Data Networks
Final project

Dr. Mohammad Reza Pakravan

Due on February 1, 2024

Introduction

This project should be implemented on Linux operating system and its command line. To use Linux on other Operating systems, you can use some Hypervisors and Virtual Machines. You can use virtualization software such as Oracle VirtualBox or VMware to create a virtual machine running Linux within your Windows environment. This allows you to run Linux as a guest operating system alongside Windows.VM software provides options to allocate and manage hardware resources for each VM, including CPU cores, memory, disk space, and network bandwidth. You can adjust these allocations based on the requirements of the guest operating systems and applications. You are also free to use dual boot. With dual boot, you can install Linux alongside Windows on your computer, allowing you to choose which operating system to use at startup. This requires partitioning your hard drive and installing Linux on a separate partition. With this approach, you can only use one operating system at a time. there are lots of tutorials and YouTube videos that can help you use Linux on Windows using VMs or dual boot.

Answer the following questions:

a. Explain about Network Hypervisors. Why are Network Hypervisors required?

b. Explain about virtualization in networks. What are the different types of virtualization? What is the role of Virtual Machines in virtualization?

Five Reasons You Need to Learn Linux

1. Linux is the future
Although Linux has been around for over 25 years, it has enjoyed a continuous rise in business-critical usage, and many see Linux as being the most popular operating system for the future.

2. Linux is on everything
Linux runs more than two-thirds of the servers on the Internet, all Android phones, most consumer network gear, such as NetGear and Linksys devices, 99 percent of the top supercomputers in the world, many Internet of Things (IoT) devices, Tesla cars, and even PlayStation gaming consoles.

3. Linux is adaptable
The very reason everything is on Linux is because it's such an adaptable operating system. Thanks to Linux's modularity and open-source nature, you can choose the pieces you need for your product or service

and develop any pieces that may not already exist. You can install tiny versions of Linux for specialized use cases, modify it to work on appliances that route packets across a large enterprise network, or use it as your desktop operating system. Your choices are practically endless.

4. Linux has a strong community and ecosystem
Linux has been so successful mainly because of the strong community and the ecosystem that surrounds it. There are Linux contributors (developers who write code to make the product better); Linux forums and communities; Linux instructors; Linux training options; Linux blogs; Linux third-party tools; Linux distributions; Linux conferences; and even Linux books.

5. Linux is open-source and sometimes free
Linux is open-source, meaning that the original source code is made freely available and may be redistributed and modified. That said, there are paid and fully supported commercial editions available, too. The open nature of Linux has made it the adaptable OS of the future, allowing it to run on everything, and has resulted in the creation of a strong ecosystem.

Basics of Linux Administration

Interaction with and navigation of the Linux file system is done with commands such as:
1. pwd: Display the directory you're currently in (short for print working directory)
2. ls: List out files that are present in the folder
3. cd: Change directory
4. rm: Remove files
5. mkdir and rmdir: Make and remove folders or directories, respectively

Use the man command (shorthand for "manual") to provide detailed documentation for just about every Linux command.

Just as you might expect with any multi-user operating system, Linux supports the concept of users with differing levels of access. By default, you'll log in as a common user and be able to view most of what's happening on the system, although you're not allowed to view log files as a standard unprivileged user. To be able to reconfigure the system or view log files, you'll need administrative rights. In Linux, these administrative privileges are referred to as superuser privileges and are equivalent to the root user, who has a user ID of 0 (zero). You may use 'sudo' command to obtain superuser privileges.

Basics of Linux Network Administration

In this part, we will get familiar with some basic commands in network administration in Linux.

Network Interfaces

a. Show different network interfaces in your system. What are the different types of interfaces and what's the role of each of them? what are the associated IP and MAC addresses?

b. Based on the output of the previous command, How your system has received its IP configuration?

probing the network between the local system and a destination

a. Use 'traceroute' command to gather information about the routers in the path when a specific web address is requested. You may use www.google.com or www.apple.com.

b. As you can see, some routers are shown by a name instead of their IP address. Why is that?

c. Why is it beneficial to be aware of the routers in the path?

d. Use the appropriate command to find the address record(IP address) for the google.com server.

Network Interface Bonding

There are times when you need more bandwidth than a single interface can provide, or you want some form of link redundancy in case of a cabling or other network problems. This link redundancy function goes by many names, depending on the vendor: EtherChannel, VMware PortGroups, Bonds, and Link Aggregation Groups (LAG) are just a few. Linux also provides this capability and calls it bonding. It allows you to create a single logical network link that is comprised of multiple physical links and that scales up as you add more interfaces, can provide load balancing across the interfaces, and can provide failover protection.

a. Create three interfaces eth0, eth1, and eth2, and put them into a bond.

b. One of the most common parameters to set when creating a bond is the "mode", which is how the bond interacts with the connected networks. Explain about different modes in a bond and discuss their differences.

Understanding Linux Internetworking

Layer 2 vs. Layer 3 Internetworking

In the OSI model, layer 1 is the physical layer that includes the physical media used to connect the network. Specifications in this area describe cable qualities and the properties of electrical and optical signals used to move bits around. Examples of layer 1 technologies include Gigabit Ethernet on category 5 cable, 100Gigabit Ethernet on parallel single mode fiber, and 802.11 wireless.
Above that is layer 2, or the data link layer; Ethernet is a broadly deployed layer 2 protocol. Ethernet networking works to encapsulate data and pass that data in the form of frames. Frames leverage the Media Access Control (MAC) addresses. An Ethernet frame includes the MAC address of the destination interface on the target system as well the MAC address of the source interface on the sending system so that the recipient device knows where the frame originated. Every Ethernet device, whether it's installed in a server, a switch, or a router, has a unique MAC address on its local network. Transparent bridges are layer 2 devices that send all frames received on one port out to the other bridge ports, based on knowledge of the frame's destination MAC address. Ethernet switches are multiport network bridges. Multiport network bridges learn of the MAC addresses in the network and intelligently forward frames based on the destination MAC address in the frame.
Layer 2 networking works in one of two ways:
1. The device has explicit knowledge of a frame's destination address, and the device sends the frame out on the port where it knows the destination exists.
2. In the event that the specific destination is unknown, the device falls back to sending the frame to every node in the layer 2 domain via what is known as a broadcast.

The problem is that these approaches limit the ability for layer 2 networks alone to operate efficiently beyond relatively small-scale locations and very simple topologies. Layer 2 networks suffer from two major limitations.
First, they allow hosts to send traffic to unknown destinations. This causes broadcasts, which impact every node in the broadcast domain. Many networks have been taken offline due to "broadcast storms," or when many hosts are broadcasting at once. In contrast, layer 3 networks do not allow for unknown communication. If a layer 3 router does not have a route to the destination IP address, it will drop the packet instead of broadcasting like layer 2 does.
Second, layer 2 networks have globally unique MAC addresses that are assigned by the manufacturer. There is no organization to these addresses across manufacturers. If you have servers with Intel and Mellanox network cards, the layer 2 MAC addresses will not have any commonality. Again, when comparing layer 2 MAC addresses to layer 3 IP addresses, companies manually plan IP addressing schemes so that there is a hierarchy to these IP addresses. An office may have all IP addresses within it as part of a single IP subnet, like 10.0.0.0, allowing the company to use a single subnet to represent the entire office. With layer 2 addressing, there is no ability to summarize or aggregate MAC addresses; every unique MAC address must be shared with every host in the layer 2 domain.
When a node sends out an IP packet, it consults its routing and neighbor (ARP) tables and sends the packet to the device most likely to get that packet where it needs to go. If the destination is in the same layer 2 network, an entry in the neighbor (or ARP) table tells the sender how to use layer 2 internetworking. When IP devices need to communicate with other IP-based addresses that are outside of their local layer 2 network, the route table may point to a specific router that will get the packet closer to the destination or fall through to the default gateway, which is then responsible for getting the packet to the destination. If no default route exists and a matching route does not exist, the packet will be dropped.

Layer 2 Internetworking on Linux Systems

Bridging

a. What do you do when you have two different Ethernet networks that need connecting? Build a bridge! Bridges have traditionally been dedicated hardware devices, but you can easily create a bridge in Linux. For example, when you have a Linux host that has two or more network interfaces, you can create a bridge to pass traffic between these interfaces. Create three interfaces eth0, eth1, and eth2 and connect them with a bridge.

b. Once a bridge is created, you can view the MAC address table, which shows which ports can reach a specific MAC address. Show the MAC address table for the network in part a. Analyze the table.

c. The downside to big networks is that you can accidentally create loops that feed upon themselves and that can ultimately bring the network down. For example, if you accidentally plug one switch port directly into another port on the same switch, you may have created a loop. What's an appropriate solution to avoid these loops in bridges? Introduce some commands in Linux to solve this problem in bridges.
Some fields in IP packets help the network to reduce the adversarial effects of the loops. Which fields are beneficial for this problem? How these fields can reduce the problems of the loops?

Layer 3 Internetworking View on Linux Systems

Neighbor Table

When an IP node wants to communicate with a system in the same layer 2 domain, it looks in its neighbor table, or ARP table, to determine how to construct the Ethernet frame. If the desired destination IP address is not in the neighbor table, the node issues an ARP request, which is broadcast to everyone in the layer 2 domain, that asks, "Please tell me the MAC address for the node with IP address X.X.X.X." Assuming the

target device is available, the node with that IP address will respond.

a. Show the neighbor table in your system. Explain each of the fields in a row of the neighbor table.

IP Routing

The routing table has knowledge of specific networks, or summaries of networks, that a node can reach.

a. Show the routing table in your system. Explain each of the fields in a row of the routing table.

b. What is the role of the default route?

c. create a static route to router 192.168.1.1 through the eth1 interface. Is this route persistent after restarting the host? How can you make it persistent?

Virtual LANs (VLANs)

You already know that a LAN is a local area network spanning a relatively small physical area. Building on that concept, a virtual LAN (or VLAN) allows LANs to span multiple switches across very large networks while still achieving traffic isolation from other networks. VLANs are used to isolate hosts or applications from each other for the purposes of security, data flow, and scale. Individual interfaces can be a part of one or more VLANs. When they are a part of more than one VLAN, in order to maintain some semblance of sanity, the frames traversing that link are tagged with an IEEE 802.1Q tag. These tags are an additional piece of information placed at the front of the frame to identify the VLAN.

a. Suppose you want a Linux system to have eth1 in one bridge (VLAN11), eth3 in a second bridge (VLAN12), and eth2 in both (a tagged trunk). use Linux commands to implement this system.
Hint: use bridges. Note that you should have a default(native) VLAN too.

b. what are the trunks?

c. Do you expect the MAC address tables in 3 VLANs to be the same? why?

d. Provide screenshots from the outputs of each of the following commands:

1. bridge link show
2. bridge vlan show
3. bridge fdb show

Namespaces in linux

In Linux, a namespace is a feature that allows processes to use the system resources they need but separately from other processes. A namespace is a way to virtualize a global system resource such that a process that needs it operates in an isolated instance of that resource.
Additionally, a group of resources and processes can refer to the same namespace, but each namespace has individual resources for each. Only processes in the same namespace can identify the changes made in a global resource.

a. What are the different namespaces in Linux? Explain each of them and their application.

In the rest of the project, we focus on the network namespace.

b. Create two network namespaces named ns1 and ns2.

c. Create and configure veth (virtual eth) cable. Veth cable is nothing but ethernet cable but it is not physical it is virtual. In physical cable, we have two heads. here, we also have two heads in the cable, but in a virtual ethernet cable, we need to name them. Name them eth1 and eth2.

d. connect the veth cable with each namespace.

e. Bind IP address with the interface of each namespace. Assign 192.168.1.1/24 to one of them(node1) and 192.168.1.2/24 to another one(node2). Ping one of the nodes from another one. Are these nodes reachable from one another? Why?

f. Activate the eth1 and eth2 interfaces and make them 'up'.

g. Write a bash script that gets the names of the nodes(node1 and node2) as input and then pings the second input from the first input.

Suppose we run this command:
./your-script.sh node2 node1

After running it node2 should start pinging the node1 and the output must be printed in the command line. Provide enough screenshots.

h. Ping 8.8.8.8 from each of the namespaces. Is 8.8.8.8(the primary DNS server for Google DNS) reachable from these namespaces? Why? How can you solve the problem? (Explain your solution, no implementation is required.)

What Should I Do?

In this project, you can use either Linux command lines or bash scripts based on your preference (in some parts, it is mentioned that codes should be in the format of bash scripts. In this case, bash script is mandatory). In the case of Linux commands, you should provide enough screenshots of your commands step by step and their outputs. You must upload (.sh) files and screenshots in each part separated into specific folders. You should also answer the questions in your report. The report should include a summary of your commands and codes too. Compress all files and rename the compressed file to STUDENT_ID_project.zip. If you have any questions regarding the problem statement or understanding the concept, feel free to ask in Telegram.