Plan using File Analytics

XCP

NetApp
May 25, 2022

Table of Contents

Plan using File Analytics

Plan your data migration using File Analytics.

XCP is a CLI, whereas File Analytics has a GUI.

Overview

XCP File Analytics uses the XCP scan API to collect data from NFS or SMB hosts. This data is then displayed on XCP File Analytics GUI. There are three main components involved in XCP File Analytics:

- XCP service
- File Analytics database
- File Analytics GUI to manage and view data

The deployment method for XCP File Analytics components depends on the solution required:

- Deploying XCP File Analytics solutions for NFS file systems:
  - You can deploy the File Analytics GUI, database, and XCP service in the same Linux host.
- Deploying XCP File Analytics solutions for SMB file systems:
  You must deploy the File Analytics GUI and database in a Linux host and deploy the XCP service on a Windows host.

Access File Analytics

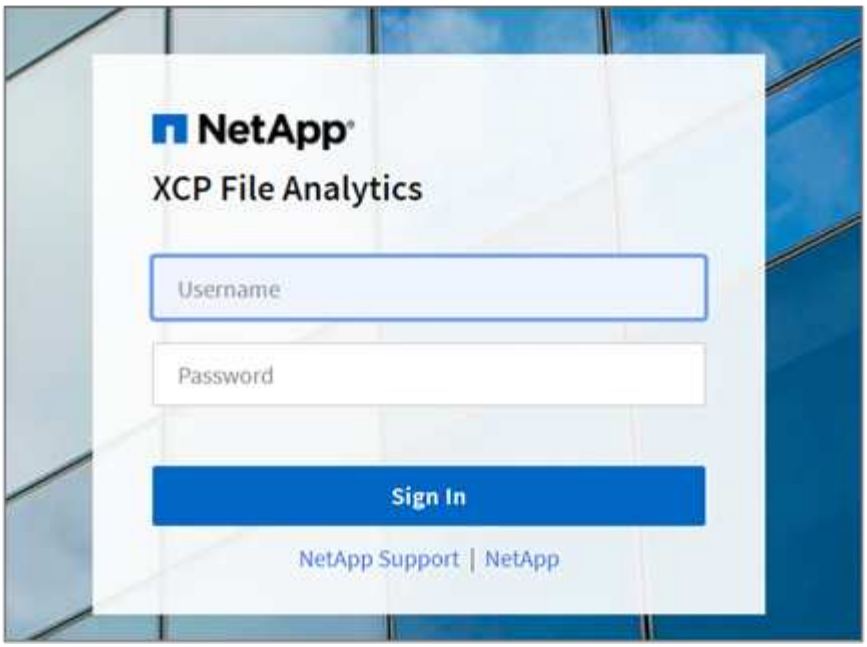
File Analytics provides a graphical view of scan results.

XCP File Analytics GUI provides a dashboard with graphs for visualizing File Analytics.
The XCP File Analytics GUI is enabled when you configure XCP on a Linux machine.

To check the supported browsers for accessing File Analytics, see the NetApp IMT.

Steps

1. Use the link `https://<IP address of linux machine>/xcp` to access the File Analytics GUI. When prompted, accept the security certificate:
   a. Select Advanced below the privacy statement.
   b. Select the proceed to `<IP address of linux machine>` link.
2. Log in to the GUI using the username “admin” and the password you set during configuration



3. On login, you can see that the NFS agent is added: a green tick is present showing minimal system configuration of the Linux system and XCP version.
4. If you have configured an SMB agent, you can see the SMB agent added in the same agent card.

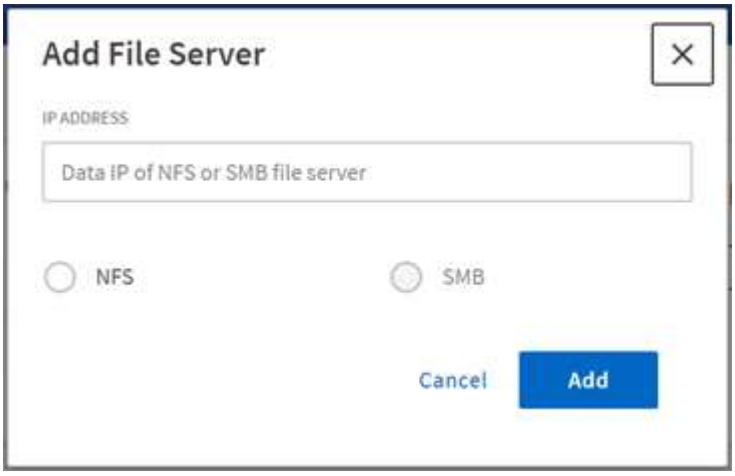
Add file servers

You can configure NFS and SMB exported file systems in the XCP File Analytics GUI.

This enables XCP File Analytics to scan and analyze data on the file system. Use the following steps to add NFS or SMB file servers.

Step

1. To add file servers, select **Add File Server**.



Add the file server IP address, select the NFS or SMB option and click **Add**.

If an SMB agent is not visible in the GUI, you will not be able to add SMB server.

After adding the file server, XCP displays:

- Total file shares available
- File shares with analytics data
  (The initial count is “0”, this updates when you run a successful scan)
- Total space utilization – the sum of space utilized by all the exports
- The data for file shares and space utilization is real-time data direct from the NFS/SMB server. Collecting and processing the data takes several seconds.

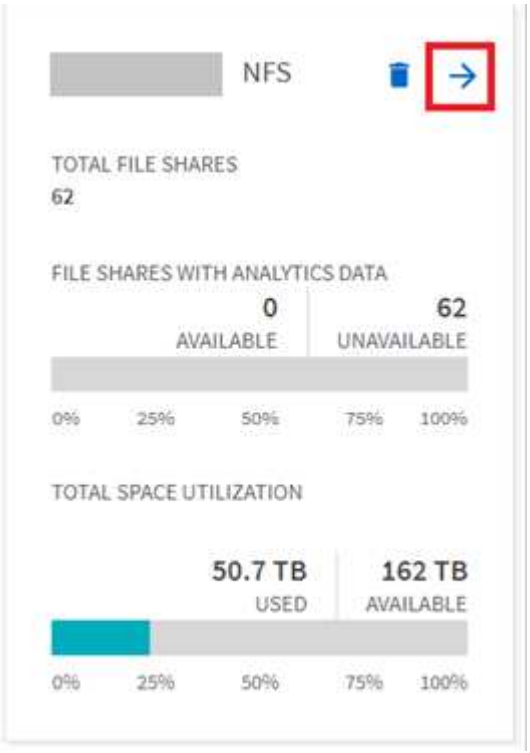
Space available versus space used in File Analytics is calculated from each exported file system available over NFS. For example, if the volumes consist of qtrees and the exports are created over a qtree, the overall space is the cumulative space of the volume size and the qtree size.

Run a scan

When the NFS/SMB files system is added to the XCP File Analytics GUI, you can start a file system scan to analyze and represent the data.

Steps

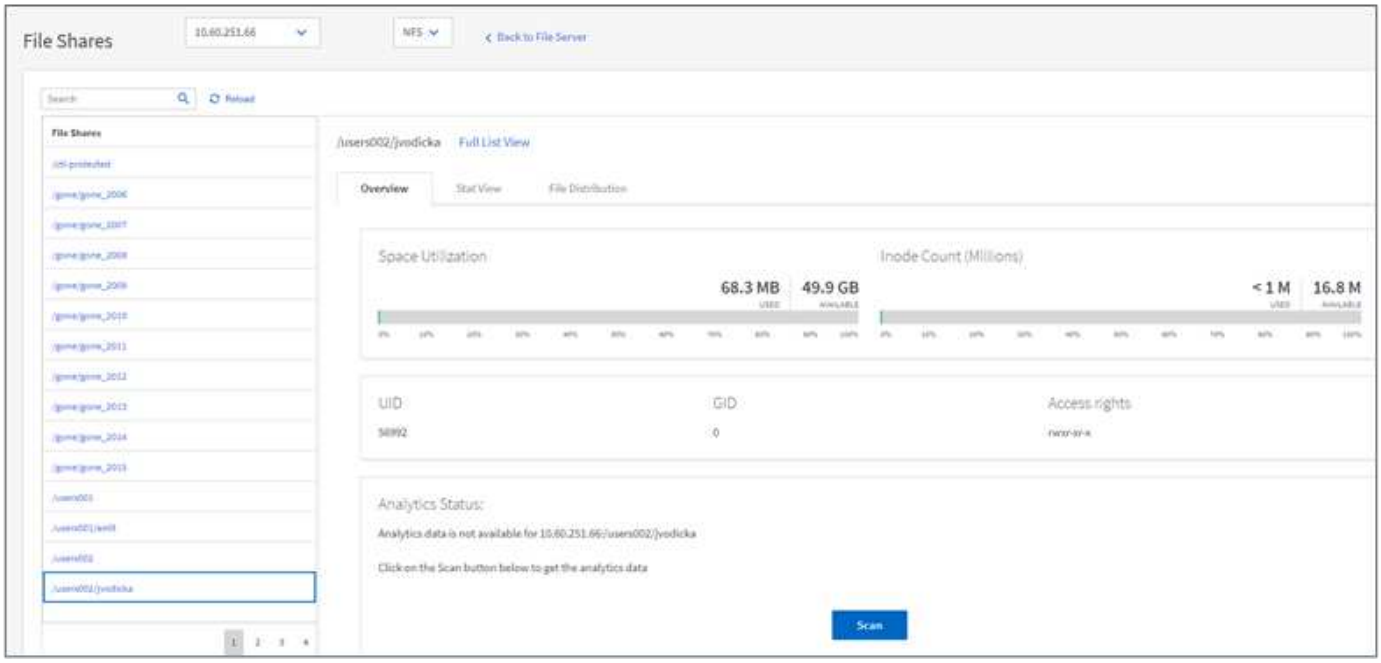
1. Select the arrow on the added file server card to view the file shares on the file server.



2. From the list of file shares, select the name of the file share to scan.
3. Select **Scan** to start the scan.

   XCP displays a progress bar for the scan.

4. When the scan is complete the **stat view** and **file distribution** tabs are enabled to allow you to view graphs.

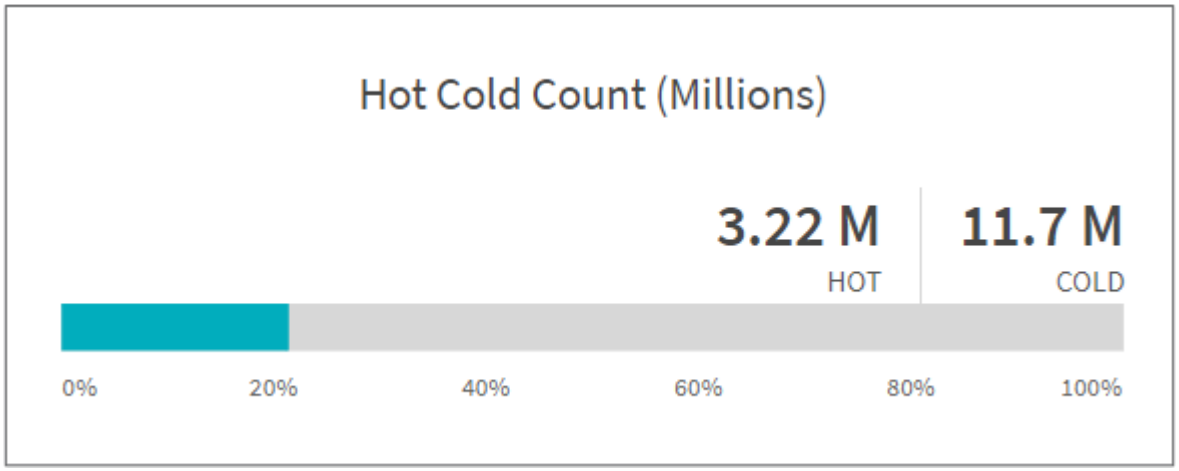


Learn about graphs

The File Analytics GUI dashboard displays multiple graphs for visualizing File Analytics.

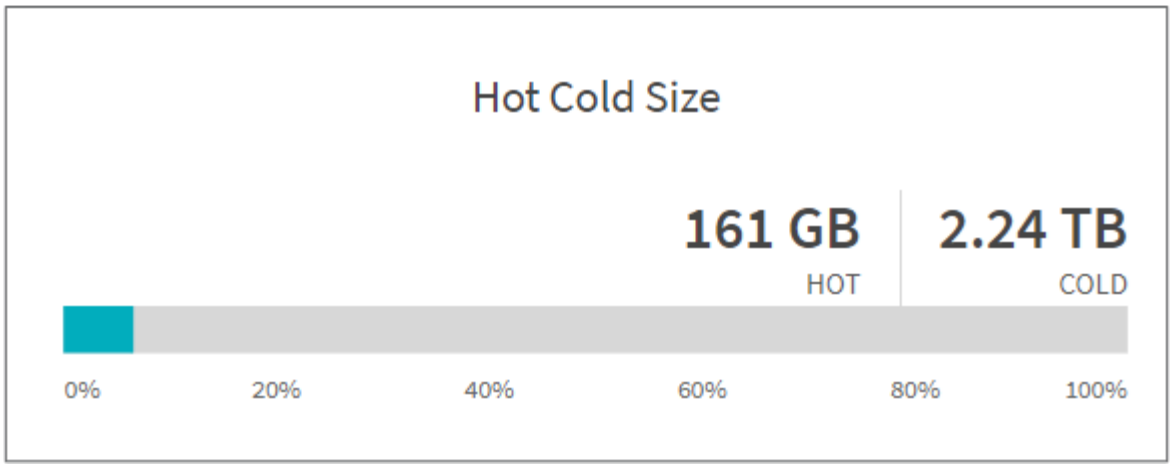
Hot Cold Count Graph

XCP File Analytics categorizes files not accessed for 90 days as cold data. Files accessed in the last 90 days are hot data. Criteria to define hot and cold data is based on access time only.



The Hot Cold Count graph displays the number of inodes (in millions) that are hot or cold in XCP NFS. In XCP SMB, this graph denotes the number of files that are hot or cold.
The colored bar represents the hot data and shows the percentage of files accessed within 90 days.

Hot Cold Size Graph

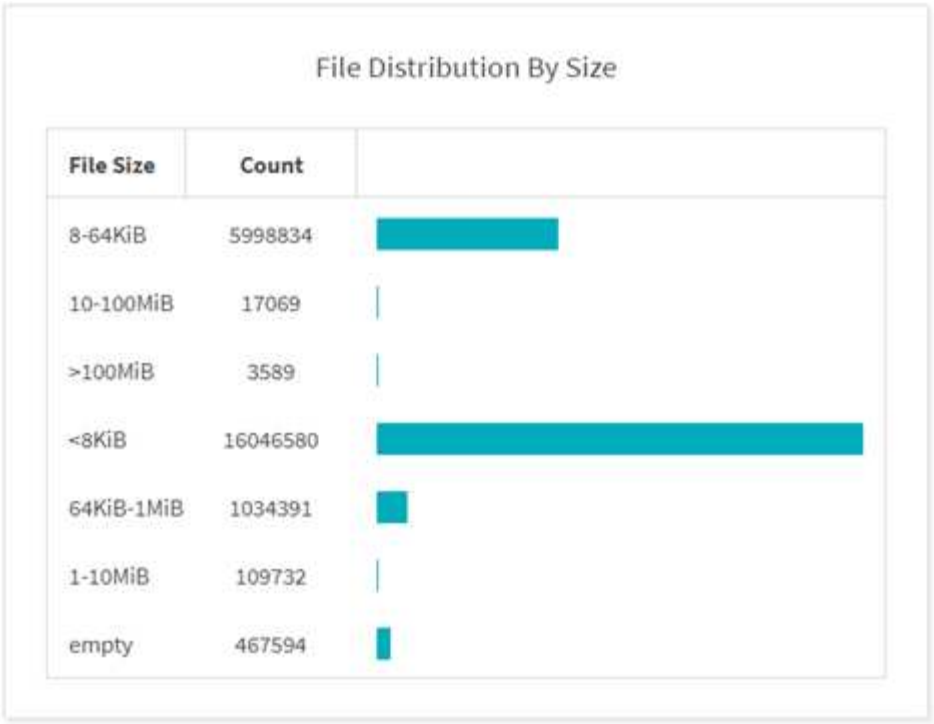


The Hot Cold Size graph displays the percentage of files that are hot and cold and the total size of the files in each category. The colored bar represents the hot data and the uncolored part represents the cold data. Criteria to define hot and cold data is based on access time only.

Entries in Directory Graph

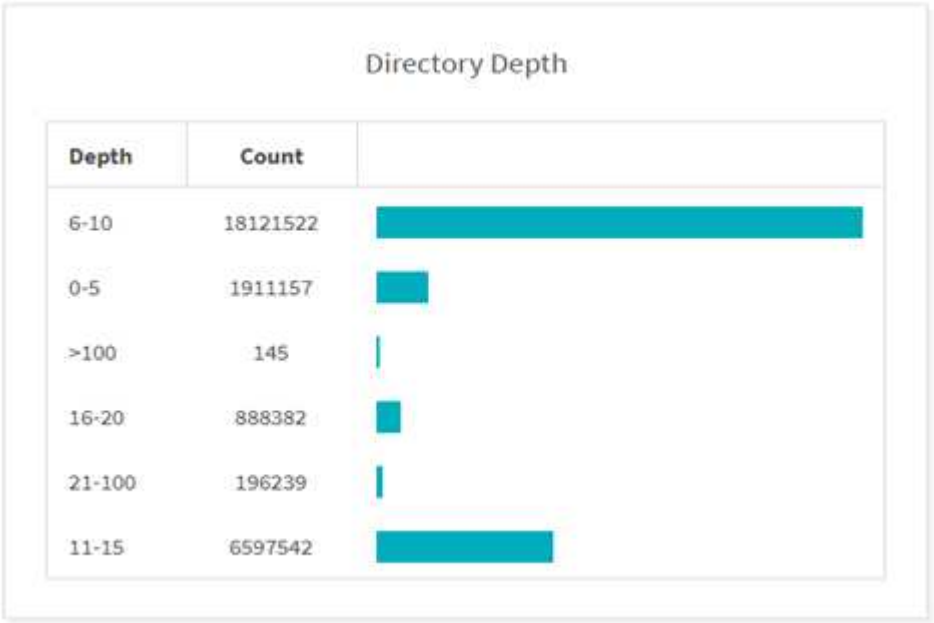
The Entries in Directories graph displays the number of entries in directories. The Depth column contains different directory sizes and the Count column indicates the number of entries in each directory depth.

File Distribution by Size Graph



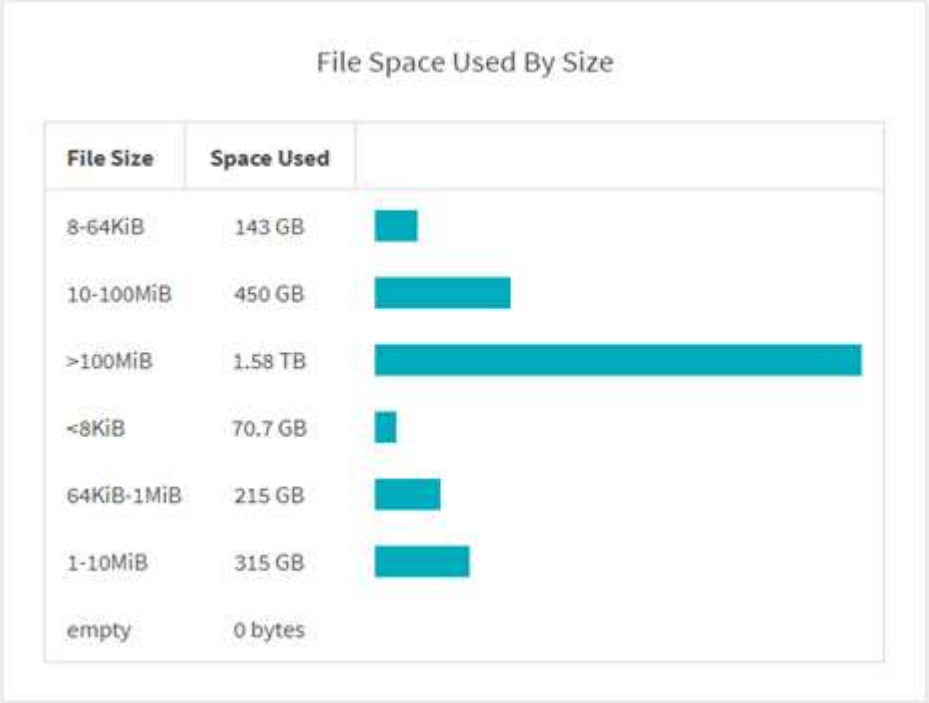
The File Distribution by Size graph displays the number of files that are under the given file sizes. The File Size column contains the categories of file size and the Count column indicates the distribution of the number of files.

Directory Depth Graph



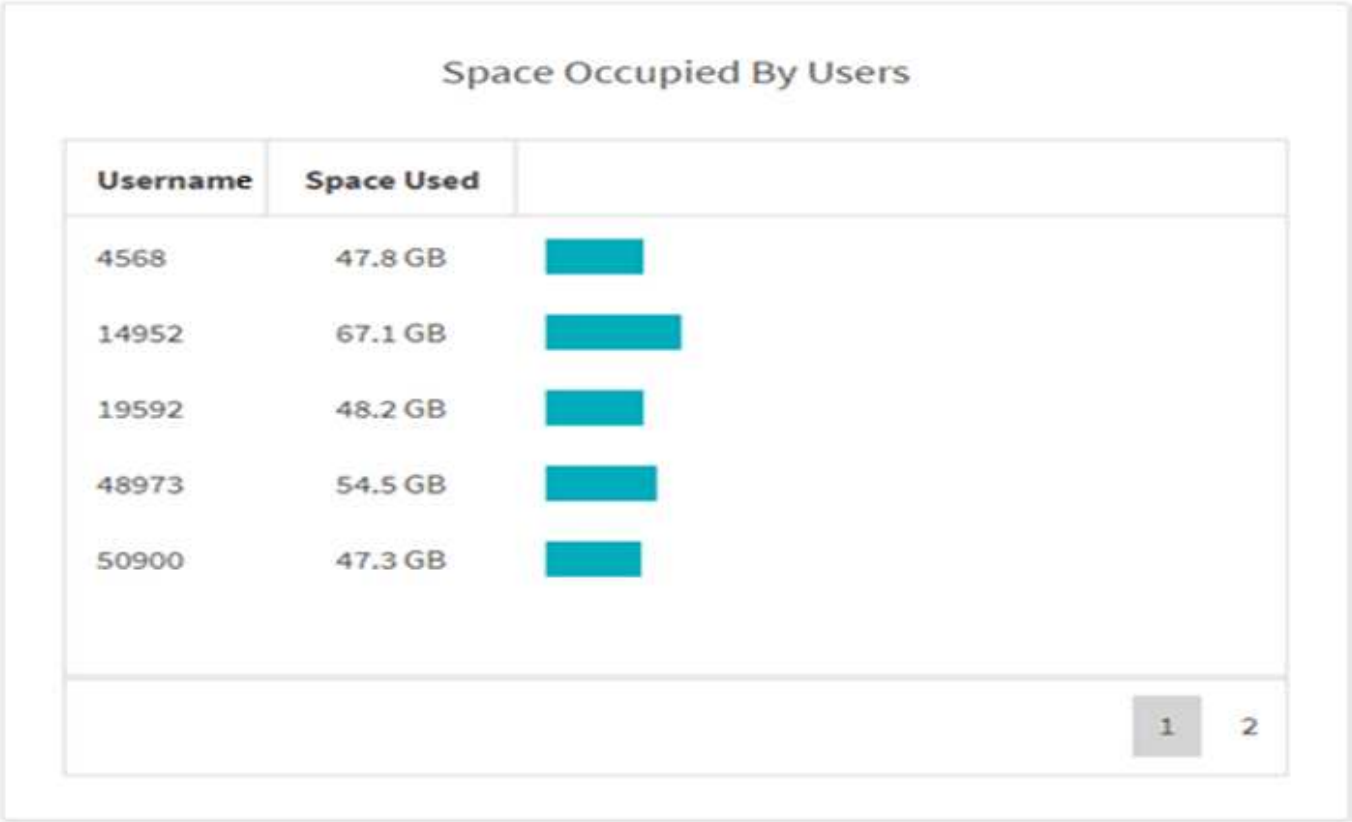
The Directory Depth graph represents the distribution of the number of directories in various directory depth ranges. The Depth column contains various directory depths and the Count column contains the count of each directory depth in the file share.

File Space Used by Size Graph



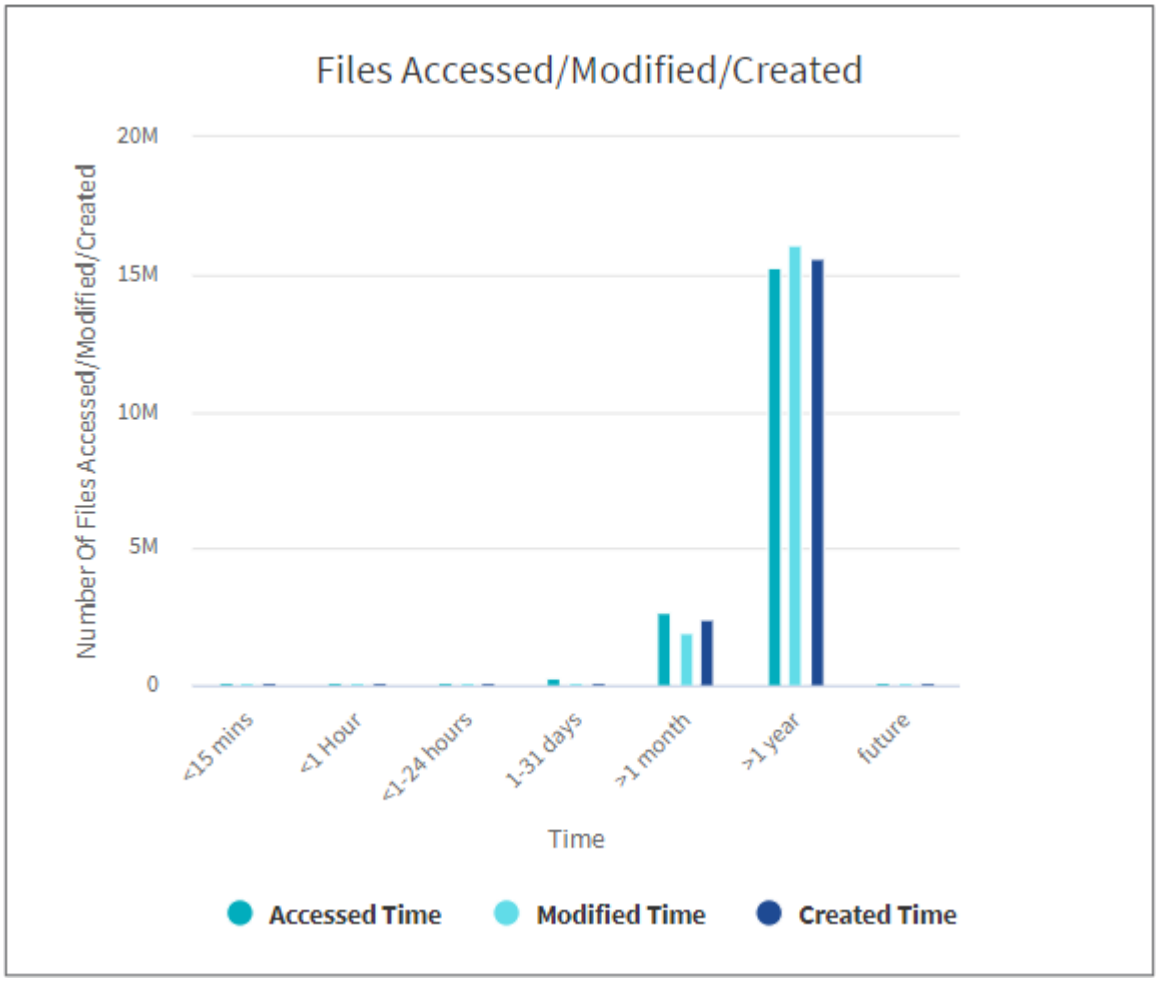
The File Space Used by Size graph displays the number of files in different file-size ranges. The File Size column contains different file size ranges and the Space Used column indicates the space used by each file size range.

Space Occupied by Users Graph



The Space Occupied by Users graph displays the space used by users. The Username column contains the names of users (UID when usernames cannot be retrieved) and the Space Used column indicates the space used by each username.

Files Accessed/Modified/Created Graph



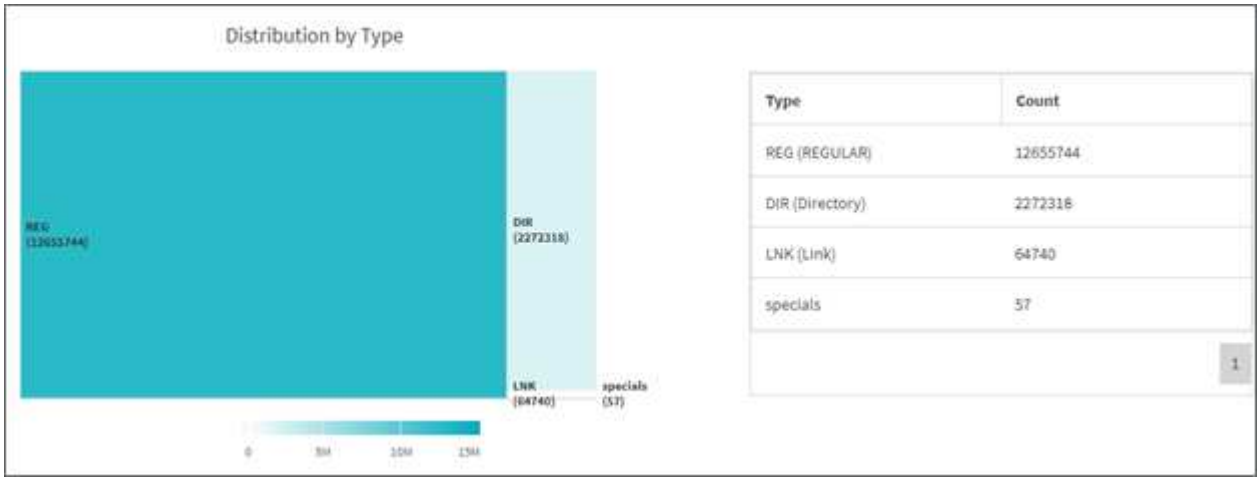
The Files Accessed/Modified/Created graph displays the count of files changed overtime. The X-axis represents the period of time within which changes were made and the y- axis represents the number of files changed.

To get the access time (atime) graph in SMB scans, check the box for preserving atime before running a scan.

File Distribution by Extension Graphic

The File Distribution by Extension graph represents the count of the different file extensions in a file share. The size of the divisions representing the extensions is based on the number of files with each extension.

File Distribution by Type Graph



The Distribution by Type graph represents the count of the following types of files:

- REG: Regular files
- LNK: Files with links
- Specials: Files with device files and character files.
- DIR: Files with directories
- Junction: Available in SMB only