HW 6

14–740 Fundamentals of Telecommunications Systems

Due Date: April 13, 2016. 11:59 PM ET

---

1. The goal of this problem is to get you some experience with software controllers for switches. For the purpose of this problem, you will be using the POX controller as the controller for the switch.

   Web services often use multiple web and application servers for scalability. Many cloud computing platforms like AWS allow the number of servers to be scaled up or down based on the load experienced by the servers. In such load balanced systems, clients typically connect to a server that acts as the load balancer and routes their requests to one of the servers. A dedicated load balancer using consistent hashing is a popular solution today, but it suffers from being an expensive additional piece of hardware and has limited customizability. Your goal in this problem is to develop a load-balancing solution that uses the OpenFlow API and a POX controller to perform load balancing using generic OpenFlow enabled switches and thereby reduce costs of a server acting as the load balancer.

   To this end, you will assume that the clients are given the IP address of the load-balancer (say, `10.0.0.254`) and send requests to this address. There is no physical host corresponding to the load-balancer's IP address and hence the switch/controller has to forward the message to one of the actual servers. For the purpose of this problem we will assume that the actual servers whose IP addresses are of the form `10.0.0.x`, where `x` is between 2 and 253, are directly connected to a switch. For the first request/message of each client/flow, the switch picks a server using a load balancing algorithm (round-robin, random,...) and sends the message to that server. All subsequent requests belonging to the same flow are sent to the same server (unless the flow is inactive for a long time).

   Here are some of the issues you might want to consider for this problem

   (A) How do you want unmodified server code to work inside your system? What are the IP addresses of the various servers and what is the relationship with the IP address on the incoming packet? What OpenFlow features are appropriate in getting the packets to the appropriate server?

   (B) How often is the controller invoked? Is it on a per-flow basis or on a per-packet basis? Can you do better than invoking it on a per-packet basis?

   (C) Since there is no actual host with the IP address of the load balancer, who handles ARP requests and how?

   (D) Bi-directional flows may need to be installed

   Testing your load-balancer: Use the following Mininet command to emulate a simple topology.
   `sudo mn --topo single,10 --mac --arp --switch ovsk --controller remote`
   We will assume that some of these hosts are clients while others are servers (hence all servers in this problem are connected by the same switch). You will run a service on multiple servers (may be the web server from Project 1) and have clients send messages. Your POX controller will load-balance the clients across all the servers.

   Deliverables: For this problem, you will submit the code for your controller. You will also submit any test scripts you use to test the correctness of your load-balancer.