



Plan data migration

XCP

NetApp
March 14, 2024

Table of Contents

- Plan data migration 1
 - Plan NFS data migration 1
 - Plan SMB data migration 2
 - Plan HDFS data migration 2
 - Plan using File Analytics 3
- Filters 16
- Logging for NFS and SMB (optional) 16

Plan data migration

You can plan your migration using the CLI or the File Analytics GUI.

Use the following commands to plan your migration:

- Show
- Scan

Use File Analytics to visualize the statistics for exports and shares.

Plan NFS data migration

Plan your NFS data migrations.

Show

The `show` command queries the RPC services and NFS exports of one or more storage servers. The command lists the available services and exports with the used and free capacity of each export, followed by the root attributes of each export.

Example:

- `xcp show <NFS file server IP/FQDN>`
- `xcp show nfs_server01.netapp.com`

Run `xcp help show` for more details.

Scan

The `scan` command recursively scans the entire source NFSv3 exported paths and prints the statistics of file structure at the end of the scan. NetApp recommends putting the source NFS export mounts in read-only mode during the scan operation.



If a file or directory name contains non-UTF-8 characters, these characters are converted to the UTF-8 format and displayed when you run the `xcp-scan` command. Depending on the translation from the source encoding to UTF-8, the characters might not display as expected.

Example:

- `xcp scan NFS [server:/export path | file://]`
- `xcp scan nfs_server01.netapp.com:/export1`
- `xcp scan file:///mnt/nfs-source`

Run `xcp help scan` for more details.

Optionally, use File Analytics to view the results graphically.

Plan SMB data migration

Plan your SMB data migrations.

Show

The `show` command shows all SMB shares available on the server with the permissions and space available.

Example:

- `xcp show \\<SMB file server IP/FQDN>`
- `xcp show smb_server01.netapp.com`

Run `xcp help show` for more details.

Scan

The `scan` command recursively scans the entire SMB share and lists all the files at the end of the scan.



During the scan operation, you can use the `-preserve-atime` flag with the `scan` command to preserve access time at the source .

Example:

- `xcp scan \\SMB server\share1`
- `xcp scan smb_server01.netapp.com:/share1`

Run `xcp help scan` for more details.

Optionally, use File Analytics to view the results graphically.

Plan HDFS data migration

Plan your HDFS data migrations.

Scan

The `scan` command recursively scans the entire source paths and prints the statistics for the file structure at the end of the scan.

- `xcp scan HDFS [hdfs://<hdfs mounted path>]`
- `xcp scan hdfs:///demo/user1`
- `xcp scan s3://my-bucket`
- `xcp scan -s3.profile <s3 profile name> -s3.endpoint <endpoint-url> s3://my-bucket`

Run `xcp help scan` for more details.

Plan using File Analytics

Plan your data migration

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.

Log in to the File Analytics GUI

The 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 File Analytics GUI.

There are two ways to log in to the File Analytics GUI:

Log in using user credentials

- a. Log in to the GUI using the user credentials obtained when you installed File Analytics.



- b. Optionally, change the password to your own password.

If you want to change the password obtained during installation to your own password, select the user icon and select **Change password**.

Your new password must be at least eight characters in length and contain at least one number, one upper case letter, one lower case letter, and one special character (! @ # \$ % ^ & * - _).



After changing the password, you are automatically logged out of the GUI and you must sign in again using the new password that you created.

Configure and enable SSO capability

You can use this login capability to set up XCP File Analytics on a particular machine and share the web UI URL enterprise-wide, allowing users to log in to the UI using their single sign-on (SSO) credentials.



SSO login is optional and can be configured and enabled permanently. To set up Security Assertion Markup Language (SAML) based SSO login, see [Configure SSO credentials](#).

3. After logging in, you can see the NFS agent; 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.

Configure SSO credentials

The SSO login functionality is implemented in XCP File Analytics using SAML and is supported with the Active

Directory Federation Services (ADFS) identity provider. SAML offloads the authentication task to the third-party identity provider (IdP) for your enterprise which can utilize any number of approaches for MFA (multifactor authentication).

Steps

1. Register the XCP File Analytics application with your enterprise identity provider.

File Analytics now runs as a service provider and therefore must be registered with your enterprise identity provider. Generally, there is a team in the enterprise that handles this SSO integration process. The first step is to find and reach out to the relevant team and share the File Analytics application metadata details with them.

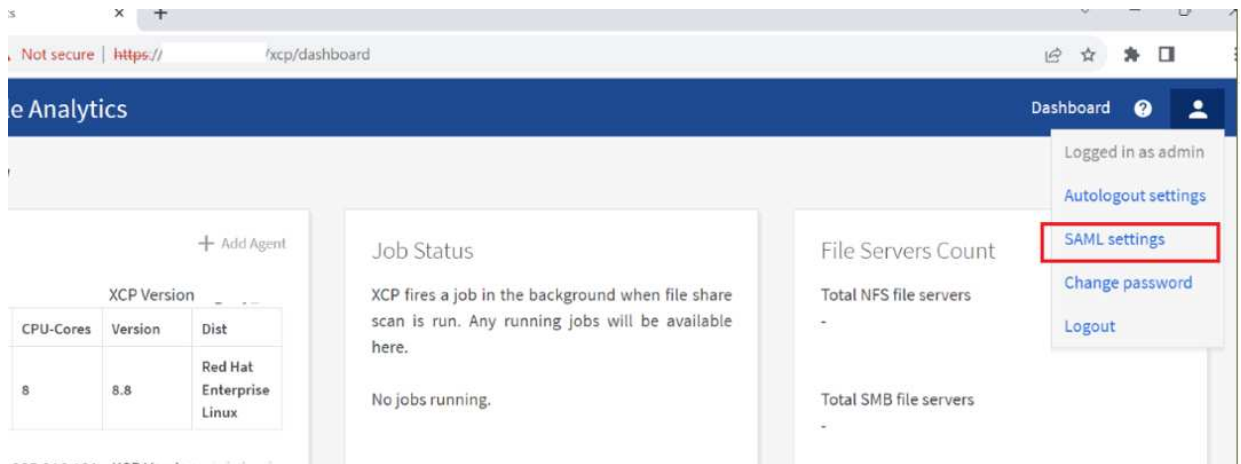
The following are the mandatory details that you must share to register with your identity provider:

- **Service provider entity ID:** `https://<IP address of linux machine>/xcp`
- **Service provider Assertion Consumer Service (ACS) URL:** `https://<IP address of linux machine>:5030/api/xcp/SAML/sp`

You can also verify these details by logging in to the File Analytics UI:

- a. Log in to the GUI using the steps described in [Log in to the File Analytics GUI](#).
- b. Select the **User** icon on the top right corner of the page, then select **SAML settings**.

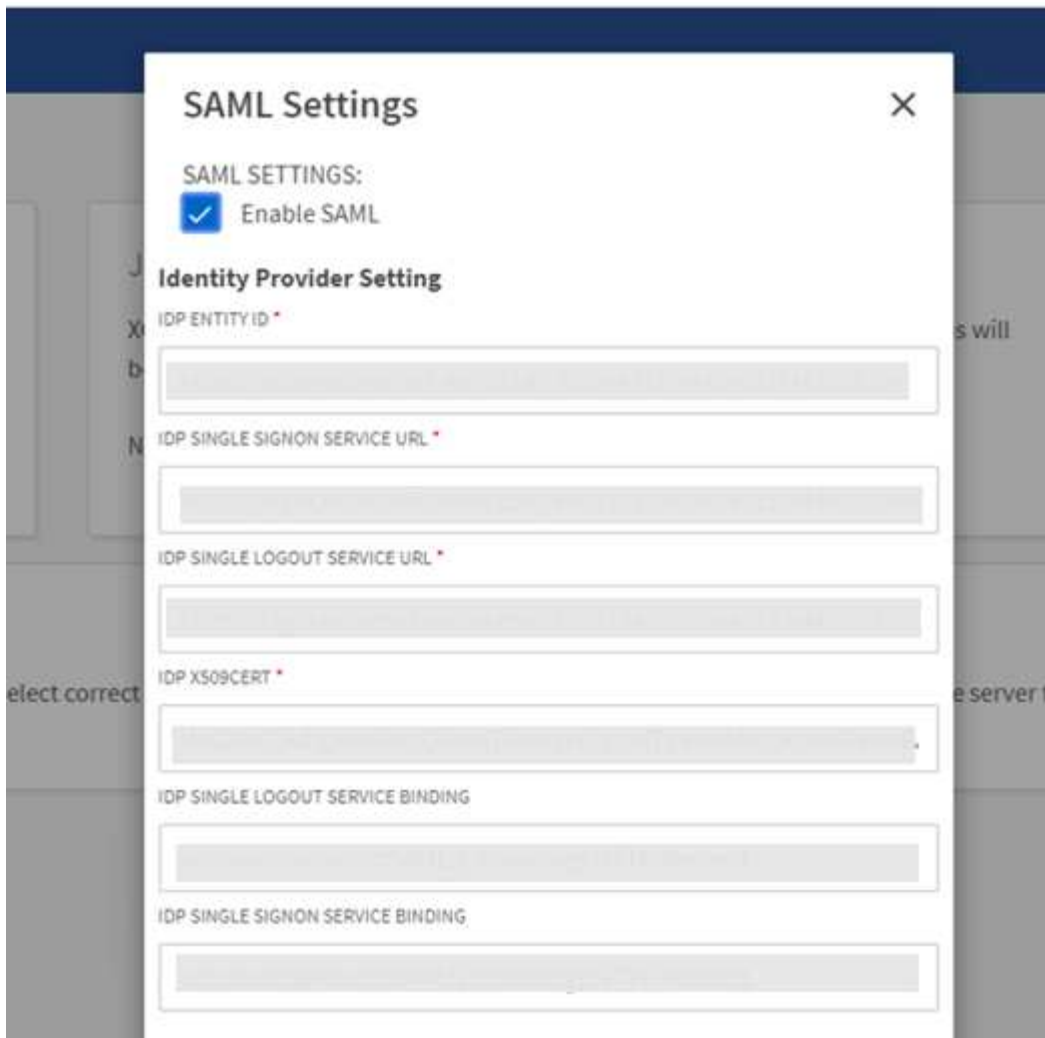
Check **Service provider settings** in the drop down menu that appears.



After registration, you receive the IdP endpoint details for your enterprise. You are required to provide this IdP endpoint metadata to the File Analytics UI.

2. Provide the IdP details:
 - a. Go to **Dashboard**. Select the **User** icon at the top right corner of the page and select **SAML settings**.
 - b. Input the IdP details that you obtained after registration.

Example



- c. Select the **Enable SAML** checkbox to permanently enable SAML-based SSO.
- d. Select **Save**.
- e. Log out of File Analytics and log back in again.

You are redirected to your enterprise SSO page.

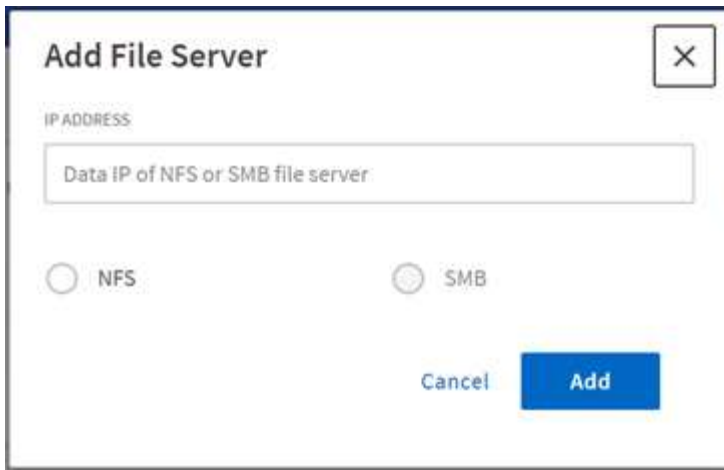
Add file servers

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**.

A dialog box titled "Add File Server" with a close button (X) in the top right corner. It contains a label "IP ADDRESS" above a text input field with the placeholder text "Data IP of NFS or SMB file server". Below the input field are two radio buttons: "NFS" (selected) and "SMB". At the bottom are two buttons: "Cancel" and "Add".

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.



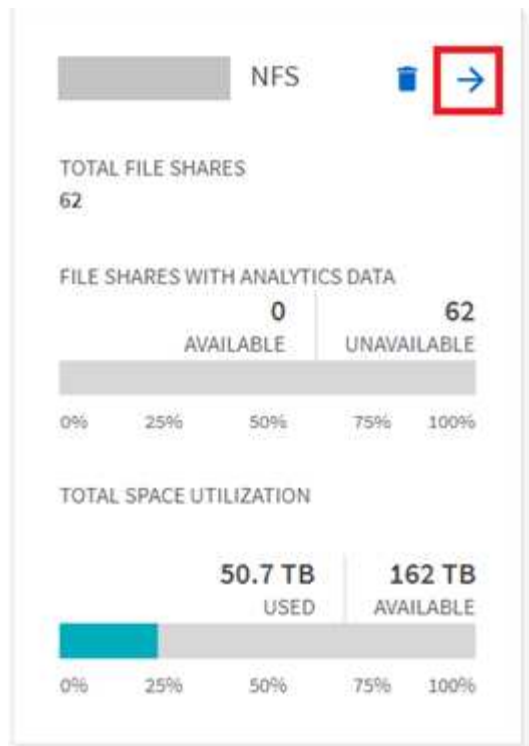
Space available versus space used in File Analytics is calculated from each exported file system available over NFS. For example, if the volumes consist of qtrees and the exports are created over a qtree, the overall space is the cumulative space of the volume size and the qtree size.

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

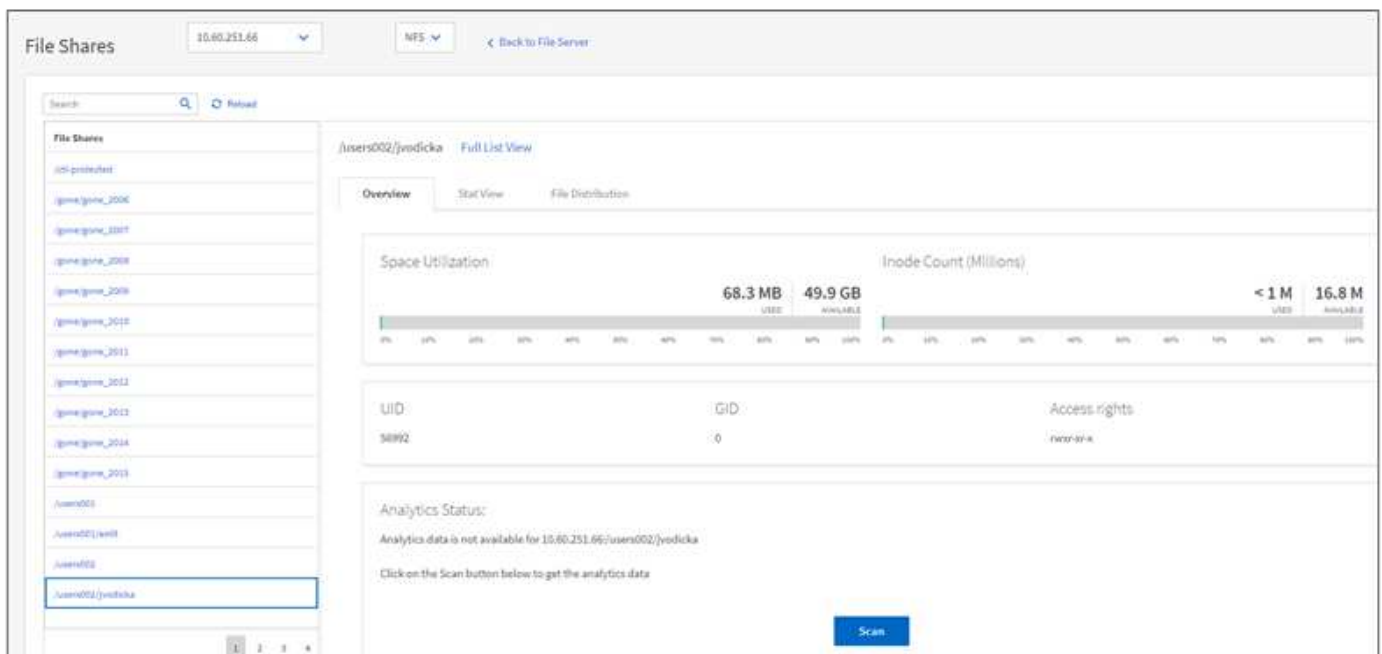
1. Select the arrow on the added file server card to view the file shares on the file server.



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

XCP displays a progress bar for the scan.

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

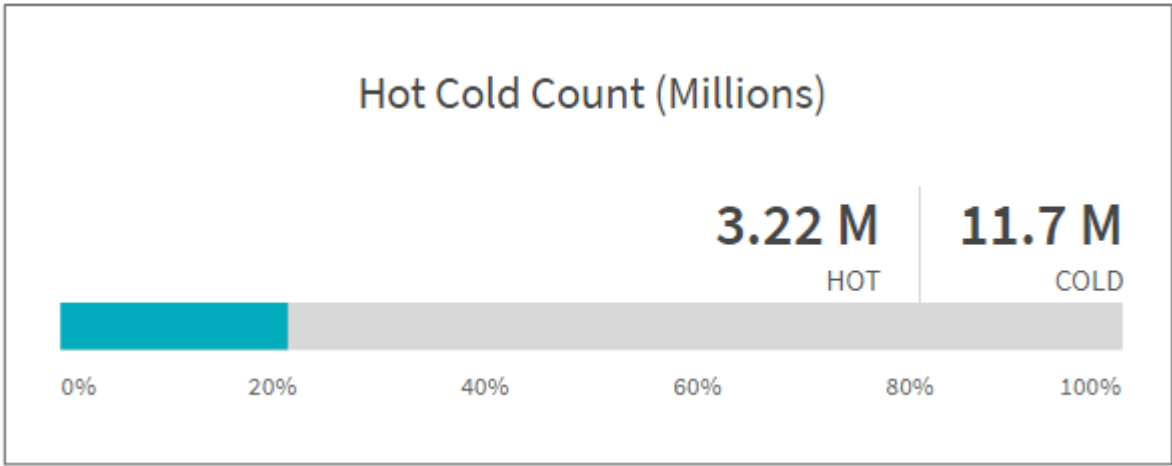


Learn about graphs

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

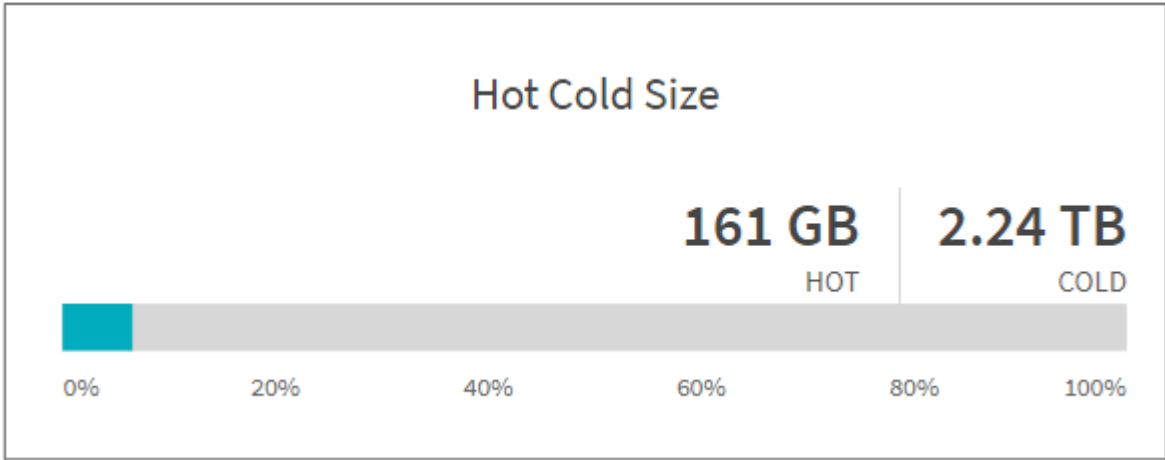
Hot Cold Count Graph

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.



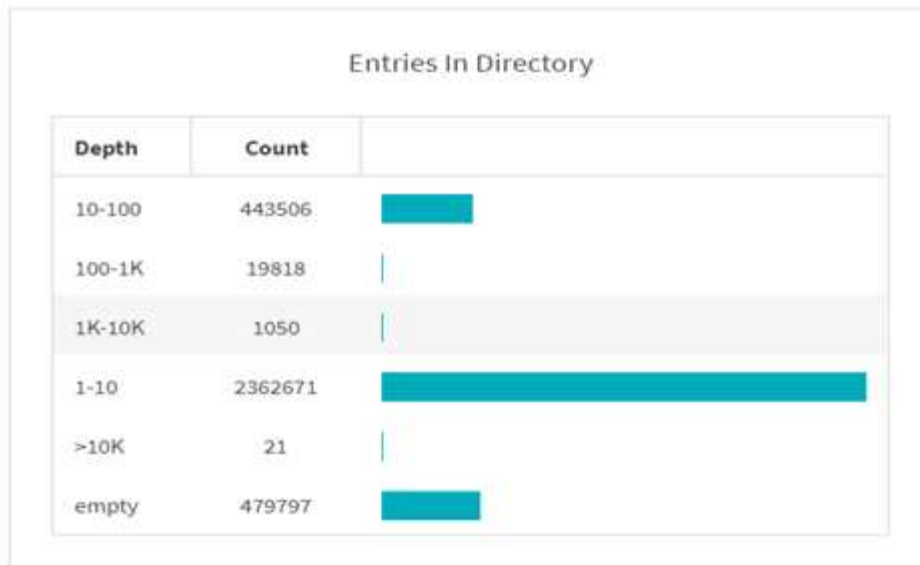
The Hot Cold Count graph displays the number of inodes (in millions) that are hot or cold in XCP NFS. In XCP SMB, this graph denotes the number of files that are hot or cold. The colored bar represents the hot data and shows the percentage of files accessed within 90 days.

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



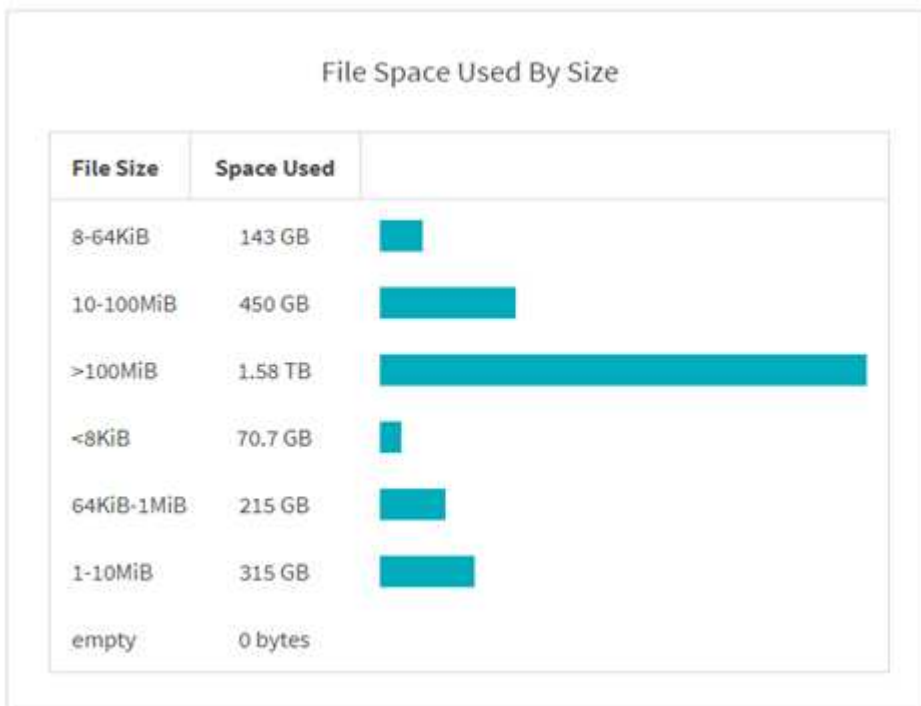
The File Distribution by Size graph displays the number of files that are under the given file sizes. The File Size column contains the categories of file size and the Count column indicates the distribution of the number of files.

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

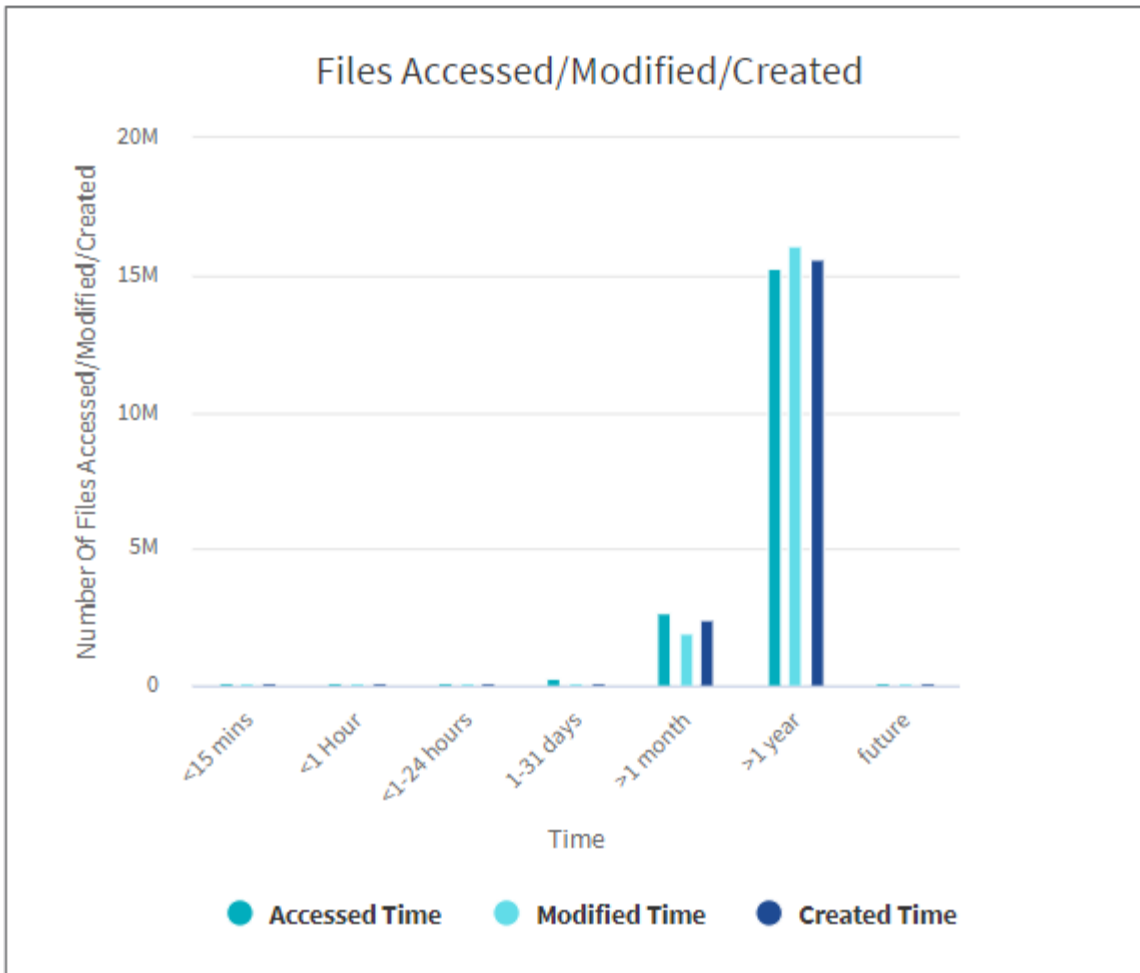
Space Occupied By Users

Username	Space Used	
4568	47.8 GB	
14952	67.1 GB	
19592	48.2 GB	
48973	54.5 GB	
50900	47.3 GB	

12

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

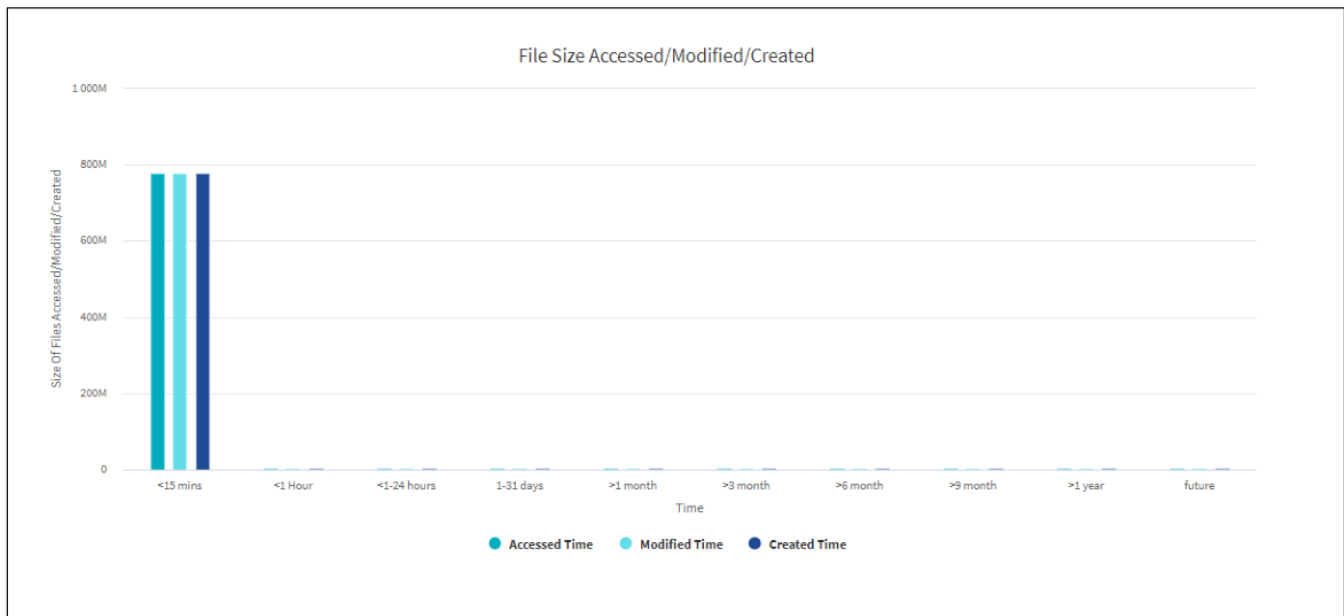


The Files Accessed/Modified/Created graph displays the count of files changed over time. The X-axis represents the period of time within which changes were made and the Y-axis represents the number of files changed.



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

File Size Accessed/Modified/Created Graph

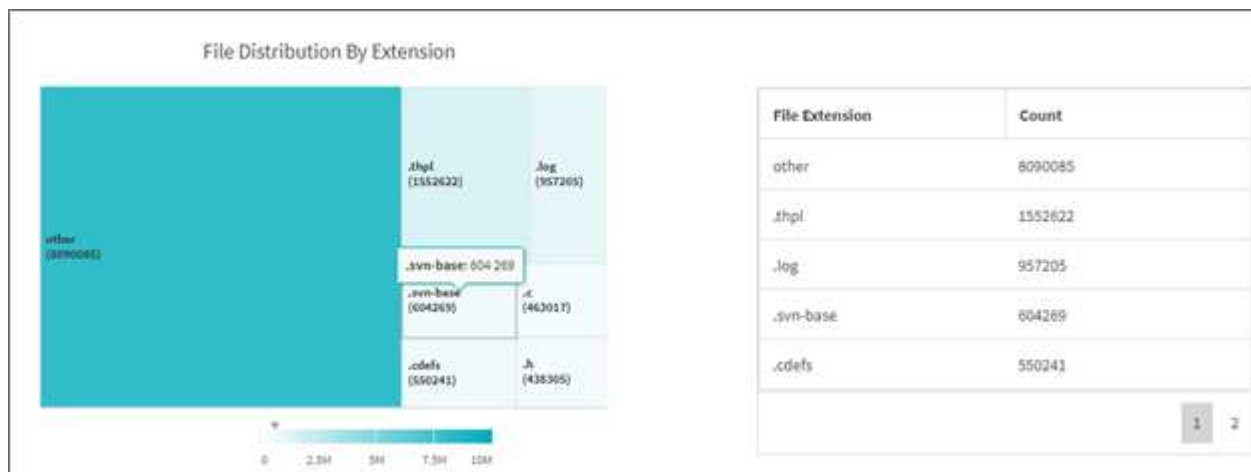


The File Size Accessed/Modified/Created graph displays the size of files changed over time. The X-axis represents the period of time within which changes were made and the Y- axis represents the size of files changed.



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

File Distribution by Extension Graph



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.

Additionally, for SMB shares, you can get the number of Alternate Data Streams (ADS) files for each file extension by checking the box for ADS before running a scan.

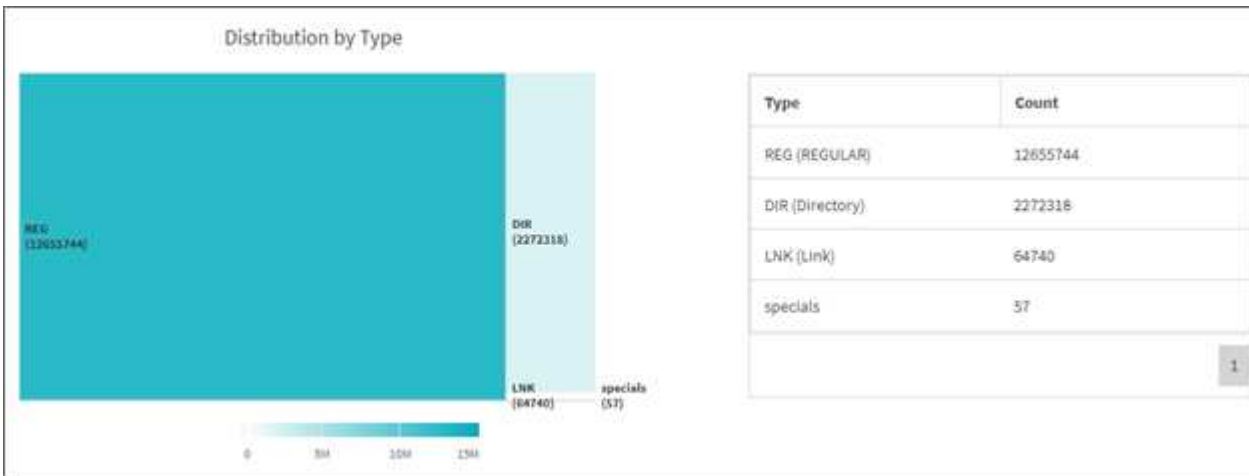


File Size Distribution by Extension Graph



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

File Distribution by Type Graph

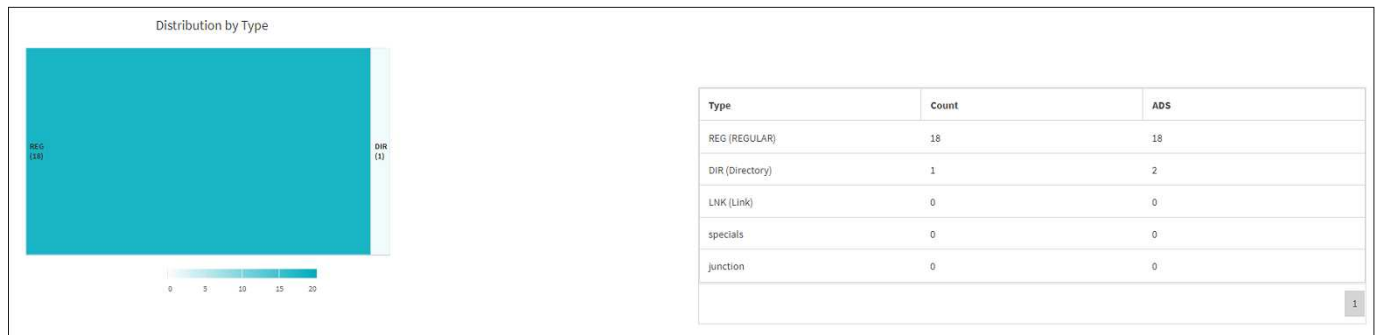


The Distribution by Type graph represents the count of the following types of files:

- REG: Regular files
- LNK: Files with links
- Specials: Files with device files and character files.
- DIR: Files with directories

- Junction: Available in SMB only

Additionally, for SMB shares, you can get the number of Alternate Data Streams (ADS) files for different types by checking the box for ADS before running a scan.



Filters

XCP provides filter options that can be used in XCP operations.

XCP uses filters for `-match` and `-exclude` options for NFS and SMB.

For NFS, run `xcp help info` and refer to the FILTERS section to see how to use `-match` and `-exclude` filters.

For SMB, run `xcp help -match` and `xcp help -exclude` to get more details on match and exclude filters.

If you want to use filters in XCP commands, run `xcp help <command>` to see if they are supported options.

Logging for NFS and SMB (optional)

Logging for XCP NFS and SMB.

XCP supports configuring multiple optional features by using the `xcpLogConfig.json` JSON configuration file. To enable only specific features, manually create the `xcpLogConfig.json` configuration file. You can use the `xcpLogConfig.json` configuration file to enable:

- event log messages
- syslog client for XCP
- custom XCP logging

Event log messages and the syslog client are disabled in the default configuration. Configuration is common for both NFS and SMB.

Config JSON file location	NFS	SMB
Config file default location	/opt/NetApp/xFiles/xcp/	C:\NetApp\XCP\ConfigFile
Custom location requires the XCP_CONFIG_DIR environment variable	Use the location you have set against the XCP_CONFIG_DIR variable	N/A

The JSON configuration file options are case sensitive. These options are the same for XCP NFS and XCP SMB.

Sub options name	JSON data type	Default	Description
logConfig			Option to customize XCP logging.
"level"	String	INFO	Log message severity filter level. XCP log messages support five severity levels in order of decreasing severity: CRITICAL, ERROR, WARNING, INFO, DEBUG (NetApp strongly recommends using INFO or DEBUG)
"maxBytes"	Integer	52428800	Size of each rotating log file. Max supported rotation files are 10.
"name"	String	xcp.log	Option to set custom log file name.
eventlog			Option to configure event log message.
"isEnabled"	Boolean	true	This boolean option is used to enable event messaging. Setting it to <code>false</code> will not generate any event messages and no event logs will be published to event log file.
"level"	String	INFO	Event message severity filter level. Event messaging support five severity levels in order of decreasing severity: CRITICAL, ERROR, WARNING, INFO, DEBUG
syslog			Option to configure syslog messaging.
"isEnabled"	Boolean	false	This boolean option is used to enable syslog client in XCP.
"level"	String	INFO	Message severity filter level. XCP event log messages support five severity levels in order of decreasing severity: CRITICAL, ERROR, WARNING, INFO, DEBUG
"serverIp"	String	None	Remote syslog server IP addresses or hostname.
"port"	Integer	514	Remote syslog receiver port. Syslog receivers accepting syslog datagrams on a different port can be configured with port option UDP port 514 but you can also configure to the desired port.

Sub options name	JSON data type	Default	Description
"sanitize"	Boolean	false	<p>A common option for XCP support; setting its value to true hides sensitive information (IP and username) in the messages going to support (logging, events, syslog, and so on).</p> <p>For example, with the <code>sanitize</code> option as false:</p> <pre>* 2020-07-17 03:10:23,779 - INFO - 12806 xcp xcp Paths: ['10.234.104.251:/cat_vol'] * 2020-07-17 03:10:23,778 - INFO - 12806 xcp xcp User Name: root</pre> <p>With the <code>sanitize</code> option as true:</p> <pre>* 2020-07-17 03:13:51,596 - INFO - 12859 xcp xcp Paths: ['IP: XX.XX.XX.XX:/cat_vol'] * 2020-07-17 03:13:51,595 - INFO - 12859 xcp xcp User Name: * * *</pre>

Create the JSON configuration file

If you want to enable event log messages, the syslog client, or customer logging, complete the following steps.

Steps

1. Open any text editor, such as notepad or vi.
2. Create a new file with the following JSON template.

```
{
  "logConfig": {
    "level": "INFO",
    "maxBytes": 52428800,
    "name": "xcp.log"
  },
  "eventlog": {
    "isEnabled": false,
    "level": "INFO"
  },
  "syslog": {
    "isEnabled": false,
    "level": "INFO",
    "serverIp": "10.234.219.87",
    "port": 514
  },
  "sanitize": false
}
```

3. For any features that you want to enable, change the `isEnabled` value to `true`.
4. Name the file `xcpLogConfig.json` and save it to the default location: `/opt/NetApp/xFiles/xcp/`

If the `XCP_CONFIG_DIR` environment variable is set, save the `xcpLogConfig.json` file in the same location that is set against the `XCP_CONFIG_DIR` variable.

Default configuration	Example json configuration file
<pre>{ "logConfig": { "level": "INFO", "maxBytes": 52428800, "name": "xcp.log" }, "sanitize": false }</pre>	<pre>{ "logConfig": { "level": "INFO", "maxBytes": 52428800, "name": "xcp.log" }, "eventlog": { "isEnabled": false, "level": "INFO" }, "syslog": { "isEnabled": false, "level": "INFO", "serverIp": "10.234.219.87", "port": 514 }, "sanitize": false }</pre>

Copyright information

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.