

aproxymate

It's a proxy, mate

v. 0.1.0

Aproxymate is a fairly simple HTTP proxy written in python that handles HTTP GET requests and does basic persistent caching.

Usage:

In the root directory, run the command `python aproxymate/aproxymate.py [port]`

Abilities:

GET Request Forwarding:

In its current state, aproxymate is able to interpret HTTP GET requests, forward the original request (with all headers intact except encoding - see [Limitations](#)) to the remote host, and return the response to that request to the original client.

GET Request Caching:

During the process outlined above, the URL requested is checked against a key-value store (cache), the keys of which are the URLs previously requested. If there is a hit in the cache, the proxy eschews requesting the document from the remote server and instead responds using the cached response.

Future Improvements:

I hope to get most, if not all, of these improvements by the time the final project is due:

- Port the caching mechanism to Memcache, or (preferably) Redis to enable persistent caching between server launches (or implement it myself). This may include the ability to specify cache settings or clear certain cached sites.
- Make the caching mechanism cache-header-aware, and either follow those directions or have its own

invalidation scheme for how long to cache a result

- Have a number of helper threads to distribute the I/O load
- Create a full-on message and task queue to completely decouple: request, checking the cache/requesting from the destination server, and responding to the client

Limitations:

As of version 0.1.0, the following limitations are explicitly in place (others may be implicit):

- The destination server is notified not to use gzip or any other compression encoding schemes (the binary stream was getting garbled) and instead use `identity` encoding
- The proxy only accepts HTTP GET requests
- Upon shutdown, the proxy's cache is emptied as it is stored in memory and does not serialize
- Unable to handle 301 and other redirect / more complex response codes