Chapter 14

Introduction to networking

A network is a group of computers and computing devices connected together through communication channels, such as cables or wireless media. The computers connected over a network may be located in the same geographical area or spread across the world.

A network is used to:

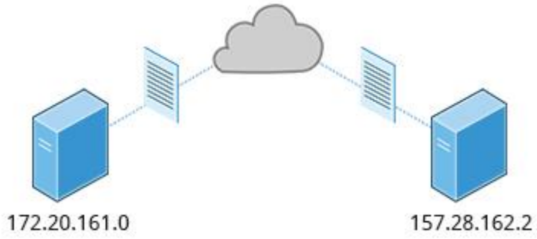
- Allow the connected devices to communicate with each other
- Enable multiple users to share devices over the network, such as printers and scanners
- Share and manage information across computers easily.

Most organizations have both an internal network and an Internet connection for users to communicate with machines and people outside the organization. The **Internet** is the largest network in the world and can be called "the network of networks".

IP Addresses

Devices attached to a network must have at least one unique network address identifier known as the **IP** (**I**nternet **P**rotocol) address. The address is essential for routing packets of information through the network.

Exchanging information across the network requires using streams of small packets, each of which contains a piece of the information going from one machine to another. These packets contain data buffers, together with headers which contain information about where the packet is going to and coming from, and where it fits in the sequence of packets that constitute the stream. Networking protocols and software are rather complicated due to the diversity of machines and operating systems they must deal with, as well as the fact that even very old standards must be supported.



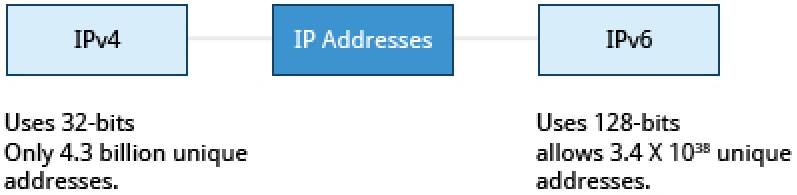
IPv4 and IPv6

There are two different types of IP addresses available: IPv4 (version 4) and IPv6 (version 6). IPv4 is older and by far the more widely used, while IPv6 is newer and is designed to get past limitations inherent in the older standard and furnish many more possible addresses.

IPv4 uses 32-bits for addresses; there are *only* 4.3 billion unique addresses available. Furthermore, many addresses are allotted and reserved, but not actually used. IPv4 is considered inadequate for meeting future needs because the number of devices available on the global network has increased enormously in recent years.

IPv6 uses 128-bits for addresses; this allows for 3.4 X 10^{38} unique addresses. If you have a larger network of computers and want to add more, you may want to move to IPv6, because it provides more unique addresses. However, it can be complex to migrate to IPv6; the two protocols do not always inter-operate well. Thus, moving equipment and addresses to IPv6 requires significant effort and has not been quite as fast as was originally intended. We will discuss IPv4 more than IPv6 as you are more likely to deal with it.

One reason IPv4 has not disappeared is there are ways to effectively make many more addresses available by methods such
as NAT (Network Address Translation). NAT enables sharing one IP address among many locally connected computers, each of which has a unique address only seen on the local network. While this is used in organizational settings, it also used in simple home networks. For example, if you have a router hooked up to your Internet Provider (such as a cable system) it gives you one externally visible address, but issues each device in your home an individual local address.



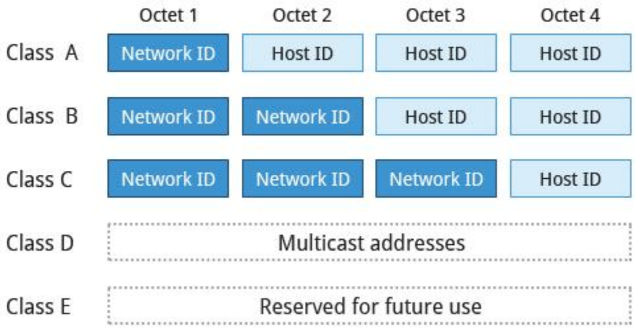
Decoding IPv4 addresses

A 32-bit IPv4 address is divided into four 8-bit sections called octets.

Example:
IP address → 172 . 16 . 31 . 46
Bit format → 10101100.00010000.00011111.00101110

Note: Octet is just another word for byte.

Network addresses are divided into five classes: A, B, C, D and E. Classes A, B and C are classified into two parts: Network addresses (Net ID) and Host address (Host ID). The Net ID is used to identify the network, while the Host ID is used to identify a host in the network. Class D is used for special multicast applications (information is broadcast to multiple computers simultaneously) and Class E is reserved for future use. In this section you will learn about classes A, B and C.



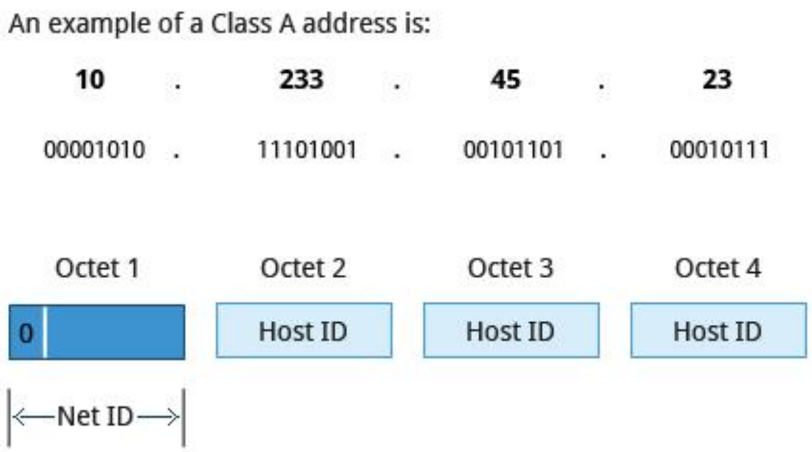
	Octet 1	Octet 2	Octet 3	Octet 4
Class A	Network ID	Host ID	Host ID	Host ID
Class B	Network ID	Network ID	Host ID	Host ID
Class C	Network ID	Network ID	Network ID	Host ID
Class D	Multicast addresses			
Class E	Reserved for future use			

Class A Network Addresses

Class A addresses use the first octet of an IP address as their Net ID and use the other three octets as the Host ID**.** The first bit of the first octet is always set to zero. So you can use only 7-bits for unique network numbers. As a result, there are a maximum of 126 Class A networks available (the addresses 0000000 and 1111111 are reserved). Not surprisingly, this was only feasible when there were very few unique networks with large numbers of hosts. As the use of the Internet expanded, Classes B and C were added in order to accommodate the growing demand for independent networks.

Each Class A network can have up to 16.7 million unique hosts on its network. The range of host address is from **1.0.0.0** to **127.255.255.255**.

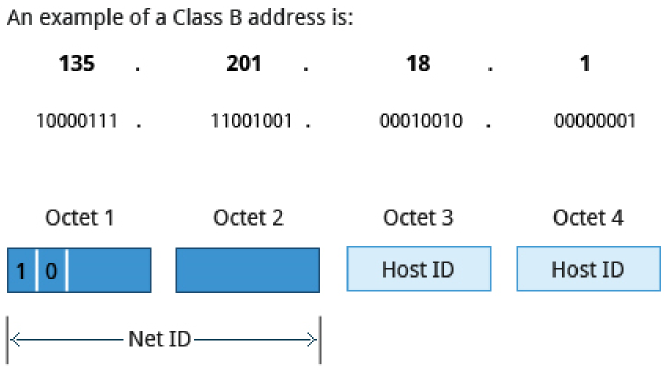
Note: The value of an octet, or 8-bits, can range from 0 to 255.



Class B Network Addresses

Class B addresses use the first two octets of the IP address as their Net ID and the last two octets as the Host ID. The first two bits of the first octet are always set to binary 10, so there are a maximum of 16,384 (14-bits) Class B networks. The first octet of a Class B address has values from 128 to 191. The introduction of Class B networks expanded the number of networks but it soon became clear that a further level would be needed.

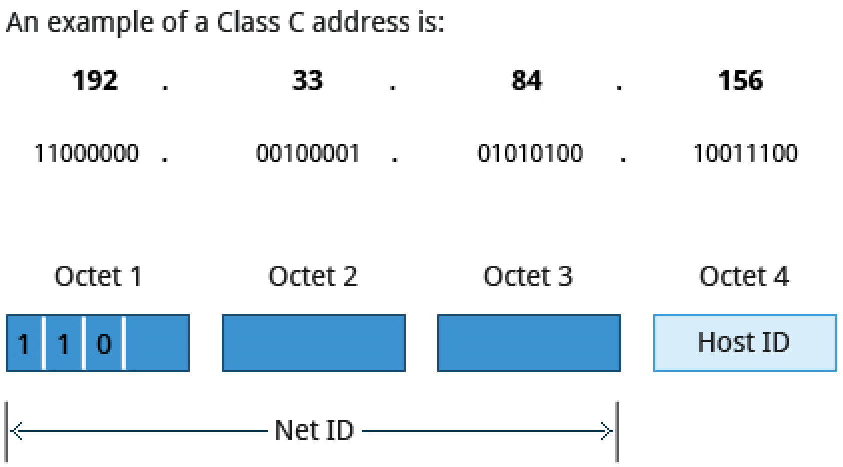
Each Class B network can support a maximum of 65,536 unique hosts on its network. The range of host address is from **128.0.0.0** to **191.255.255.255**.



Class C Network Addresses

Class C addresses use the first three octets of the IP address as their Net ID and the last octet as their Host ID. The first three bits of the first octet are set to binary 110, so almost 2.1 million (21-bits) Class C networks are available. The first octet of a Class C address has values from 192 to 223. These are most common for smaller networks which don't have many unique hosts.

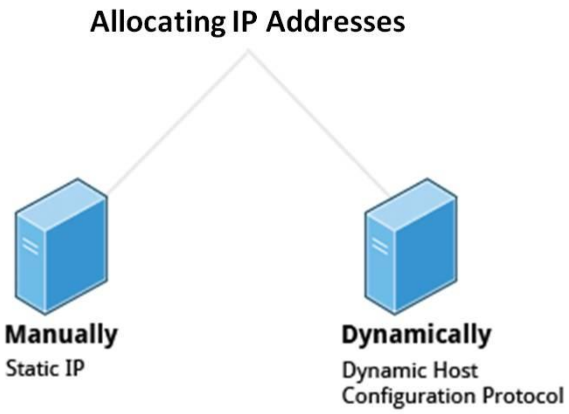
Each Class C network can support up to 256 (8-bits) unique hosts. The range of host address is from **192.0.0.0** to **223.255.255.255**.



IP Address Allocation

Typically, a range of IP addresses are requested from your Internet Service Provider (ISP) by your organization's network administrator. Often, your choice of which class of IP address you are given depends on the size of your network and expected growth needs. If NAT is in operation, such as in a home network, you only get one externally visible address!

You can assign IP addresses to computers over a network either manually or dynamically. Manual assignment adds static (never changing) addresses to the network. Dynamically assigned addresses can change every time you reboot or even more often; the **D**ynamic **H**ost **C**onfiguration **P**rotocol (**DHCP**) is used to assign IP addresses.



Name Resolution

Name Resolution is used to convert numerical IP address values into a human-readable format known as the hostname. For example, **104.95.85.15** is the numerical IP address that refers to the hostname whitehouse.gov. Hostnames are much easier to remember!

Given an IP address, you can obtain its corresponding hostname. Accessing the machine over the network becomes easier when you can type the hostname instead of the IP address.

You can view your system's hostname simply by typing **hostname** with no argument.

Note: If you give an argument, the system will try to change its hostname to match it, however, only root users can do that.

The special hostname localhost is associated with the IP address **127.0.0.1**, and describes the machine you are currently on (which normally has additional network-related IP addresses).

Network Configuration Files

Network configuration files are essential to ensure that interfaces function correctly. They are located in the **/etc** directory tree. However, the exact files used have historically been dependent on the particular Linux distribution and version being used.

For Debian family configurations, the basic network configuration files could be found under **/etc/network/**, while for Fedora and SUSE family systems one needed to inspect **/etc/sysconfig/network**.

Modern systems emphasize the use of Network Manager, which we briefly discussed when we considered graphical system administration, rather than try to keep up with the vagaries of the files in **/etc**. While the graphical versions of Network Manager do look somewhat different in different distributions, the **nmtui** utility (shown in the screenshot) varies almost not at all, as does the even more sparse **nmcli** (command line interface) utility. If you are proficient in the use of the GUIs, by all means, use them. If you are working on a variety of systems, the lower level utilities may make life easier.

Network Interfaces

Network interfaces are a connection channel between a device and a network. Physically, network interfaces can proceed through a network interface card (NIC), or can be more abstractly implemented as software. You can have multiple network interfaces operating at once. Specific interfaces can be brought up (activated) or brought down (de-activated) at any time.

Information about a particular network interface or all network interfaces can be reported by the **ip** and **ifconfig** utilities, which you may have to run as the superuser, or at least, give the full path, i.e. **/sbin/ifconfig**, on some distributions. **ip** is newer than **ifconfig** and has far more capabilities, but its output is uglier to the human eye. Some new Linux distributions do not install the older **net-tools** package to which **ifconfig** belongs, and so you would have to install it if you want to use it.

The IP utility

To view the IP address:

```
$ /sbin/ip addr show
```

To view the routing information:

```
$ /sbin/ip route show
```

ip is a very powerful program that can do many things. Older (and more specific) utilities such as **ifconfig** and **route** are often used to accomplish similar tasks. A look at the relevant man pages can tell you much more about these utilities.

Command: ping

ping is used to check whether or not a machine attached to the network can receive and send data; i.e. it confirms that the remote host is online and is responding.

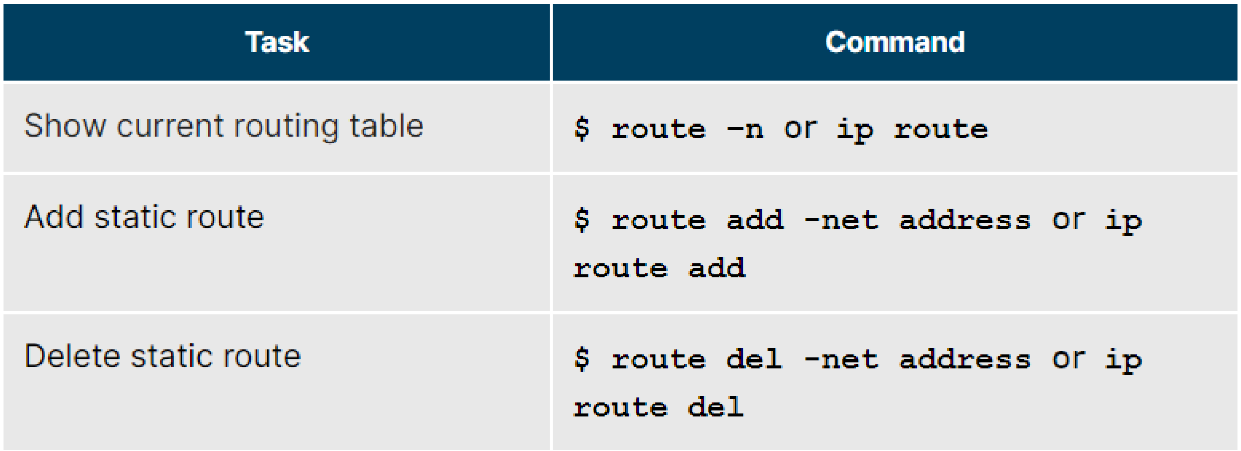
To check the status of the remote host, at the command prompt, type **ping <hostname>**.

ping is frequently used for network testing and management; however, its usage can increase network load unacceptably. Hence, you can abort the execution of **ping** by typing **CTRL-C**, or by using the **-c** option, which limits the number of packets that **ping** will send before it quits. When execution stops, a summary is displayed.

Command: route

A network requires the connection of many nodes. Data moves from source to destination by passing through a series of routers and potentially across multiple networks. Servers maintain routing tables containing the addresses of each node in the network. The IP routing protocols enable routers to build up a forwarding table that correlates final destinations with the next hop addresses.

One can use the **route** utility or the newer **ip route** command to view or change the IP routing table to add, delete, or modify specific (static) routes to specific hosts or networks. The table explains some commands that can be used to manage IP routing:



Task	Command
Show current routing table	`$ route -n` or `ip route`
Add static route	`$ route add -net address` or `ip route add`
Delete static route	`$ route del -net address` or `ip route del`

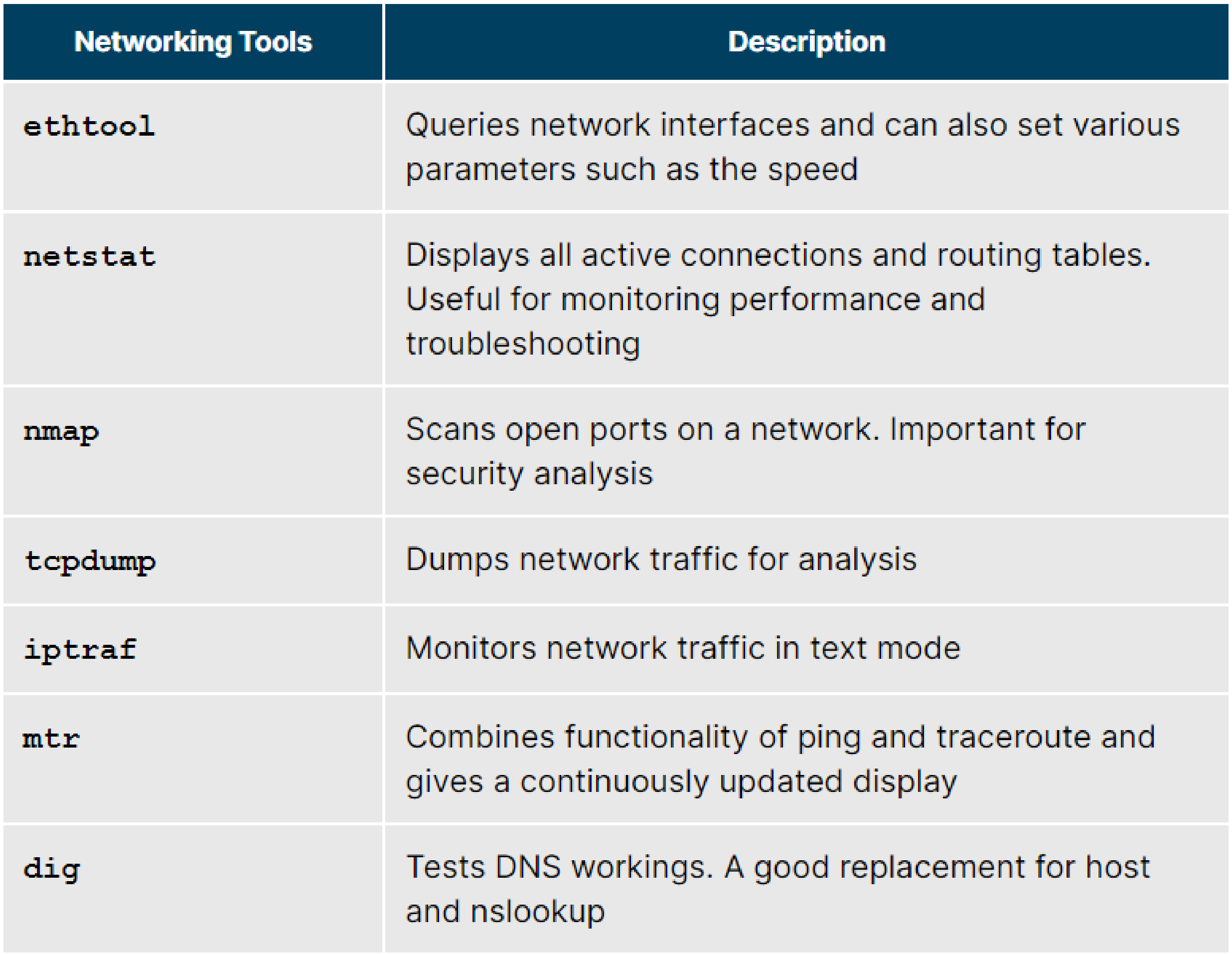
Command: traceroute

traceroute is used to inspect the route which the data packet takes to reach the destination host, which makes it quite useful for troubleshooting network delays and errors. By using **traceroute**, you can isolate connectivity issues between hops, which helps resolve them faster.

To print the route taken by the packet to reach the network host, at the command prompt, type **traceroute <address>**.

More networking tools

Now, let’s learn about some additional networking tools. Networking tools are very useful for monitoring and debugging network problems, such as network connectivity and network traffic.



Networking Tools	Description
`ethtool`	Queries network interfaces and can also set various parameters such as the speed
`netstat`	Displays all active connections and routing tables. Useful for monitoring performance and troubleshooting
`nmap`	Scans open ports on a network. Important for security analysis
`tcpdump`	Dumps network traffic for analysis
`iptraf`	Monitors network traffic in text mode
`mtr`	Combines functionality of ping and traceroute and gives a continuously updated display
`dig`	Tests DNS workings. A good replacement for host and nslookup

Graphical and non-graphical browsers

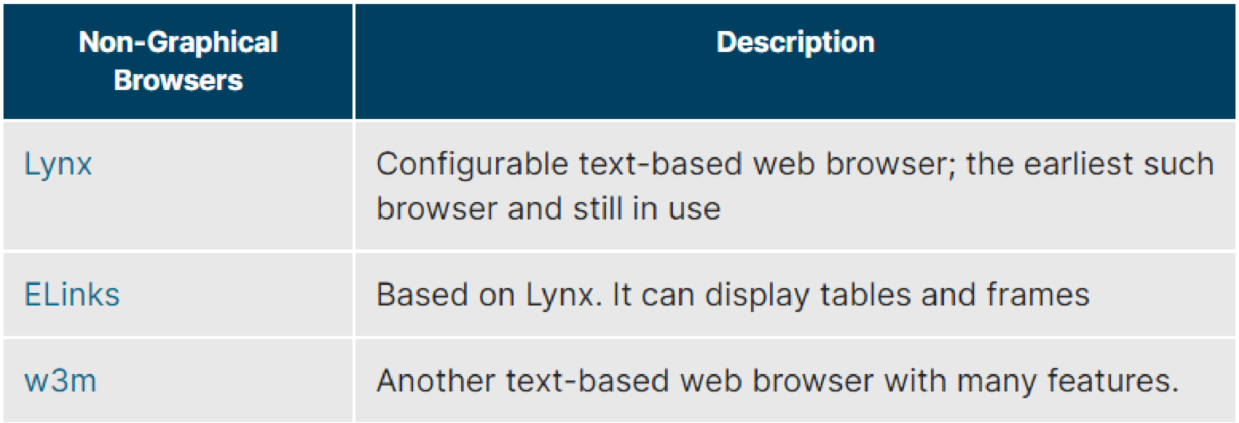
Browsers are used to retrieve, transmit, and explore information resources, usually on the World Wide Web. Linux users commonly use both graphical and non-graphical browser applications.

The common graphical browsers used in Linux are:

- Firefox
- Google Chrome
- Chromium

- Konqueror
- Opera

Sometimes, you either do not have a graphical environment to work in (or have reasons not to use it) but still need to access web resources. In such a case, you can use non-graphical browsers, such as the following:



Non-Graphical Browsers	Description
Lynx	Configurable text-based web browser; the earliest such browser and still in use
ELinks	Based on Lynx. It can display tables and frames
w3m	Another text-based web browser with many features.

Command: wget

Sometimes, you need to download files and information, but a browser is not the best choice, either because you want to download multiple files and/or directories, or you want to perform the action from a command line or a script. **wget** is a command line utility that can capably handle the following types of downloads:

- Large file downloads
- Recursive downloads, where a web page refers to other web pages and all are downloaded at once
- Password-required downloads
- Multiple file downloads.

To download a web page, you can simply type **wget <url>**, and then you can read the downloaded page as a local file using a graphical or non-graphical browser.

Command: curl

Besides downloading, you may want to obtain information about a URL, such as the source code being used. **curl** can be used from the command line or a script to read such information. **curl** also allows you to save the contents of a web page to a file, as does **wget.**

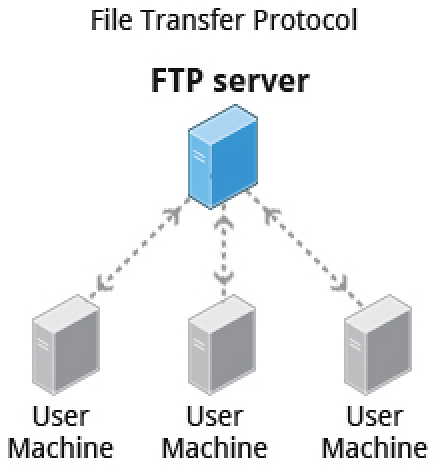
You can read a URL using **curl <URL>**. For example, if you want to read http://www.linuxfoundation.org, type **curl http://www.linuxfoundation.org**.

To get the contents of a web page and store it to a file, type **curl -o saved.html http://www.mysite.com**. The contents of the main index file at the website will be saved in **saved.html**.

FTP (File Transfer Protocol)

When you are connected to a network, you may need to transfer files from one machine to another. **F**ile **T**ransfer **P**rotocol (**FTP**) is a well-known and popular method for transferring files between computers using the Internet. This method is built on a client-server model. FTP can be used within a browser or with stand-alone client programs.

FTP is one of the oldest methods of network data transfer, dating back to the early 1970s. As such, it is considered inadequate for modern needs, as well as being intrinsically insecure. However, it is still in use and when security is not a concern (such as with so-called anonymous FTP) it can make sense. However, many websites, such as kernel.org, have abandoned its use.



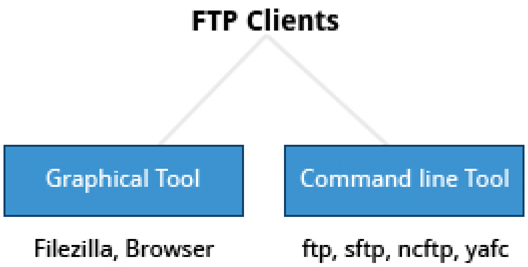
FTP Clients

FTP clients enable you to transfer files with remote computers using the FTP protocol. These clients can be either graphical or command line tools. Filezilla, for example, allows use of the drag-and-drop approach to transfer files between hosts. All web browsers support FTP, all you have to do is give a URL like **ftp://ftp.kernel.org** where the usual **http://** becomes **ftp://**.

Some command line FTP clients are:

- **ftp**
- **sftp**
- **ncftp**
- **yafc** (Yet Another FTP Client).

FTP has fallen into disfavor on modern systems, as it is intrinsically insecure, since passwords are user credentials that can be transmitted without encryption and are thus prone to interception. Thus, it was removed in favor of using **rsync** and web browser https access for example. As an alternative, **sftp** is a very secure mode of connection, which uses the Secure Shell (**ssh**) protocol, which we will discuss shortly. **sftp** encrypts its data and thus sensitive information is transmitted more securely. However, it does not work with so-called anonymous FTP (guest user credentials).

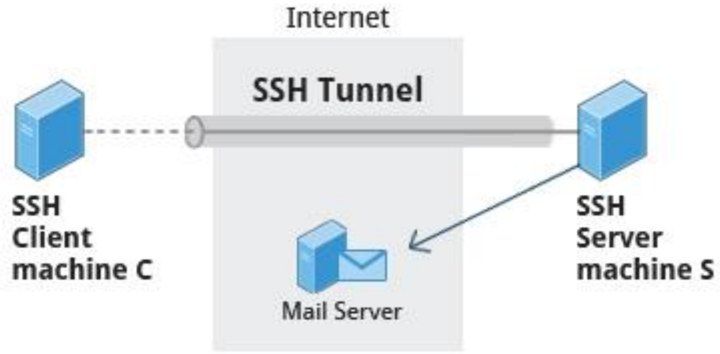


SSH: Executing Commands Remotely

Secure Shell (SSH) is a cryptographic network protocol used for secure data communication. It is also used for remote services and other secure services between two devices on the network and is very useful for administering systems which are not easily available to physically work on, but to which you have remote access.

To login to a remote system using your same user name you can just type **ssh some_system** and press **Enter**. **ssh** then prompts you for the remote password. You can also configure ssh to securely allow your remote access without typing a password each time.

If you want to run as another user, you can do either **ssh -l someone some_system** or **ssh someone@some_system**. To run a command on a remote system via SSH, at the command prompt, you can type **ssh some_system my_command**.

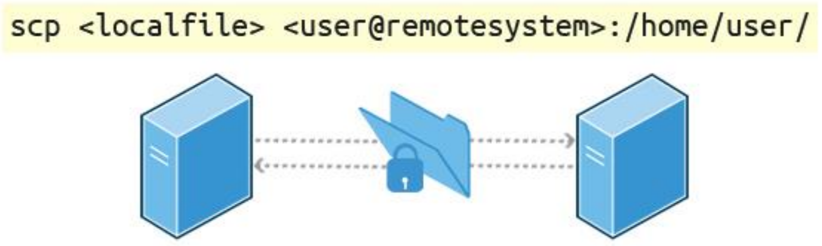


Copying files securely with scp

We can also move files securely using Secure Copy (scp) between two networked hosts. scp uses the SSH protocol for transferring data.

To copy a local file to a remote system, at the command prompt, type **scp <localfile> <user@remotesystem>:/home/user/** and press **Enter**.

You will receive a prompt for the remote password. You can also configure **scp** so that it does not prompt for a password for each transfer.



```
scp <localfile> <user@remotesystem>:/home/user/
```

Summary

- The IP (Internet Protocol) address is a unique logical network address that is assigned to a device on a network.
- IPv4 uses 32-bits for addresses and IPv6 uses 128-bits for addresses.
- Every IP address contains both a network and a host address field.
- There are five classes of network addresses available: A, B, C, D & E.
- DNS (Domain Name System) is used for converting Internet domain and host names to IP addresses.
- The **ifconfig** program is used to display current active network interfaces.
- The commands **ip addr show** and **ip route show** can be used to view IP address and routing information.
- You can use **ping** to check if the remote host is alive and responding.
- You can use the **route** utility program to manage IP routing.
- You can monitor and debug network problems using networking tools.
- Firefox, Google Chrome, Chromium, and Epiphany are the main graphical browsers used in Linux.
- Non-graphical or text browsers used in Linux are Lynx, Links, and w3m.
- You can use **wget** to download webpages.
- You can use **curl** to obtain information about URLs.
- FTP (File Transfer Protocol) is used to transfer files over a network.
- ftp, sftp, ncftp, and yafc are command line FTP clients used in Linux.
- You can use **ssh** to run commands on remote systems.