Infrastructure - Bug #1144

/**cn**/ escaped slash mishandled by server

2010-12-06 21:49 - Rob Nahf

Status:	Closed		**Start date:**	
Priority:	Urgent		**Due date:**	
Assignee:			**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	Sprint-2010.49			
Milestone:			**Story Points:**	
Product Version:	*			

Description

escaped slashes (in identifiers) are being decoded and misdirecting in cn calls, returning a 404 (Not Found) error from the Apache Server. It should preserve the pct-encoding and pass the encoded URL to the servlet. An example:

http://cn-dev.dataone.org/cn/object/foo%2F and
http://cn-dev.dataone.org/cn/object/foo%2Fxxx

both give the following (differing as you would expect at the URL):
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">

404 Not Found

Not Found
The requested URL /cn/object/foo/bar was not found on this server.

Apache/2.2.12 (Ubuntu) Server at cn-dev.dataone.org Port 80

Update Log time Watch Duplicate

History

#1 - 2010-12-06 22:28 - Rob Nahf

Log file excerpts from cn-dev:
from apache/error.log

[Mon Dec 06 13:56:22 2010] [error] [client 129.24.0.10] File does not exist: /var/www/favicon.ico, referer:
http://cn-dev.dataone.org/cn/object/foo%2Fbar
[Mon Dec 06 13:56:43 2010] [error] [client 129.24.0.10] File does not exist: /var/www/favicon.ico, referer: http://cn-dev.dataone.org/cn/object/foo%2F

from apache2/other_vhosts_access.log:
cn-dev.dataone.org:80 129.24.0.10 - - [06/Dec/2010:13:56:22 -0800] "GET /cn/object/foo%2Fbar HTTP/1.1" 404 511 "-" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_5; en-us) AppleWebKit/533.19.4 (KHTML, like Gecko) Version/5.0.3 Safari/533.19.4"
cn-dev.dataone.org:80 129.24.0.10 - - [06/Dec/2010:13:56:22 -0800] "GET /favicon.ico HTTP/1.1" 404 506 "
http://cn-dev.dataone.org/cn/object/foo%2Fbar" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_5; en-us) AppleWebKit/533.19.4 (KHTML, like Gecko) Version/5.0.3 Safari/533.19.4"
cn-dev.dataone.org:80 129.24.0.10 - - [06/Dec/2010:13:56:38 -0800] "GET /cn/object/foo%2F HTTP/1.1" 404 510 "-" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_5; en-us) AppleWebKit/533.19.4 (KHTML, like Gecko) Version/5.0.3 Safari/533.19.4"
cn-dev.dataone.org:80 129.24.0.10 - - [06/Dec/2010:13:56:43 -0800] "GET /favicon.ico HTTP/1.1" 404 506 "
http://cn-dev.dataone.org/cn/object/foo%2F" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_5; en-us) AppleWebKit/533.19.4 (KHTML, like Gecko) Version/5.0.3 Safari/533.19.4"

/var/log/tomcat6/cn.log doesn't contain entries for these calls, but does for normal cases.
Found:
[INFO] 2010-12-06 14:34:18,001 (MetacatProxyObjectServiceImpl:debugWrapper:194) proxy.wrapper javax.servlet.forward.request_uri:
/cn/object/knb:testid:201033641222165

and not found:
[INFO] 2010-12-06 14:13:23,534 (MetacatProxyObjectServiceImpl:debugWrapper:194) proxy.wrapper javax.servlet.forward.request_uri:
/cn/object/foofoo

similarly tomcat6/localhost.2010-12-06.log contains urlrewrite logging info for normal cases only.

/var/log/syslog does not contain entries for these requests either, but is less informative, as it only logs caught exceptions for bogus id situations;

nothing for successfully found ids.

#2 - 2010-12-06 23:34 - Rob Nahf

see http://httpd.apache.org/docs/2.2/mod/core.html#allowencodedslashes

#3 - 2010-12-09 23:48 - Rob Nahf

- Status changed from New to Closed

- Priority changed from Normal to Urgent

thorough troubleshooting ruled out possible solutions to this problem. Apache's AllowEncodedSlashes directive does not seem to work as it should (https://issues.apache.org/bugzilla/show_bug.cgi?id=35256). Tests on cn-dev with against all permutations of AllowEncodedSlashes [on|off] and AcceptPathInfo [on|off|default] still leave %2F blocked in the first stage of request processing. Giving the error :

[Thu Dec 09 14:16:47 2010] [info] [client 127.0.0.1] found %2f (encoded '/') in URI (decoded='/cn/object/xxhttpAESon/APInotset'), returning 404

in the apache2/error.log

All similar tests with the %2F in the query portion of the URI make it through tomcat and to the /cn/ service endpoint. I used URLs in the form of: localhost/cn/object/get?id=ooo%2Fooo

localhost/cn/object?id=ooo%2Fooo

Regarding the apache2 bug, it looks like a longstanding issue that's not getting much attention.

#4 - 2010-12-13 17:33 - Rob Nahf

- File escapedSlashesTests.xlsx added

Files

escapedSlashesTests.xlsx	46.2 KB	2010-12-13		Rob Nahf