HTTP 500 - squid vs java.net.URL

January 4, 2011 / Gavin Jackson

development java squid tomcat

Ran into an interesting bug today that I thought I should share with the interwebs. We have a web application that supports RESTful http requests to interact with the system.

Seems to work fine using browsers/wget etc. But when trying to use java.net.URL (as used by several JUnit tests), I kept getting an http 500 (internal server) error message.

Tomcat logs showed no interaction with the web server, indicating that it must be something wrong with the Squid reverse proxy server (which we have set up to cache outbound web traffic). SQUID LOGS

```
From java test harness (java.net.URL)
1294109122.110 1175 127.0.0.1 TCP_MISS/500 280 GET
http://localhost:3128/geoserver/rest/gazetteer/latlon/-35.0,145.3 - FIRST_UP_PARENT/myAccel -
1294109146.216 0 127.0.0.1 TCP_NEGATIVE_HIT/500 306 GET
http://localhost:3128/geoserver/rest/gazetteer/latlon/-35.0,145.3 - NONE/- -
1294109191.214 0 127.0.0.1 TCP_NEGATIVE_HIT/500 306 GET
http://localhost:3128/geoserver/rest/gazetteer/latlon/-35.0,145.3 - NONE/- -
...
```

Always comes back with 500 error code.

From browser:

```
1294109466.962 1226 127.0.0.1 TCP_MISS/200 698 GET
http://localhost:3128/geoserver/rest/gazetteer/latlon/-35.0,145.3 - FIRST_UP_PARENT/myAccel
application/xml
```

Note that the error code is 200 (not 500). 200 == HTTP OK 500 == HTTP Internal Server Error Subsequent harness requests come back ok (once it is in the cache). WHAT IS GOING OVER THE WIRE? So the question is, how do the http GET requests differ? Wireshark dump (from harness):

```
GET /gazetteer/latlon/-35.0,145.4 HTTP/1.1
User-Agent: Java/1.6.0_22
Host: spatial-dev.ala.org.auAccept: text/html, image/gif, image/jpeg, *; q=.2, */*; q=.2
Connection: keep-alive
HTTP/1.1 500 Internal Server Error
Date: Tue, 04 Jan 2011 02:58:29 GMT
Server: Apache-Coyote/1.1
Age: 129
Content-Length: 0
X-Cache: HIT from ala-devmaps.vm.csiro.auX-Cache-Lookup: HIT from ala-devmaps.vm.csiro.au:3128Via:
1.1 ala-devmaps.vm.csiro.au:3128 (squid/2.7.STABLE5)
Connection: close
Content-Type: text/plain
```

Wireshark dump (from browser):

```
GET /gazetteer/latlon/-35.0,145.5 HTTP/1.1
Host: spatial-dev.ala.org.auUser-Agent: Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10.6; en-US;
rv:1.9.2.6) Gecko/20100625 Firefox/3.6.6
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 115
Connection: keep-alive
Cookie: __utma=100819395.1000872318.1288999346.1292997025.1293056871.6;
__utmz=100819395.1288999346.1.1.utmcsr=(direct)|utmccn=(direct)|utmcmd=(none)
HTTP/1.1 200 OK
Date: Tue, 04 Jan 2011 03:04:26 GMT
Server: Noelios-Restlet-Engine/1.0..8
Expires: Mon, 04 Jan 2021 03:04:26 GMT
Content-Encoding: gzip
Content-Type: application/xml;charset=ISO-8859-1
Content-Length: 286
X-Cache: MISS from ala-devmaps.vm.csiro.auX-Cache-Lookup: MISS from ala-devmaps.vm.csiro.au:3128Via:
1.1 ala-devmaps.vm.csiro.au:3128 (squid/2.7.STABLE5)
Keep-Alive: timeout=15, max=100
Connection: Keep-Alive
.............N.0.E.~E....W...q?.].$....N;y-
2..B....,..k..mr.f...Kv:(m....}....S......4T@.$..4.7.e..R....b>OK..M...B...De#.b.?.....Ph.7kt..o
...).).....g.8 ..~.G..D..;m[..z...9...a4/...........
!..........K-..`...\dkB/[...l.....uKKb.U...........(?...s_g...19.j..j..MW{\...........xnt....
```

Note that there are additional fields set in the GET request, after a bit more investigation, this is the GET request header that you need to set for this to work properly: Accept: text/html,application/xhtml+xml,application/xml;q=0.9,/;q=0.8 So the following code works (note that if you drop the gzip,deflate option, you don't need the GZIPInputStream stuff):

```
...
 URL url = new URL("http://spatial.ala.org.au/gazetteer/search?lon=148.866□=-36.195");
 URLConnection connection = url.openConnection();
 connection.setRequestProperty("Accept",
"text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8");
 connection.setRequestProperty("Accept-Encoding", "gzip,deflate");
 connection.setDoInput(true);
 InputStream inStream = connection.getInputStream();
 GZIPInputStream gis = new GZIPInputStream(inStream);
 BufferedReader input =
 new BufferedReader(new InputStreamReader(gis));
 String line = "";
 while ((line = input.readLine()) != null)
 System.out.println(line);
...
```

Downloaded from https://www.gavinj.net/post/http-500-squid-vs-javaneturl
Generated September 13, 2026. Copyright Gavin Jackson. All rights reserved.