

# Linux Firewall Exploration Lab

Copyright © 2006 - 2014 Wenliang Du, Syracuse University.

The development of this document is/was funded by three grants from the US National Science Foundation: Awards No. 0231122 and 0618680 from TUES/CCLI and Award No. 1017771 from Trustworthy Computing. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation. A copy of the license can be found at <http://www.gnu.org/licenses/fdl.html>.

## 1 Overview

The learning objective of this lab is for students to gain the insights on how firewalls work by playing with firewall software and implement a simplified packet filtering firewall. Firewalls have several types; in this lab, we focus on two types, the *packet filter* and application firewall. Packet filters act by inspecting the packets; if a packet matches the packet filter's set of rules, the packet filter will either drop the packet or forward it, depending on what the rules say. Packet filters are usually *stateless*; they filter each packet based only on the information contained in that packet, without paying attention to whether a packet is part of an existing stream of traffic. Packet filters often use a combination of the packet's source and destination address, its protocol, and, for TCP and UDP traffic, port numbers. Application firewall works at the application layer. A widely used application firewall is web proxy, which is primarily used for egress filtering of web traffic. In this lab, students will play with both types of firewalls, and also through the implementation of some of the key functionalities, they can understand how firewalls work.

## 2 Lab Tasks

### 2.1 Task 1: Using Firewall

Linux has a tool called `iptables`, which is essentially a firewall. It has a nice front end program called `ufw`. In this task, the objective is to use `ufw` to set up some firewall policies, and observe the behaviors of your system after the policies become effective. You need to set up at least two VMs, one called Machine A, and other called Machine B. You run the firewall on your Machine A. Basically, we use `ufw` as a personal firewall. Optionally, if you have more VMs, you can set up the firewall at your router, so it can protect a network, instead of just one single computer. After you set up the two VMs, you should perform the following tasks:

- Prevent A from doing telnet to Machine B.
- Prevent B from doing telnet to Machine A.
- Prevent A from visiting an external web site. You can choose any web site that you like to block, but keep in mind, some web servers have multiple IP addresses.

You can find the manual of `ufw` by typing "`man ufw`" or search it online. It is pretty straightforward to use. Please remember that the firewall is not enabled by default, so you should run a command to specifically enable it. We also list some commonly used commands in Appendix A.

Before start the task, go to default policy file `/etc/default/ufw`. If `DEFAULT_INPUT_POLICY` is `DROP`, please change it to `ACCEPT`. Otherwise, all the incoming traffic will be dropped by default.

## 2.2 Task 2: Evading Egress Filtering

Many companies and schools enforce egress filtering, which blocks users inside of their networks from reaching out to certain web sites or Internet services. They do allow users to access other web sites. In many cases, this type of firewalls inspect the destination IP address and port number in the outgoing packets. If a packet matches the restrictions, it will be dropped. They usually do not conduct deep packet inspections (i.e., looking into the data part of packets) due to the performance reason. In this task, we show how such egress filtering can be bypassed using the tunnel mechanism. There are many ways to establish tunnels; in this task, we only focus on SSH tunnels.

You need two VMs A and B for this task (three will be better). Machine A is running behind a firewall (i.e., inside the company or school's network), and Machine B is outside of the firewall. Typically, there is a dedicated machine that runs the firewall, but in this task, for the sake of convenience, you can run the firewall on Machine A. You can use the firewall program that you implemented in the previous task, or directly use `ufw`. You need to set up the following two block rules:

- Block all the outgoing traffic to external telnet servers. In reality, the servers that are blocked are usually game servers or other types of servers that may affect the productivity of employees. In this task, we use the telnet server for demonstration purposes. You can run the telnet server on Machine B (run `sudo service openbsd-inetd start`). If you have a third VM, Machine C, you can run the telnet server on Machine C.
- Block all the outgoing traffic to `www.facebook.com`, so employees (or school kids) are not distracted during their work/school hours. Social network sites are commonly blocked by companies and schools. After you set up the firewall, launch your Firefox browser, and try to connect to Facebook, and report what happens. If you have already visited Facebook before using this browser, you need to clear all the caches using Firefox's menu: History -> Clear Recent History; otherwise, the cached pages may be displayed. If everything is set up properly, you should not be able to see the Facebook pages. It should be noted that Facebook has many IP addresses, it can change over the time. Remember to check whether the address is still the same by using `ping` or `dig` command. If the address has changed, you need to update your firewall rules. You can also choose web sites with static IP addresses, instead of using Facebook. For example, most universities' web servers use static IP addresses (e.g. `www.syr.edu`); for demonstration purposes, you can try block these IPs, instead of Facebook.

In addition to set up the firewall rules, you also need the following commands to manage the firewall:

```
$ sudo ufw enable           // this will enable the firewall.
$ sudo ufw disable          // this will disable the firewall.
$ sudo ufw status numbered // this will display the firewall rules.
$ sudo ufw delete 3         // this will delete the 3rd rule.
```

**Task 2.a: Telnet to Machine B through the firewall** To bypass the firewall, we can establish an SSH tunnel between Machine A and B, so all the telnet traffic will go through this tunnel (encrypted), evading the inspection. The following command establish an SSH tunnel from the localhost's port 8000 and machine B, and when packets come out of B's end, it will be forwarded to Machine C's port 23 (telnet port). If you only have two VMs, you can replace Machine C with Machine B.

```
$ ssh -L 8000:Machine_C_IP:23 seed@Machine_B_IP
```

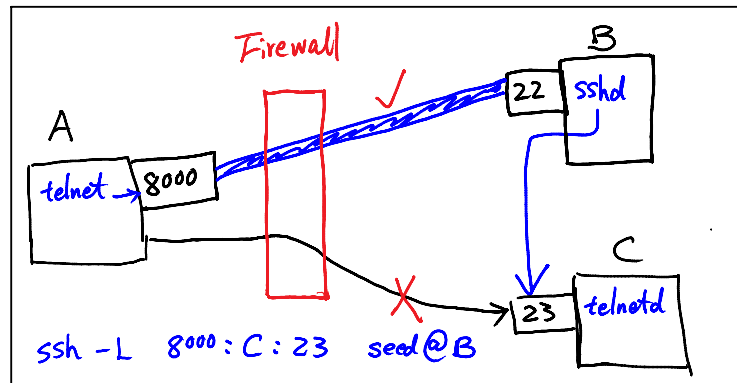


Figure 1: SSH Tunnel Example

After establishing the above tunnel, you can telnet to your localhost using port 8000: `telnet localhost 8000`. SSH will transfer all your TCP packets from your end of the tunnel (localhost:8000) to Machine B, and from there, the packets will be forwarded to Machine C:23. Replies from Machine C will take a reverse path, and eventually reach your telnet client. This resulting in your telnetting to Machine C. Please describe your observation and explain how you are able to bypass the egress filtering. You should use Wireshark to see what exactly is happening on the wire.

**Task 2.b: Connecting to Facebook using SSH Tunnel.** To achieve this goal, we can use the approach similar to that in Task 3.a, i.e., establishing a tunnel between your localhost:port and Machine B, and ask B to forward packets to Facebook. To do that, you can use the following command to set up the tunnel: "`ssh -L 8000:FacebookIP:80 ...`". We will not use this approach, and instead, we use a more generic approach, called dynamic port forwarding, instead of a static one like that in Task 3.a. To do that, we only specify the local port number, not the final destination. When Machine B receives a packet from the tunnel, it will dynamically decide where it should forward the packet to based on the destination information of the packet.

```
$ ssh -D 9000 -C seed@machine_B
```

Similar to the telnet program, which connects localhost:9000, we need to ask Firefox to connect to localhost:9000 every time it needs to connect to a web server, so the traffic can go through our SSH tunnel. To achieve that, we can tell Firefox to use localhost:9000 as its proxy. The following procedure achieves this:

```
Edit -> Preference -> Advanced tab -> Network tab -> Settings button.

Select Manual proxy configuration
SOCKS Host: 127.0.0.1      Port: 9000
SOCKS v5
No Proxy for: localhost, 127.0.0.1
```

After the setup is done, please do the followings:

1. Run Firefox and go visit the Facebook page. Can you see the Facebook page? Please describe your observation.

2. After you get the facebook page, break the SSH tunnel, clear the Firefox cache, and try the connection again. Please describe your observation.
3. Establish the SSH tunnel again and connect to Facebook. Describe your observation.
4. Please explain what you have observed, especially on why the SSH tunnel can help bypass the egress filtering. You should use Wireshark to see what exactly is happening on the wire. Please describe your observations and explain them using the packets that you have captured.

**Question 1:** If `ufw` blocks the TCP port 22, which is the port used by SSH, can you still set up an SSH tunnel to evade egress filtering?

### 2.3 Task 3: Web Proxy (Application Firewall)

There is another type of firewalls, which are specific to applications. Instead of inspecting the packets at the transport layer (such as TCP/UDP) and below (such as IP), they look at the application-layer data, and enforce their firewall policies. These firewalls are called application firewalls, which controls input, output, and/or access from, to, or by an application or service. A widely used category of application firewalls is web proxy, which is used to control what their protected browsers can access. This is a typical egress filtering, and it is widely used by companies and schools to block their employees or students from accessing distracting or inappropriate web sites.

In this task, we will set up a web proxy and perform some tasks based on this web proxy. There are a number of web proxy products to choose from. In this lab, we will use a very well-known free software, called `squid`. If you are using our pre-built VM (Ubuntu 12.04), this software is already installed. Otherwise, you can easily run the following command to install it.

```
$ sudo apt-get install squid

Here are several commands that you may need:

$ sudo service squid3 start      // to start the server
$ sudo service squid3 restart    // to restart the server
```

Once you have installed `squid`, you can go to `/etc/squid3`, and locate the configuration file called `squid.conf`. This is where you need to set up your firewall policies. Keep in mind that every time you make a change to `squid.conf`, you need to restart the `squid` server; otherwise, your changes will not take effect.

**Task 3.a: Setup.** You need to set up two VMs, Machine A and Machine B. Machine A is the one whose browsing behaviors need to be restricted, and Machine B is where you run the web proxy. We need to configure A's Firefox browser, so it always uses the web proxy server on B. To achieve that, we can tell Firefox to use B:3128 as its proxy (by default, `squid` uses port 3128, but this can be changed in `squid.conf`). The following procedure configures B:3128 as the proxy for Firefox.

```
Edit -> Preferences -> Advanced tab -> Network tab -> Settings button.

Select "Manual proxy configuration"
```

```
Fill in the following information:  
HTTP Proxy: B's IP address      Port: 3128
```

Note: to ensure that the browser always uses the proxy server, the browser's proxy setting needs to be locked down, so users cannot change it. There are ways for administrators to do that. If you are interested, you can search the Internet for instructions.

After the setup is done, please perform the following tasks:

1. Try to visit some web sites from Machine A's Firefox browser, and describe your observation.
2. By default, all the external web sites are blocked. Please Take a look at the configuration file `squid.conf`, and describe what rules have caused that (hint: search for the `http_access` tag in the file).
3. Make changes to the configuration file, so all the web sites are allowed.
4. Make changes to the configuration file, so only the access to `google.com` is allowed.

You should turn on your Wireshark, capture packets while performing the above tasks. In your report, you should describe what exactly is happening on the wire. Using these observations, you should describe how the web proxy works.

**Task 3.b: Using Web Proxy to evade Firewall.** Ironically, web proxy, which is widely used to do the egress filtering, is also widely used to bypass egress filtering. Some networks have packet-filter type of firewall, which blocks outgoing packets by looking at their destination addresses and port numbers. For example, in Task 1, we use `ufw` to block the access of Facebook. In Task 3, we have shown that you can use a SSH tunnel to bypass that kind of firewall. In this task, you should do it using a web proxy. Please demonstrate how you can access Facebook even if `ufw` has already blocked it. Please describe how you do that and also include evidence of success in your lab report.

**Question 2:** If `ufw` blocks the TCP port 3128, can you still use web proxy to evade the firewall?

**Task 3.c: URL Rewriting/Redirection.** Not only can `squid` block web sites, they can also rewrite URLs or redirect users to another web site. For example, you can set up `squid`, so every time a user tries to visit Facebook, you redirect them to a web page, which shows a big red stop sign. `Squid` allows you to make any arbitrary URL rewriting: all you need to do is to connect `squid` to an URL rewriting program. In this task, you will write a very simple URL-rewriting programming to demonstrate such a functionality of web proxy. You can use any programming language, such as Perl, Java, C/C++, etc.

Assume that you have written a program called `myprog.pl` (a perl program). You need to modify `squid.conf` to connect to this program. You need to add the following to the configuration file:

```
url_rewrite_program /home/seed/myprog.pl  
url_rewrite_children 5
```

In case `squid` cannot find `myprog.pl` under `/home/seed`, you can put your url rewrite program under `/etc/squid3`.

This is how it works: when `squid` receives an URL from browsers, it will invoke `myprog.pl`, which can get the URL information from the standard input (the pipe mechanism is used). The rewriting program

can do whatever it wants to the URL, and eventually print out a new URL or an empty line to the standard output, which is then piped back to squid. Squid uses this output as the new URL, if the line is not empty. This is how an URL gets rewritten. The following Perl program gives you an example:

```
#!/usr/bin/perl -w
use strict;
use warnings;

# Forces a flush after every write or print on the STDOUT
select STDOUT; $| = 1;

# Get the input line by line from the standard input.
# Each line contains an URL and some other information.
while (<>)
{
    my @parts = split;
    my $url = $parts[0];
    # If you copy and paste this code from this PDF file,
    # the ~ (tilde) character may not be copied correctly.
    # Remove it, and then type the character manually.
    if ($url =~ /www\.cis\.syr\.edu/) {
        # URL Rewriting
        print "http://www.yahoo.com\n";
    }
    else {
        # No Rewriting.
        print "\n";
    }
}
```

Please conduct the following tasks:

1. Please describe what the above program does (e.g. what URLs got rewritten and what your observations are).
2. Please modify the above program, so it replaces all the Facebook pages with a page that shows a big red stop sign.
3. Please modify the above program, so it replaces all the images (e.g. .jpg or .gif images) inside any page with a picture of your choice. When an HTML web page contains images, the browser will identify those image URLs, and send out an URL request for each image. You should be able to see the URLs in your URL rewriting program. You just need to decide whether a URL is trying to fetch an image file of a particular type; if it is, you can replace the URL with another one (of your choice).

**Note:** If you have a syntax error in the above program, you will not be able to see the error message if you test it via the web proxy. We suggest that you run the above perl program on the command line first, manually type an URL as the input, and see whether the program functions as expected. If there is a syntax error, you will see the error message. Do watch out for the tilde sign if you copy and paste the above program from the PDF file.

**Task 3.d ): A Real URL Redirector.** The program you wrote above is a toy URL rewriting program. In reality, such a program is much more complicated because of performance issues and many other functionalities. Moreover, they work with real URL datasets (the URLs that need to be blocked and rewritten) that contains many URLs. These datasets can be obtained from various sources (some are free and some are not), and they need to be routinely updated. If you are interested in playing with a real URL rewriting program, you can install SquidGuard and also download a blacklist URL dataset that works with it. Here is some related information:

- You can download SquidGuard from <http://www.squidguard.org/>, compile and install it. You can get the instruction from the following URL <http://www.squidguard.org/Doc/>.
- You also need to download Oracle Berkeley DB, because SquidGuard depends on it. You should do this step first.
- Download a blacklist data set from the following URL:  
<http://www.squidguard.org/blacklists.html>.

### 3 Submission and Demonstration

You will need to submit a detailed lab report to describe what you have done and what you have observed, and provide detailed explanations. Reports should include the evidences to support the observations for each task. Evidences include packet traces, screendumps, etc. For each task, please answer all the questions in the lab description. Please submit your report as a PDF file to Blackboard.

## A Firewall Lab Cheat Sheet

**Using ufw.** The default firewall configuration tool for Ubuntu is ufw, which is developed to ease iptables firewall configuration. By default UFW is disabled, so you need to enable it first.

```
$ sudo ufw enable           // Enable the firewall
$ sudo ufw disable          // Disable the firewall
$ sudo ufw status numbered  // Display the firewall rules
$ sudo ufw delete 2         // Delete the 2nd rule
```

**Using squid.** The following commands are related to squid.

```
$ sudo service squid3 start      // start the squid service
$ sudo service squid3 restart    // restart the squid service
$ sudo service squid3 stop       // stop the squid service

/etc/squid3/squid.conf: This is the squid configuration file.
```