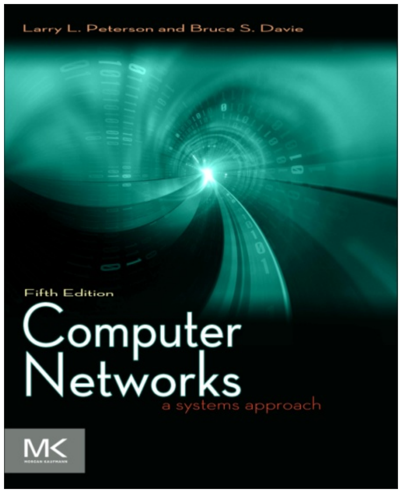


Chapter 4

Advanced Internetworking

Problems

- How do we build a routing system that can handle hundreds of thousands of networks and billions of end nodes?
- How to handle address space exhaustion of IPv4?

Chapter Goal

- Understanding the scalability of routing in the Internet
- Discussing IPv6

NSFNET

- How do we build a routing system that can handle hundreds of thousands of networks and billions of end nodes?
- How to handle address space exhaustion of IPv4?

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

National Science Foundation Network

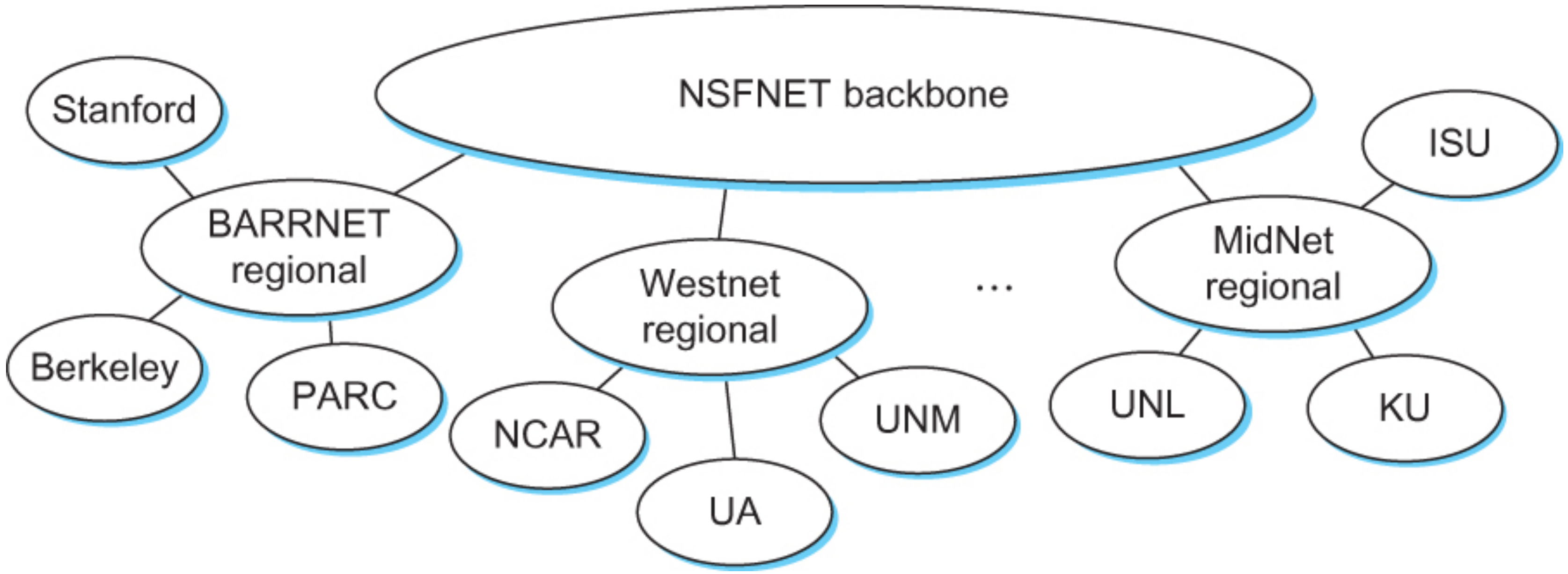
From Wikipedia, the free encyclopedia

See also: History of the Internet

The **National Science Foundation Network** (**NSFNET**) was a program of coordinated, evolving projects sponsored by the National Science Foundation (NSF) beginning in 1985 to promote advanced research and education networking in the United States.[1] NSFNET was also the name given to several nationwide backbone networks that were constructed to support NSF's networking initiatives from 1985 to 1995. Initially created to link researchers to the nation's NSF-funded supercomputing centers, through further public funding and private industry partnerships it developed into a major part of the Internet backbone.

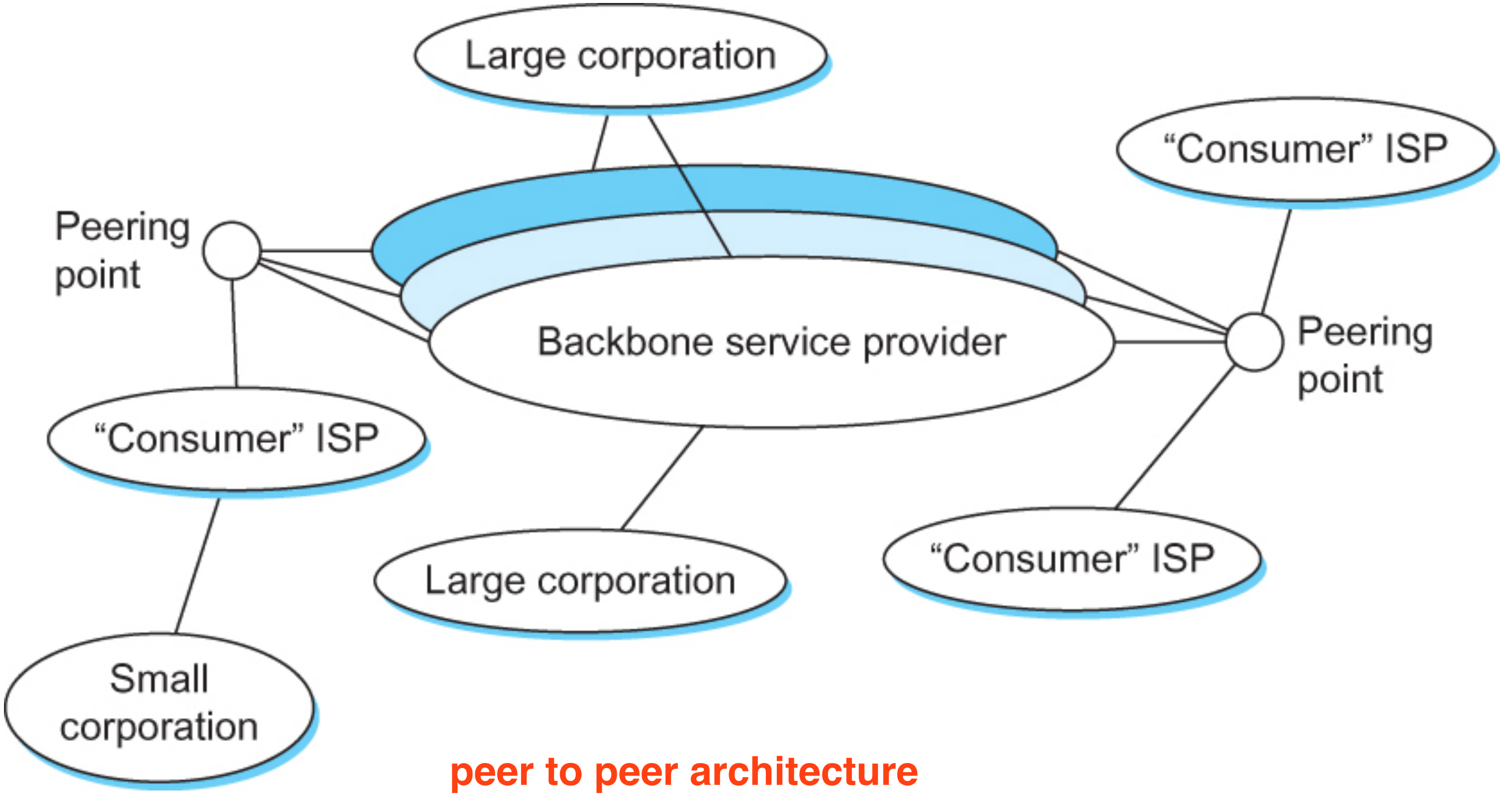
Chapter Goal:
- Understanding the scalability of routing in the Internet
- Discussing IPv6

The Global Internet



The tree structure of the Internet in 1990

The Global Internet



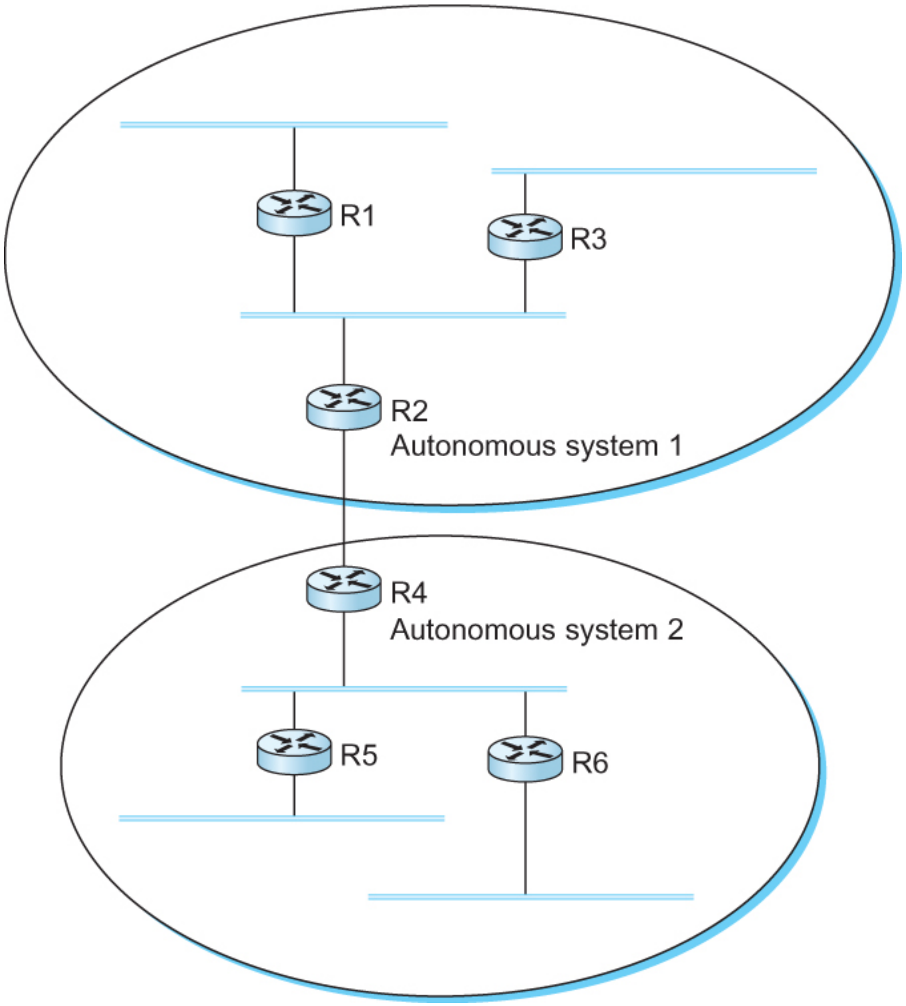
peer to peer architecture

A simple multi-provider Internet

Interdomain Routing (BGP)

- Internet is organized as autonomous systems (AS) each of which is under the control of a single administrative entity
- Autonomous System (AS)
  - corresponds to an administrative domain
  - examples: University, company, backbone network
- A corporation's internal network might be a single AS, as may the network of a single Internet service provider **(ISP)**

Interdomain Routing



A network with two autonomous system

Route Propagation

- Idea: Provide an additional way to hierarchically aggregate routing information in a large internet.
  - Improves scalability

- Divide the routing problem in two parts:
  1. Routing within a single autonomous system (intra-domain)
  2. Routing between autonomous systems (inter-domain)

- Another name for autonomous systems in the Internet is routing domains
  - Two-level route propagation hierarchy
    - Inter-domain routing protocol (Internet-wide standard)
    - Intra-domain routing protocol (each AS selects its own)

Intra or Inter Autonomous System Routing

Routing Protocols

The protocols of routers,

Routers Talking to Routers on the Network

Sharing Routing Information With Each Other

IGP - Interior Gateway Protocols

RIP RIPv2 IGRP EIGRP OSPF ISIS

EGP - Exterior Gateway Protocols

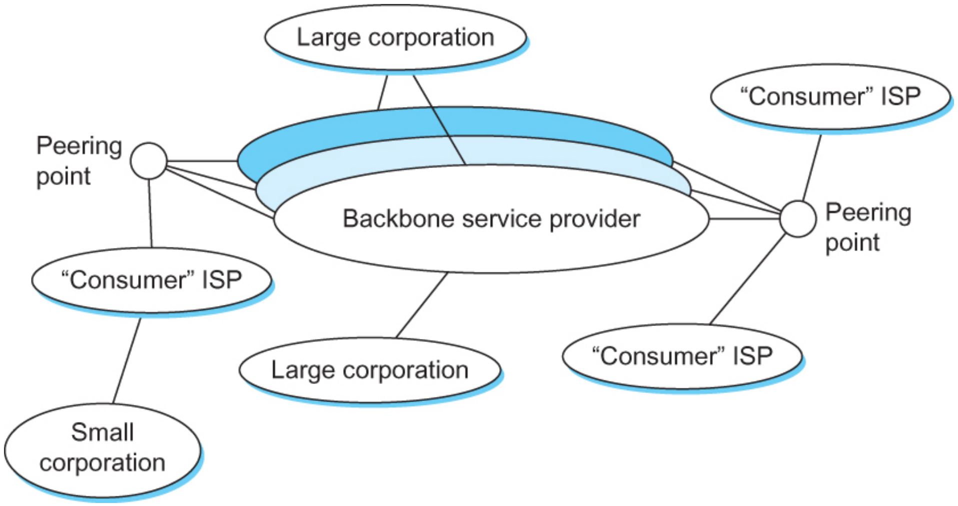
BGP

EGP and BGP

- Inter-domain Routing Protocols
  - Exterior Gateway Protocol (EGP)
    - Forced a tree-like topology onto the Internet (a limitation)
  - Border Gateway Protocol (BGP), a replacement for EGP
    - Assumes that the Internet is an arbitrarily interconnected set of ASs.
      - A peer-to-peer relationship between all ASs
    - Sites are connected to each other in arbitrary ways
    - BGP4 is the latest version

BGP

- Some large corporations connect directly to one or more of the backbone, while others connect to smaller, non-backbone service providers.
- Many service providers exist mainly to provide service to "consumers" (individuals with PCs in their homes), and these providers must connect to the backbone providers
- Often many providers arrange to interconnect with each other at a single "peering point"



https://www.youtube.com/watch?v=z8INzy9E628

Route Propagation

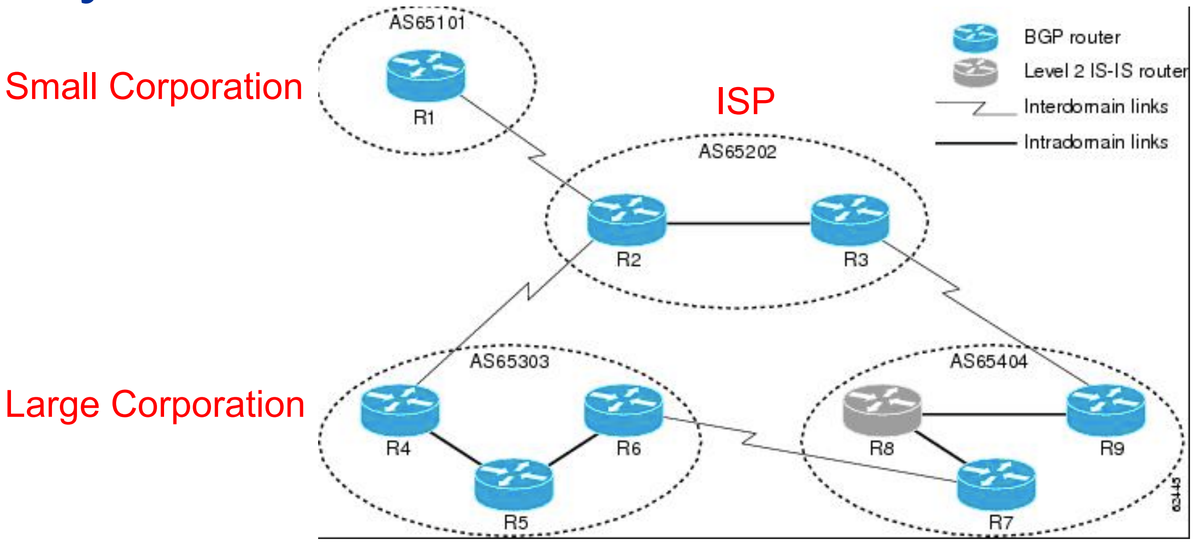
- Idea: Provide an additional way to hierarchically aggregate routing information in a large internet.
  - Improves scalability

- Divide the routing problem in two parts:
  1. Routing <u>within</u> a single autonomous system (intra-domain)
  2. Routing <u>between</u> autonomous systems (inter-domain)

- Another name for autonomous systems in the Internet is routing domains
  - Two-level route propagation hierarchy
    - Inter-domain routing protocol (Internet-wide standard)
    - Intra-domain routing protocol (each AS selects its own)

Atonomous Systems

- Assumes the Internet
  is an arbitrarily interconnected set of AS's.

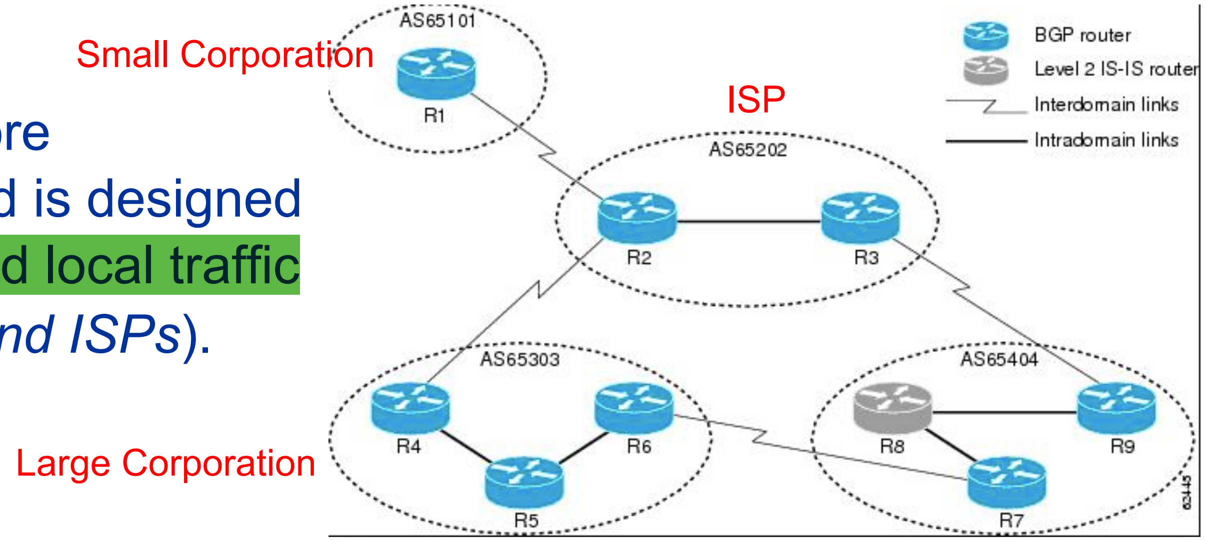


- Let's define:
  - *Local traffic* (intra-AS) as traffic that originates at or terminates on nodes within an AS.
  - *Transit traffic* (inter-AS) as traffic that passes through an AS.

Atonomous Systems

We can classify AS's into three types:

- *Stub AS*: an AS that has only a single connection to one other AS; such an AS will only carry local traffic (*small corporations*).
- *Multihomed AS*: an AS that has connections to more than one other AS, but refuses to carry transit traffic (*large corporations*).
- *Transit AS*: an AS that has connections to more than one other AS, and is designed to carry both transit and local traffic (*backbone providers and ISPs*).



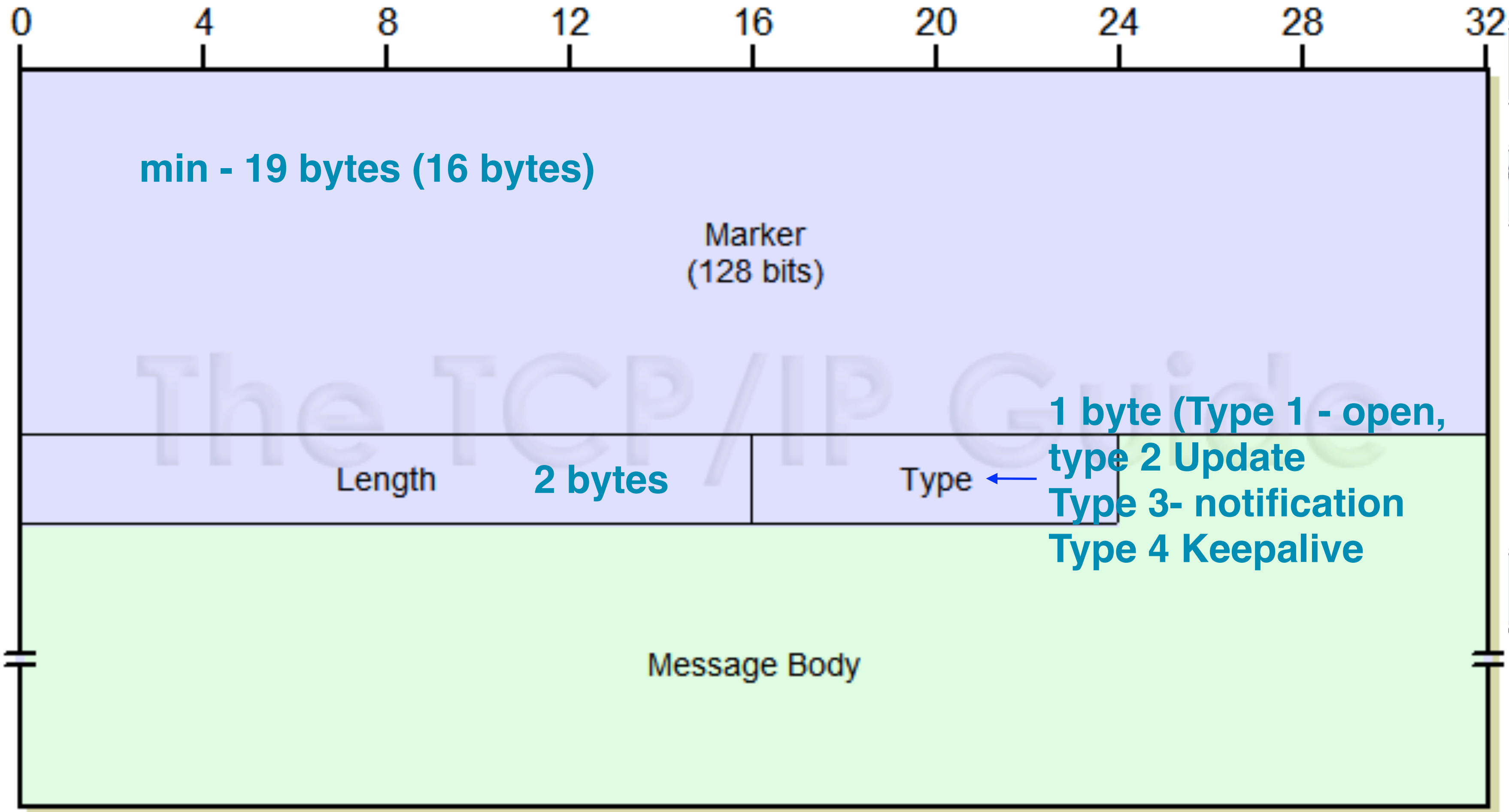
BGP

- BGP does not belong to either of the two main classes of routing protocols (distance vectors and link-state protocols)
  - It acts like RIP (an intra-AS distance vector)
  - That is why sometimes it is considered a DV protocol, but it is really not.
- BGP is considered a Path Vector based protocol.
  - Advertises *complete paths* as an enumerated lists of ASs to reach a particular network

BGP

- The current version of BGP is BGP version 4, based on RFC4271.
- BGP is the path-vector protocol that provides routing information for autonomous systems on the Internet via its AS-Path attribute. unlike UDP where packets can be lost, with TCP, it guarantee the delivery without loss
- BGP uses TCP as its transport protocol so it doesn't have to worry about the things TCP will handle.
- Peers that have been manually configured to exchange routing information will form a TCP connection and begin speaking BGP
  - Will send updated router table information only when one host has detected a change.
- An important aspect of BGP is that the AS-Path itself is an anti-loop mechanism.
  - Routers will not import any routes that contain themselves in the AS-Path.

BGP


0
4
8
12
16
20
24
28
32
min - 19 bytes (16 bytes)
Marker
(128 bits)
The TCP/IP Guide
Length
2 bytes
Type
1 byte (Type 1 - open,
type 2 Update
Type 3- notification
Type 4 Keepalive
Message Body

BGP

- The *open message* opens a BGP communications session between peers
  - Is the first message sent by each side after a transport-protocol connection is established.
  - Open messages are confirmed using a keep-alive message sent by the peer device.
- An *update* message is used to provide routing updates to other BGP systems.
  - Updates are sent using the Transmission-Control Protocol (TCP) to ensure reliable delivery.
  - Update messages can withdraw one or more unfeasible routes from the routing table and simultaneously can advertise a route while withdrawing others.

BGP

- The *notification message* is sent when an error condition is detected.
  - Notifications are used to close an active session and to inform any connected routers of why the session is being closed.
- The *keep-alive* message notifies BGP peers that a device is active.
  - Keep-alives are sent often enough to keep the sessions from expiring.

Advertisement Example slide photo in Notes

BGP

- The goal of Inter-domain routing is to find any path to the intended destination that is loop free
  - Designed to exchange routing and reachability information among autonomous systems (AS) on the Internet.
  - We are concerned with reachability than optimality
  - Finding path anywhere close to optimal is considered to be a great achievement

BGP

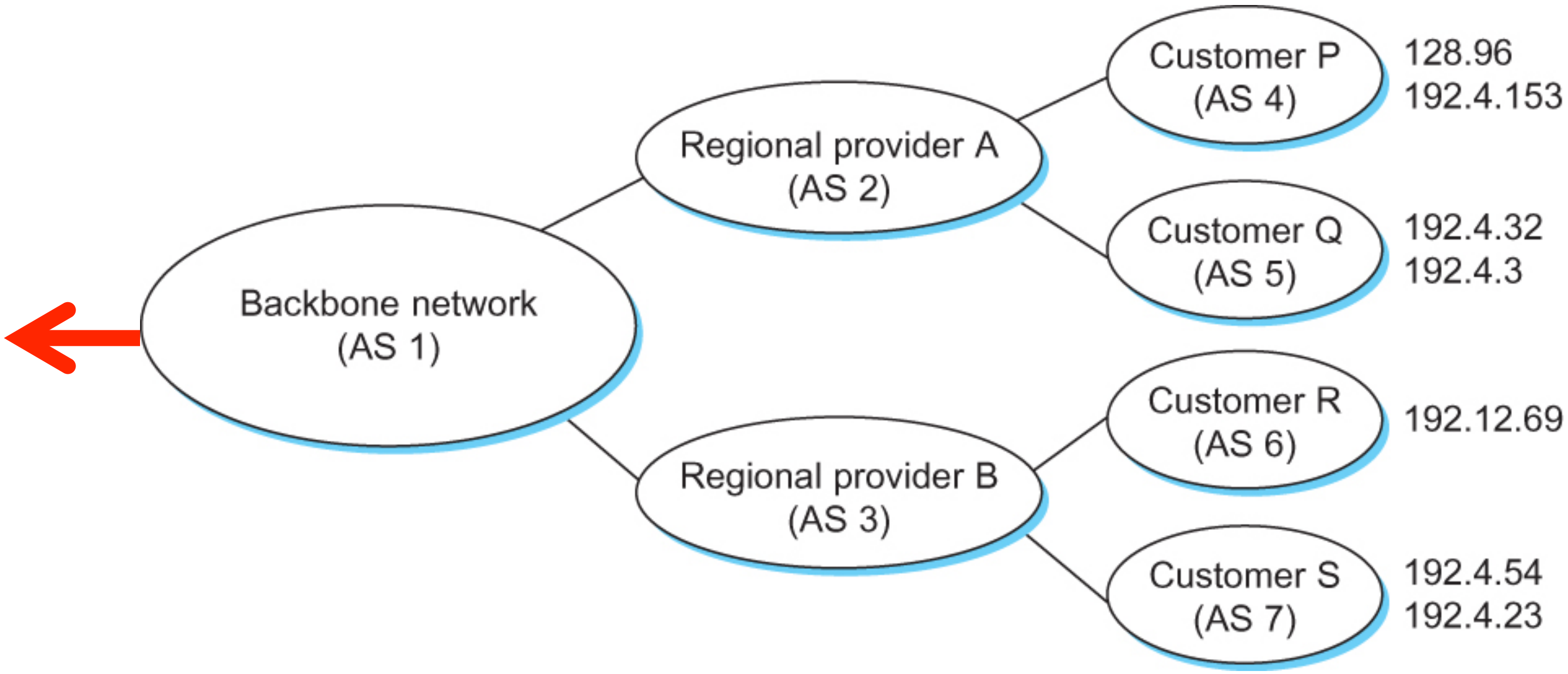
- Scalability: An Internet backbone router must be able to forward any packet destined anywhere in the Internet
  - Having a routing table that will provide a match for any valid IP address
- Autonomous nature of the domains
  - It is impossible to calculate meaningful path costs for a path that crosses multiple ASs
  - A cost of 1000 across one provider might imply a great path but it might mean an unacceptable bad one from another provide
- Issues of trust
  - Provider A might be unwilling to believe certain advertisements from provider B

BGP Advertisement

Each AS has:

- One BGP *speaker* that advertises the following:
  - local networks (within that AS)
  - other reachable networks (transit AS only)
  - gives *path* information
- In addition to the BGP speakers, the AS has one or more border "gateways" which need not be the same as the speakers
- The border gateways are the routers through which packets enter and leave the AS

BGP Advertisement Example

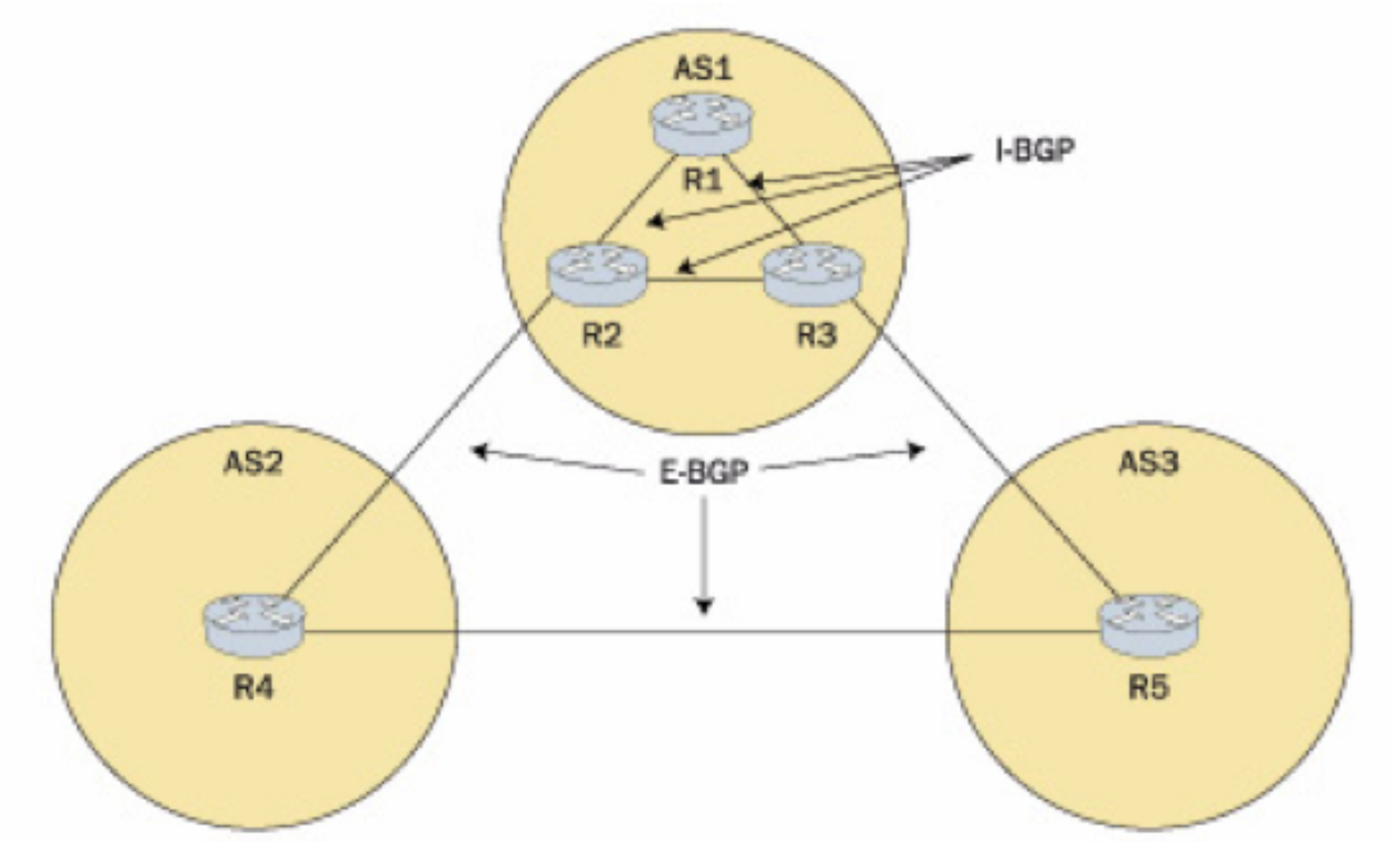


Example of a network running BGP

BGP Assumptions

- It should be apparent that the AS numbers carried in BGP need to be unique
  - A unique ASN is allocated to each AS for use in BGP routing
  - AS numbers are assigned in blocks by the Internet Assigned Numbers Authority (IANA) to Regional Internet Registries (RIRs).
  - The appropriate RIR then assigns AS numbers to entities within its designated area from the block assigned by the IANA.
- Until 2007 AS numbers were 16-bit numbers assigned by a central authority
  - a maximum of 65536 assignments
- A 32-bit number option also available as of 2007

Integrating Interdomain and Intradomain Routing



With internal neighbors we run iBGP with them and with external neighbors we run eBGP

Next Generation IP (IPv6)

Major Features - IPv6

- 128-bit addresses 16x8
- Multicast done by MLD , in IPv4 done by IGMP
- Real-time service
- Authentication and security
- Auto-configuration
- End-to-end fragmentation diff w how IPv4 does
- Enhanced routing functionality, including support for mobile hosts

nothing in the middle unlike ICMP alert in IPv4 in d middle of the process of sending to send the fragment, end to end detect error

IPv6 Addresses

- Classless addressing/routing (similar to CIDR)
- Notation: x:x:x:x:x:x:x:x (where each x represents a 16-bit hex number)
  - contiguous 0s are compressed:

    47CD:0000:0000:0000:0000:0000:A456:0124

    47CD::A456:0124
  - IPv6 compatible IPv4 address: ::128.42.1.87
    - The last 4 bytes (32 bits) are written in IPv4 notation

IPv6 Addresses

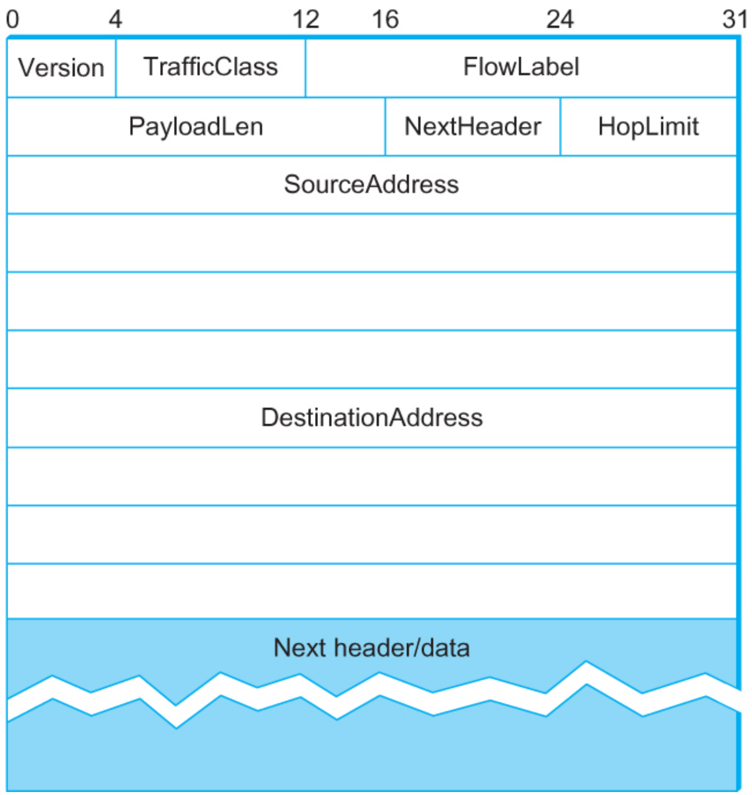
Prefix	Use
00…0 (128 bits of 0) 0:0:0:0:0:0:0:0 OR ::	Unspecified
00…1 (128 bits) 0:0:0:0:0:0:0:1 OR ::1	Loopback **only for testing purposes**
1111 1111	Multicast addresses
1111 1110 10	Link-local unicast
Everything else	Global Unicast Addresses

IPv6 Address Space

Breakpoint

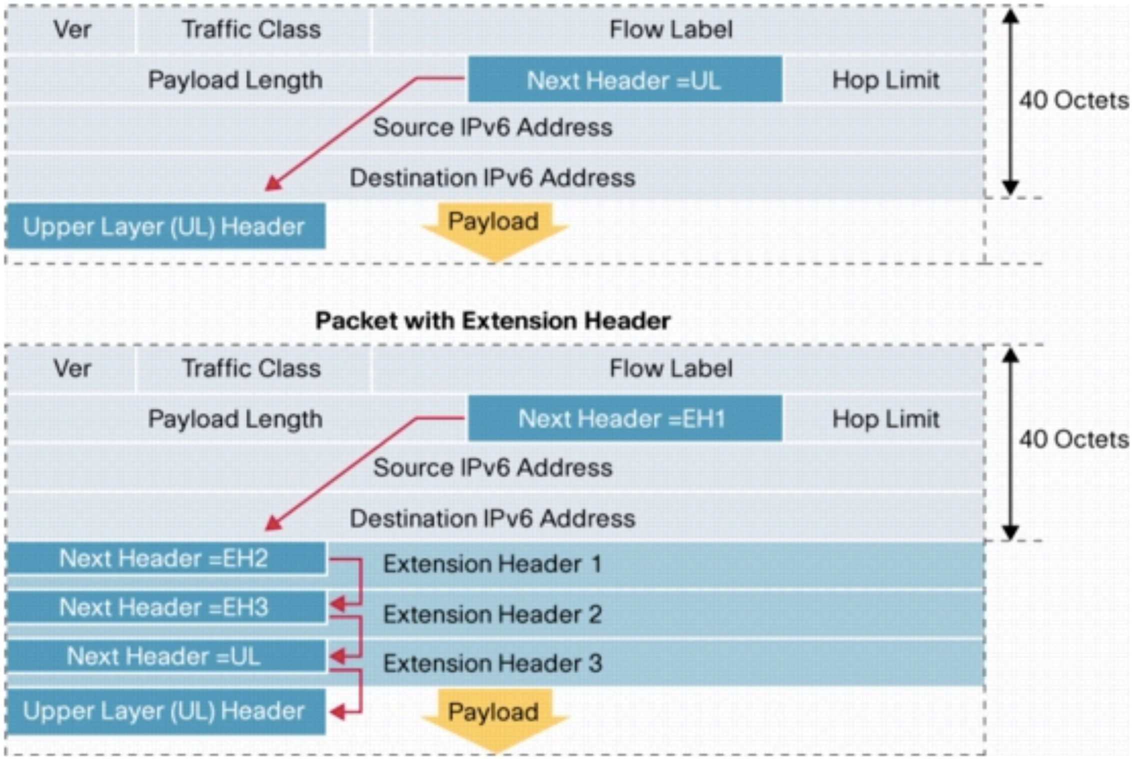
IPv6 Header

- 40-byte “base” header
- Extension headers (fixed order, mostly fixed length)
  - fragmentation
  - source routing
  - authentication and security
  - other options



IPv6 Header

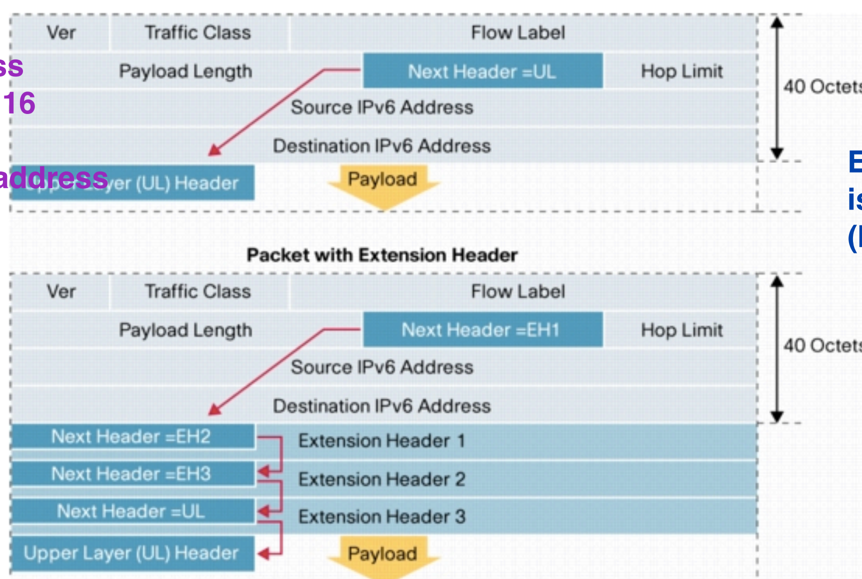
- Extension headers (fixed order, mostly fixed length)
  - fragmentation
  - source routing
  - authentication and security
  - other options



IPv6 Header

Extension Header	Type	Description
Hop-by-Hop Options	0	Options that need to be examined by all devices on the path.
Destination Options (before routing header)	60	Options that need to be examined only by the destination of the packet.
Routing	43	Methods to specify the route for a datagram (used with Mobile IPv6).
Fragment	44	Contains parameters for fragmentation of datagrams.
Authentication Header (AH)	51	Contains information used to verify the authenticity of most parts of the packet.
Encapsulating Security Payload (ESP)	50	Carries encrypted data for secure communication.
Destination Options (before upper-layer header)	60	Options that need to be examined only by the destination of the packet.
Mobility (currently without upper-layer header)	135	Parameters used with Mobile IPv6.

layer 3 IP address has link local address along with IPv6 address (:: host address last 4 tuple of 16 bits)
FE80: is reserved for link local address



EUI-64 Extended Unique Identifier is used to extend the Mac address (host ID) in IPv6 to be 128 bits