

Write Up Document Assignment 3: httpproxy

Testing

httpproxy, should be similar to an http protocol server.

- i. When Client sends curl requests (HEAD) to *httpproxy*, *httpproxy* should perform the operation and respond back to client properly.

Testing Code for the spec:

Tests were done locally ..

[illegible]

Using various Server Terminals

Using a Proxy Terminal

Using a Client Terminal:

various curl commands for *httpproxy* which try to fulfill requests for files in proxy / server directory:

Proxy Responds

- 1) request for !GET curl commands
- 2) requests which is invalid
- 3) requests when all servers are down

Proxy Handles:

- 4) requests where server healthcheck returns a bad status
- 5) requests when a server goes down which would have been the optimal server to port forward

- 6) Checking port forwarding is done correctly for load balancing scheme

Server Responds

- 7) request for files that don't exist
- 8) request for files which are forbidden

Compares to expected outputs from server.

Each Request should run in independent threads on server

Healthchecks are done in dispatcher thread

Example Client Sends:

1. curl -T small.txt localhost:1234/small.txt
2. curl localhost:1234/-badrequest.txt
3. curl localhost:1234/dne.txt
4. curl localhost:1234/forbidden.txt
5. curl localhost:1234/small.txt
6. curl localhost:1234/Makefile

The screenshot displays a terminal window on the left and three Sublime Text editor windows on the right. The terminal shows the output of a C program running a multi-threaded HTTP proxy server. It logs the start of the server, the selection of port 8880, and the completion of jobs for three workers. The Sublime Text windows show the contents of three files: '80', '81', and '82'. These files contain HTTP request logs, including the method (GET), the path (e.g., /healthcheck, /dne.txt, /forbidden.txt, /small.txt, /Makefile), the source IP, and the status code (e.g., 200, 404, 403, 401, 405, 408, 415, 429, 500, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000).

For Assignment 3, please answer the following questions:

- For this assignment, your proxy distributed load based on number of requests the servers had already serviced, and how many failed.

A more realistic implementation would consider performance attributes from the machine running the server. Why was this not used for this assignment?

The performance balancing was not used because we are not able to easily determine the performance of the server due to factors we can't control. Since the server may be receiving external requests, the server speed is variable as opposed to a server which is dedicated to our proxy server.

To implement a performance-based approach on a server which takes various requests from any connecting client or server, would require us to compute the performance / request and consider the type of request and the size of transactions in the log.

It is not straightforward and therefore, it has been omitted.

- Using the provided httpserver, do the following:
 - Place eight different large files in a single directory. The files should be around 400 MiB long, adjust according to your computer's capacity.

- Start two instances of your httpserver in that directory with four worker threads for each.
- Start your httpproxy with the port numbers of the two running servers and maximum of eight connections and disable the cache.
- Start eight separate instances of the client at the same time, one GETting each of the files and measure

(using `time(1)`) how long it takes to get the files. The best way to do this is to write a simple shell script (command file) that starts eight copies of the client program in the background, by using `&` at the end.

Attempt 1: (8) Binary Files 400mb

My proxy doesn't crash, but the servers both end up crashing and curl looks corrupted.

The screenshot displays a Kali Linux desktop environment with a multi-pane terminal window. The top pane shows a file explorer with folders like 'bin', 'etc', 'lib', 'usr'. The middle pane shows a list of files and directories. The bottom pane shows a terminal window with a netstat command output, displaying network connections for the 'sshd' process.

- Repeat the same experiment, but turn off one of the servers. Is there any difference in performance? What do you observe?

The one server is able to process some of the responses, while others return an internal server error. Presumably the server is maxing out threads / busy and cannot process the concurrent threads while it is doing other processes? My program does not sit and wait for the thread which is trying to process the curl from the server, it simply says internal server error due to timeout? The server appear unstable and unable to handle the requests... the server crashed when I ran this.

```

Time Current
100 22 100 22
100 22 100 22

real 0m0.038s
user 0m0.000s
sys 0m0.005s

real 0m0.038s
user 0m0.005s
sys 0m0.000s
100 3952k 100 3952k

real 0m0.042s
user 0m0.003s
sys 0m0.006s
100 22 100 22

real 0m0.045s
user 0m0.000s
sys 0m0.006s
100 3952k 100 3952k
100 3952k 100 3952k

real 0m0.060s
user 0m0.000s
sys 0m0.010s

real 0m0.062s
user 0m0.000s
sys 0m0.011s
100 3952k 100 3952k

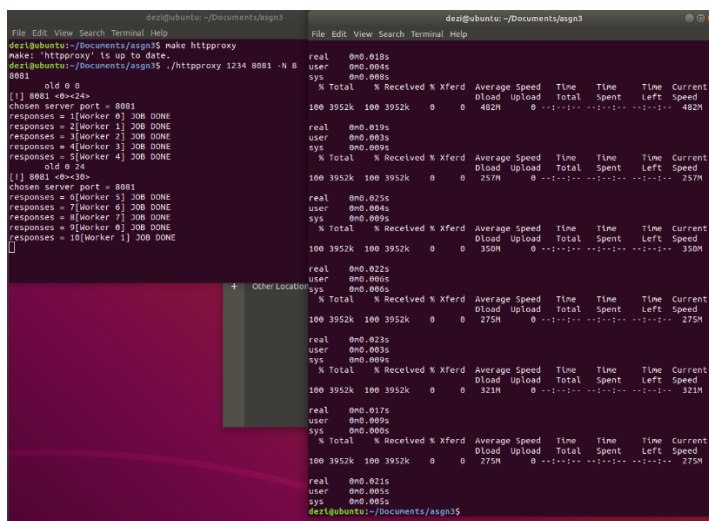
real 0m0.062s
user 0m0.000s
sys 0m0.009s

```

the times are not bad probably (because half of them are returning 500)

- Using one of the files that you created for the previous question, start the provided server with only one thread, then do the following:

- Start your proxy without caching.
- Request the file ten times. How long does it take?



```

dezi@ubuntu: ~/Documents/asn35
File Edit View Search Terminal Help
dezi@ubuntu:~/Documents/asn35$ make httpproxy
make: 'httpproxy' is up to date.
dezi@ubuntu:~/Documents/asn35$ ./httpproxy 1234 8081 -H 8
8081
old @ 0
[1] 8081 <0x24>
chosen server port = 8081
responses = 1[worker 0] JOB DONE
responses = 2[worker 1] JOB DONE
responses = 3[worker 2] JOB DONE
responses = 4[worker 3] JOB DONE
responses = 5[worker 4] JOB DONE
old @ 24
[1] 8081 <0x30>
chosen server port = 8081
responses = 6[worker 5] JOB DONE
responses = 7[worker 6] JOB DONE
responses = 8[worker 7] JOB DONE
responses = 9[worker 8] JOB DONE
responses = 10[worker 9] JOB DONE
[1]
dezi@ubuntu:~/Documents/asn35$ time curl -x http://1234:8081 -H 8081 -d @ 0
real 0m0.018s
user 0m0.000s
sys 0m0.000s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 482M 0 --:--:-- --:--:-- --:--:-- 482M

real 0m0.019s
user 0m0.003s
sys 0m0.000s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 257M 0 --:--:-- --:--:-- --:--:-- 257M

real 0m0.025s
user 0m0.004s
sys 0m0.000s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 350M 0 --:--:-- --:--:-- --:--:-- 350M

real 0m0.022s
user 0m0.000s
sys 0m0.005s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 275M 0 --:--:-- --:--:-- --:--:-- 275M

real 0m0.023s
user 0m0.003s
sys 0m0.000s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 321M 0 --:--:-- --:--:-- --:--:-- 321M

real 0m0.017s
user 0m0.000s
sys 0m0.000s
% Total % Received % Xferd Average Speed Time Time Time Current
Dload Upload Total Spent Left Speed
100 3952k 100 3952k 0 0 275M 0 --:--:-- --:--:-- --:~:~:~ 275M

real 0m0.021s
user 0m0.000s
sys 0m0.005s
dezi@ubuntu:~/Documents/asn35$

```

Each runtime was about 0.20 seconds so the total was around ~2 seconds

– Now stop your proxy and start again, this time with caching configured so it can store the selected file.

I have not done caching yet...

– Request the same file ten times again. How long does it take now?

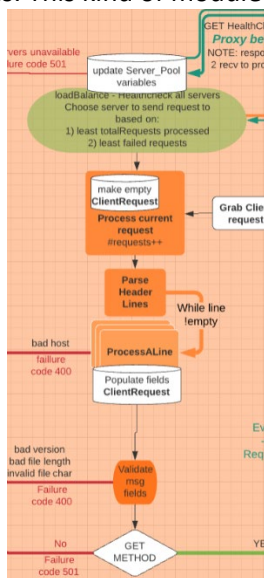
- What did you learn about system design from this class? In particular, describe how each of the basic techniques (abstraction, layering, hierarchy, and modularity) helped you by simplifying your design, making it more efficient, or making it easier to design.

Abstraction –

Having very generic functions help increase abstraction. Tolerant of inputs and strict on outputs ie: making a file descriptor be defined as anything from a client to a server, or even a file or stdout. They are all treated similarly. The result is that an operation will take place on that file descriptor no matter what it points to. Very powerful technique to reducing repetition and extraneous bugs caused by unnecessary code that has the same functionality but used in a different part of the code. All buffer transfer logic remained consistent throughout projects and it was interesting to see how a function could be used in both instances of communication server<->client or GET<->PUT. Very similar functionality but with roles being reversed.

Layering –

Learned to make a design more linear by layering modules in order to reduce the interactions. This kind of control flow makes it easy to detect when a portion breaks and cannot progress in the line of modules. This kind of module organization makes the program easier to trace as well.



Hierarchy –

Using structures and enumeration to produce hierarchical organization. Threads are organized in structures which contain information about the threads.

Using structures nested with enumerated or defined structures make it easy to store, and contain information which can be defined in a fault tolerant way.

Modularity –

learned to group operations better into functions and make functions more GENERIC to produce functionality in broad terms. When there is a logical jump in flow, it is best to put that code in a function to isolate it from other operations. This way, when a block of code is misbehaving, one can ascertain the issue lies in the parameters passed in or the logic within the code which gives an incorrect output or outcome. It's easier to narrow down bugs using modularity since it provides a indication to the piece of logic that has failed based on the function that performed it.