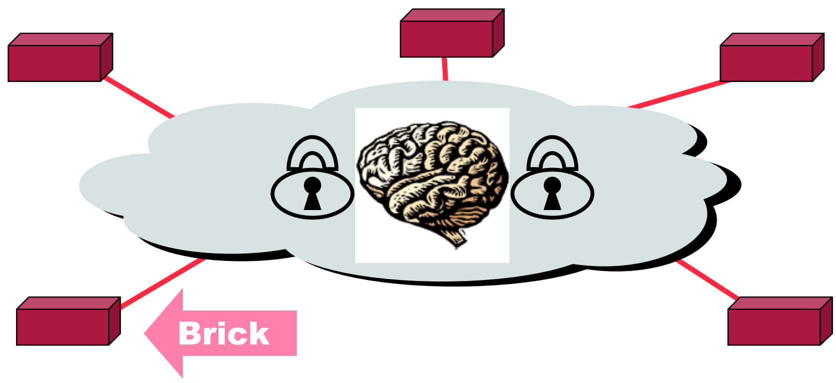
Internet Routing Protocols Lecture 01

Advanced Systems Topics

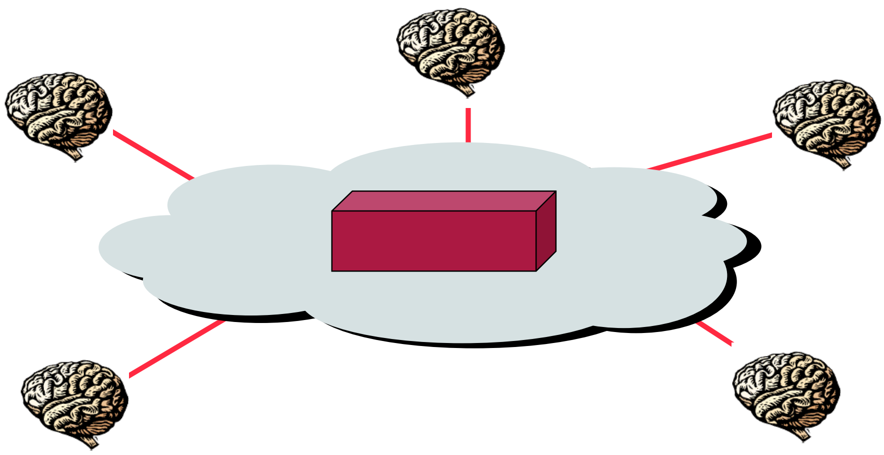
Lent Term, 2008

Timothy G. Griffin
Computer Lab
Cambridge UK

Common View of the Telco Network



Common View of the IP Network (Layer 3)

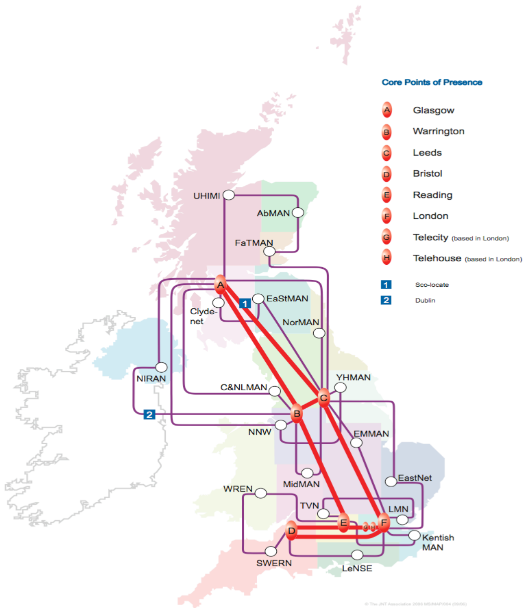


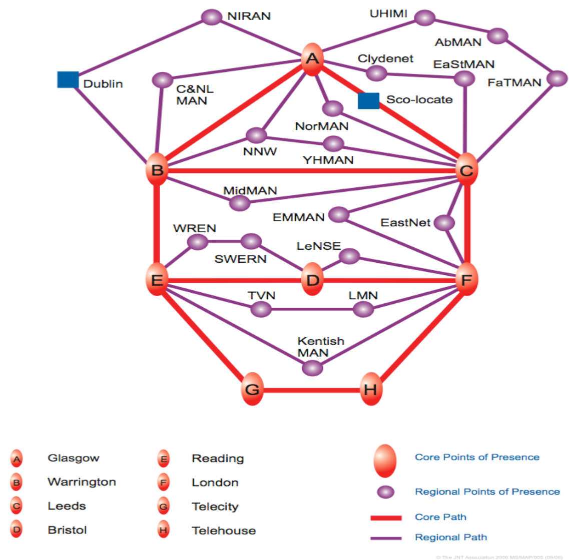
IP routing is the little bit-o-smarts left in the IP network layer

- **Dynamic Routing protocols are used to implement and maintain connectivity in the Internet.**
- **Which protocols are used?**
- **How do they work?**
- **How do they behave?**
- **What are some of the fundamental tradeoffs in the design space of routing protocols?**

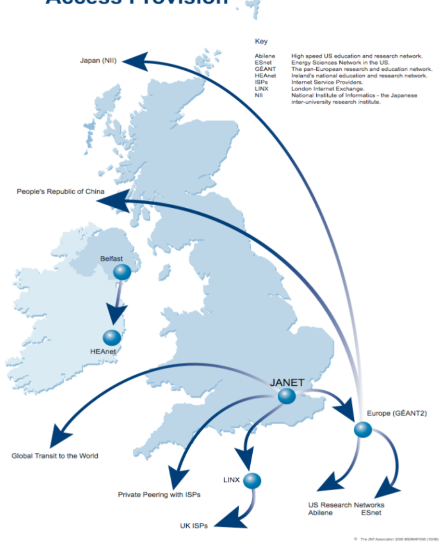
Outline

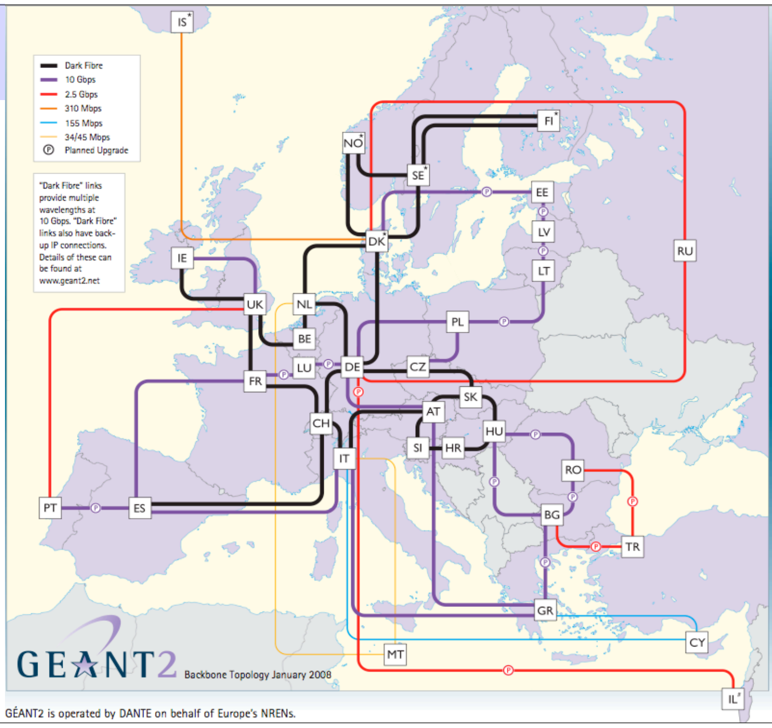
- **Lecture 1 : Routing vs. Forwarding. Internet routing architecture**
- **Lecture 2: Intra-domain routing with "shortest paths". Link-state vs. distance-vector.**
- **Lecture 3 : Inter-domain routing. The Border Gateway Protocol (BGP)**
- **Lecture 4 : BGP continued**
- **Lecture 5 : BGP dynamics**
- **Lecture 6 : BGP routing anomalies**



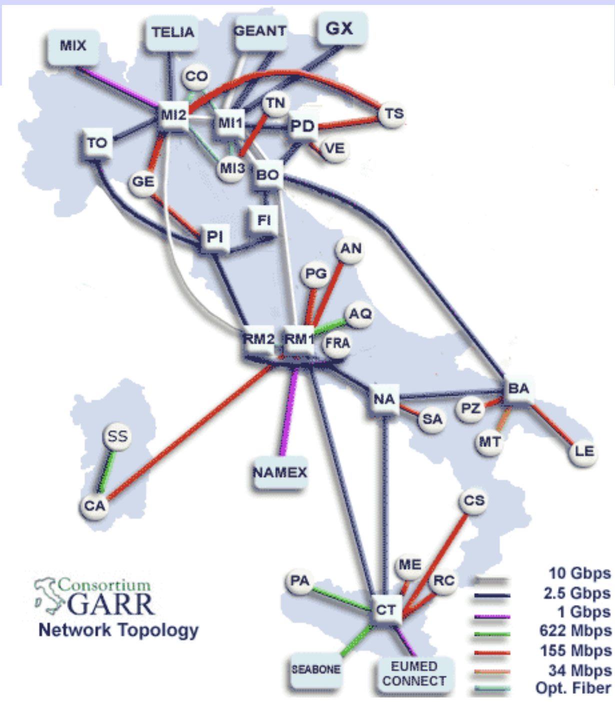


JANET External Network Access Provision

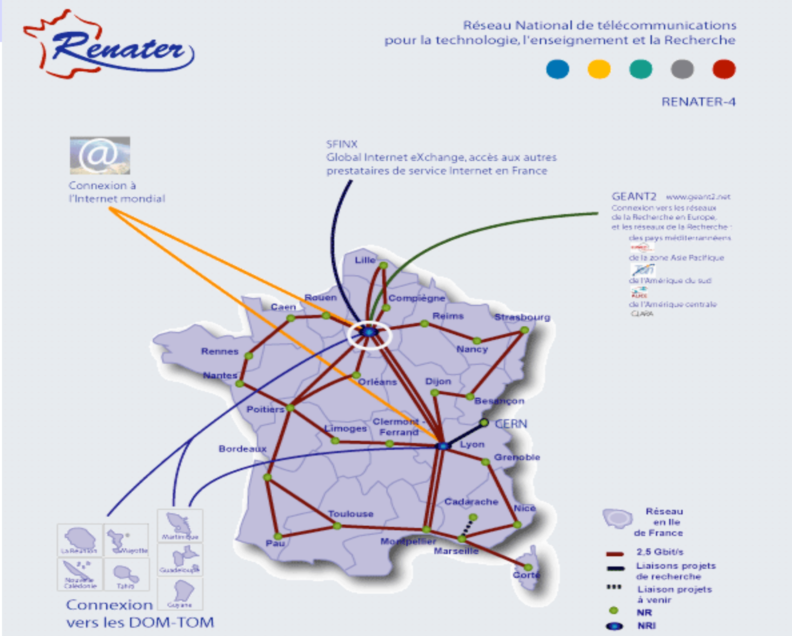




GÉANT2 is operated by DANTE on behalf of Europe's NRENs.

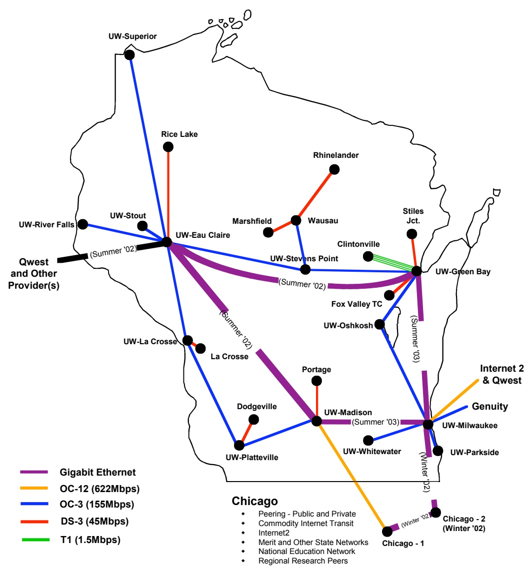


RENATER-4 is deployed since september 2005

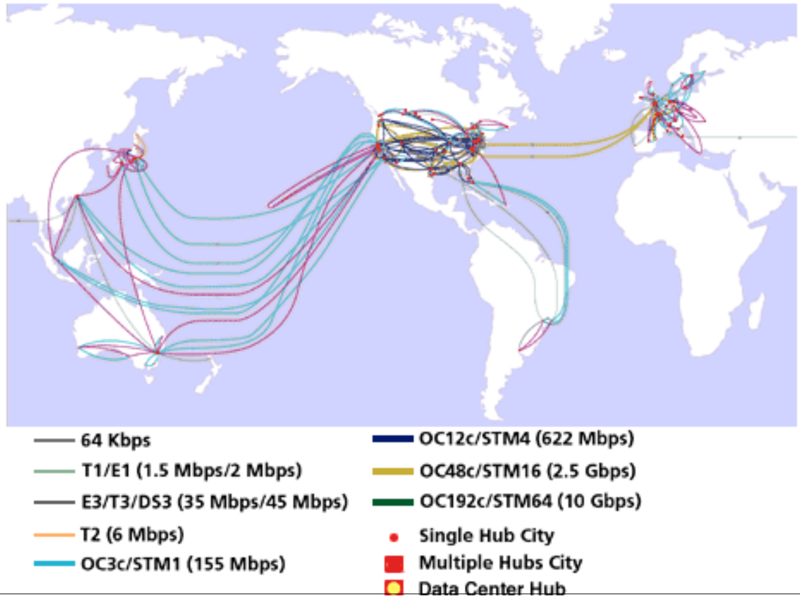


wiscnet.net

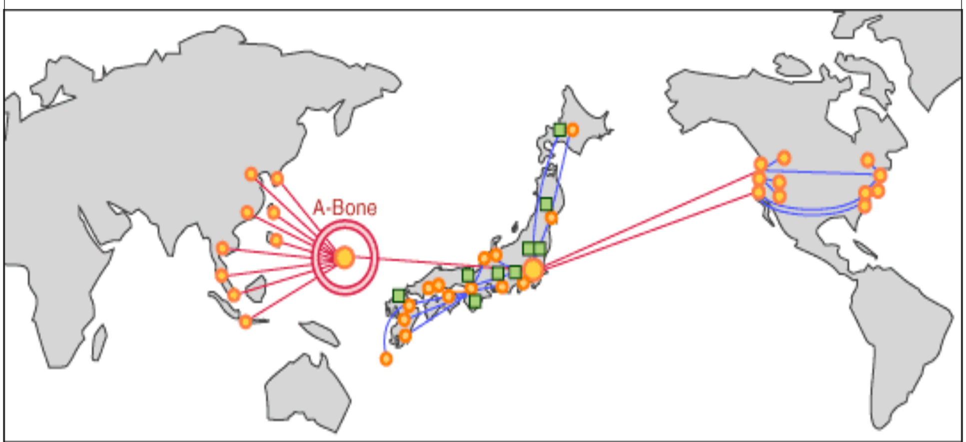
GO BUCKY!



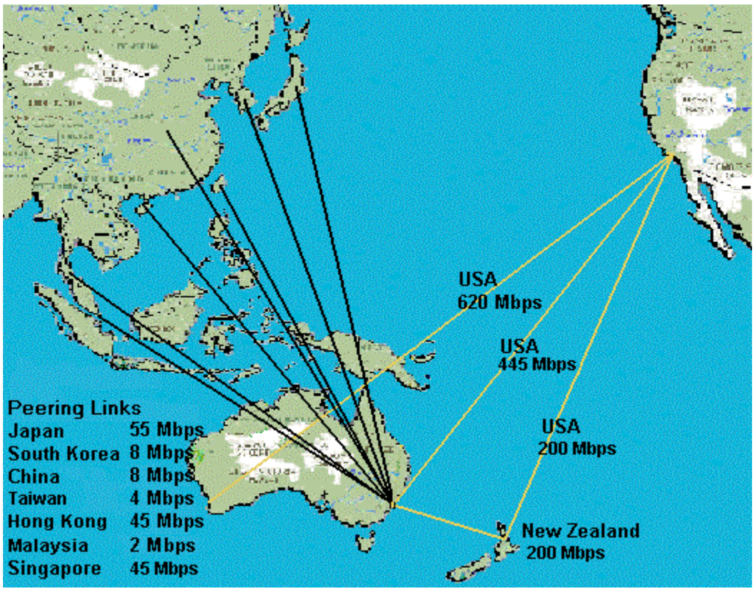
WorldCom (UUNet)



Internet Initiative Japan (IIJ)



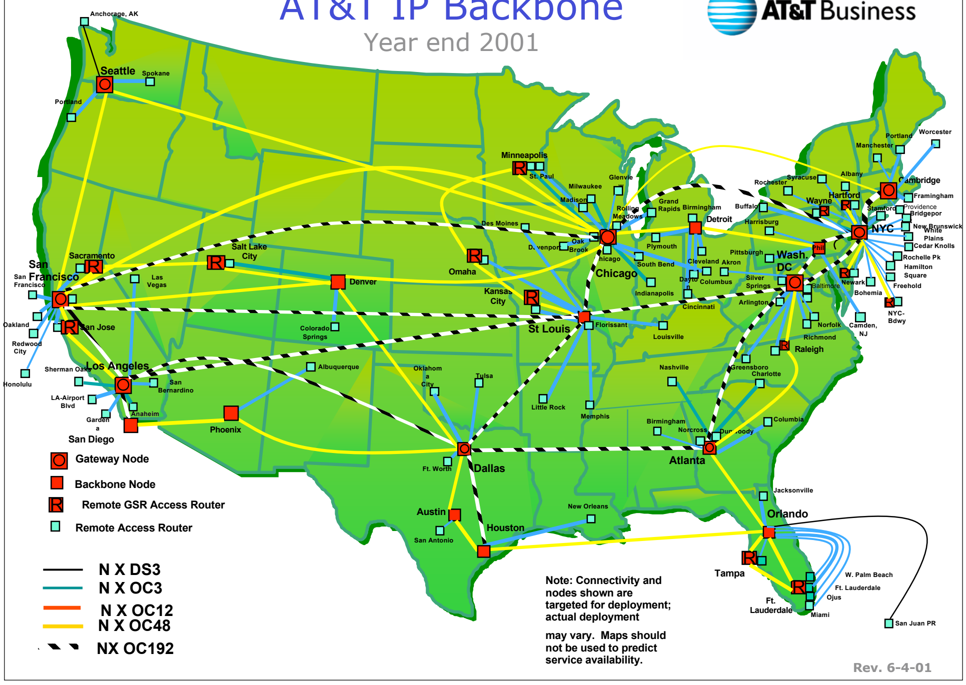
Telstra international



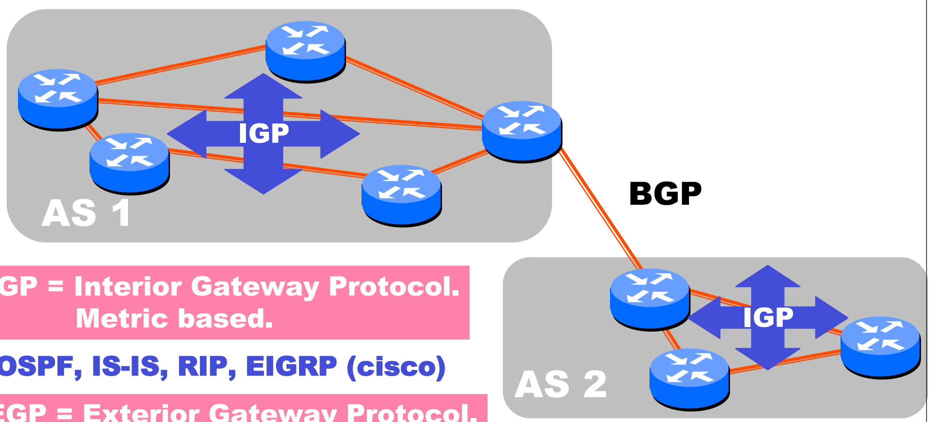
AT&T IP Backbone

Year end 2001

AT&T Business



Architecture of Dynamic Routing



IGP = Interior Gateway Protocol. Metric based.

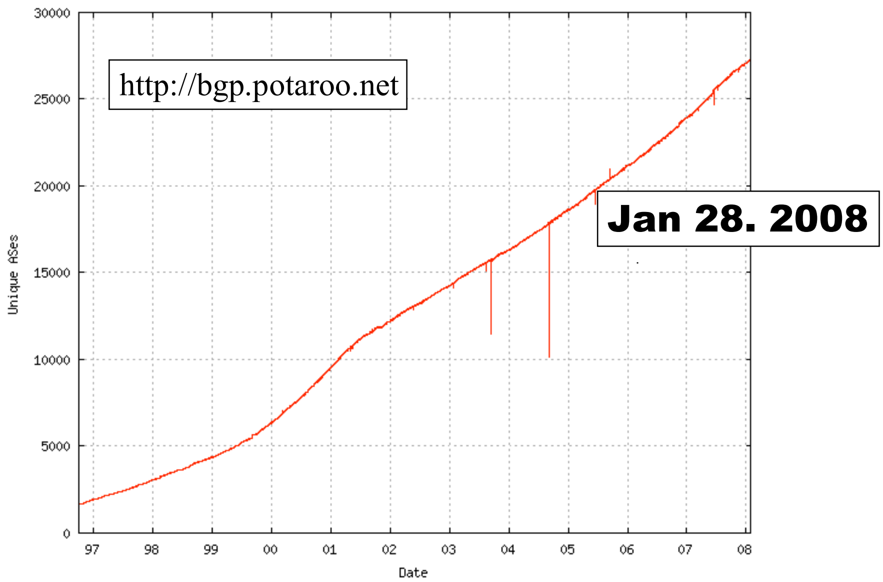
OSPF, IS-IS, RIP, EIGRP (cisco)

EGP = Exterior Gateway Protocol. Policy Based.

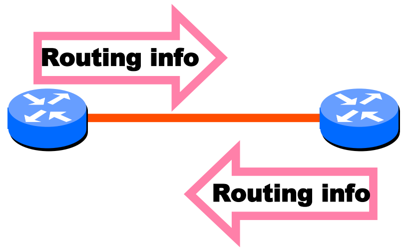
Only one: BGP

The Routing Domain of BGP is the entire Internet

How many ASN are used today?



Routers Talking to Routers (The "control plane")



- **Routing computation is distributed among routers within a routing domain**
- **Computation of best next hop based on routing information is the most CPU/memory intensive task on a router**
- **Routing messages are *usually* not routed, but exchanged via layer 2 between physically adjacent routers (internal BGP and multi-hop external BGP are exceptions)**

Technology of Distributed Routing

Link State

- Topology information is flooded within the routing domain
- Best end-to-end paths are computed locally at each router.
- Best end-to-end paths determine next-hops.
- Based on minimizing some notion of distance
- Works only if policy is shared and uniform
- Examples: OSPF, IS-IS

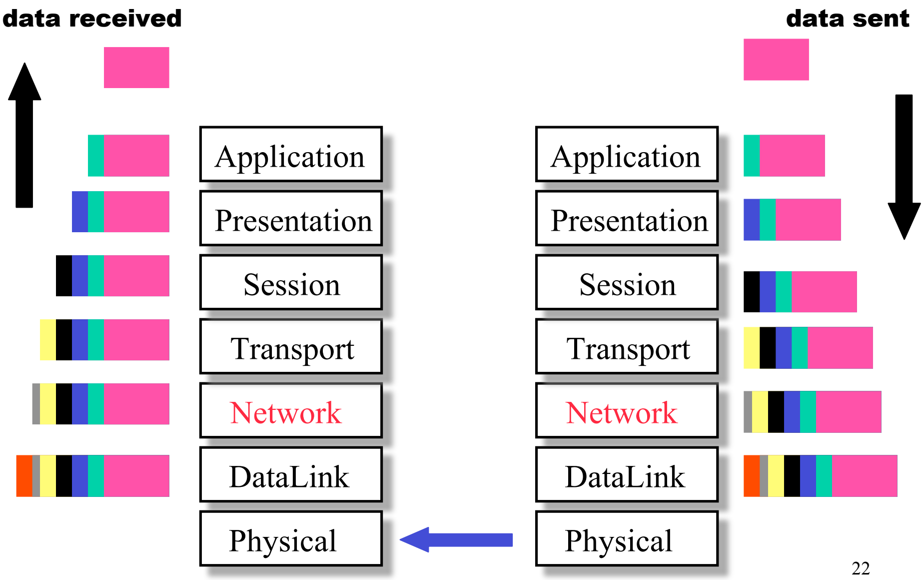
Vectoring

- Each router knows little about network topology
- Only best next-hops are chosen by each router for each destination network.
- Best end-to-end paths result from composition of all next-hop choices
- Does not require any notion of distance
- Does not require uniform policies at all routers
- Examples: RIP, BGP

The Gang of Four

	Link State	Vectoring
IGP	OSPF IS-IS	RIP
EGP		BGP

The standard model



22

All Hail the IP Datagram!

HEADER

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version|  IHL  |Service Type |          Total Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Identification        |Flags|    Fragment Offset    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Time to Live |    Protocol   |         Header Checksum       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Destination Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

DATA

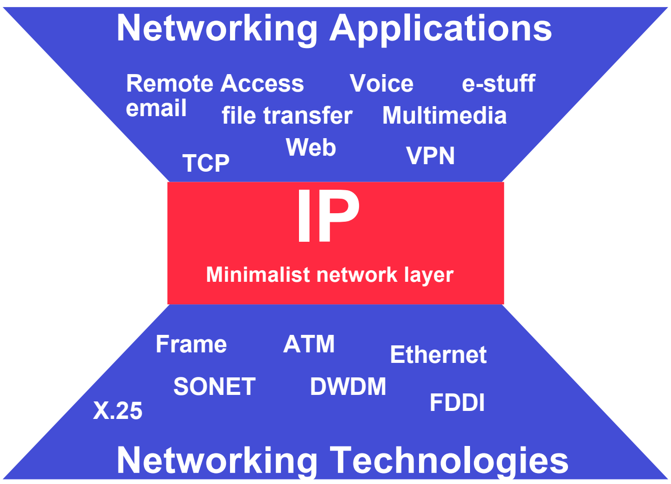
```
  :                                                             :
  :                                                             :
  |                                                             |
  +  ... up to 65,515 octets of data ...                        +
  |                                                             |
  +                                                             +
  |                                                             |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

shaded fields little-used today

1981, RFC 791

23

IP Hour Glass



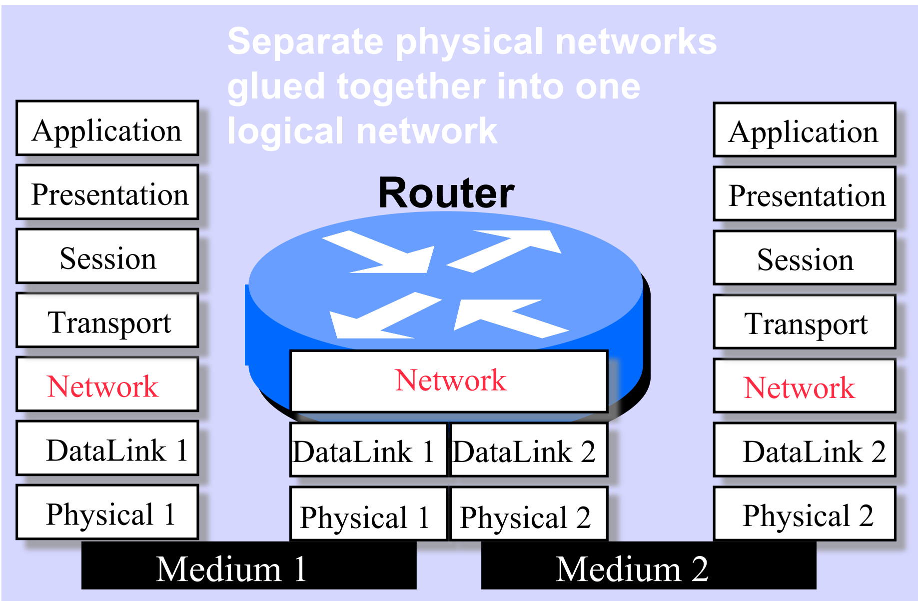
24

Best Effort, Connectionless, Connectivity

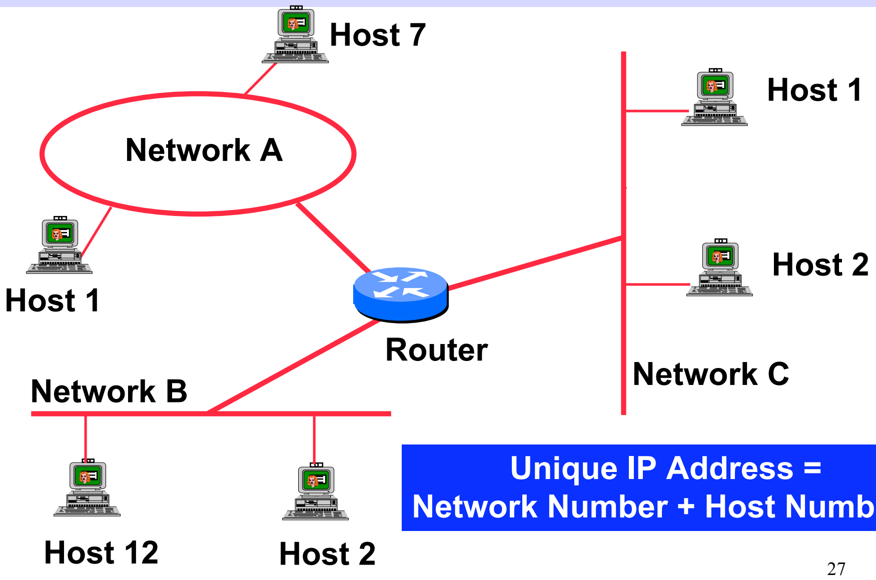
The is the fundamental service provided by Internet Service Providers (ISPs)

All other IP services depend on connectivity: DNS, email, VPNs, Web Hosting, ...

IP is a Network Layer Protocol

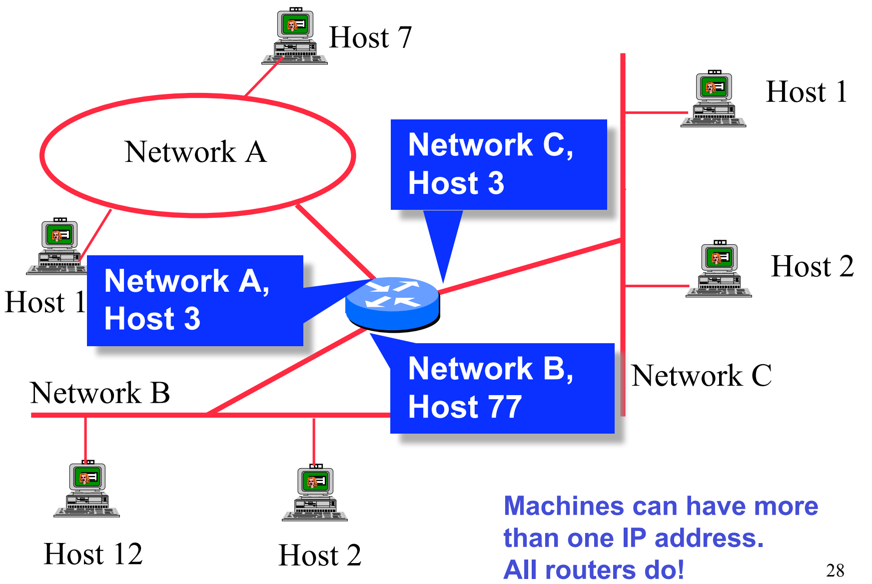


Hosts, Networks, and Routers



Unique IP Address = Network Number + Host Number

Actually, IP addresses identify Interfaces



Machines can have more than one IP address.
All routers do!

IP Forwarding Table

Destination	Next Hop	Interface
Net A	Router 1	INT 7
Net B	Direct	INT 4
Net C, Host 3	Router 2	INT 3
Net C	Router 1	INT 7

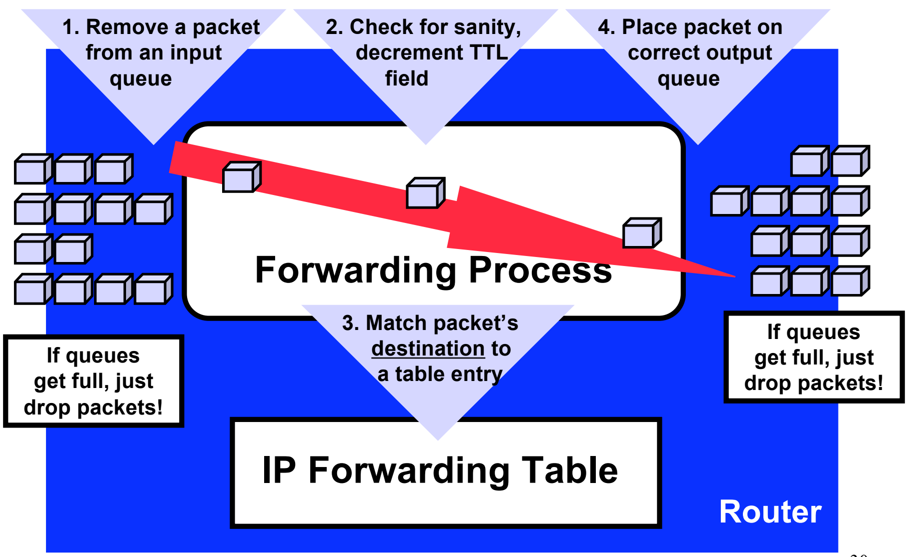
A destination is usually a network. May also be a host, or a "gateway of last resort" (default)

The next hop is either a directly connected network or a router on a directly connected network

A physical interface

29

IP Forwarding Process



30

IPv4 Addresses are 32 Bit Values

11111111	00010001	10000111	00000000
255	17	134	0

255.17.134.0

Dotted quad notation

IPv6 addresses have 128 bits

31

IP Addresses come in two parts

11111111	00010001	10000111	00000000

|← Network Number →|← Host Number →|

Where is this dividing line?
Well, that depends ....

32

Classful Addresses

Class A	0nnnnnnn	hhhhhhhh	hhhhhhhh	hhhhhhhh
Class B	10nnnnnn	nnnnnnnn	hhhhhhhh	hhhhhhhh
Class C	110nnnnn	nnnnnnnn	nnnnnnnn	hhhhhhhh

n = network address bit   h = host identifier bit

Leads to a rigid, flat, inefficient use of address space …

33

RFC 1519: Classless Inter-Domain Routing (CIDR)

Use two 32 bit numbers to represent a network.
Network number = IP address + Mask

IP Address : 12.4.0.0   IP Mask: 255.254.0.0

Address	00001100	00000100	00000000	00000000
Mask	11111111	11111110	00000000	00000000

|← Network Prefix →|← for hosts →|

Usually written as 12.4.0.0/15

34

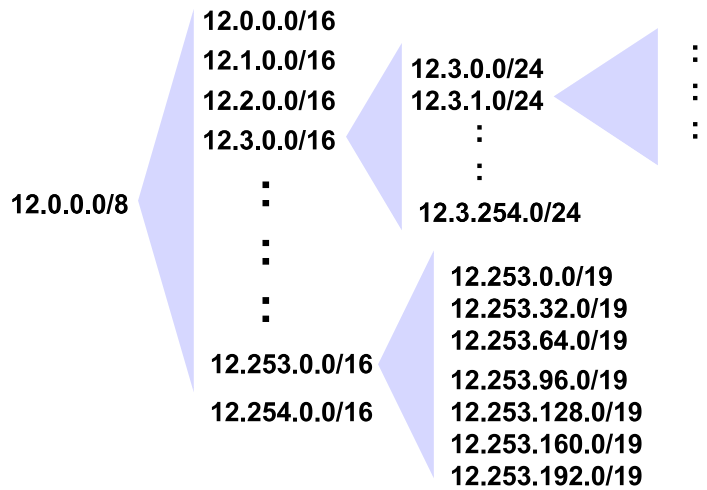
Which IP Addresses are Covered by a Prefix?

12.5.9.16 is covered by prefix 12.4.0.0/15

12.5.9.16	00001100	00000101	00001001	00010000
12.4.0.0/15	00001100	00000100	00000000	00000000
	11111111	11111110	00000000	00000000
12.7.9.16	00001100	00000111	00001001	00010000

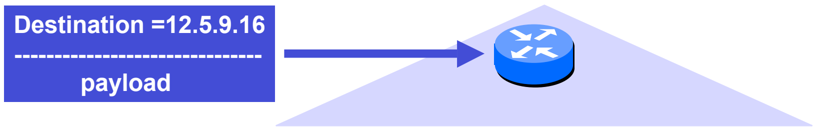
12.7.9.16 is not covered by prefix 12.4.0.0/15

CIDR allows Hierarchy in Addressing



36

Classless Forwarding



	Prefix	Next Hop	Interface
OK	0.0.0.0/0	10.14.11.33	ATM 5/0/9
better	12.0.0.0/8	10.14.22.19	ATM 5/0/8
even better	12.4.0.0/15	10.1.3.77	Ethernet 0/1/3
best!	12.5.8.0/23	attached	Serial 1/0/7

IP Forwarding Table

How Are Forwarding Tables Populated to implement Routing?

Statically

Administrator manually configures forwarding table entries

- \+ More control
- \+ Not restricted to destination-based forwarding
- \- Doesn't scale
- \- Slow to adapt to network failures

Dynamically

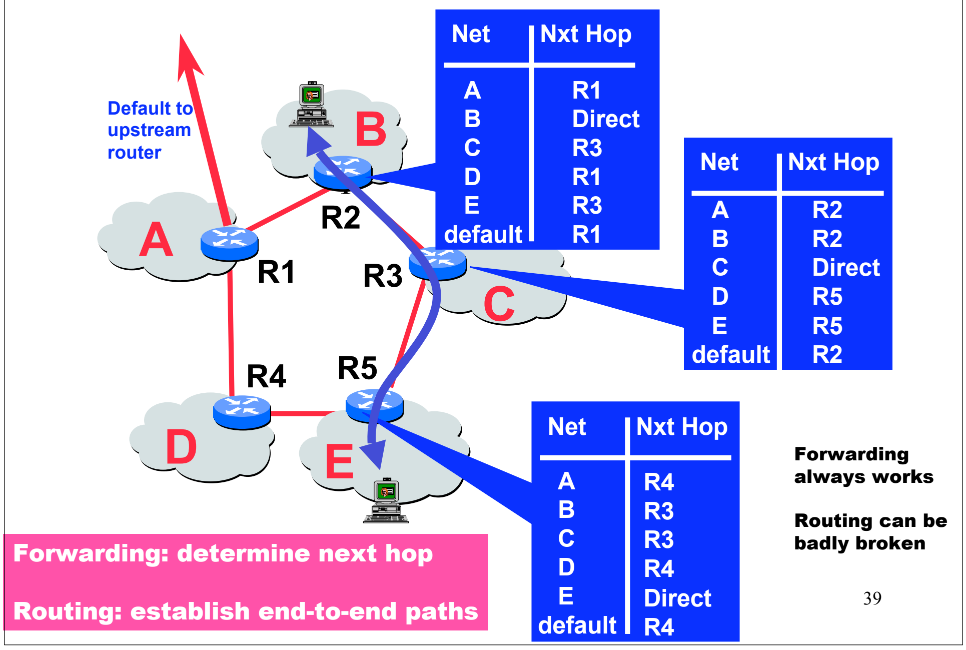
Routers exchange network reachability information using ROUTING PROTOCOLS. Routers use this to compute best routes

- \+ Can rapidly adapt to changes in network topology
- \+ Can be made to scale well
- \- Complex distributed algorithms
- \- Consume CPU, Bandwidth, Memory
- \- Debugging can be difficult
- \- Current protocols are destination-based

In practice : a mix of these. Static routing mostly at the "edge"

38

Routing vs. Forwarding



Forwarding: determine next hop

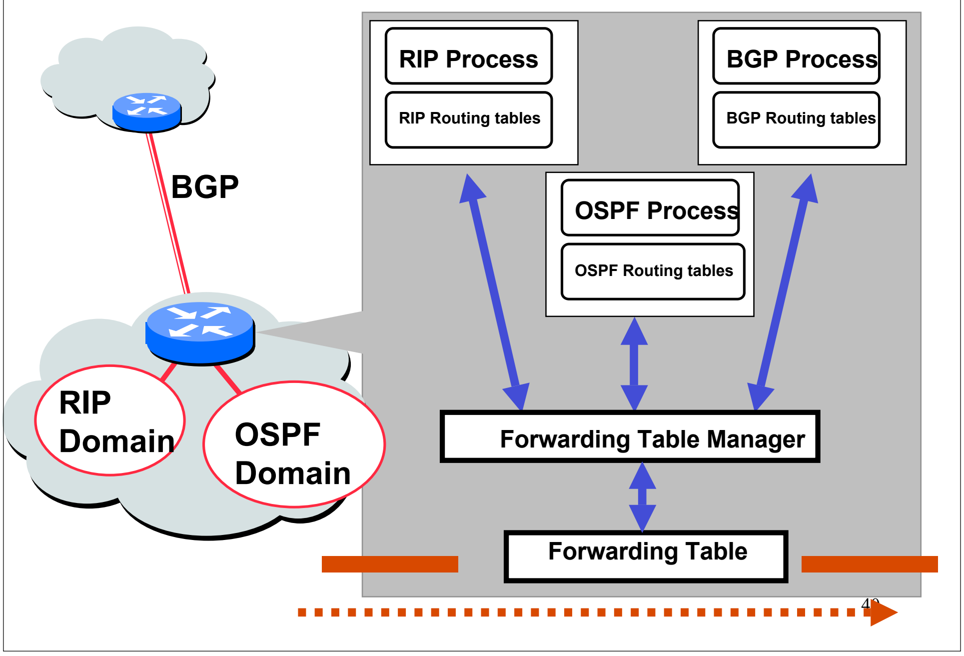
Routing: establish end-to-end paths

Forwarding always works

Routing can be badly broken

39

Happy Packets: The Internet Does Not Exist Only to Populated Routing Tables



40

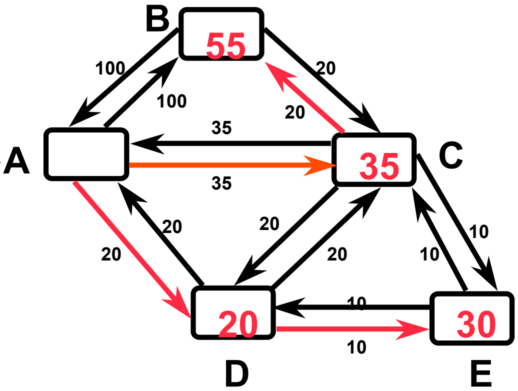
Before We Go Any Further ...

IP ROUTING PROTOCOLS DO NOT DYNAMICALLY ROUTE AROUND NETWORK CONGESTION

- **IP traffic can be very bursty**
- **Dynamic adjustments in routing typically operate more slowly than fluctuations in traffic load**
- **Dynamically adapting routing to account for traffic load can lead to wild, unstable oscillations of routing system**

Next Lecture: Shortest Path Routing

Dest.	Nxt Hop
B	C
C	C
D	D
E	D



This is what IS-IS, OSPF, and RIP do, more or less.

42