Appendix: NFS and SMB file and directory naming dependencies

ONTAP 9

NetApp
December 12, 2022

This PDF was generated from https://docs.netapp.com/us-en/ontap/nfs-admin/nfs-smb-file-directory-naming-concept.html on December 12, 2022. Always check docs.netapp.com for the latest.

Table of Contents

Appendix: NFS and SMB file and directory naming dependencies . . . 1
Appendix: NFS and SMB file and directory naming dependencies . . . 1
Characters a file or directory name can use . . . 1
Case-sensitivity of file and directory names in a multiprotocol environment . . . 1
How ONTAP creates file and directory names . . . 2
How ONTAP handles multi-byte file, directory, and qtree names . . . 2
Configure character mapping for SMB file name translation on volumes . . . 4
Commands for managing character mappings for SMB file name translation . . . 6

Appendix: NFS and SMB file and directory naming dependencies

Appendix: NFS and SMB file and directory naming dependencies

File and directory naming conventions depend on both the network clients' operating systems and the file-sharing protocols, in addition to language settings on the ONTAP cluster and clients.

The operating system and the file-sharing protocols determine the following:

- Characters a file name can use
- Case-sensitivity of a file name

ONTAP supports multi-byte characters in file, directory, and qtree names, depending on the ONTAP release.

Characters a file or directory name can use

If you are accessing a file or directory from clients with different operating systems, you should use characters that are valid in both operating systems.

For example, if you use UNIX to create a file or directory, do not use a colon (:) in the name because the colon is not allowed in MS-DOS file or directory names. Because restrictions on valid characters vary from one operating system to another, see the documentation for your client operating system for more information about prohibited characters.

Case-sensitivity of file and directory names in a multiprotocol environment

File and directory names are case-sensitive for NFS clients and case-insensitive but case-preserving for SMB clients. You must understand what the implications are in a multiprotocol environment and the actions you might need to take when specifying the path while creating SMB shares and when accessing data within the shares.

If an SMB client creates a directory named `testdir`, both SMB and NFS clients display the file name as `testdir`. However, if an SMB user later tries to create a directory name `TESTDIR`, the name is not allowed because, to the SMB client, that name currently exists. If an NFS user later creates a directory named `TESTDIR`, NFS and SMB clients display the directory name differently, as follows:

- On NFS clients, you see both directory names as they were created, for example `testdir` and `TESTDIR`, because directory names are case-sensitive.
- SMB clients use the 8.3 names to distinguish between the two directories. One directory has the base file name. Additional directories are assigned an 8.3 file name.
  - On SMB clients, you see `testdir` and `TESTDI~1`.

- ONTAP creates the `TESTDI~1` directory name to differentiate the two directories.

  In this case, you must use the 8.3 name when specifying a share path while creating or modifying a share on a storage virtual machine (SVM).

Similarly for files, if an SMB client creates `test.txt`, both SMB and NFS clients display the file name as `text.txt`. However, if an SMB user later tries to create `Test.txt`, the name is not allowed because, to the SMB client, that name currently exists. If an NFS user later creates a file named `Test.txt`, NFS and SMB clients display the file name differently, as follows:

- On NFS clients, you see both file names as they were created, `test.txt` and `Test.txt`, because file names are case-sensitive.
- SMB clients use the 8.3 names to distinguish between the two files. One file has the base file name. Additional files are assigned an 8.3 file name.
  - On SMB clients, you see `test.txt` and `TEST~1.TXT`.
  - ONTAP creates the `TEST~1.TXT` file name to differentiate the two files.

If you have enabled or modified character mapping using the Vserver CIFS character-mapping commands, a normally case-insensitive Windows lookup becomes case-sensitive.

How ONTAP creates file and directory names

ONTAP creates and maintains two names for files or directories in any directory that has access from an SMB client: the original long name and a name in 8.3 format.

For file or directory names that exceed the eight character name or the three character extension limit (for files), ONTAP generates an 8.3-format name as follows:

- It truncates the original file or directory name to six characters, if the name exceeds six characters.
- It appends a tilde (~) and a number, one through five, to file or directory names that are no longer unique after being truncated.

  If it runs out of numbers because there are more than five similar names, it creates a unique name that bears no relation to the original name.

- In the case of files, it truncates the file name extension to three characters.

For example, if an NFS client creates a file named `specifications.html`, the 8.3 format file name created by ONTAP is `specif~1.htm`. If this name already exists, ONTAP uses a different number at the end of the file name. For example, if an NFS client then creates another file named `specifications_new.html`, the 8.3 format of `specifications_new.html` is `specif~2.htm`.

How ONTAP handles multi-byte file, directory, and qtree names

Beginning with ONTAP 9.5, support for 4-byte UTF-8 encoded names enables the creation and display of file, directory, and tree names that include Unicode supplementary characters outside the Basic Multilingual Plane (BMP). In earlier releases, these

supplementary characters did not display correctly in multiprotocol environments.

To enable support for 4-byte UTF-8 encoded names, a new *utf8mb4* language code is available for the `vserver` and `volume` command families.

- You must create a new volume in one of the following ways:
- Setting the volume `-language` option explicitly:

  ```
  volume create -language utf8mb4 {…}
  ```

- Inheriting the volume `-language` option from an SVM that has been created with or modified for the option:

  ```
  vserver [create|modify] -language utf8mb4 {…}``volume create {…}
  ```

- You cannot modify existing volumes for utf8mb4 support; you must create a new utf8mb4-ready volume, and then migrate the data using client-based copy tools.

  You can update SVMs for utf8mb4 support, but existing volumes retain their original language codes.

  LUN names with 4-byte UTF-8 characters are not currently supported.

- Unicode character data is typically represented in Windows file systems applications using the 16-bit Unicode Transformation Format (UTF-16) and in NFS file systems using the 8-bit Unicode Transformation Format (UTF-8).

  In releases prior to ONTAP 9.5, names including UTF-16 supplementary characters that were created by Windows clients were correctly displayed to other Windows clients but were not translated correctly to UTF-8 for NFS clients. Similarly, names with UTF-8 supplementary characters by created NFS clients were not translated correctly to UTF-16 for Windows clients.

- When you create file names on systems running ONTAP 9.4 or earlier that contain valid or invalid supplementary characters, ONTAP rejects the file name and returns an invalid file name error.

  To avoid this issue, use only BMP characters in file names and avoid using supplementary characters, or upgrade to ONTAP 9.5 or later.

Unicode characters are allowed in qtree names.

- You can use either the `volume qtree` command family or System Manager to set or modify qtree names.
- qtree names can include multi-byte characters in Unicode format, such as Japanese and Chinese characters.
- In releases before ONTAP 9.5, only BMP characters (that is, those that could be represented in 3 bytes) were supported.

  In releases before ONTAP 9.5, the junction-path of the qtree's parent volume can contain qtree and directory names with Unicode characters. The `volume show` command displays these names correctly when the parent volume has a UTF-8 language setting. However, if the parent volume language is not one of the UTF-8 language settings, some parts of the junction-path are displayed using a numeric NFS alternate name.

- In 9.5 and later releases, 4-byte characters are supported in qtree names, provided that the qtree is in a volume enabled for utf8mb4.

Configure character mapping for SMB file name translation on volumes

NFS clients can create file names that contain characters that are not valid for SMB clients and certain Windows applications. You can configure character mapping for file name translation on volumes to allow SMB clients to access files with NFS names that would otherwise not be valid.

About this task

When files created by NFS clients are accessed by SMB clients, ONTAP looks at the name of the file. If the name is not a valid SMB file name (for example, if it has an embedded colon ":" character), ONTAP returns the 8.3 file name that is maintained for each file. However, this causes problems for applications that encode important information into long file names.

Therefore, if you are sharing a file between clients on different operating systems, you should use characters in the file names that are valid in both operating systems.

However, if you have NFS clients that create file names containing characters that are not valid file names for SMB clients, you can define a map that converts the invalid NFS characters into Unicode characters that both SMB and certain Windows applications accept. For example, this functionality supports the CATIA MCAD and Mathematica applications as well as other applications that have this requirement.

You can configure character mapping on a volume-by-volume basis.

You must keep the following in mind when configuring character mapping on a volume:

- Character mapping is not applied across junction points.

  You must explicitly configure character mapping for each junction volume.

- You must make sure that the Unicode characters that are used to represent invalid or illegal characters are characters that do not normally appear in file names; otherwise, unwanted mappings occur.

  For example, if you try to map a colon (:) to a hyphen (-) but the hyphen (-) was used in the file name correctly, a Windows client trying to access a file named "a-b" would have its request mapped to the NFS name of "a:b" (not the desired outcome).

- After applying character mapping, if the mapping still contains an invalid Windows character, ONTAP falls back to Windows 8.3 file names.
- In FPolicy notifications, NAS audit logs, and security trace messages, the mapped file names are shown.
- When a SnapMirror relation of type DP is created, the source volume's character mapping is not replicated on the destination DP volume.
- Case sensitivity: Because the mapped Windows names turn into NFS names, the lookup of the names follows NFS semantics. That includes the fact that NFS lookups are case-sensitive. This means that the applications accessing mapped shares must not rely on Windows case-insensitive behavior. However, the 8.3 name is available, and that is case-insensitive.
- Partial or invalid mappings: After mapping a name to return to clients doing directory enumeration ("dir"), the resulting Unicode name is checked for Windows validity. If that name still has invalid characters in it, or

if it is otherwise invalid for Windows (e.g. it ends in "." or blank) the 8.3 name is returned instead of the invalid name.

Step

1. Configure character mapping:

   ```
   vserver cifs character-mapping create -vserver vserver_name -volume
   volume_name -mapping mapping_text, ...
   ```

   The mapping consists of a list of source-target character pairs separated by “:”. The characters are Unicode characters entered using hexadecimal digits. For example: 3C:E03C.

   The first value of each `mapping_text` pair that is separated by a colon is the hexadecimal value of the NFS character you want to translate, and the second value is the Unicode value that SMB uses. The mapping pairs must be unique (a one-to-one mapping should exist).

   - Source mapping

     The following table shows the permissible Unicode character set for source mapping:

Unicode character	Printed character	Description
0x01-0x19	Not applicable	Non-printing control characters
0x5C	\	Backslash
0x3A	:	Colon
0x2A	*	Asterisk
0x3F	?	Question mark
0x22	"	Quotation mark
0x3C	<	Less than
0x3E	>	Greater than
0x7C	|	Vertical line
0xB1	±	Plus-minus sign

   - Target mapping

     You can specify target characters in the “Private Use Area” of Unicode in the following range: U+E0000…U+F8FF.

Example

The following command creates a character mapping for a volume named “data” on storage virtual machine (SVM) vs1:

```
cluster1::> vserver cifs character-mapping create -volume data -mapping
3c:e17c,3e:f17d,2a:f745
cluster1::> vserver cifs character-mapping show

Vserver         Volume Name  Character Mapping
--------------  -----------  ----------------------------------------
vs1             data         3c:e17c, 3e:f17d, 2a:f745
```

Commands for managing character mappings for SMB file name translation

You can manage character mapping by creating, modifying, displaying information about, or deleting file character mappings used for SMB file name translation on FlexVol volumes.

If you want to…	Use this command…
Create new file character mappings	`vserver cifs character-mapping create`
Display information about file character mappings	`vserver cifs character-mapping show`
Modify existing file character mappings	`vserver cifs character-mapping modify`
Delete file character mappings	`vserver cifs character-mapping delete`

For more information, see the man page for each command.

Copyright information

Copyright © 2022 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.