ECE 333: Introduction to Communication Networks
Fall 2001

Lecture 24: Routing and Addressing III

- Routing & addressing in IP

In Lectures 22 and 23 we discussed general issues and approaches to routing in communication networks. In this lecture, we will examine how routing and addressing are done in TCP/IP networks, particularly in the Internet. First we give some background on TCP/IP and Internet, then we discuss IP addresses, and finally we will look at how routing is done in IP. In the next lecture, we will look at some other issues related to IP.

TCP/IP

We have seen examples of various networking technologies for both LAN and WAN applications. Different organizations and different parts of one organization may use different types of networks, for example an Ethernet LAN may be used within an office and a Frame Relay WAN may be used to connect several offices in different locations. Various networks use different packet formats, different addresses and different protocols. Thus, a host attached to one type of network could not directly send a packet to a host connected to another type.

The TCP/IP protocols were designed to allow hosts connected to different networks to communicate with each other; this is referred to as ***internetworking***. Internetworking in TCP/IP is accomplished by adding a new protocol layer, the internet layer above the top layers in the other networks. Any collection of heterogeneous networks that are connected together in this way is referred to as an ***internet*;** when all of the networks are owned by a single company, it is also referred to as an ***intranet*.** The global Internet, which you use to web browse, etc., is commonly distinguished by capitalizing the initial "I". In an internet, two or more networks are connected together at special nodes, called routers[1]. This is similar in some ways to using a bridge to connect two LANs. However, a router operates at

[1] Sometimes these nodes are also called Gateways or multi-protocol routers.

the internet layer, instead of the data link layer. The figure below shows several networks connected by routers.

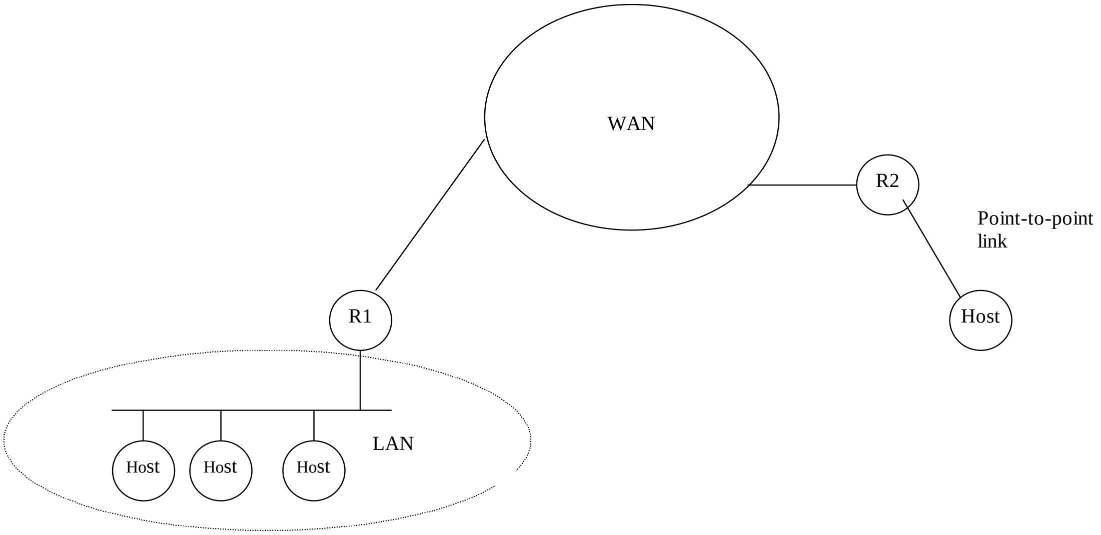
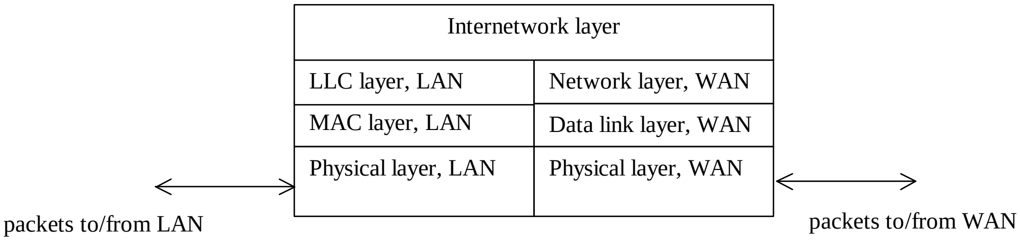


Figure: Example of an internet; routers are labeled with an R.

Each router must implement the lower layer functions of each network to which it is attached. For example, in this figure, router R1, will implement both the protocols that are used on the LAN and the protocols that are used in the WAN. From the point-of-view of each of these networks, the router looks like another host.



Internetwork layer	
LLC layer, LAN	Network layer, WAN
MAC layer, LAN	Data link layer, WAN
Physical layer, LAN	Physical layer, WAN

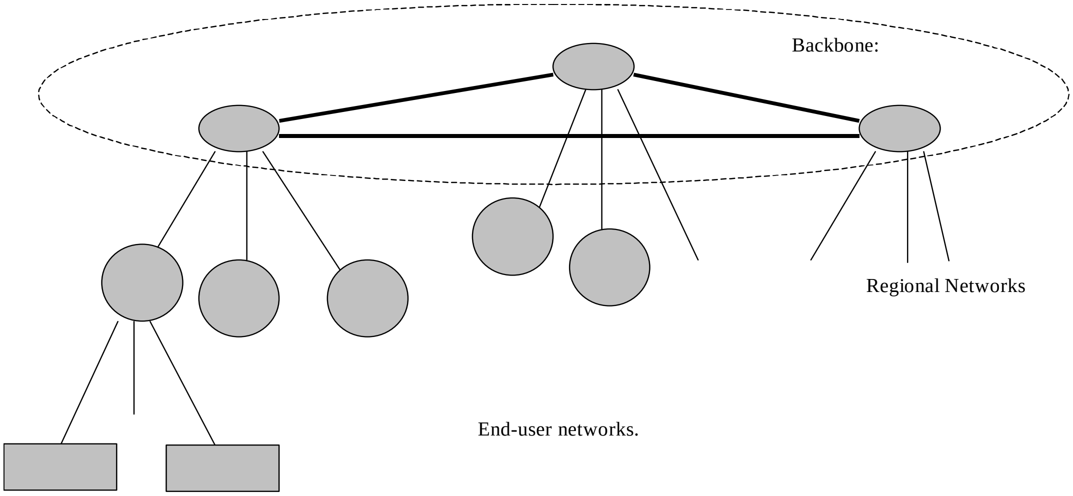
Figure: Protocol stack implemented in router R1.

A router may also be attached to more than two networks.

Conceptually, an internet can be viewed as a collection of routers and hosts connected by "virtual links". These "links" are an abstraction of the underlying physical networks that connect two routers - a link in the internet may actually be a point-to-point link or may be a LAN or WAN. The internet layer hides the details of these lower layer technologies from higher layer applications and provides the abstraction of all hosts being connected to one seamless network.

The Internet topology

Next, we look at the topology of the Internet. In the early 1990's, the Internet consisted of a single *backbone*, managed by NSF (NSFNET). Most end-user networks (e.g. at companies and universities) connected to the backbone through one of three regional provider networks. The regional and backbone networks were primarily routers connected by point-to-point links. End-user networks are themselves often a collection of multiple networks linked together by bridges and other routers.



In 1995, NSFNET was decommissioned, and the management of the Internet backbone was commercialized. Today there are multiple (international) backbones (also called service provider networks) operated by private companies (e.g., UUNET operated by WorldCom or SprintLink). These backbone networks are connected together at points called Network Access Points (NAPs) - in the U.S the regional telephone companies often run these NAPs. Two service provider networks may also connect directly together at private *peering points*.

Large companies may connect directly to backbone or to a smaller (regional or local) ISP's.

Additionally two smaller service providers may arrange to connect directly together without going through the backbone.

The resulting network is still loosely hierarchical but has much less discernible structure than the NSFNET.

IP Addresses

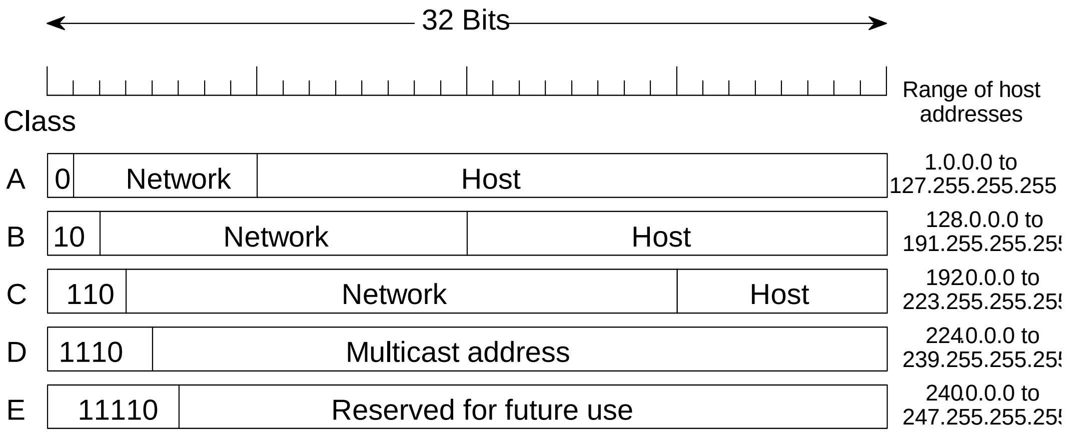
The version of IP that is widely used today is version 4 (IPv4). A new version, IPv6 has been standardized and is used in some places. In the following, we will mainly discuss IPv4.

All IP addresses are 32 bits (this becomes 128 bits in IPv6). Each host in the Internet must have a globally unique IP address.

To make routing efficient, hierarchical addressing is used in the Internet. An IP address is divided into two parts. The first part of the address (the prefix) identifies the network to which a host is attached. The second part of the address identifies the host on the network. The assignment of network addresses is coordinated globally, so that no two networks have the same network address. In North America, network addresses are assigned by **ARIN** (American Register of Internet Networks). The assignment of host addresses within a network is done locally, for example by a system administrator.

The 32 bits in an IP address are divided between the network and host addresses. Originally, IP used a “class-based” approach for this division. In this approach, the available IP addresses where divided into different classes. Each class specifies a different number of bits to the network and host parts. These classes are shown below.

IP Addresses


32 Bits
Class
Range of host addresses
A
0
Network
Host
1.0.0.0 to 127.255.255.255
B
10
Network
Host
128.0.0.0 to 191.255.255.25
C
110
Network
Host
1920.0.0 to 223.255.255.25
D
1110
Multicast address
2240.0.0 to 239.255.255.25
E
11110
Reserved for future use
2400.0.0 to 247.255.255.25

IP class-based address formats.

The first few bits in an address indicate the class of the address. Class A, B, and C addresses are used for individual hosts. For class A networks, there are 32 bits for host addresses and 7 bits for network addresses. Thus, there are 2^{32} different host addresses for each class A network, and there can be at most 2^{7} different class A networks. There can be more class B and C networks, but each can have fewer hosts per network.

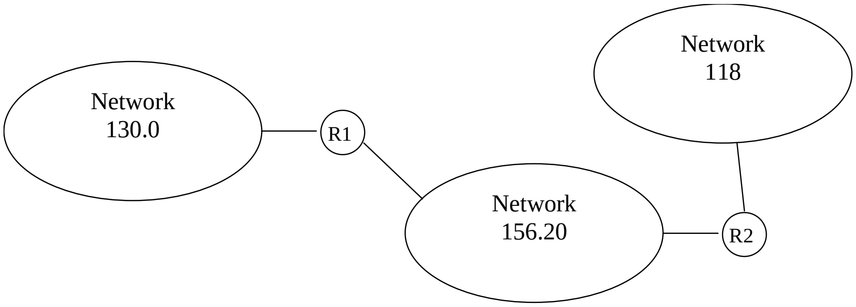
IP Addresses

IP addresses are written in ***dotted decimal notation***, where each byte is represented as a decimal value and the value for different bytes are separated by periods. For example, the IP address

11000000 00011001 00000110 00010100

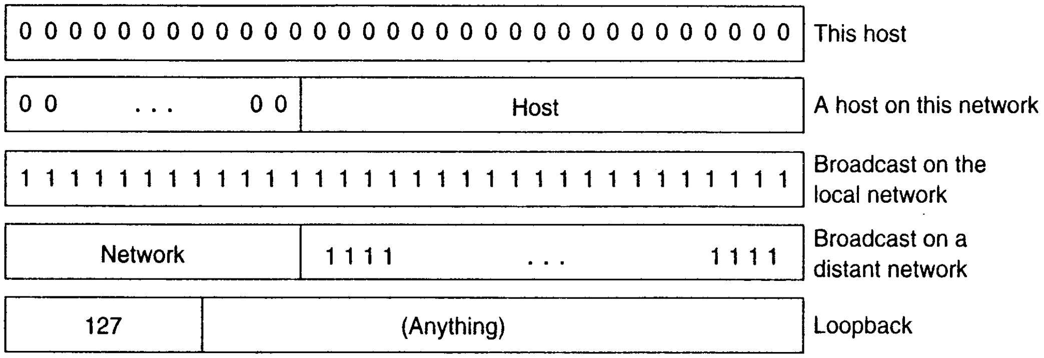
is written 192.41.6.20. (Notice this is a class C network.).

As an example of class-based addressing, consider the internet shown below. There are three networks, two class B networks and one class A network, with the prefix for each network shown. Note the routing table at router R1 need only contain one entry for every host on network 118. Any IP packet received with this prefix will be sent on network 156.20 to router R2.



Special IP addresses

Several IP addresses are reserved for special meanings. In either the network or host sub-field, all 0's and all 1's are reserved and have the meanings shown below. The prefix 127 is also reserved and used for "loop back." Packets sent with this prefix never leave the host; these are used for testing applications.



Special IP addresses

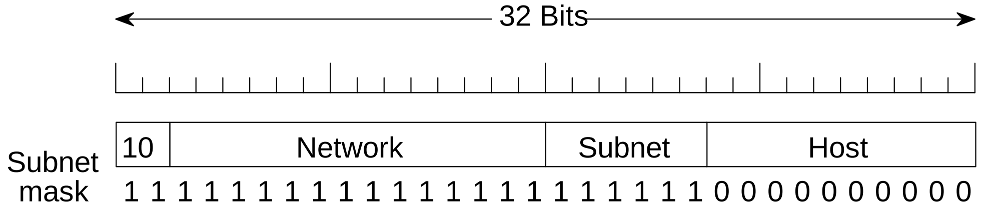
Problems with class-based addresses:

As the Internet grew, the class-based addressing scheme described above led to several problems. It was intended that the network part of the address correspond to one physical network. Thus an organization with many internal networks would need many different network addresses; each of these network addresses would need to be stored in all global routing tables, leading to routing inefficiencies. To solve this problem, the concept of ***subnetting*** was introduced. With subnetting an organization could get one network address, and then divide this address internally into several **subnets**[2]. Each subnet corresponds to one internal network. This adds another layer to the address hierarchy. Routers outside of the organization do not need to know about the subnets and route only based on the network address.

[2] Note this is a different use of "subnet" than we have been using.

Subnets

To specify a subnet the host part of an IP address is divided into two parts, the first part specifying the subnet address, and the second part specifying the host address within the subnet. The figure below shows an example of this for a class B address, where 5 bits are used for the subnet address. The number of bits used for the subnet address is indicated by a ***subnet mask.*** This is a 32 bit string with a zero in each position corresponding to the host part of the address and a one everywhere else. The corresponding subnet mask is also shown in the figure below. Using the dotted-decimal notation, this subnet mask would be written as 255.255.252.0.



Routers determine the subnet address by adding mod 2 the destination address with subnet mask.

Note only routers within subnet need to understand subnetting. Local network administrators can decide the on number of bits in mask.

Classless Interdomain Routing (CIDR)

Another problem with class based addressing is that it led to an inefficient use of the available addresses. For example, an organization with 1,000 hosts is too large for a single class C network. However, if this organization was given a class B address, then it would use up 65,536 available IP addresses. This organization could be assigned several class C addresses, but this complicates routing tables. Another problem with the class-based addressing scheme is that class A networks are too large for most organizations and class C networks are too small. This led to a shortage of class B addresses.

To overcome these problems class-based addressing was replaced with a "classless" approach called Classless InterDomain Routing (CIDR). CIDR allows for an arbitrary division of an IP address into network and host parts. This is accomplished by essentially the same idea as used in subnetting, described above. With CIDR an address mask is used to specify the network part of the address. Routers then exchange these masks along with the destination addresses. The notation 129.20.0.0 /16 is used to specify that this address consists of a 16 bit network address.

CIDR example

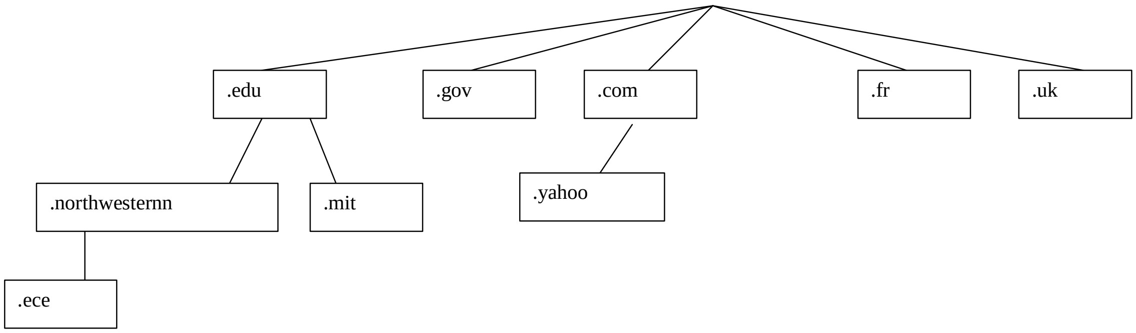
Consider the above example where an organization required addresses for 1000 hosts. Instead of allocating this organization a class B address, using CIDR this organization could be allocated a block of 1,024 addresses, with the form w.x.y.z/22. In this case the first 22 bits of the address indicate the network address.

CIDR can be combined with subnetting to provide multiple levels of hierarchy.

Domain Names

IP only deals with numerical IP addresses as discussed above. Most applications written for humans deal with ***domain names***, e.g. **ece.northwestern.edu**

The translation between domain names and IP addresses is done in the application layer using the ***Domain Name System (DNS).*** DNS provides a distributed hierarchical database of domain names and IP addresses. Names are arranged in this hierarchy as shown below.



Names are assigned by ICANN, (Internet Corporation for assigned Names and Numbers).

In Nov, 2000 ICANN approved several new names for the top-level hierarchy, including .biz, .name, .info, .pro, .museums, .coop, .aero.

Prior to this there were 6 top-level domains with in the U.S., and country specific domain names for other countries.

The tree is divided into disjoint collection of “zones”. In each zone there is a name server (actually several for reliability). Each name server knows the address for the name servers of any of its children and the names and IP addresses of all the hosts in that zone.

E.g. A name server at .edu needs to know the addresses of the name server for .northwestern.edu and .mit.edu.

The DNS servers communicate by sending IP packets.

To improve efficiency, recent DNS responses are cached locally for a given length of time.

ARP (Address Resolution Protocol)

When transmitting IP datagrams over another network, need to be able to translate IP address into network address. (e.g. Ethernet address).

One approach - have system administrator maintain table at each host.

This becomes difficult for large networks.

2nd approach - automatically update tables. ARP is a protocol used for this on broadcast LANs.

Idea – Each host periodically broadcasts an ARP query with destination IP address, source IP address, and source link layer address. Destination replies with its link layer address.

Other hosts who receive the query can also update their tables.

Management Hierarchy

In lecture 23, we noted that to address problems of scale in routing, some form of hierarchical structure was helpful. In the Internet, this structure is provided by grouping the routers into collections of ***Autonomous Systems* (*AS*)**. An autonomous system loosely refers to a portion of the Internet controlled by a single administrative entity. This might be a single company, a university, or an ISP (however a large company, could also divide its network into several autonomous systems). Within each autonomous system, routers all exchange messages and run the same routing protocol. Different autonomous systems may use different routing protocols. Routing between AS's is accomplished using a different set of routing protocols.

Each AS is responsible for its own internal routing. The routing algorithm used within an AS is referred to as an ***interior gateway protocol*** (IGP). Two examples of IGP's are

- **RIP** - this is a distance vector protocol, based on Bellman-Ford. It was the original IGP used in the Internet.
- **OSPF** - this is a link state IGP, which became an Internet standard in 1990.

The Internet must also deal with routing between AS's. This is addressed by an ***exterior gateway protocol*** (EGP). The primary EGP used in the Internet is **BGP-4.**

Inter-AS Routing: the Exterior Gateway Protocol BGP

All routers in an AS run some IGP for interior routing (e.g. RIP or OSPF)

Inter-AS routing information is exchanged and processed using the **BGP** (Border Gateway Protocol). The routers that implement BGP are called ***border routers*** or ***BGP routers***. BGP routers communicate routing information to each other over (reliable) TCP connections.

With inter-AS routing, *policy and politics* now become an important part of the routing decision.

BGP routers view the network as collection of BGP routers connected together by “links''. Two BGP routers are linked if they are connected to a common AS.

BGP is similar to a distance vector protocol except routers exchange exact paths to each destinations (as opposed to lengths) – also routers can propagate multiple paths to a destination.

This allows routers to make policy decisions on which path to take. For example routers owned by one company may not want to forward any packets over a paths that goes through an AS owned by a competitor.