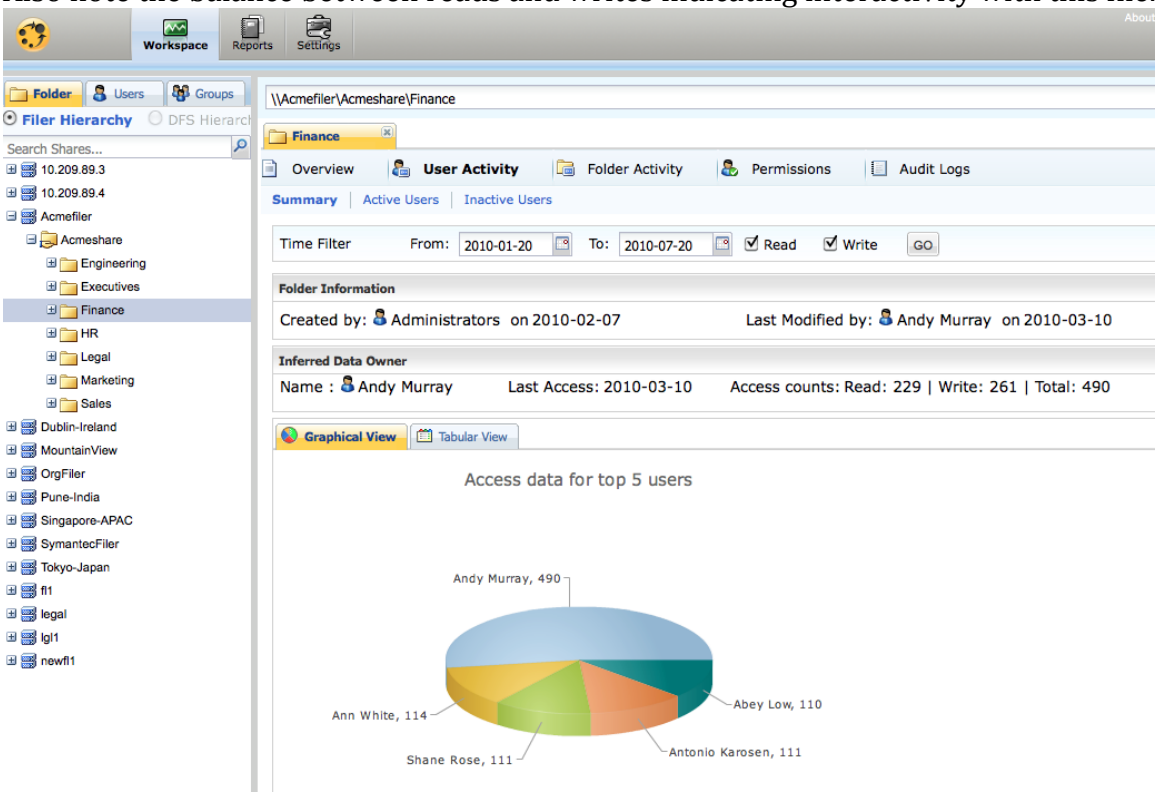
Note folder overview. This shows basic data on size, and access information. Also lists files in the parent folder



Next click on User Activity

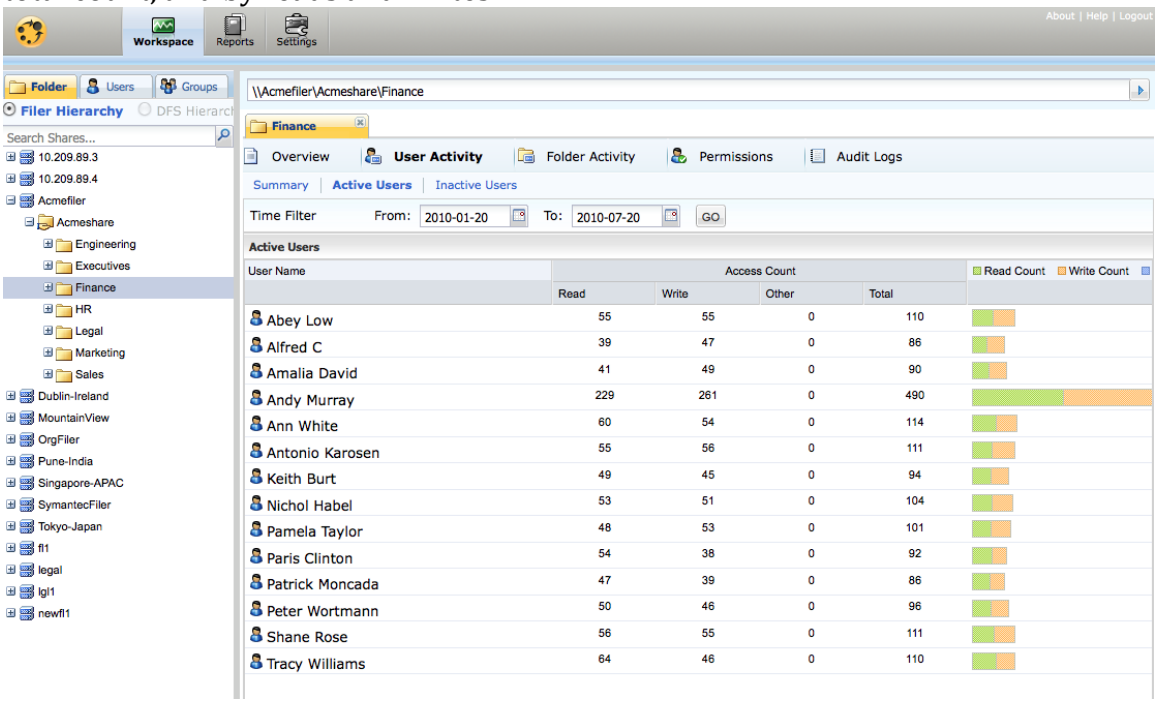
This screen gives you an activity overview with top users, and the inferred data owner. This screen also illustrates if there's a high confidence in the inferred data owner based on the large proportions of accesses by the user.

Also note the balance between reads and writes indicating interactivity with this file.



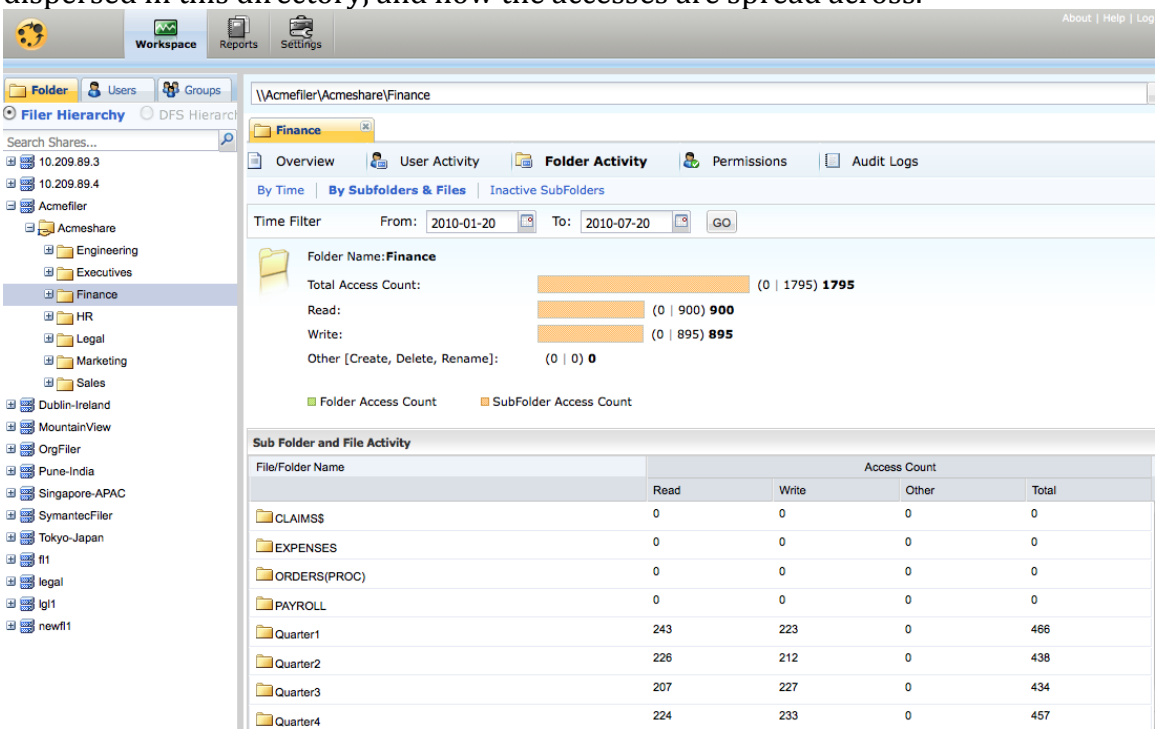
Next select Active Users sub tab.

This gives you a larger user list with a quick glance of access proportions in relation to the total count, and by reads and writes.

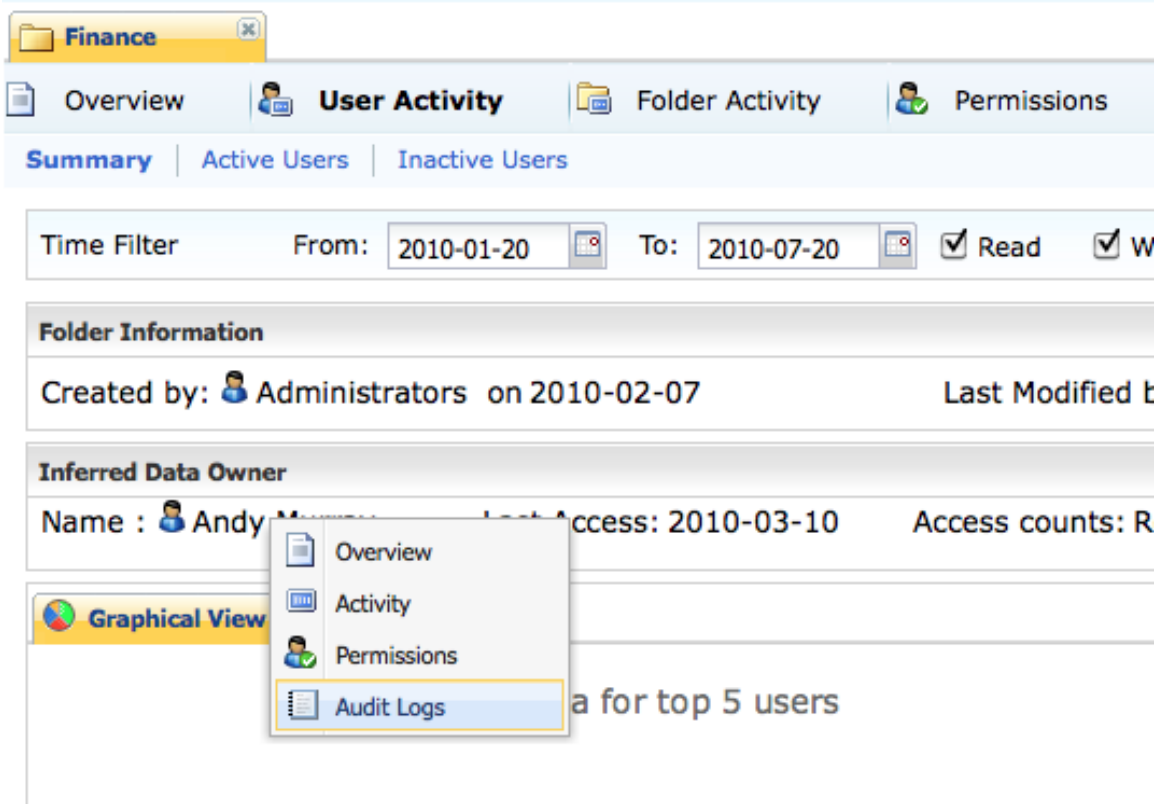


Next click on Folder Activity, then select by subfolders and files.

This screen gives you a breakdown by subfolder. You can get a sense of what data is dispersed in this directory, and how the accesses are spread across.

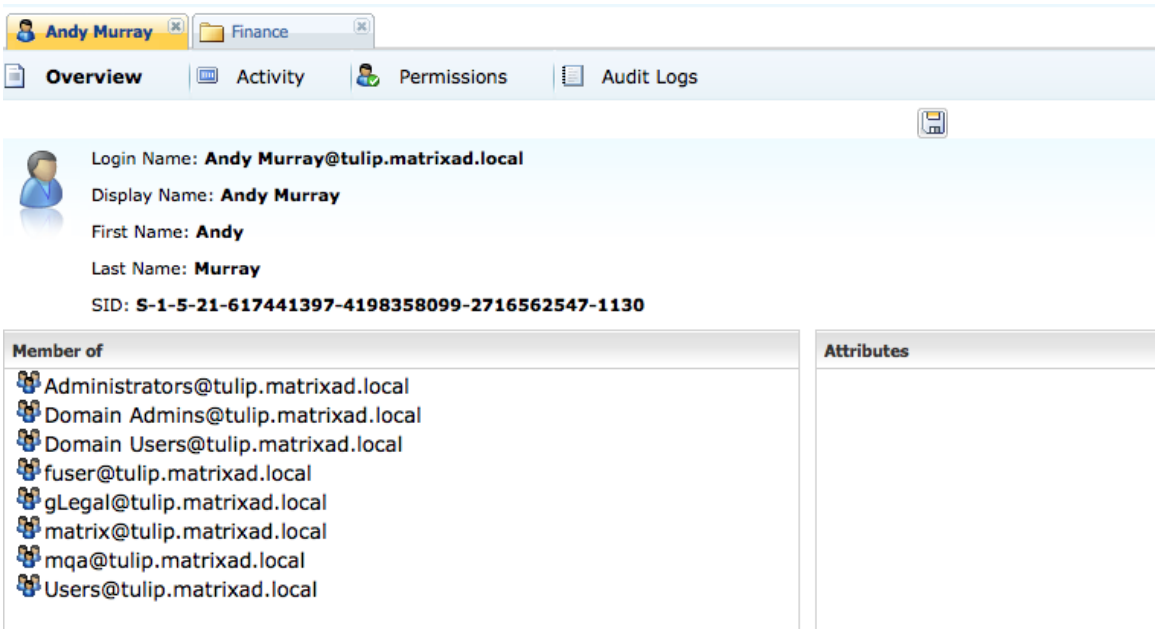


Now that we have gotten more data about the owner, we can go back to the user activity screen and right click on the Data Owner's name.
We are given the option of getting an overview, global activity, global permissions, and raw access history from the audit logs.

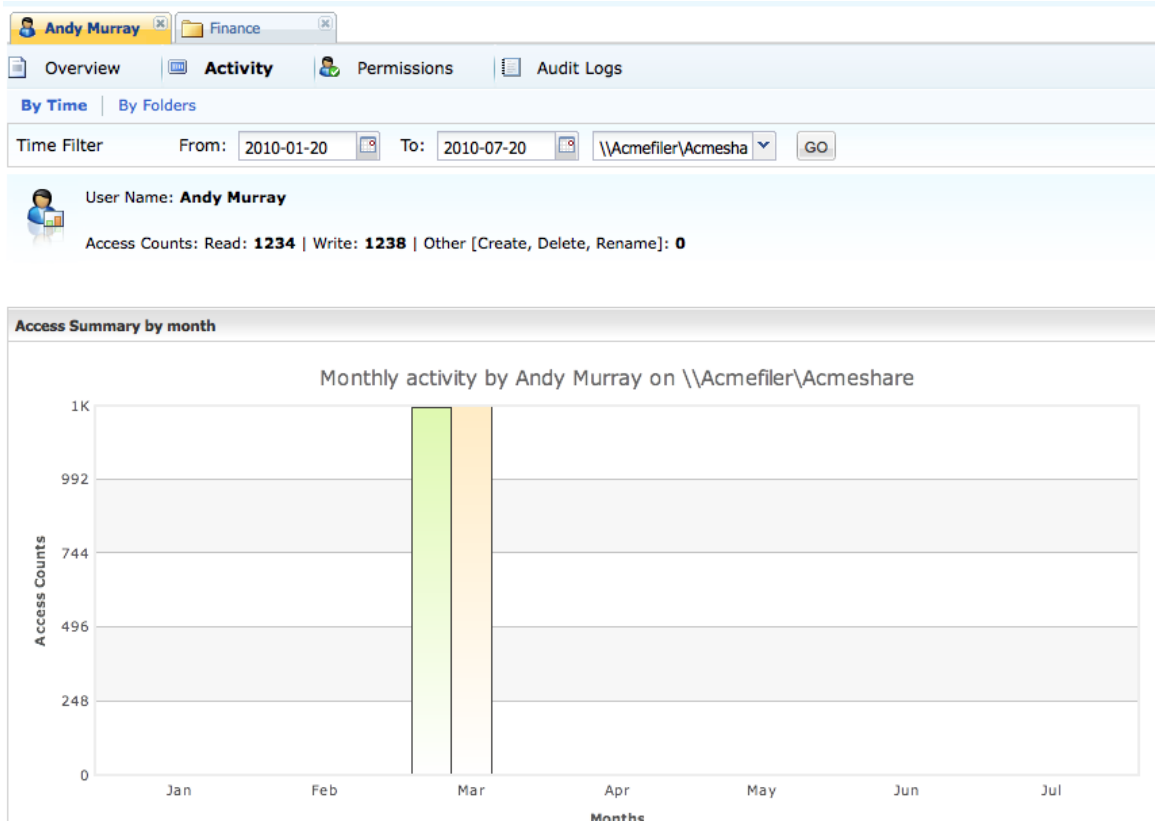


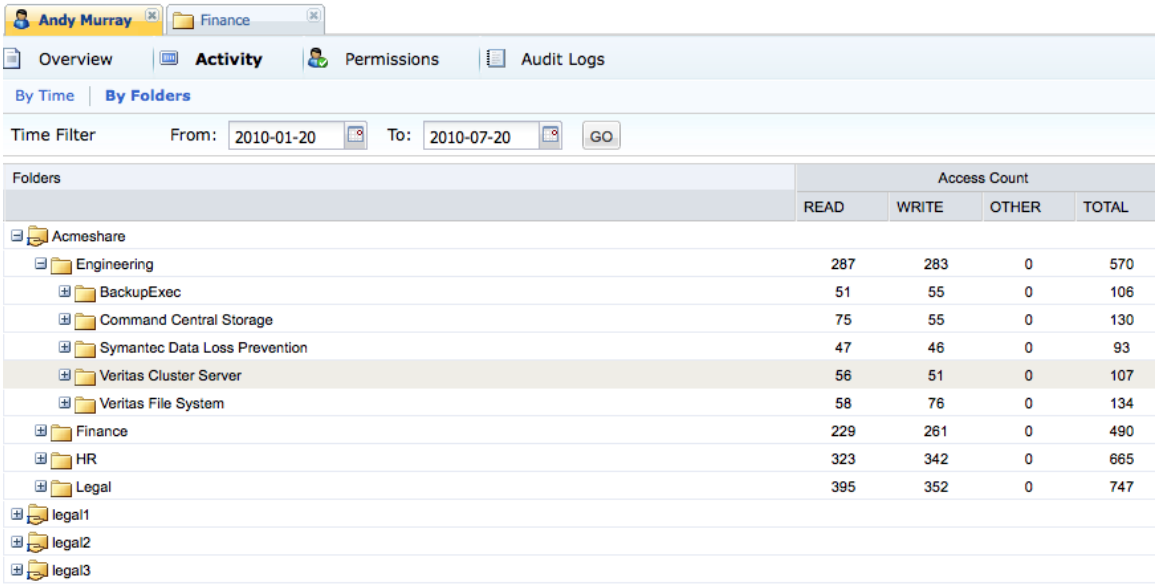
Clicking on overview will give you details on the Data Owner including security groups, and any custom AD attributes that have been configured to be gathered.

This info is useful in determining some details about any such as dept, job function, etc.



You can also click on the Activity tab to get a view of global activity; as a time series, or by folder.





Folders	Access Count			
	READ	WRITE	OTHER	TOTAL
Acmeshare				
Engineering	287	283	0	570
BackupExec	51	55	0	106
Command Central Storage	75	55	0	130
Symantec Data Loss Prevention	47	46	0	93
Veritas Cluster Server	56	51	0	107
Veritas File System	58	76	0	134
Finance	229	261	0	490
HR	323	342	0	665
Legal	395	352	0	747
legal1				
legal2				
legal3				

Stale Data Use Cases

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

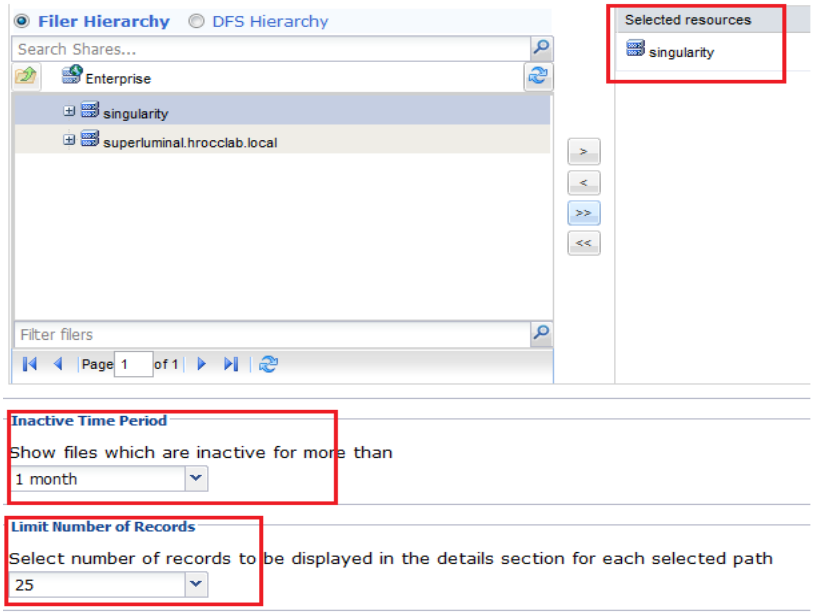
Stale Data
Evaluate age of data and understand potential areas for remediation. Improve storage utilization through file management.

Questions

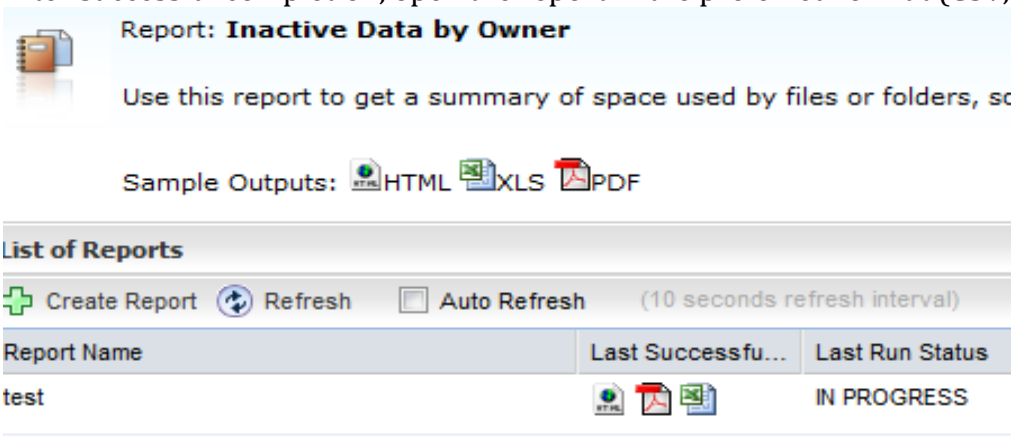
- How old are all the files or folders owned by a certain user or group?
- How much can be saved with stale file remediation techniques?

The main idea of this use case is to try to get an overview of the age of your data, to identify stale files and folders for remediation.

Navigate to reports / Inactivity Reports
Select "Inactive Data by Owner" report
Choose Create Report
Here you can choose the output of the report, when it's run, and what areas of the NAS environment it looks at. To ensure data in the report, select an entire NAS device and set the "Inactive Time Period" to "1 Month" and
"Limit Number of Records" to 25



Select "Save and Run" to see the new report listed.
After successful completion, open the report in the preferred format (CSV, HTML, PDF)



The top area of the report calls out all owners of files considered stale based on your report settings. The number of files as well as what % of all files owned by the user are considered stale.

User	User Account	Size (MB)	% of Size	File Count	BU Name	Custodian
Administrators	Administrators	9273.31	7.87	2393		
Total		**9273.31**		**2393**		

The remainder of the report calls out individual files, locations, and dates of modification and access.

User	Data Owner Policy	File Server	Share Name	File (Path)	Size (MB)	Last Modified	Last Accessed
Administrators (Administrators)							
	Parent Owner	singularity	AppleBackup	\\singularity\AppleBackup\Data Insight Use Case Test Scenarios.docx	1.90	Oct 25, 2010 12:54 AM	Nov 12, 2010 6:13 PM
	Parent Owner	singularity	DataNTAPCIFS	\\singularity\DataNTAPCIFS\vms\DataInsightx64\DataInsightx64.vmem	1024.00	Oct 26, 2010 2:23 PM	Nov 12, 2010 12:08 PM
	Parent Owner	singularity	DataNTAPCIFS	\\singularity\DataNTAPCIFS\vms\DataInsightx64\DataInsightx64.vmss	19.71	Oct 26, 2010 2:31 PM	Nov 12, 2010 12:42 PM
	Parent Owner	singularity	SoftwareNTAPCIFS	\\singularity\SoftwareNTAPCIFS\Windows ISOs\SW_DVD5_Windows_Svr_DC_EE_SE_Web_2008R -bit_English_X15-59754.ISO	2857.97	Feb 1, 2010 11:27 PM	Apr 11, 2010 2:13 AM

Chargeback Enablement Use Cases

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

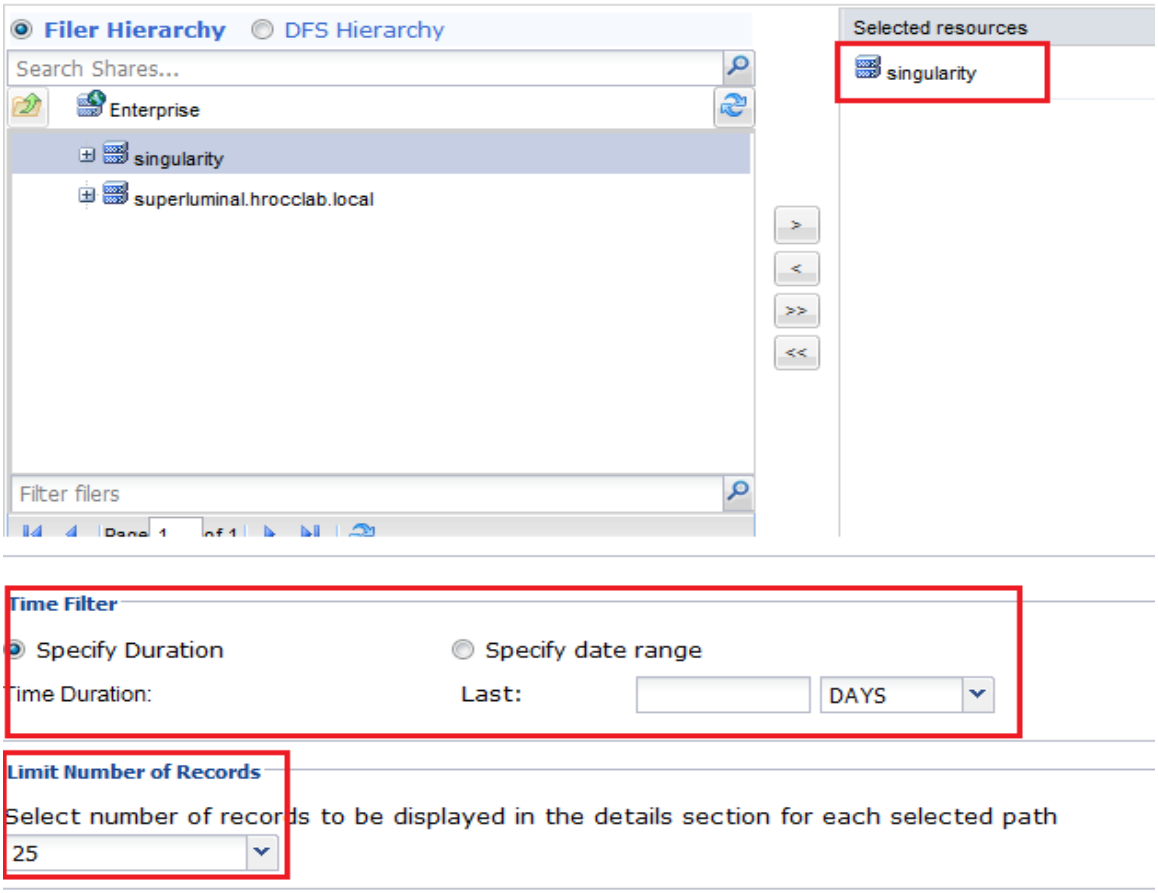
Enable Chargeback
Evaluated size of files owned by particular users and groups. Improve storage utilization through accountability

Questions

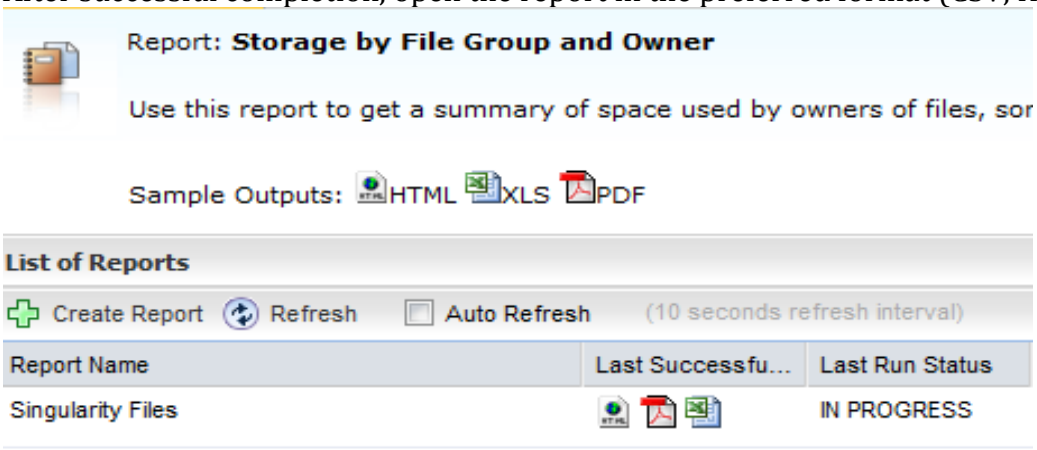
- What users and groups are using the most storage for unstructured data?
- How much could each user or group pay for access to that storage?

The main idea of this use case is to try to get an overview of the sizes of your unstructured data files, and assign ownership for accountability and chargeback

Navigate to reports / Utilization Reports
Select "Storage by file group and owner" report
Choose Create Report
Here you can choose the output of the report, when it's run, and what areas of the NAS environment it looks at.
To ensure data in the report, select an entire NAS device and
"Specify duration" to the length of time DataInsight has been running.
"Limit Number of Records" to 25



Select "Save and Run" to see the new report listed.
After successful completion, open the report in the preferred format (CSV, HTML, PDF)



The report can be broken down by user, filetype, and any Active Directory attributes like Business Unit:

User	User Account	File Group Type	Space Used (MB)	File Count	BU Name	Custodian
John Dodds	john_dodds@hrocclab.local	All Files	108488.35	937	SAMG	
Administrators	Administrators	All Files	7734.18	2300		
Ryan Jancaitis	ryan_jancaitis@hrocclab.local	All Files	0.43	3	SAMG	
Creator Owner	Creator Owner	All Files	0.01	1		
John D'Alto	john_dalto@hrocclab.local	All Files	0.01	3	SAMG	
Tom Harwood	tom_harwood@hrocclab.local	All Files	0.01	2	SAMG	

File usage by user and type of file:

User	Type	Data Owner Policy	File Server	File (Path)	Size (MB)	Access Count
John Dodds (john_dodds@hrocclab.local)						
	VMWare Server Files	Read+Write Count	singularity	\\singularity\DataNTAPCIFS\vms\Enforcedemo_V11\Enforcedemo_V11-Snapshot1.vmem	3096.00	28
	VMWare Server Files	Read+Write Count	singularity	\\singularity\DataNTAPCIFS\vms\Enforcedemo_V11\Enforcedemo_V11-Snapshot2.vmsn	135.78	2
	VMWare Server Files	Read+Write Count	singularity	\\singularity\DataNTAPCIFS\vms\Enforcedemo_V11\Enforcedemo_V11-Snapshot1.vmsn	134.35	2
	VMWare Server Files	Read+Write Count	singularity	\\singularity\DataNTAPCIFS\vms\Enforcedemo_V11\Enforcedemo_V11.vmx	0.01	1
	VMWare Server Files	Read+Write Count	singularity	\\singularity\DataNTAPCIFS\vms\Netapp_Filer (v11)\Netapp_Filer (v11).vmx	0.01	3

Exporting and manipulating this data through spreadsheets allows complete flexibility in chargeback reports.

ILM Use Cases

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

ILM: Data Aging
Evaluated age of files owned by particular users and groups. Improve storage utilization through accountability

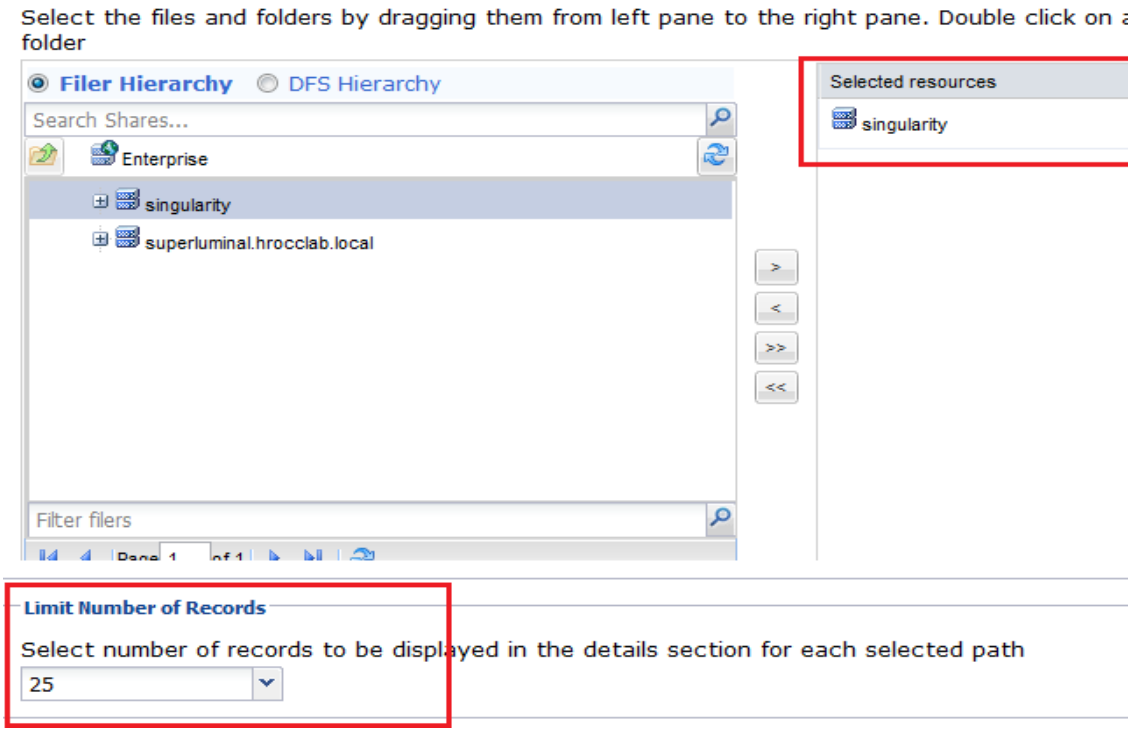
Questions

- What users and groups are using the most storage for unstructured data?
- Are the proper users and groups keeping the right files for extended periods of time?

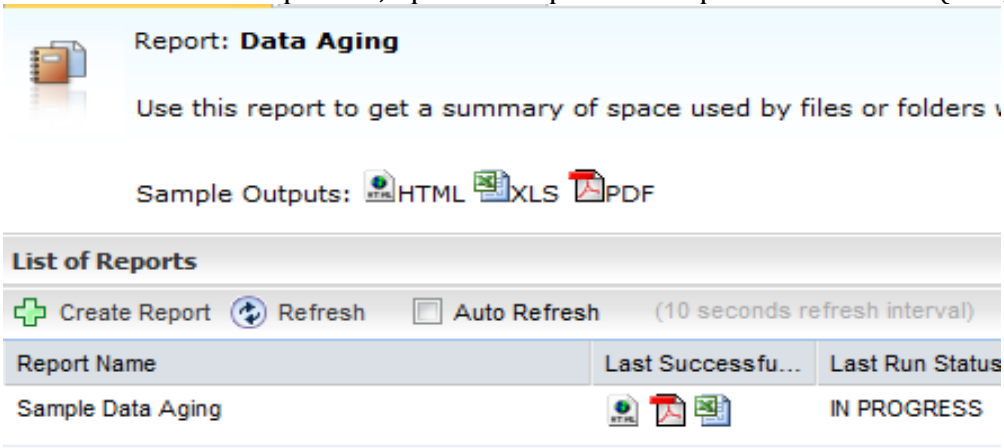
The main idea of this use case is to try to get an overview of the age of your unstructured data files, and assign ownership for remediation.

Navigate to reports / Inactivity Reports

Select "Data Aging" report
Choose Create Report
Here you can choose the output of the report, when it's run, and what areas of the NAS environment it looks at.
To ensure data in the report, select an entire NAS device and
"Limit Number of Records" to 25

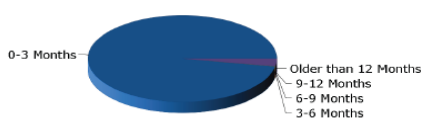


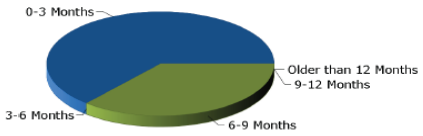
Select "Save and Run" to see the new report listed.
After successful completion, open the report in the preferred format (CSV, HTML, PDF)



This report is broken down into aging "buckets" of last access to quickly identify count and size of all files.

Data (Count) Aging by last access date	Data (Size) Aging by last access date





Age	File Count	Percent
0-3 Months	2472	97.13
3-6 Months	0	0.00
6-9 Months	2	0.08
9-12 Months	0	0.00
Older than 12 Months	71	2.79

Age	Size (MB)	Percent
0-3 Months	5252.80	63.73
3-6 Months	0.00	0.00
6-9 Months	2989.17	36.27
9-12 Months	0.00	0.00
Older than 12 Months	0.01	0.00

This is broken down further into NAS device, share, and eventually file.

Overview of aging per NAS device.

Age	File Server	Size (MB)	File Count
0-3 Months			
	singularity	3714.10	2327
	superluminal.hrocclab.local	1538.70	145
6-9 Months			
	singularity	2989.17	2
Older than 12 Months			
	singularity	0.01	17
	superluminal.hrocclab.local	0.00	54

Overview of aging per NAS Share and device

Age	File Server	Share Name	Size (MB)	File Count
0-3 Months				
	singularity	SoftwareNTAPCIFS	3710.71	2134
	singularity	DataNTAPCIFS	3.39	193
	superluminal.hrocclab.local	SuperluminalShare	1538.70	145
6-9 Months				
	singularity	SoftwareNTAPCIFS	2989.17	2
Older than 12 Months				
	singularity	HOME	0.01	1
	singularity	DataNTAPCIFS	0.00	16
	superluminal.hrocclab.local	SuperluminalShare	0.00	54

Overview of aging per share and file

Age	Filer Name	Share Name	File (Path)	Space Used (MB)	Last Accessed
Older than 12 Months					
	singularity	DataNTAPCIFS	\\singularity\DataNTAPCIFS\DILoad\ORDERS(PROC)\2007\Feb\order8_20.txt	0.00	Jan 1, 1970 12:00 AM
	singularity	DataNTAPCIFS	\\singularity\DataNTAPCIFS\DILoad\CLAIMS$\2008\ClaimsJan2_17.txt	0.00	Jan 1, 1970 12:00 AM
	singularity	DataNTAPCIFS	\\singularity\DataNTAPCIFS\DILoad\CLAIMS$\2008\ClaimsJan2_18.txt	0.00	Jan 1, 1970 12:00 AM

Data Forensics Use Cases

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

Data Forensics
Investigate and solve data breaches; monitor sensitive data usage

Questions

- Who as been touching sensitive files?
- How did this person get access to a file?
- Who wrote to a file and potential was responsible for the inclusion of sensitive information?
- What has a person been up to lately?
- What time an ip has a person of interest been using to get to a file?

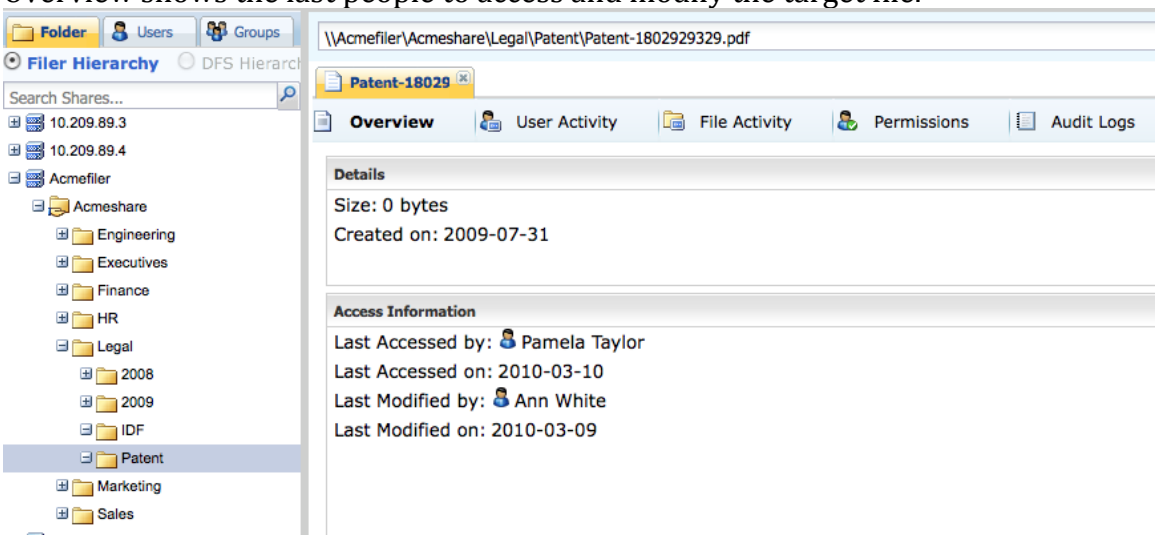
The main idea of this use case is to try to get a root cause for the introduction of sensitive data, or the scope of a potential data leak.

Navigate to workspace / folders

Select a registered filer and drill down in the tree.
Pick a file or folder of interest. (Must have activity; pick something higher in the tree to increase odds of meaningful access history)

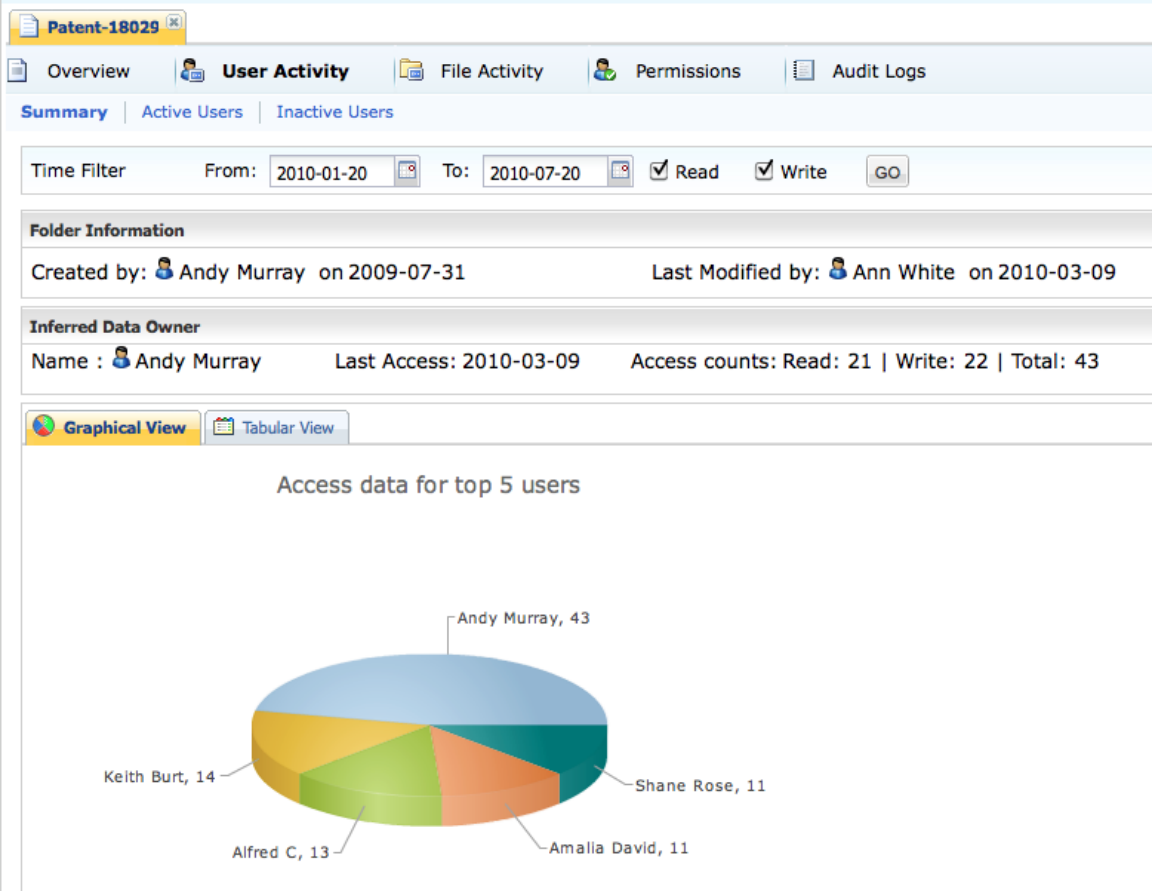
You can continue to use the same \\Singularity\DataNTAPCIFS share that was used for the previous Data Ownership lab module.

Overview shows the last people to access and modify the target file.

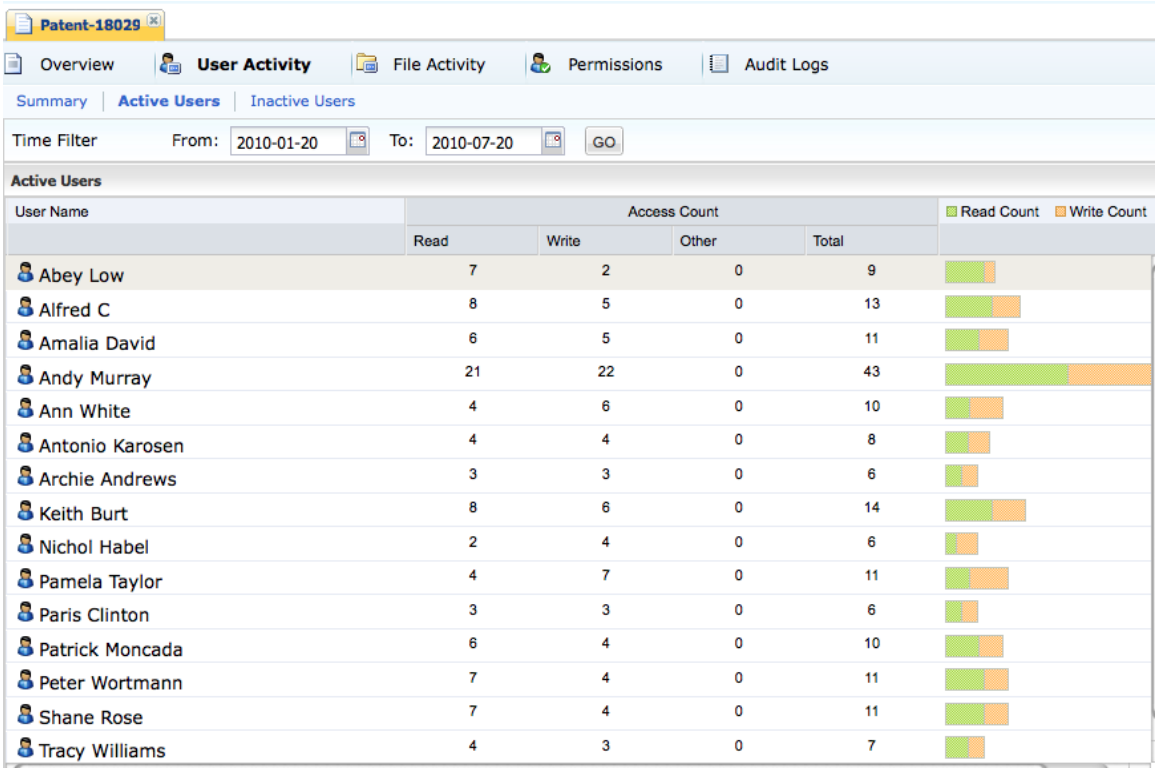


Click on user activity to see the top users of this file as well as the inferred data owner. The inferred data owner would be the primary contact for remediation, or the best person to notify if there's a suspected issue, or incident created on this file.

You can use this screen to also potentially identify any users that have a suspiciously large amount of file accesses, or people that possibly should be using this resource at a high level.

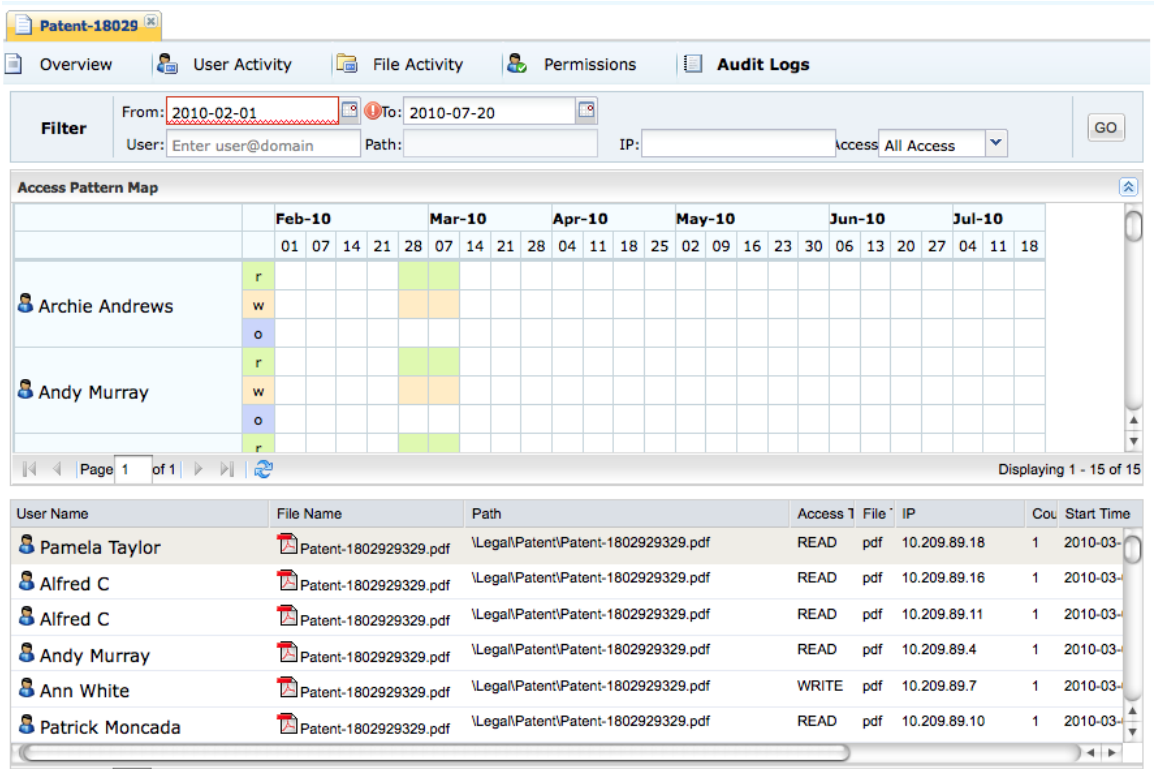


Click on active users to get more detail on usage, as well as an overview of the file's most active writers.

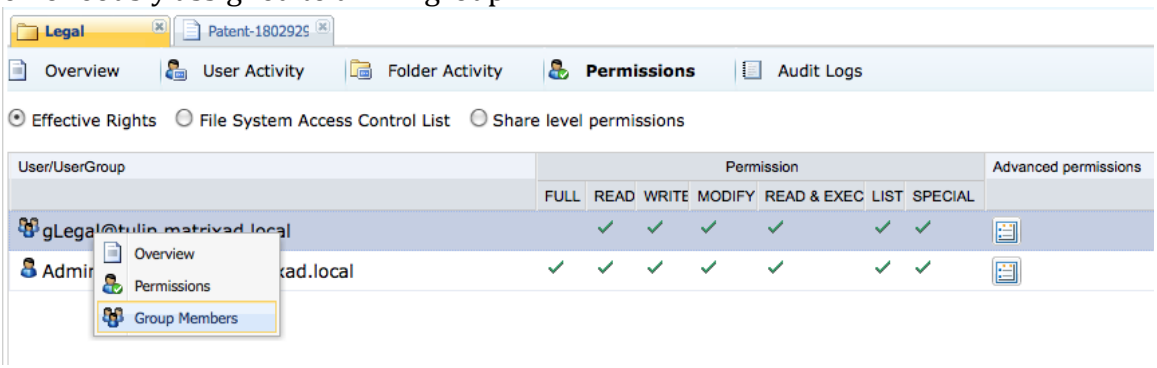


At this point, we have a good idea of people that have been using this file; we want to know who's touched it recently. If the data breach was in the last 24 hours we can modify the Time filer of the above screen to see who's written or read this file.

Next we are interested in the raw activities that have happened on this file. We may have a person or time of interest, click on audit lots to see the raw access history and choose the appropriate time frame. We can also filter this list by user, path, ip, or type of access. If I wanted to see all writes in a certain time window by a user, this is possible with the filters. We are presented with a time series access heat map as well as the raw events that can pinpoint the suspicious activity.



You can select permissions and evaluate the raw permissions at the share and file level to determine if the ACL is valid. You can also start here to see where the suspicious user got access to this file. You can right click to see memberships.
It's possible for an invalid explicit ACE to be applied to a file or folder, or for a user to be erroneously assigned to an AD group.



Data Protection Use Cases

Listing of main Data Insight Use Cases, and methods to validate data and show functionality.

Data Protection

Evaluate access and understand potential access breaches and evaluate ACL changes. Improve security though tighter security.

Questions

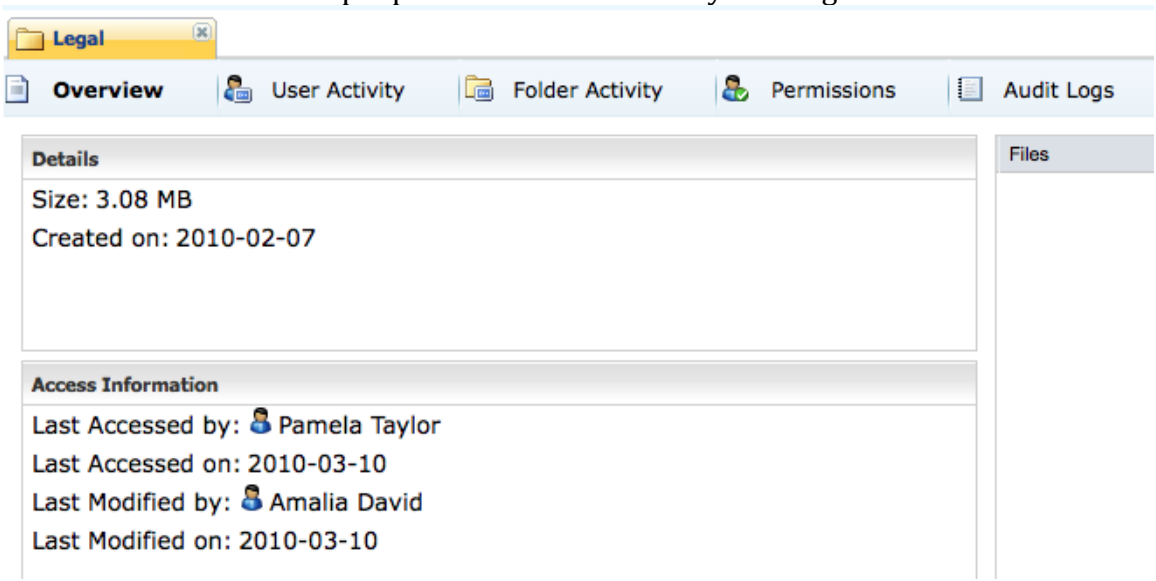
- How did a person or group inherit access to a folder or file
- At what point does inheritance break and the ACL tree change?
- How many have access, but may not need it based on inactivity?

The main idea of this use case is to try to get a root cause for the introduction of sensitive data, or the scope of a potential data leak.

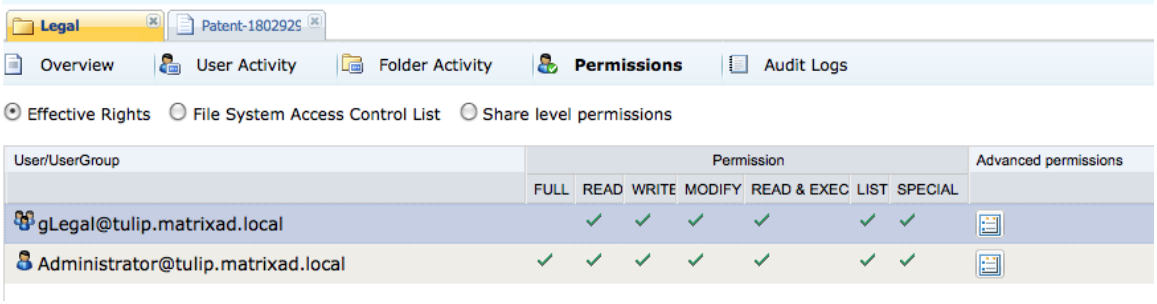
Navigate to workspace / folders
Select a registered filer and drill down in the tree.
Pick a file or folder of interest. (Must have activity; pick something higher in the tree to increase odds of meaningful access history)

You can continue to use the same \\Singularity\DataNTAPCIFS share that was used for the previous lab module.

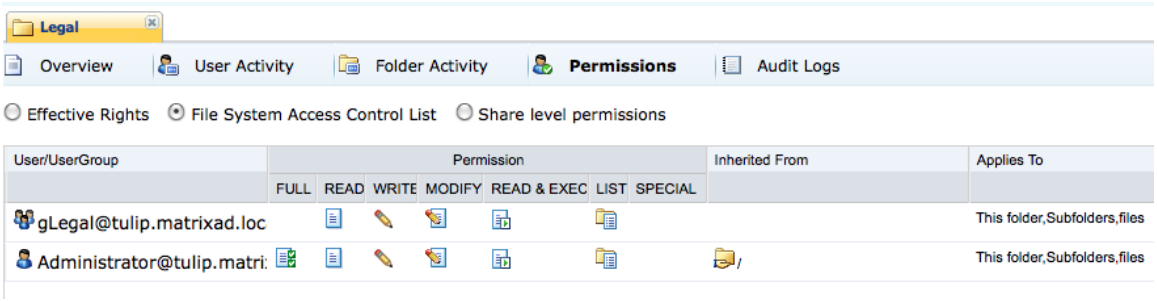
Overview shows the last people to access and modify the target file.



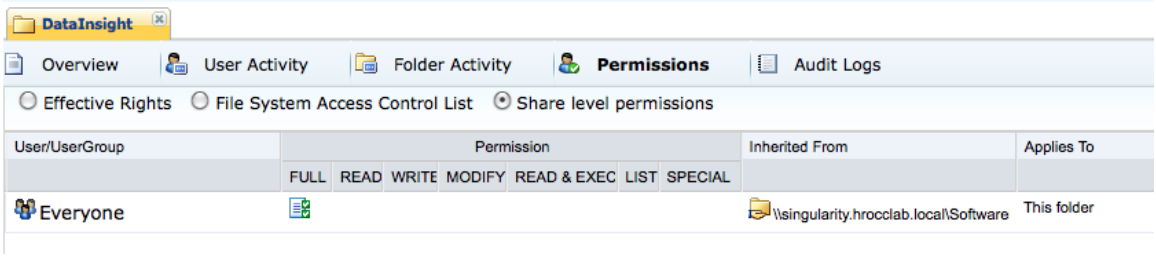
Next Click on Permissions. This gives us an effective permissions view that includes folder and share level ACL calculations.



For a raw ACE list (ntfs style permissions) with inheritance, click on radio button next to File System Access Control List. Note the inheritance data. No inheritance means a direct ACE on that Object. If you are using a share such as \\Singularity\DataNTAPCIFS, there should be no inheritance as this is the root of the share. To see inheritance, navigate to a sub folder such as DILoad.

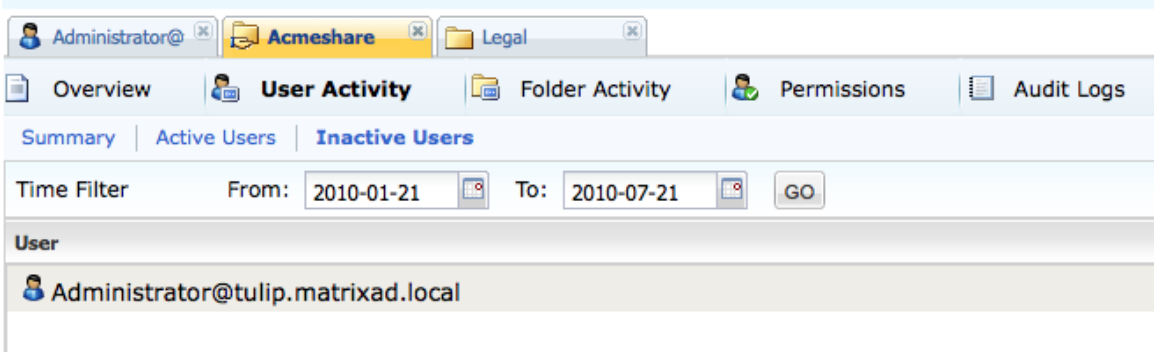


Next evaluate the permissions granted at the share level. Effective access is calculated by evaluating NTFS style ACL's and masking it with the restrictions at the share point.

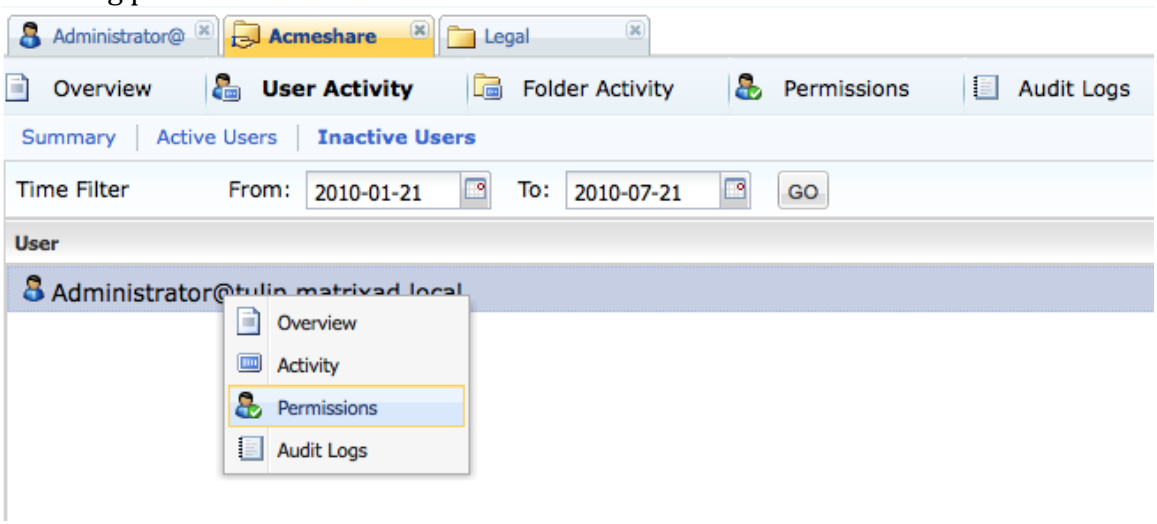


Next we can attempt to make recommendations on tightening up the access list by seeing who has not been using their access to this folder. Click on User Activity, and then select inactive users.

In a real world environment, there is likely to be many users that aren't using the permissions granted to a share. Larger numbers of inactive users may indicate an overly permissive share.



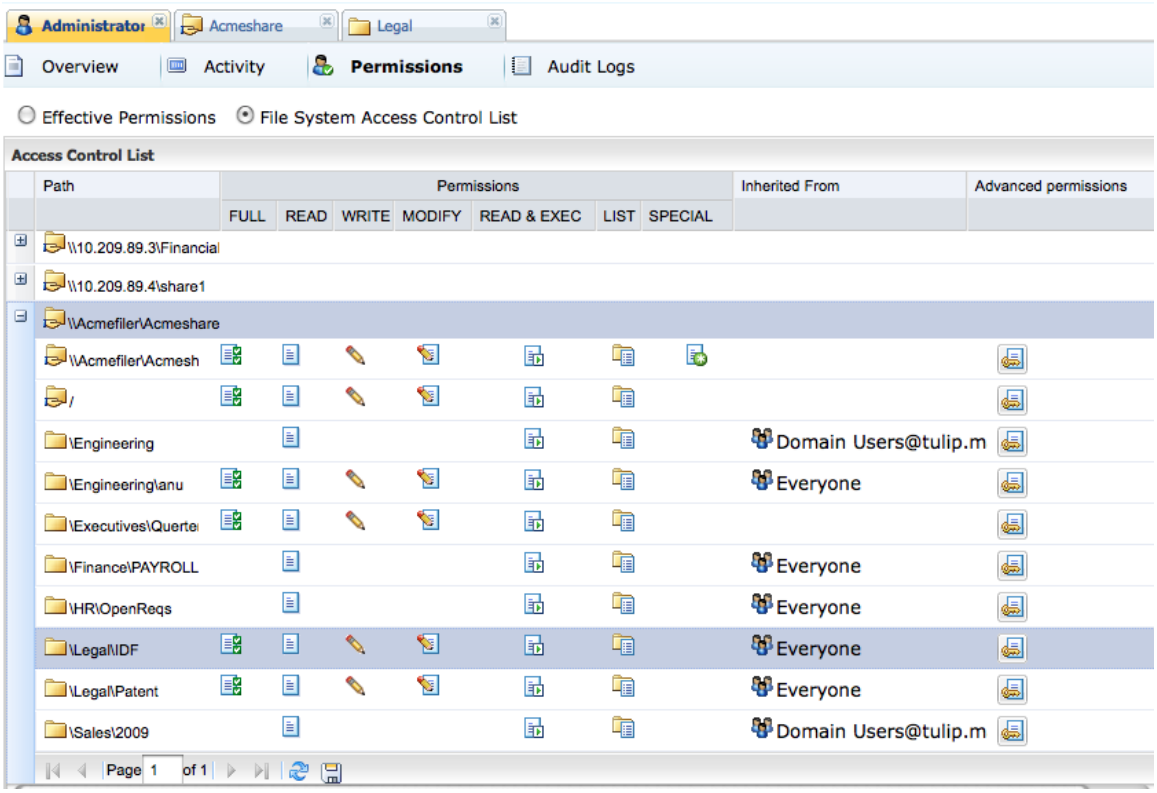
We can then determine how a user got access to this folder by right clicking on the user and selecting permissions.



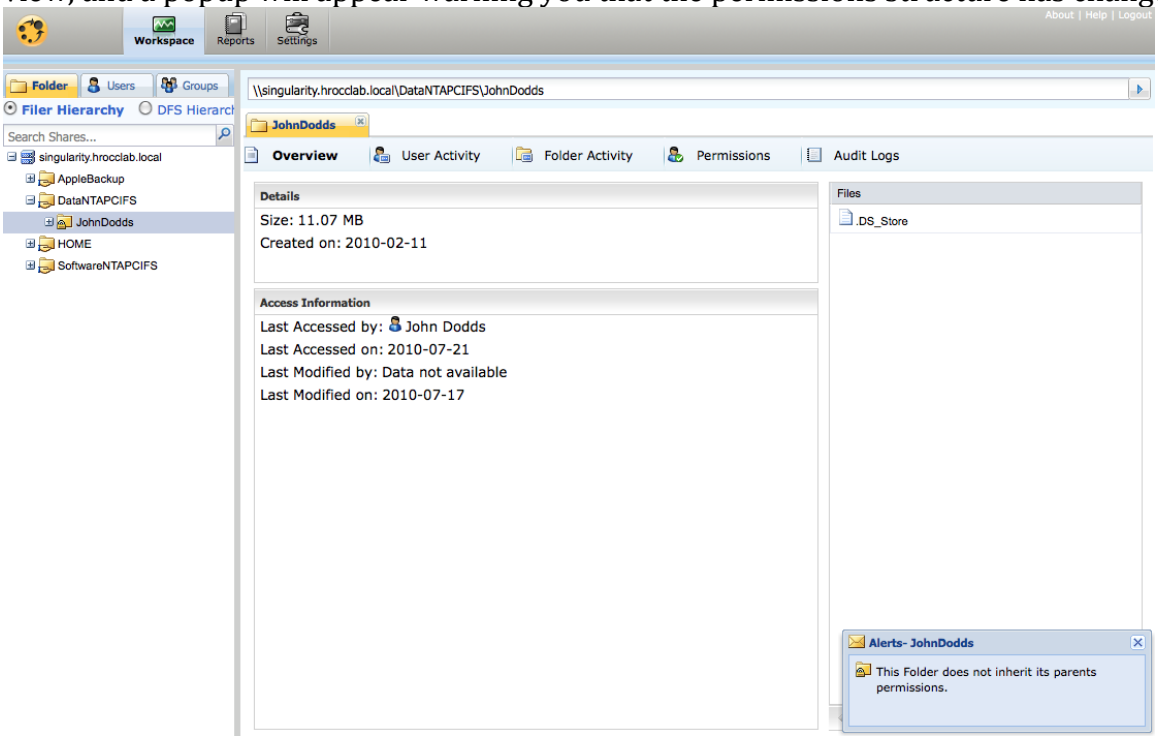
Once on the user's global permissions screen, we can scroll to the file and share in question. This view will show direct ACLs as well as permissions that have been inherited by virtue of group membership.

Instances such as the Everyone group can be a potentially useful red flag.

It's also possible that the permissions are valid, but there are people that are erroneously in groups that have valid access. This would pinpoint the error in membership.



If there is a break in ACE inheritance, you will see a lock next to the folder icon in the tree view, and a popup will appear warning you that the permissions structure has changed.



Identifying unique control points can also be a useful tool in evaluating permissions, or finding locations where high-level users have broken standards and change permissions.