# IP Routing Tutorial

## Introduction to IP Routing Protocols

**Stefan Fouant**

*ShortestPathFirst Consulting Services*

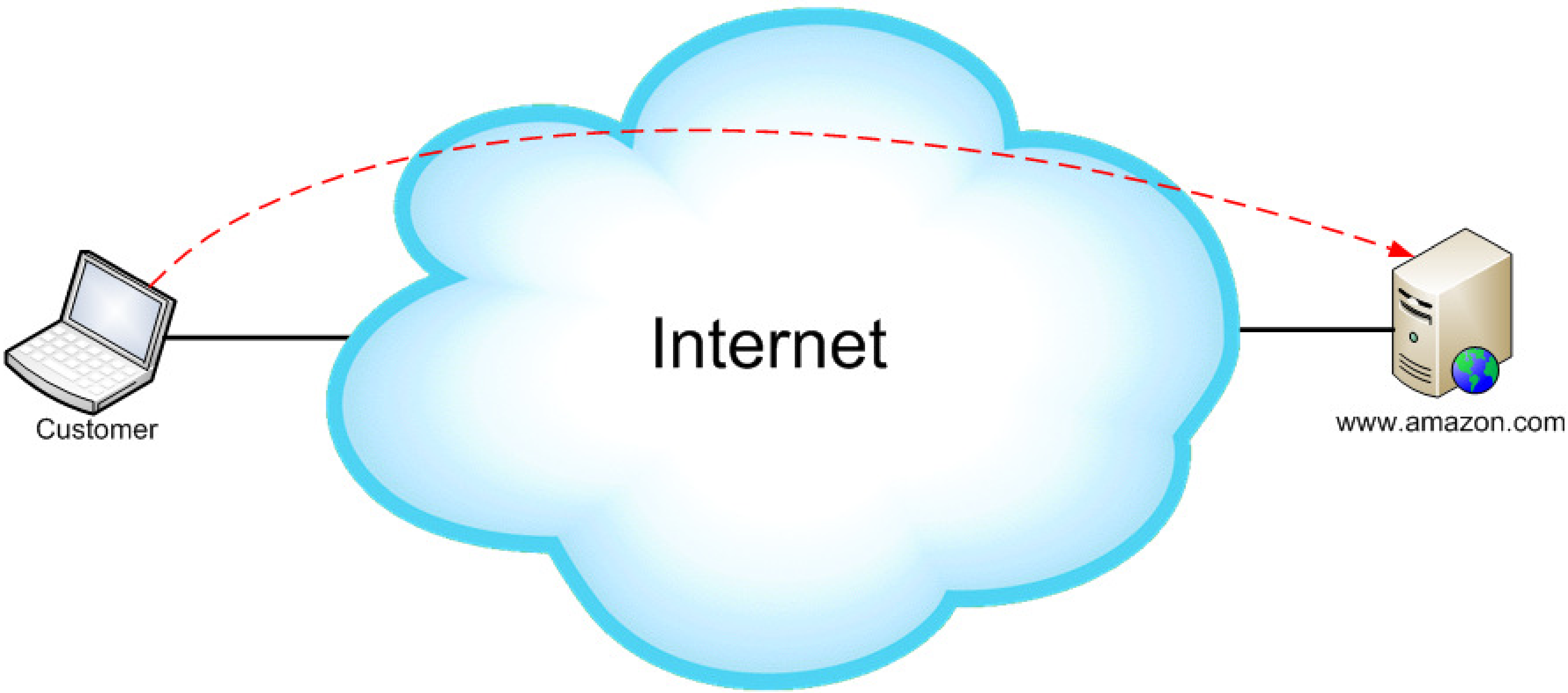
*www.shortestpathfirst.com*

# Agenda

- Routing – The 10,000 Foot View
- Routing in the Internet: A Brief History
  - Static Routing and the evolution of Dynamic Routing Protocols
- Interior Gateway Protocols
  - Purpose
  - Distance-Vector vs. Link-State
- Exterior Gateway Protocols
  - Purpose
  - Autonomous Systems
  - Policy
- Multi-Protocol Label Switching
  - Shortest-Cost Routing vs. Traffic-Engineering
  - MPLS VPNs
- Quality of Service
- The Future of Internet Routing

# Introduction

- How does data go from one place on the Internet to another?



# Routing – The Easy Explanation?

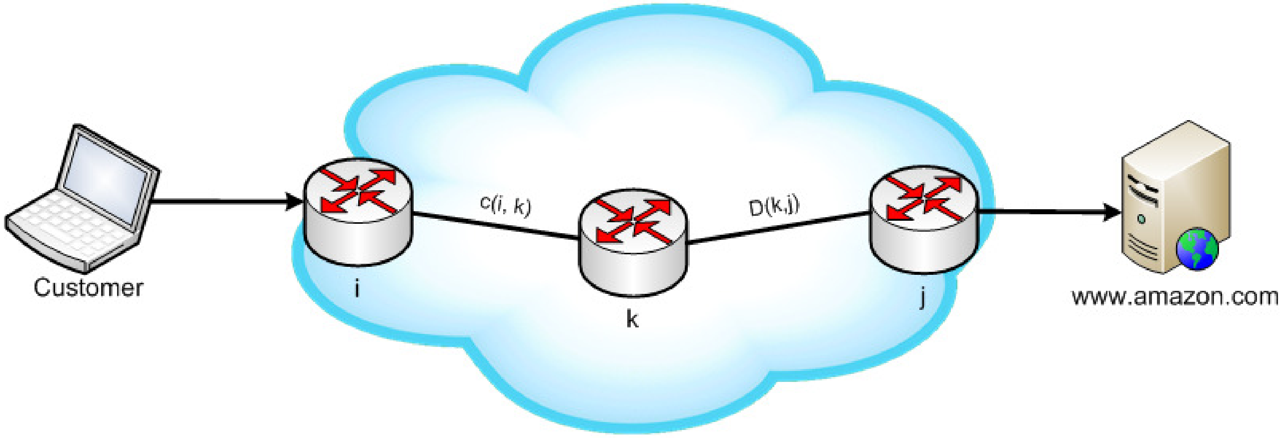
- The "Easy" answer:
  - » If the shortest path from node i to node j, with distance D(i,j) passes through neighbor k, with link cost c(i,k), then:

**D(i,j) = c(i,k) + D(k,j) (Distance-Vector Protocols)**

**or**

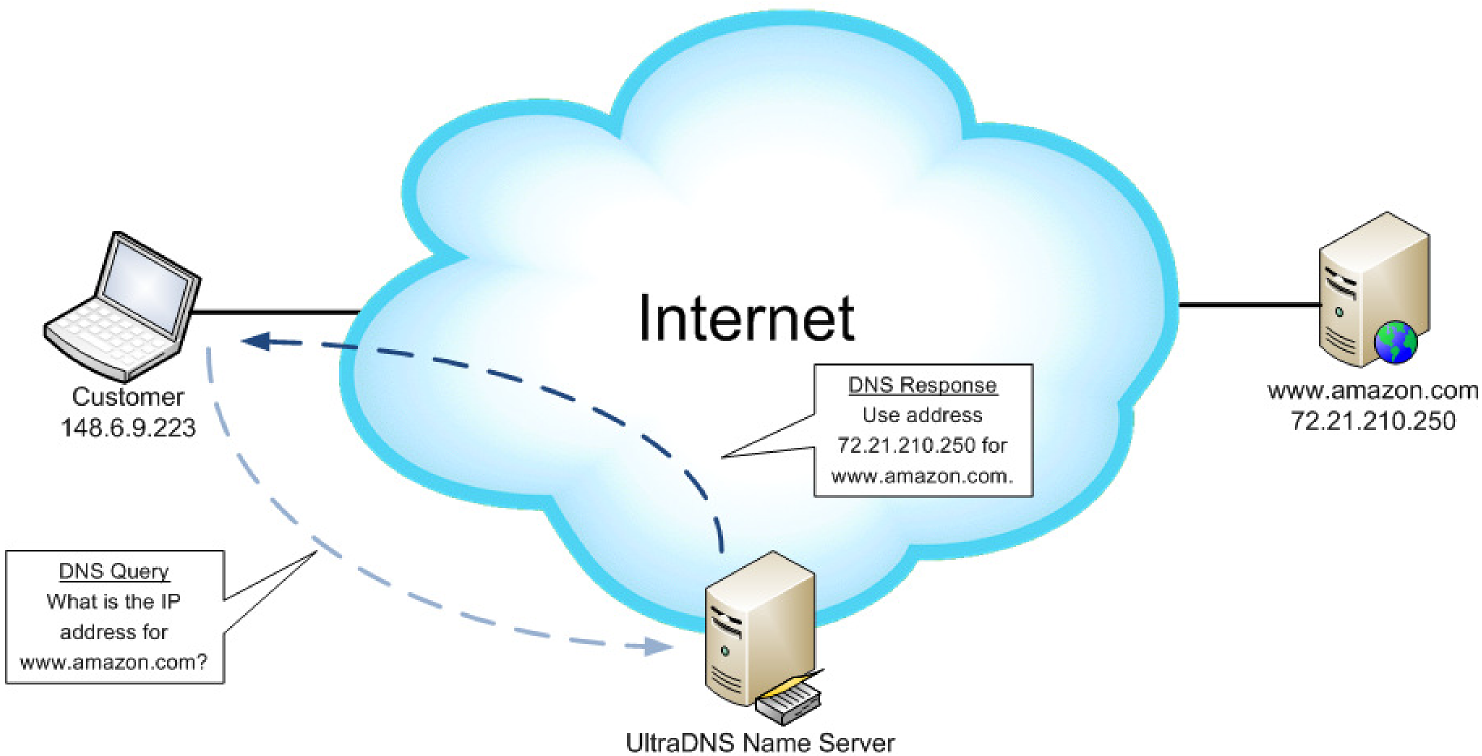
**D(i,j) = D(i,k) + c(k,j) (Link-State Protocols)**

Confused???



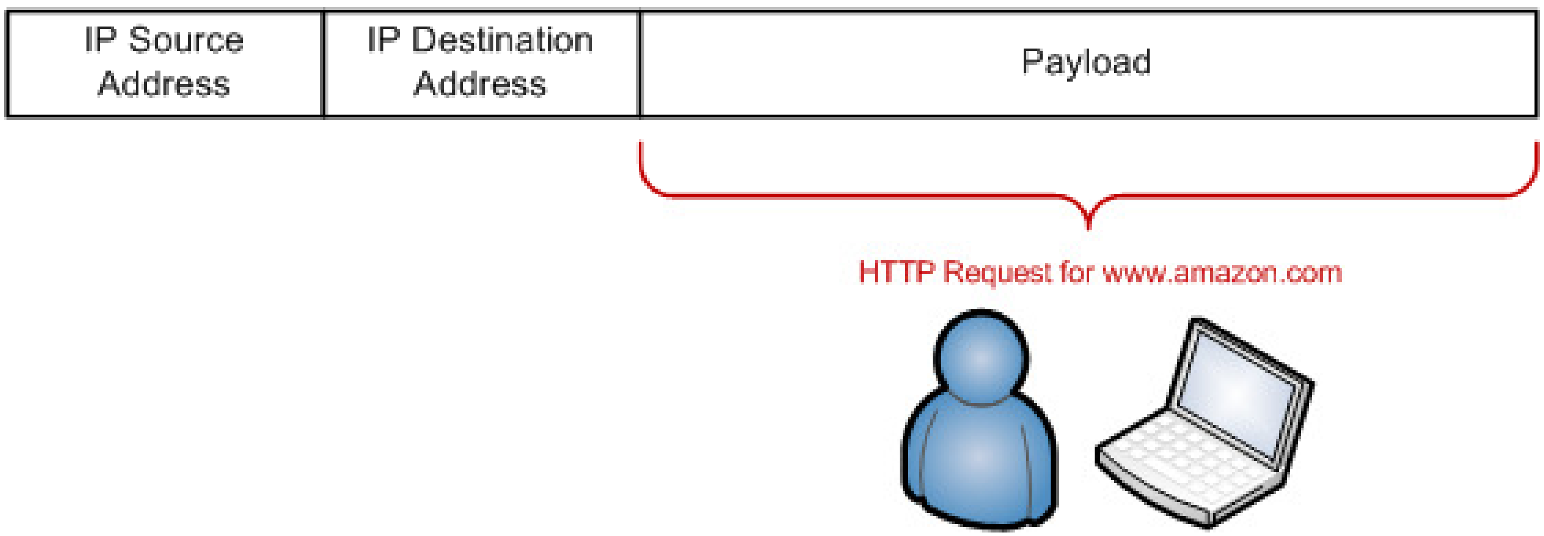
# Routing – the 10,000 Foot View

- Computers use numerical Internet Protocol (IP) addresses to reach each other
- DNS provides a mechanism whereby a user-friendly DNS name (i.e. www.amazon.com) can be translated to a numerical IP address (i.e. 72.21.210.250) which can than be used to reach the intended destination
- Once we have the destination IP address, we need to package our data into IP packets



# Routing – the 10,000 Foot View

- Whenever data is transmitted from one location to another, that data is broken into chunks and wrapped into structures called IP packets
- Each packet contains information about the IP address of the source and the destination nodes, sequence numbers and some other control information



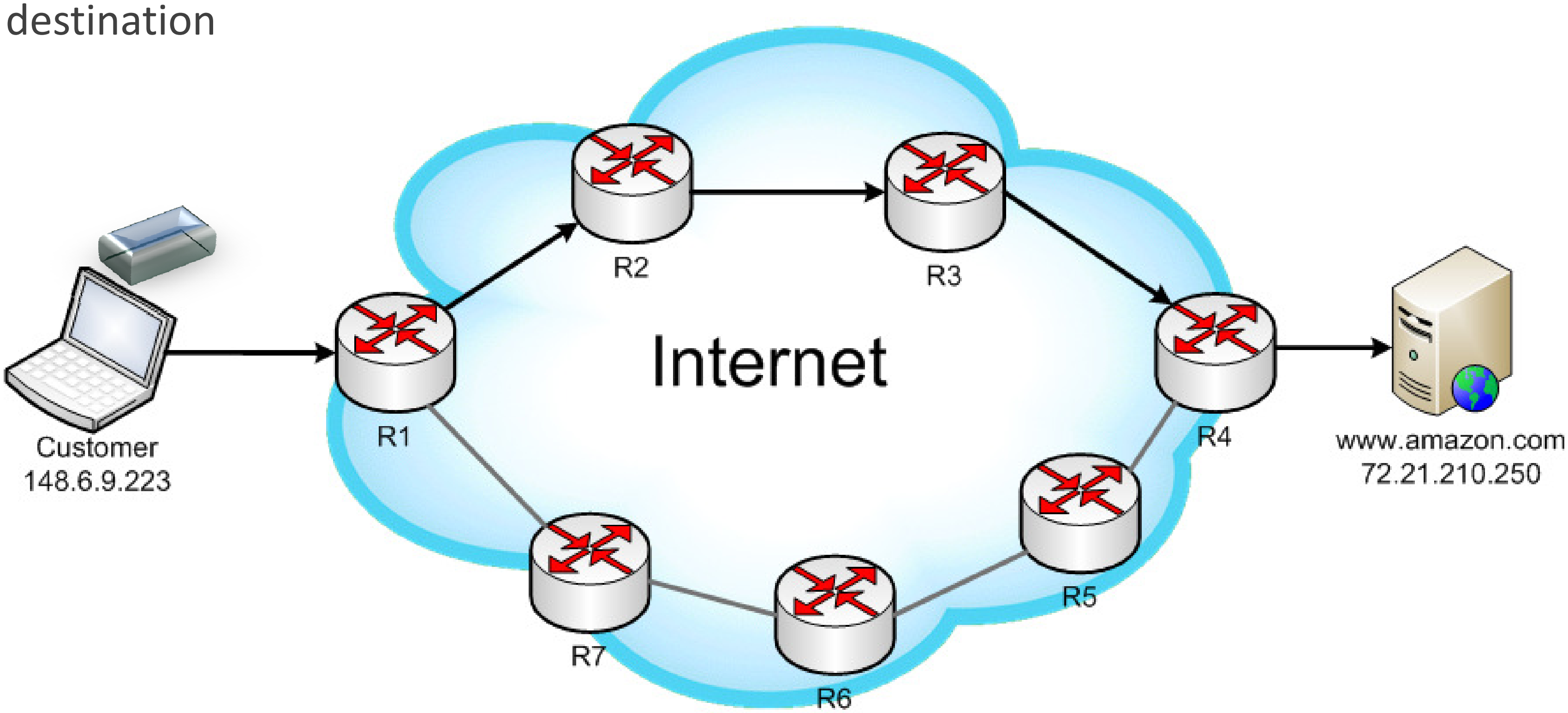

- Once they reach the destination, the IP headers are removed and packets are reassembled to make up the original data again

# Routing – the 10,000 Foot View

- The IP addressing information in the packet is what is used to forward packets towards the intended destination
- Source address information is included in the IP packet so that the destination knows how to direct packets back to the source
- Once the IP packet is created, routers forward the packet towards the destination

# Routing – the 10,000 Foot View

- Routers are the intelligent glue which are used to direct traffic from one point on the Internet to another
- Typically, this is done by looking at the destination IP address of the data packet and consulting a routing table in order to determine where to route the traffic
- Each router makes a **local** decision as to how to forward the packet towards it intended destination



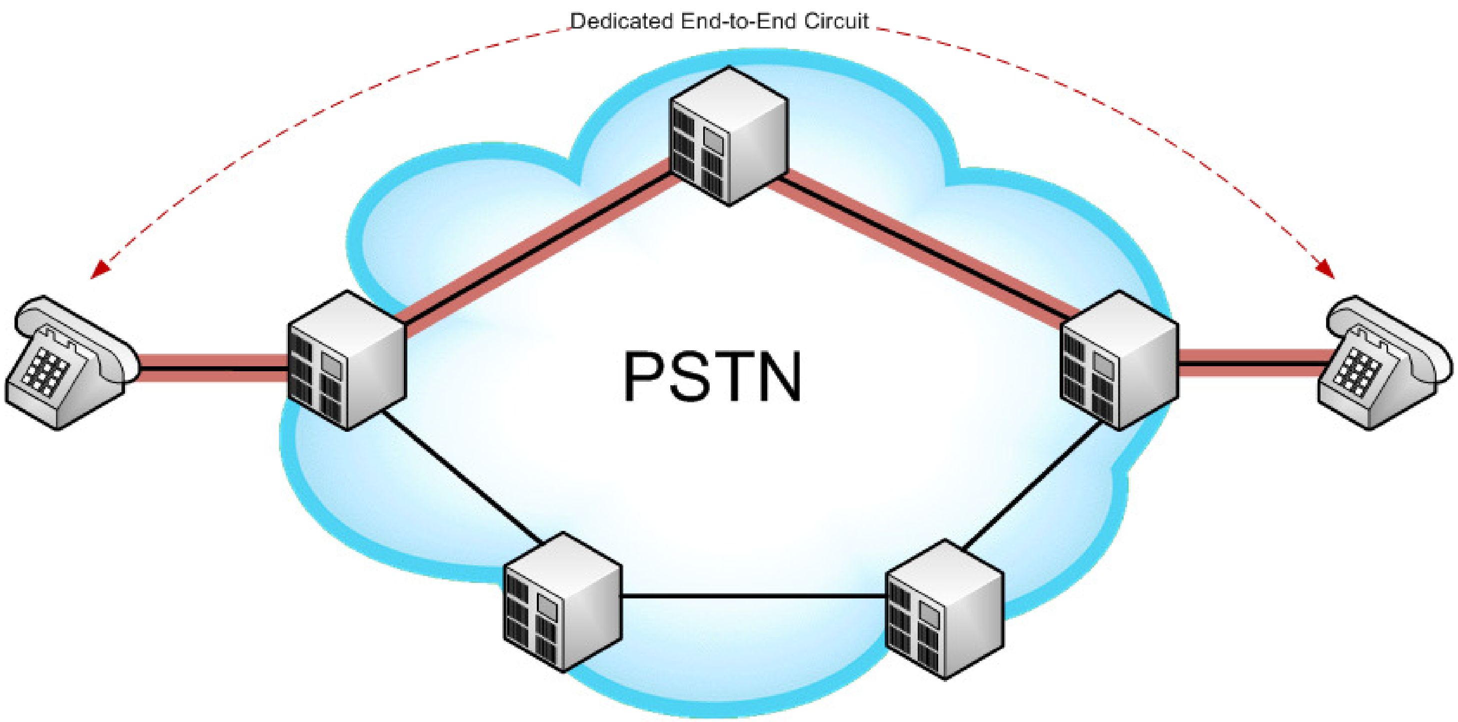
# Routing – the 10,000 Foot View

- Sending information from one location to another via electronic means is not something new, in fact the Public Switched Telephone Network (PSTN) has been doing this for a long time!

- Routing of phone calls on the PSTN uses a connection-oriented mechanism known as circuit-switching
- The process for routing of IP packets over the Internet is termed connectionless-based forwarding via a mechanism known as packet-switching

# Circuit Switching Overview

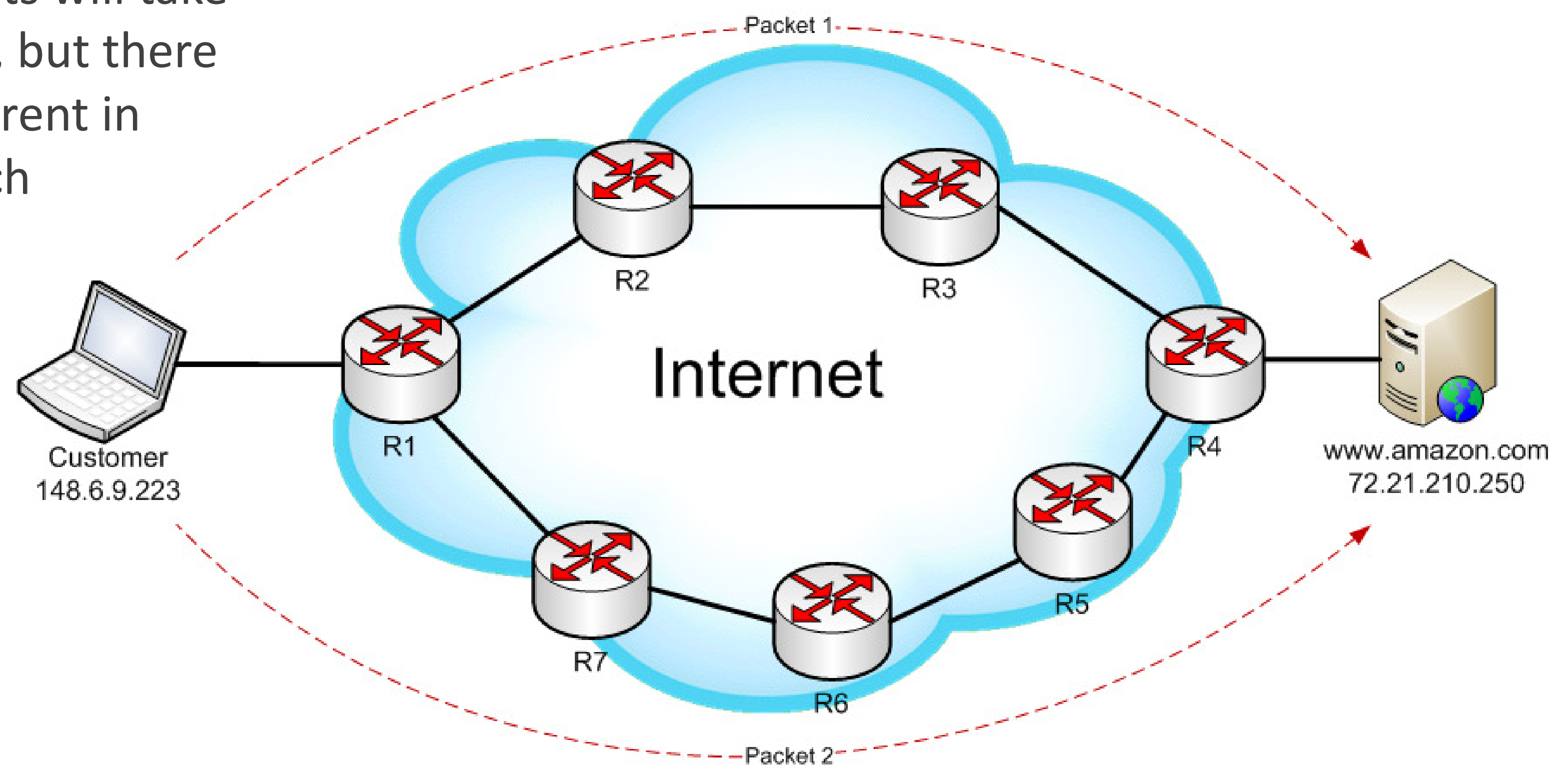
- In **circuit-switching**, the path is decided upon <u>before</u> the data transmission starts. The system decides on which route to follow, based on a resource-optimizing algorithm, and transmission goes according to the path
- For the whole length of the communication session between the two communicating bodies, the route is dedicated and exclusive, and released only when the session terminates




ShortestPathFirst Consulting Services
www.shortestpathfirst.net

# Packet Switching Overview

- In **packet-switching**, the packets are sent towards the destination irrespective of each other
- There is no predetermined path; each router makes a **local** decision as to how to forward the packet towards it intended destination
- Typically packets will take the same path, but there is nothing inherent in IP routing which dictates that each packet must take the same path



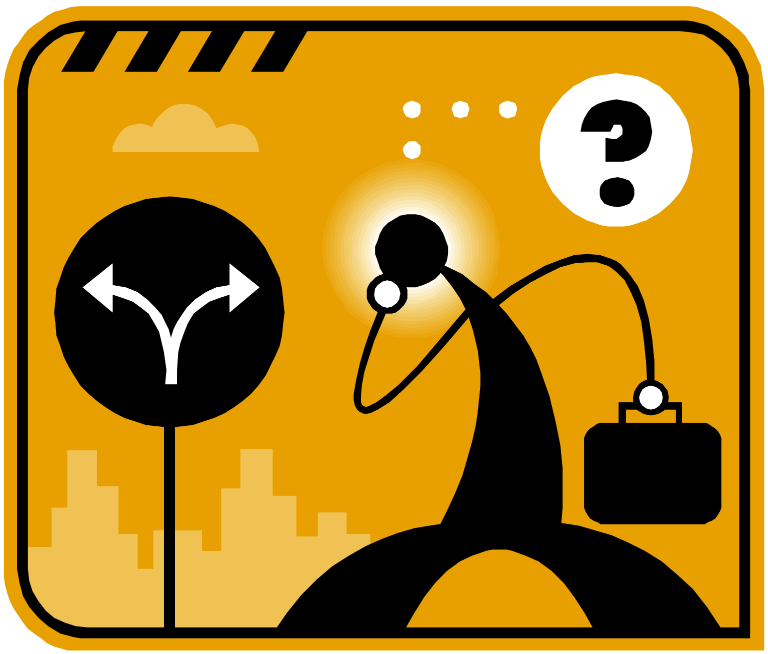
# Circuit vs. Packet Switching

- Circuit switching is old and expensive, and it is what PSTN uses
- Packet switching is more modern, and allows traffic to be rerouted around failures
- As you may have figured out already, the PSTN uses circuit switching whereas emerging technologies such as VoIP uses packet switching

| Circuit Switching | Packet Switching |
| --- | --- |
| End-to-end circuit | Packets represent pieces of data |
| Sequenced communication | Unordered transmission |
| PSTN | Internet |
| Traditional telephony | VoIP |
| Not cost efficient | Shared cost model |
| Less Delay | Higher Delay |
| Highly Reliable, but failures typically mean the circuit must be rebuilt | Less reliable, but can reroute around failures |

# Distribution of Routing Information in Packet Switched Networks

- How do individual routers know how to perform the correct forwarding decision for a given destination address?
  - Through knowledge of the topology state of the network
  - This knowledge is maintained and distributed via routing protocols



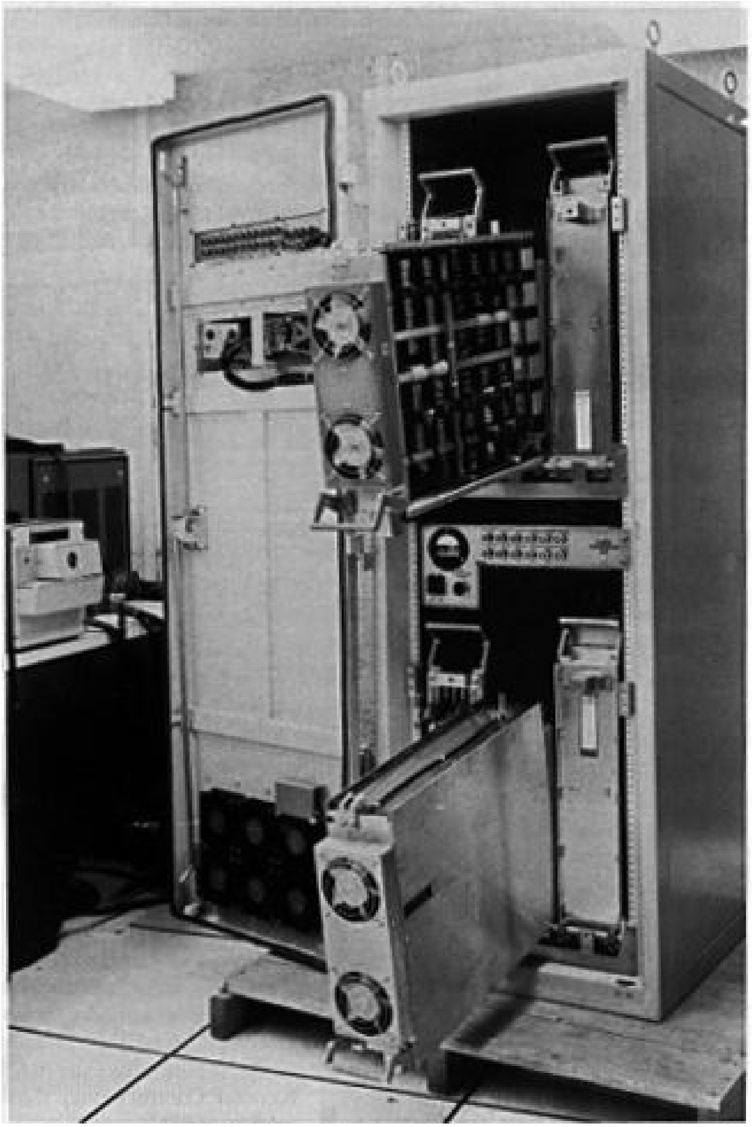
# Routing in the Internet: A Brief History

**Way back in the day…**

**when Vint Cerf still had hair…**

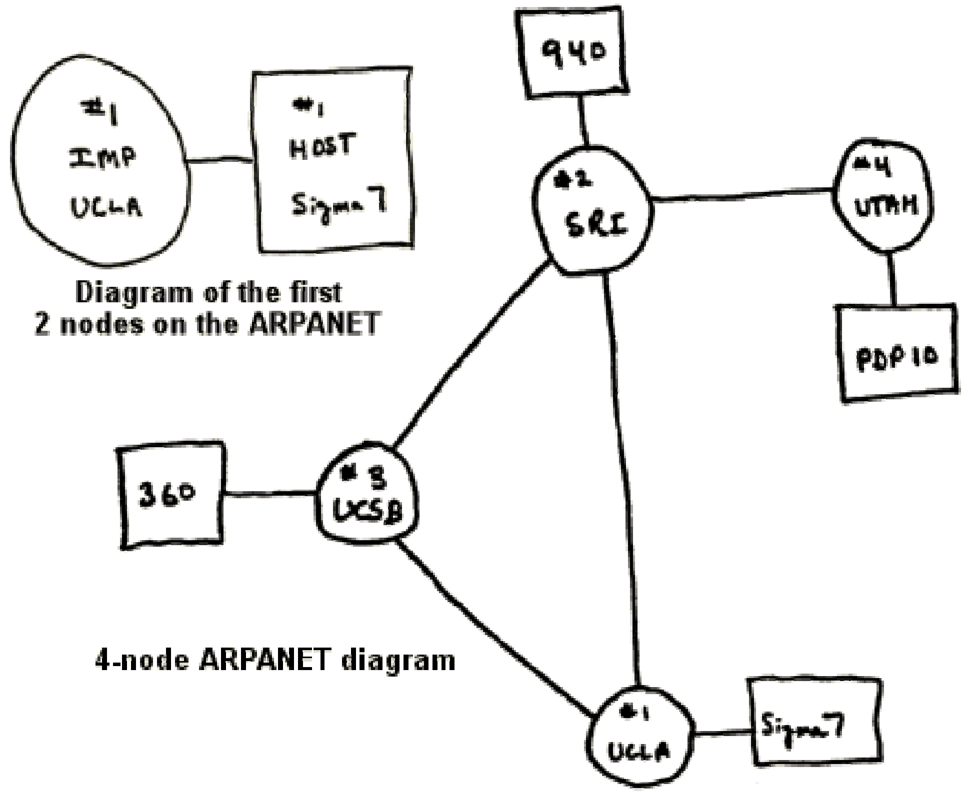
# Routing in the Internet: A Brief History

- In the early days of the ARPAnet, when what would later become known as 'The Internet' contained only a few hosts, a device called an Interface Message Processor (IMP) was used to route traffic between various destinations
- This device had fundamentally the same functionality as a router does today, i.e. switching packets based on destination IP addressing



# Routing in the Internet: A Brief History

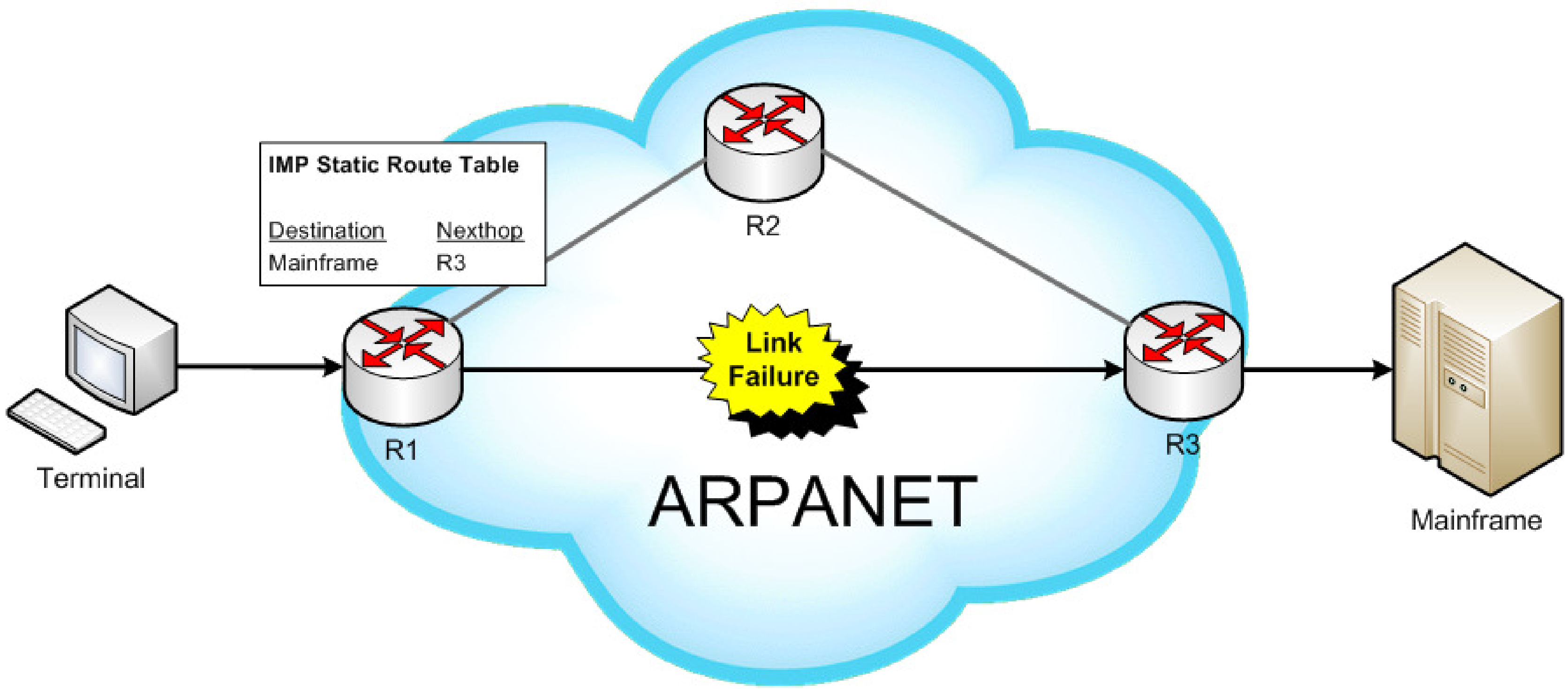
- The creation of the ARPANET began in December 1969, when four nodes were interconnected via 56 Kbps circuits
- With connectivity to only a few hosts, static routing was sufficient
- With static routing, an administrator would manually configure individual routers/IMPs with specific routing instructions
  - Example Instruction Set: From UCLA, in order to get to UTAH, go to SRI




4-node ARPANET diagram

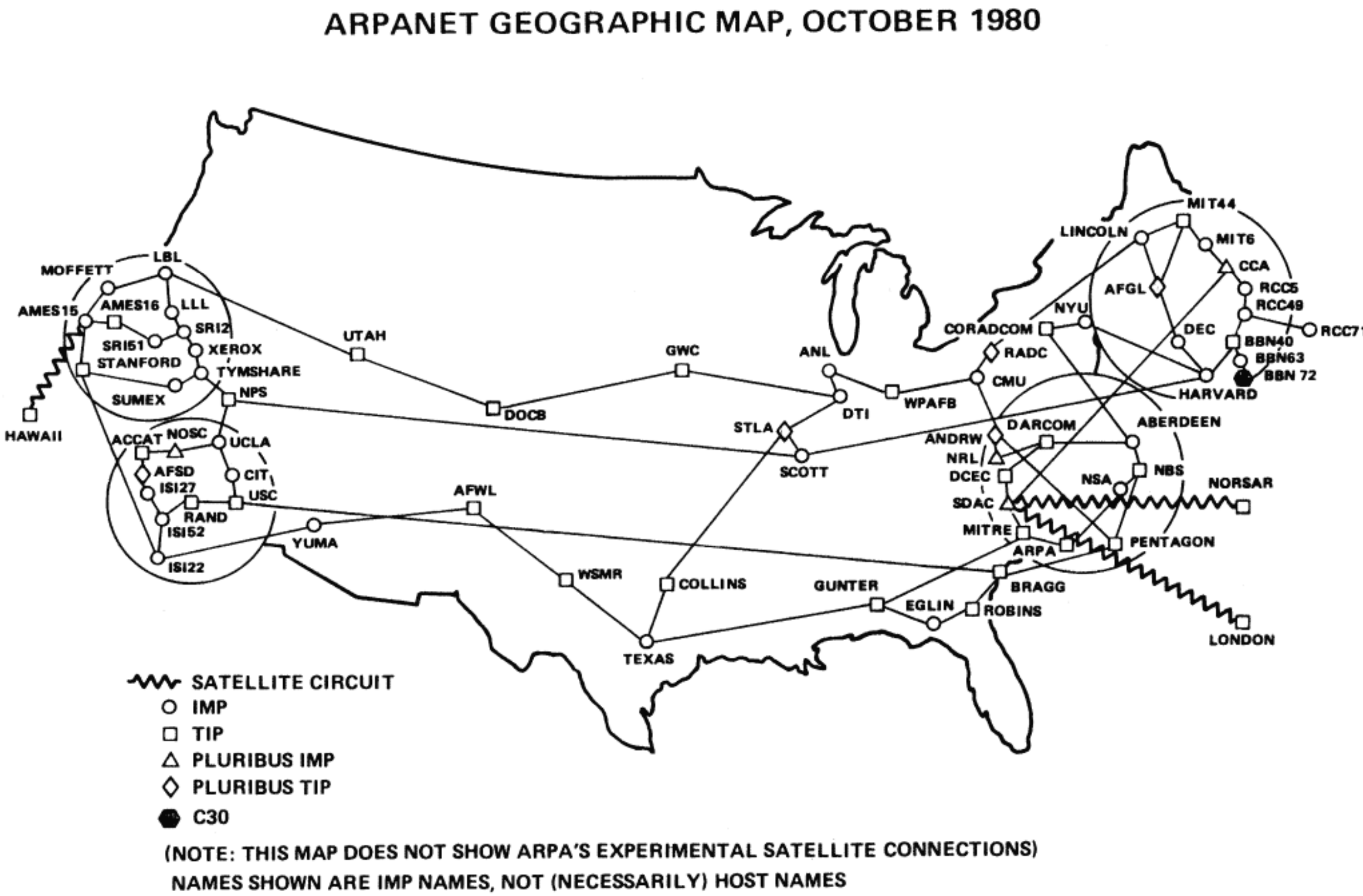
# Routing in the Internet: A Brief History

- Over time, as the Internet began to grow, and a few hosts grew to several thousand hosts, maintaining static routes on individual routers became unwieldy, didn't scale, and was prone to operator error
- Furthermore, static routing could create routing loops and provided little resiliency in the event of network failures

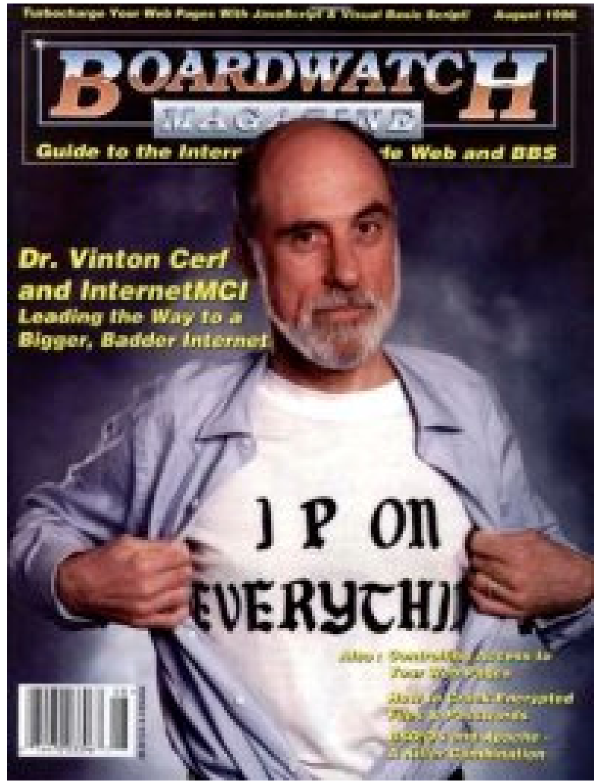


# Routing in the Internet: A Brief History

- By 1980, static routing was no longer a viable option and network operators were already utilizing dynamic routing protocols to distribute routing information and adjust to network failures



# Routing in the Internet: The Present Day




**Fast forward to the present day…**

**… Vint Cerf would like to put "IP on Everything" and has a little less hair**

# Dynamic Routing Protocols

- So… What the heck is a protocol anyway?

  **pro·to·col [proh-tuh-kawl, -kol, -kohl]**

  Definition: *Computers*. a set of rules governing the format of messages that are exchanged between computers.

- Dynamic routing protocols allow routers to exchange network reachability information automatically, as opposed to manually using the static route approach
  - Routers use the information gleaned from routing protocols to compute the best routes to a given destination
  - Allows routers to rapidly adapt to changes in the network topology
  - Scale much better than static routes

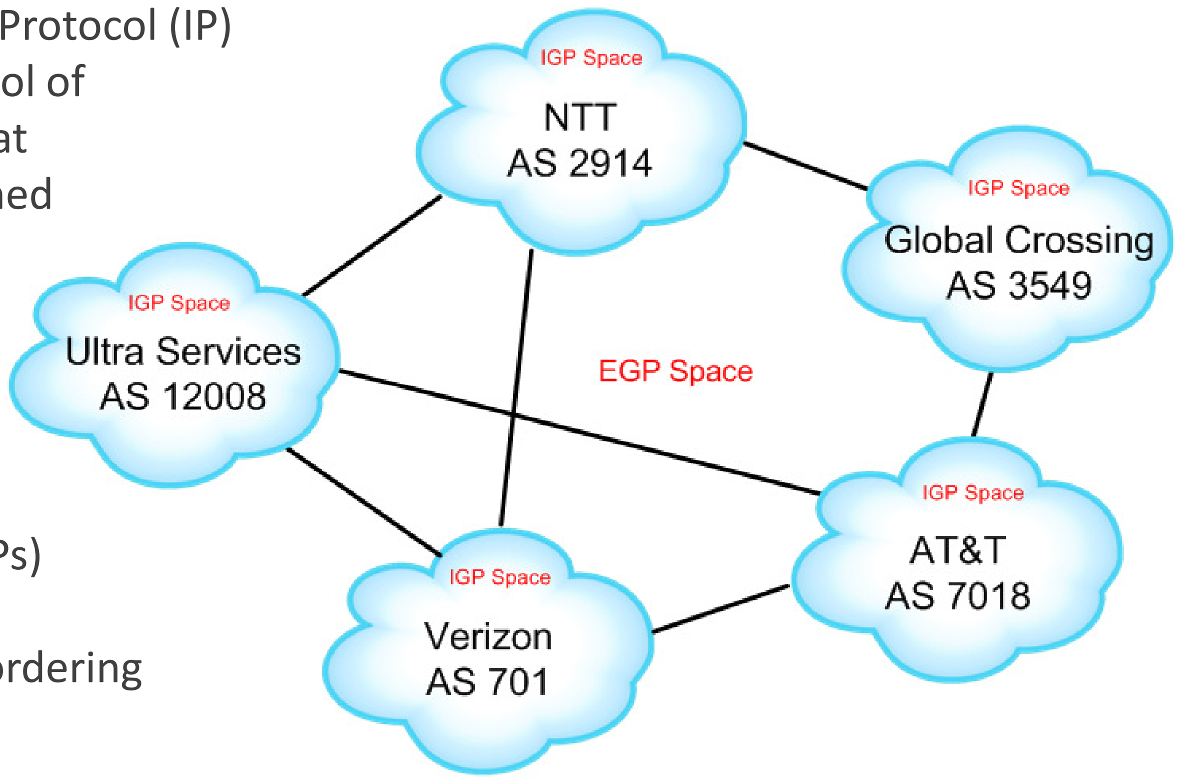
*… even though C3P0 is a protocol droid fluent in over six million forms of communication, his programming took place a long, long time ago, before Internet routing protocols were prevalent*

# Dynamic Routing Protocols

- Dynamic Routing Protocols encompass a diverse set of mechanisms which distribute the topology of the network
  - » Once topology information is distributed, paths can be calculated between two given points, using any set of constraints
  - » The algorithms in use in Dynamic Routing Protocols stem from an area of mathematics known as graph theory
  - » Several applied sciences use graph theory, namely:
    - Highway Engineering
    - Aviation
    - GPS
    - Etc.

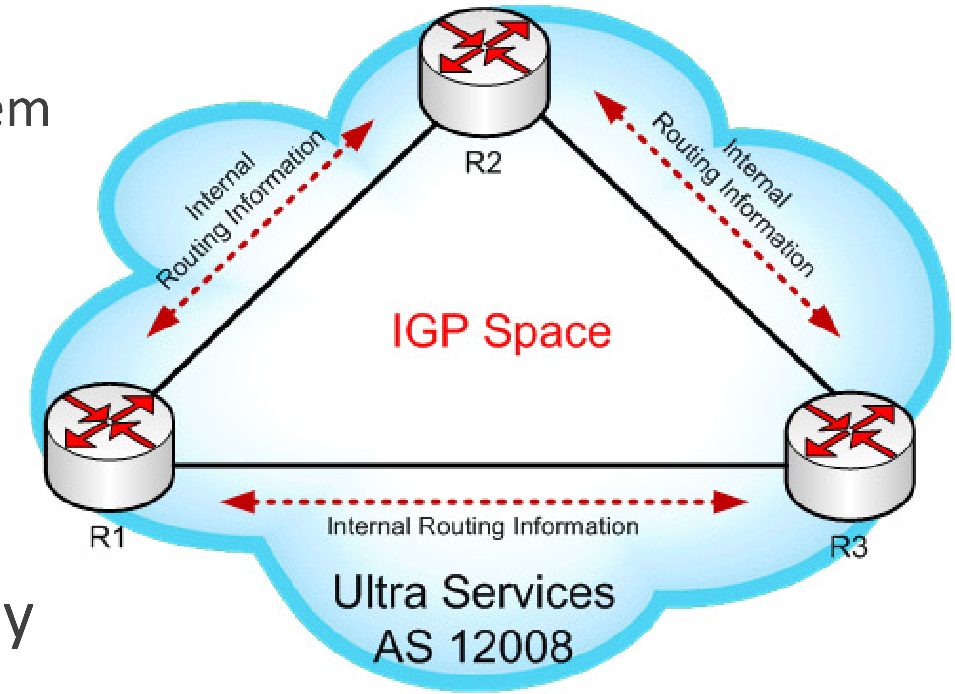
# Dynamic Routing Protocols

- Networks within a given organization are grouped into Autonomous Systems (AS)
  - An AS is a collection of Internet Protocol (IP) routing prefixes under the control of a single administrative entity that present a common, clearly defined routing policy to the Internet
  - Interior Gateway Protocols (IGPs) are used to exchange routing information within a given AS
  - Exterior Gateway Protocols (EGPs) are used to exchange routing information between routers bordering two networks (i.e. between two Autonomous Systems)



# Interior Gateway Protocols

- Interior Gateway Protocols
  - » Used within an Autonomous System
  - » Distribute internal infrastructure prefixes only, not external routing information
  - » Examples –OSPF, IS-IS, RIP
- IGPs are designed to route packets within an AS and rapidly adapt to network failures




- All IGPs function to identify the shortest cost path between two endpoints, typically via summation of the metrics of all the individual links

# Interior Gateway Protocols

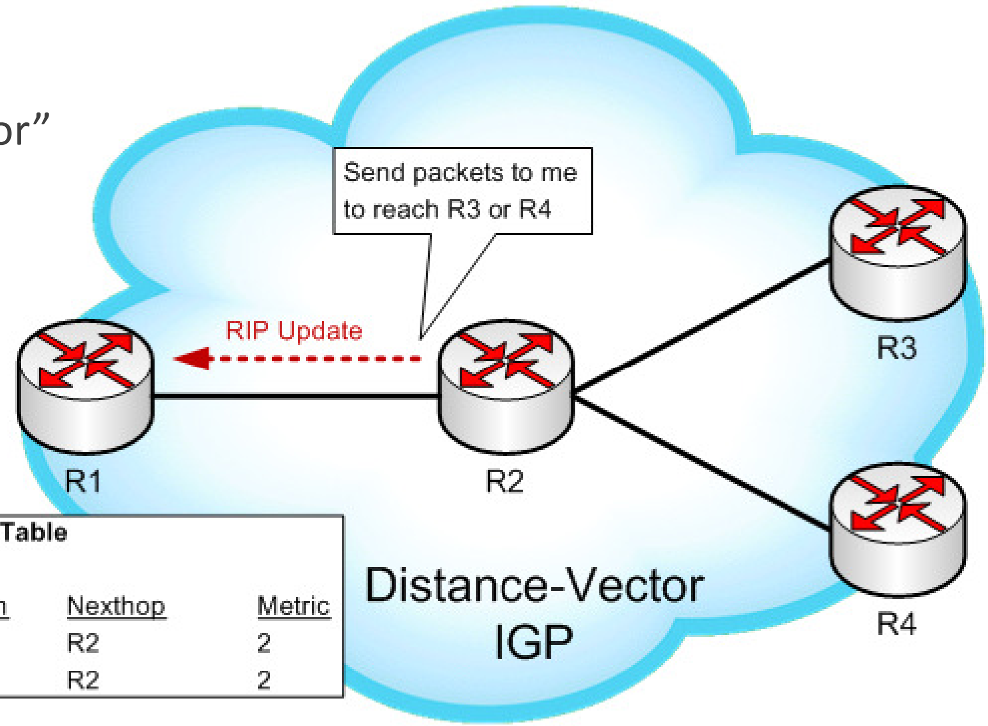
- Two classifications of IGPs in widespread use today: **Distance-Vector** and **Link-State**

- **Distance-Vector** is akin to calling Bill and asking them to pass a message along to Ted
  - Lower overhead in this approach, but it means that the message Ted receives from Bill is subject to some degree of uncertainty and interpretation
  - Routers running Distance-Vector protocols don't see the entire network, rather they simply trust what their neighbor tells them is reachable

- **Link-State** is akin to calling Bill and Ted separately, and giving each of them the exact same message
  - Obviously, this induces a higher degree of overhead, but it removes uncertainty
  - Routers running Link-State protocols have a topological view of the entire network, because each router informs every other router of their connectivity

# IGPs – Distance-Vector

- Distance-Vector
  - Distance-Vector routing protocols are commonly referred to as “routing by rumor”
  - Uses the Bellman-Ford mathematical algorithm
  - Examples: RIP
  - Each routers knows very little about the overall network topology, they only see their directly connected neighbors and “trust” that what is being advertised is reachable
  - Uses a distance, i.e. metric, and direction or vector to describe the next-hop router to use to reach the intended destination
  - Susceptible to routing loops

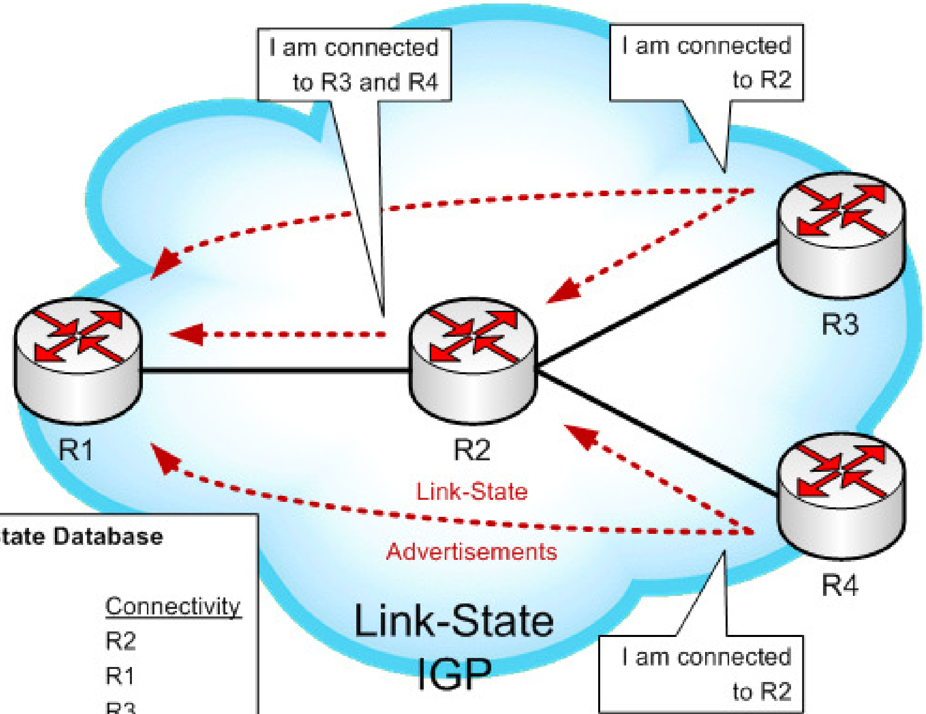



**R1 Route Table**

| Destination | Nexthop | Metric |
|---|---|---|
| R3 | R2 | 2 |
| R4 | R2 | 2 |

# IGPs – Link-State

- Link-State
  - » Topology information is flooded throughout the entire network, allowing each Link-State router to builds an identical topology database and have a consistent view of the network
  - » Each router runs a separate algorithm, known as the Dijkstra algorithm, allowing each router to come to their own conclusion as to what path is the best path through the network for a given destination
  - » Examples, OSPF, IS-IS
  - » Quicker to converge and eliminates loop susceptibility

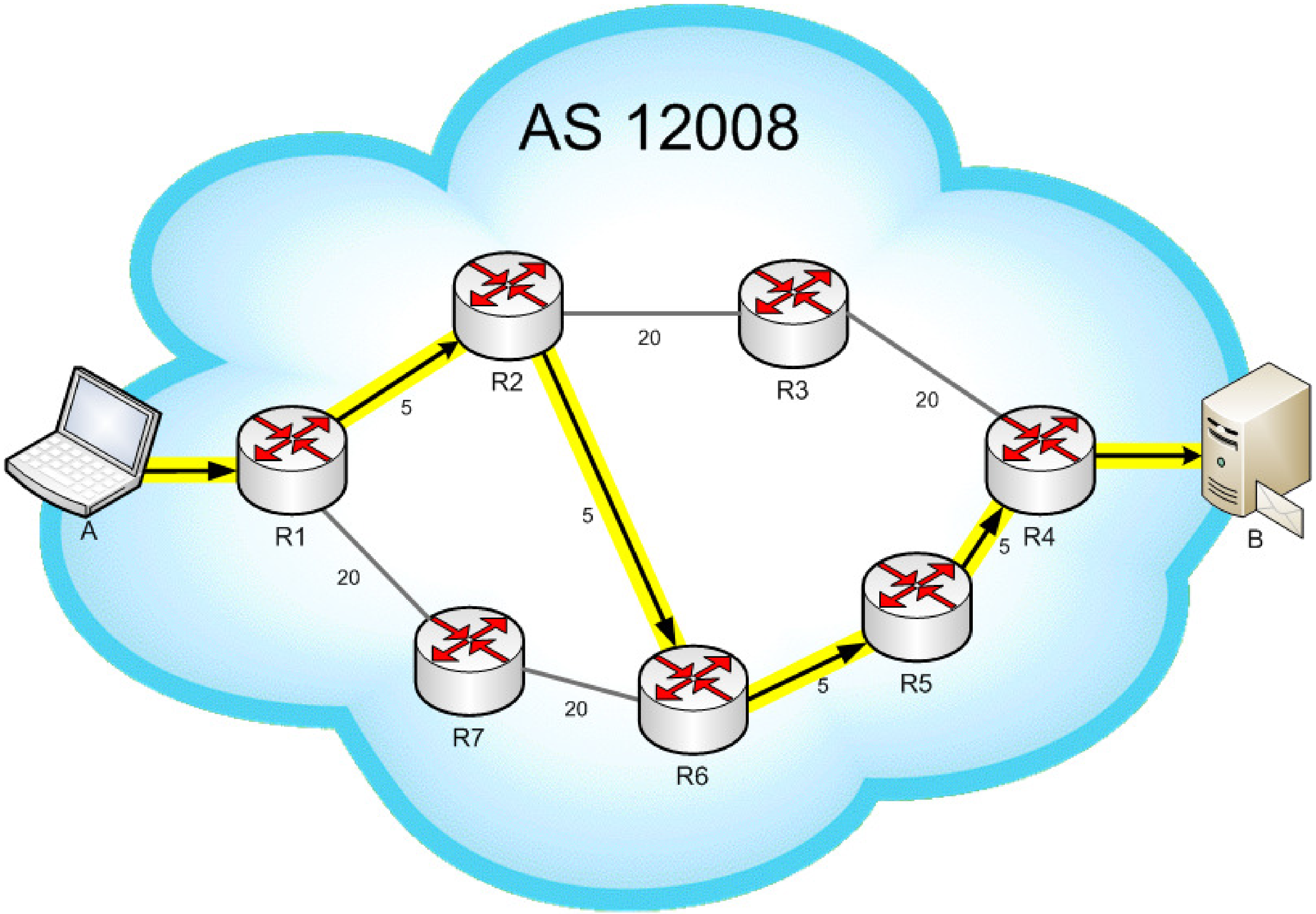



**Link-State Database**

| Node | Connectivity |
|---|---|
| R1 | R2 |
| R2 | R1 |
| R2 | R3 |
| R2 | R4 |
| R3 | R2 |
| R4 | R2 |

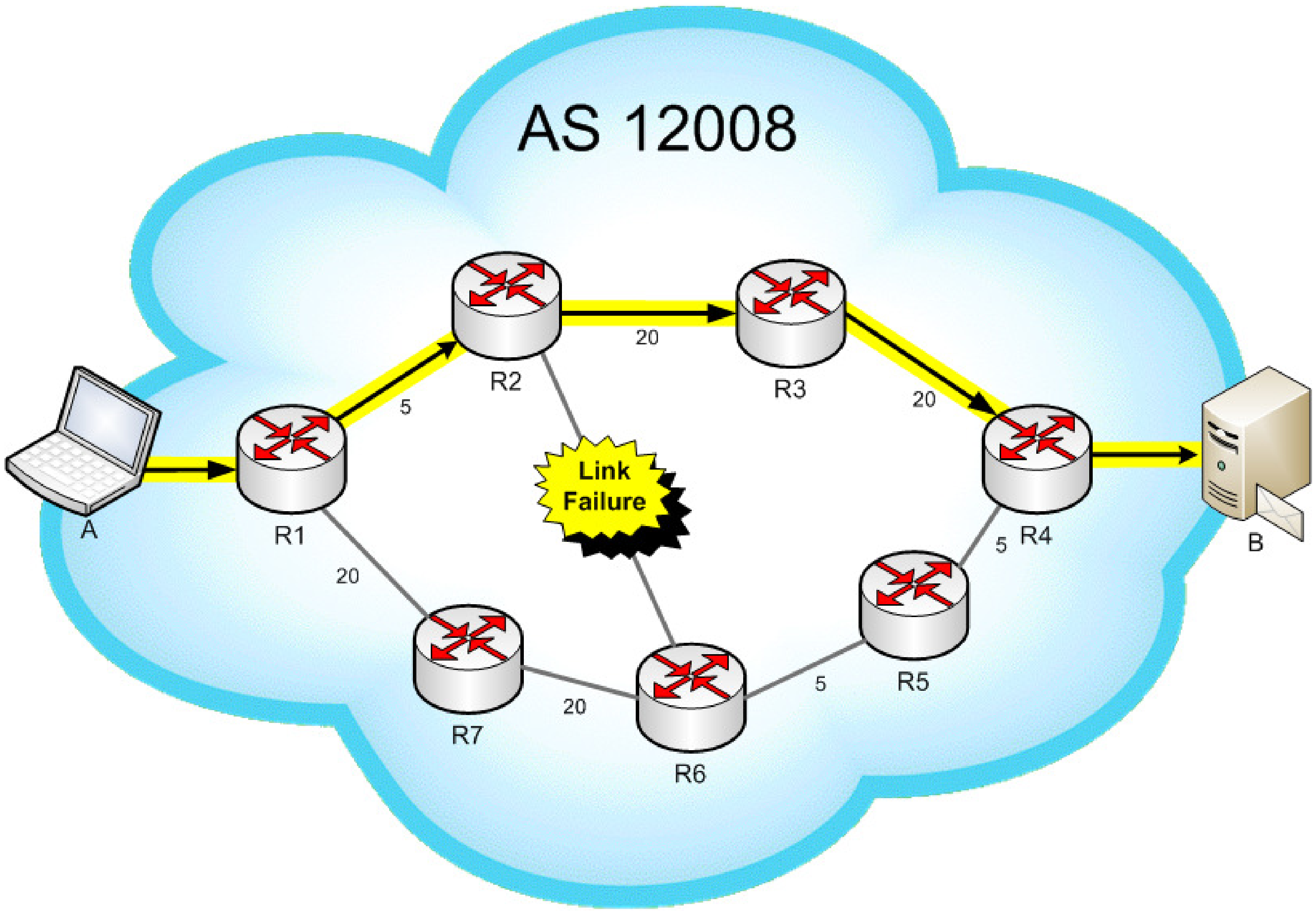
# IGP Path Selection

- Typically paths are compared by computing the metrics end-to-end. In the example below, the shortest cost path from A to B is 20 units



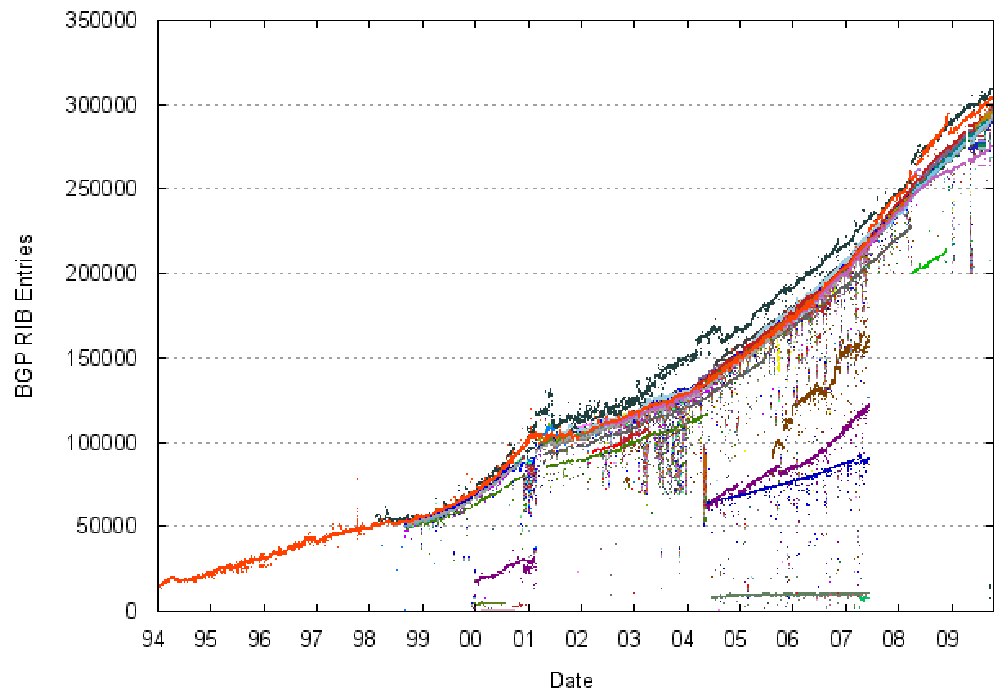
# IGP Path Selection

- If the link between R2 and R6 breaks, the minimum cost path from A to B is now computed at 45 metric units



# Internet Growth

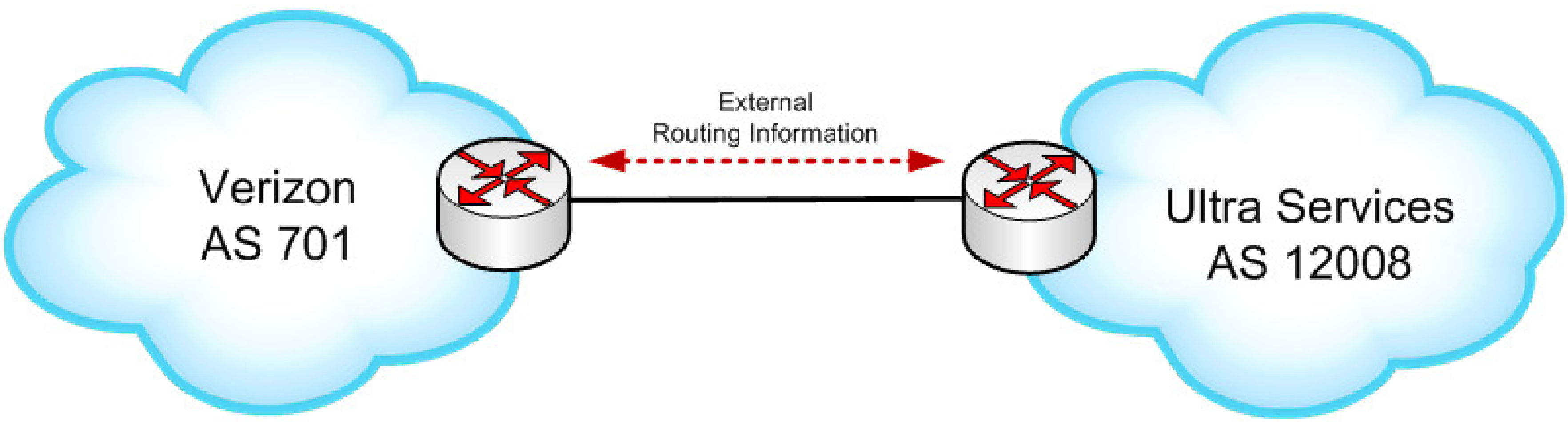
- The number of routes in the Internet has steadily grown since its inception, putting a heavy strain on IGPs
  - » IGPs were designed for fast convergence but not for very large routing tables
- As routing table sizes grew and more and more organizations connected to the Internet, a separate protocol was needed to support much larger routing tables and to exchange routing information between Autonomous Systems




**Internet Routing Table Growth 1994-2009**

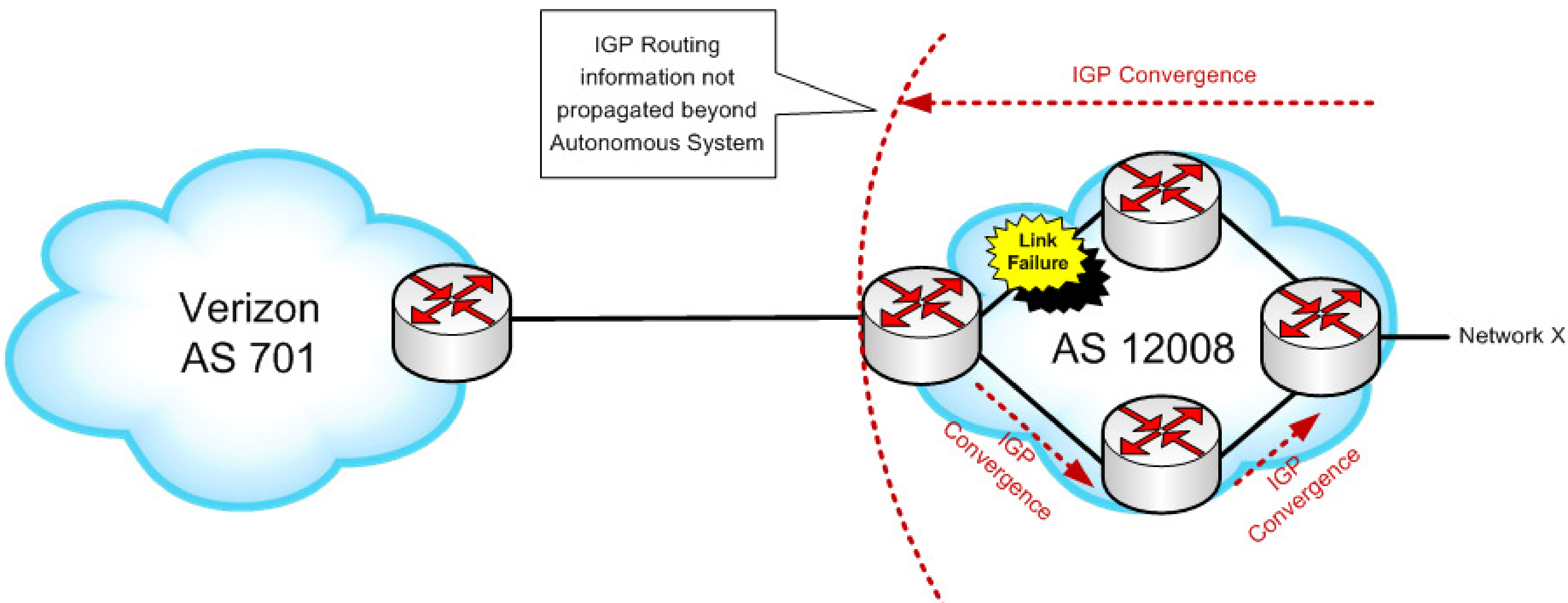
# Exterior Gateway Protocols

- Exterior Gateway Protocols
  - Used between Autonomous Systems to convey external routing information
  - Used to apply Routing Policy, i.e. Policy Based Control over routing vs. shortest-cost path
  - Examples – EGP, BGP
- EGPs are designed to support very large routing tables and scale to support very large networks



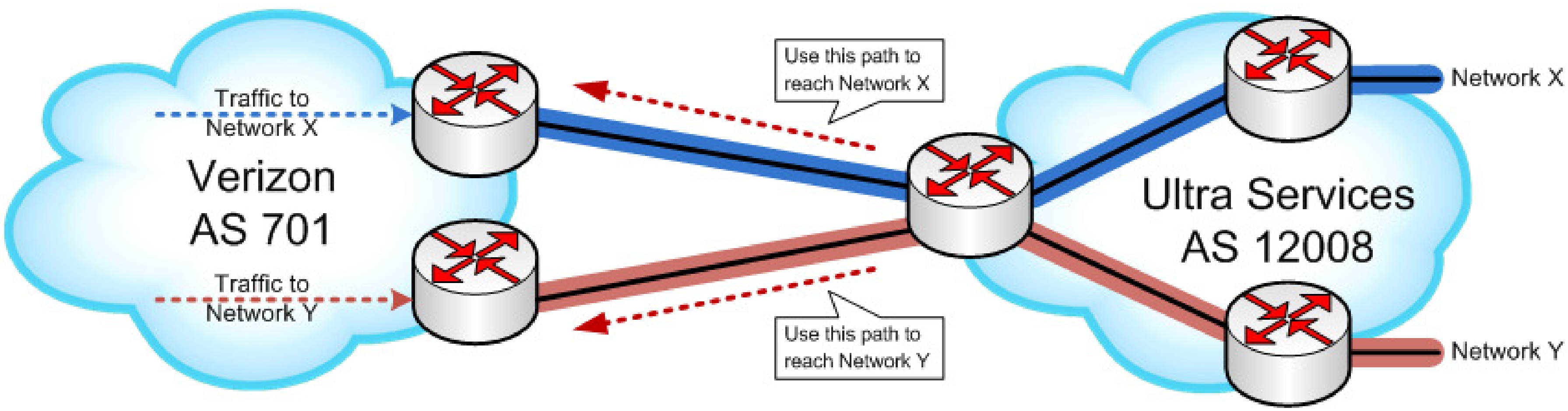
# Exterior Gateway Protocols

- The reason EGPs can support very large networks is because they are designed to hide the internal topology of the networks they are connecting
- Separation of IGPs from EGPs allows networks to grow because it constrains the amount of churn between networks and constrains convergence
  - Link failures in an IGP cause a lot of routing updates to be distributed, and ensuing reconvergence consumes significant CPU cycles



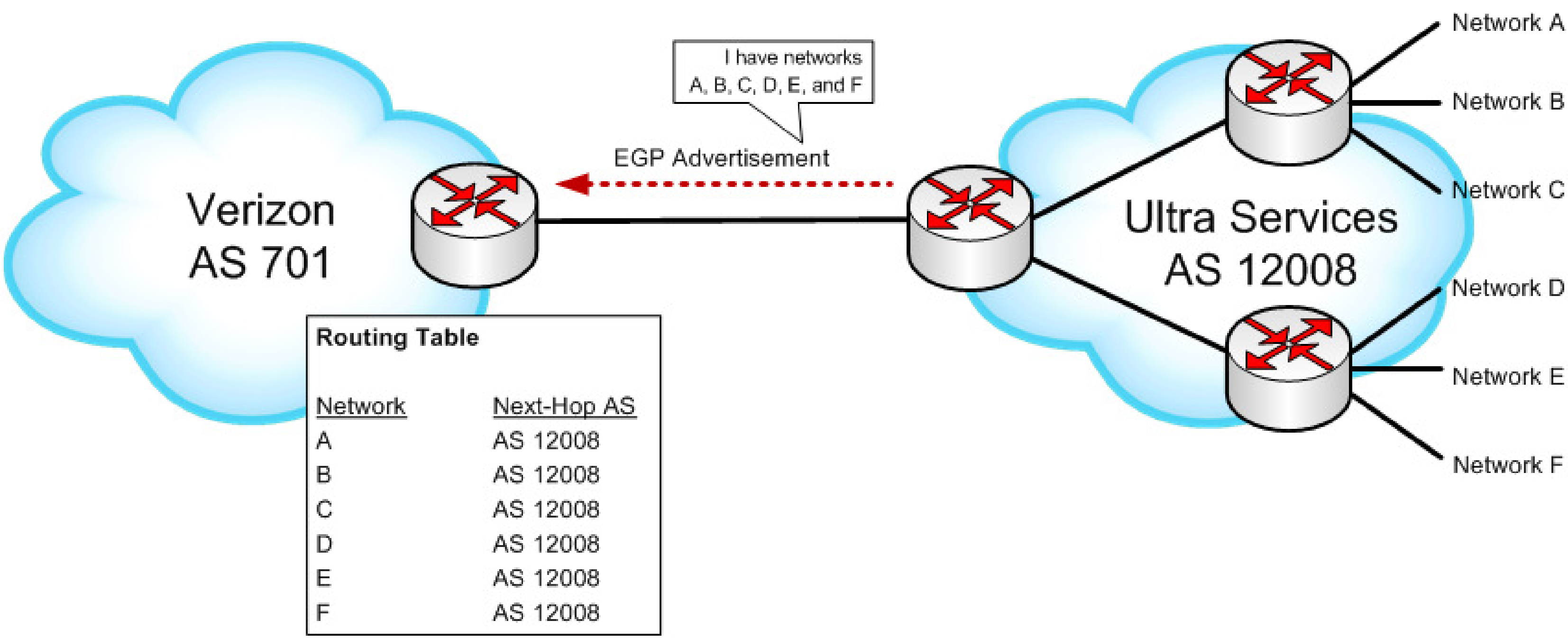
# Exterior Gateway Protocols

- EGPs allow the administrator to define administrative boundaries between Autonomous Systems such that policy can be enforced for routing decisions
  - » Routing policy allows network operators to aggregate routing information such that only a superset of routes are advertised to a peer, rather than every component network within the superset
  - » Allows for much greater scalability in Internet core routers due to the reduction in routing table sizes
  - » Routing policy also allows the network operator to choose the best path based on a wide variety of attributes that might be associated with a given route



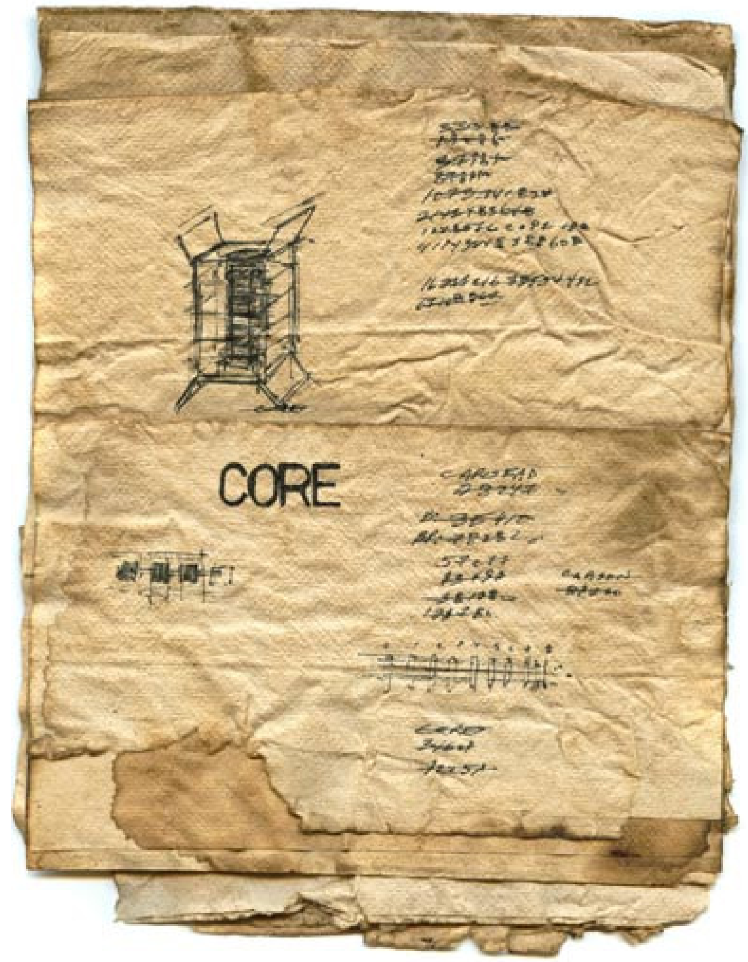
# Exterior Gateway Protocol

- The first instantiation of an exterior gateway protocol, the Exterior Gateway Protocol (EGP) version 3, was used to interconnect autonomous systems
- EGP was originally specified in 1982 by Eric C. Rosen of Bolt, Beranek and Newman, and David L. Mills
- Although EGP supported some policy-based controls over routing, it was susceptible to routing loops, lacked a diverse feature set and did not support CIDR, therefore address aggregation was difficult



# Border Gateway Protocol

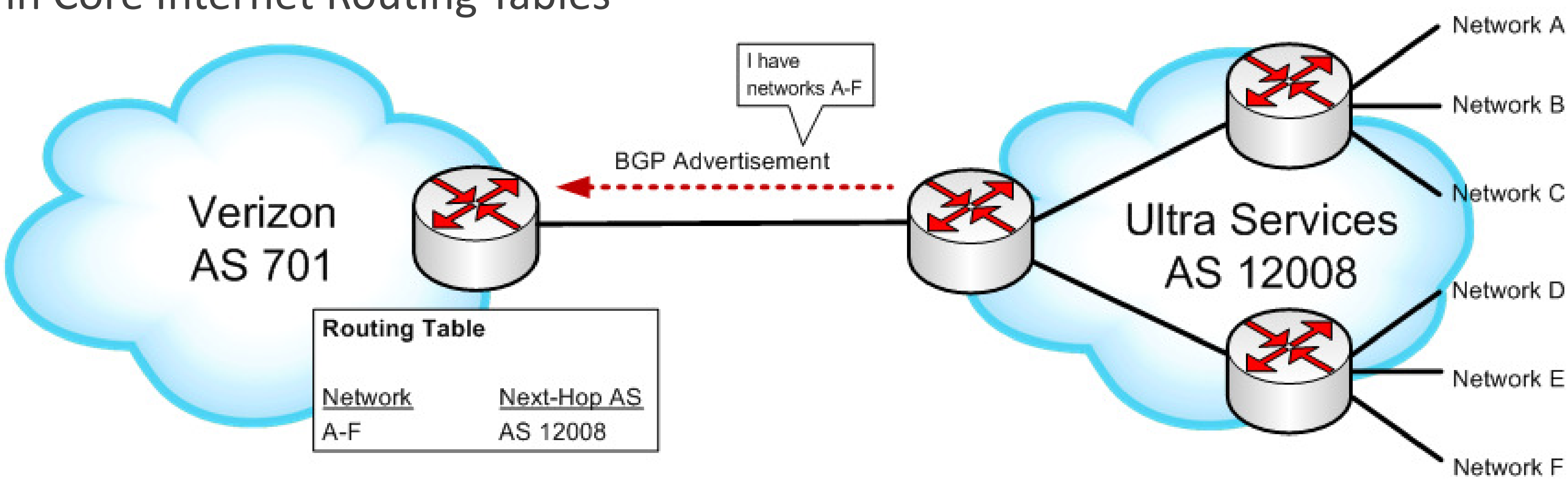
- Routing table sizes continued to grow
- Loop susceptibility and lack of address aggregation capabilities in EGP necessitated the need for a new protocol between Autonomous Systems
- Over lunch at the 12th IETF meeting in January 1989, Len Bosack, Kirk Lougheed, and Yakov Rekhter devised a protocol called the "Border Gateway Protocol"
  - The design goals behind BGP were to develop a protocol capable of providing policy control, loop detection, and the scalability required to support hundreds of thousands of networks through address aggregation techniques




*…the original idea for BGP was written on three napkins, giving BGP its unofficial title as the "Three Napkins Protocol"*

# Border Gateway Protocol

- BGP is an example of a Path-Vector protocol, which are a class of Distance-Vector protocols using the Bellman-Ford routing algorithm
- In BGP, routing tables maintain the autonomous systems that are traversed in order to reach the destination network system
  - » This is the primary mechanism which provides loop avoidance
  - » As a network is advertised from one AS to another, the AS Path is updated to include the AS it has just passed through
- BGP also supports address aggregation, allowing a large reduction in Core Internet Routing Tables

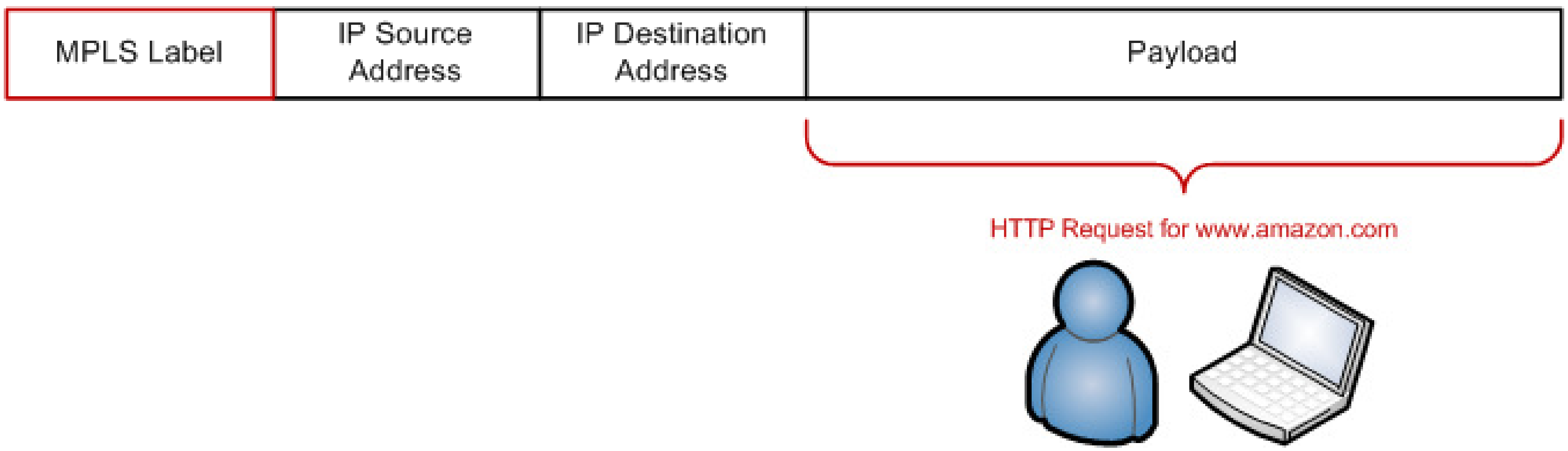


# Evolution of MPLS

- In the late 90s, all the big providers were using Asynchronous Transfer Mode (ATM) in the core of their networks
  - Cell-switching using ATM was fast, routing was slow!
  - ATM also provided Traffic Engineering capabilities!
- Service Providers employed the IP over ATM overlay model to provide IP services over their ATM cores
- By 1998, ATM vendors couldn’t produce cost effective Segmentation and Reassembly (SAR) chips capable of segmenting IP packets into ATM cells at speeds above OC-48 (2.488 Gbps)
- Service Providers wanted to move to POS OC-192 (10 Gbps) links, but couldn’t use their ATM cores due to the SAR chip issues
- Without their ATM cores, Service Providers had to give up their Traffic Engineering capabilities!

# Evolution of MPLS

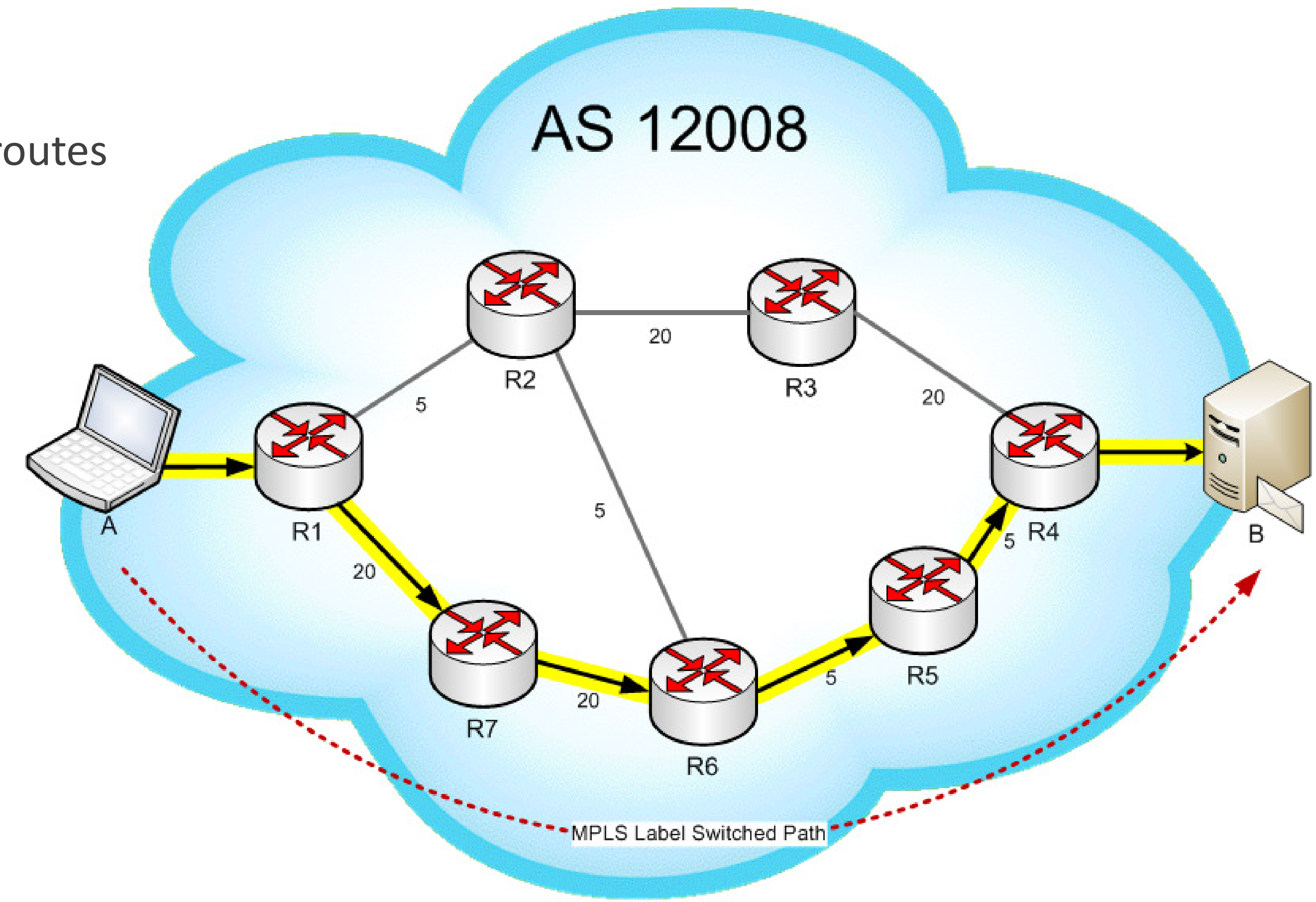
- Multi-Protocol Label Switching evolved as a method to make routers faster by allowing routers to look at short fixed-length labels (4 Bytes) as opposed to IP headers (20+ Bytes)




- Switching based on labels allowed routers to approach the speeds of ATM switches

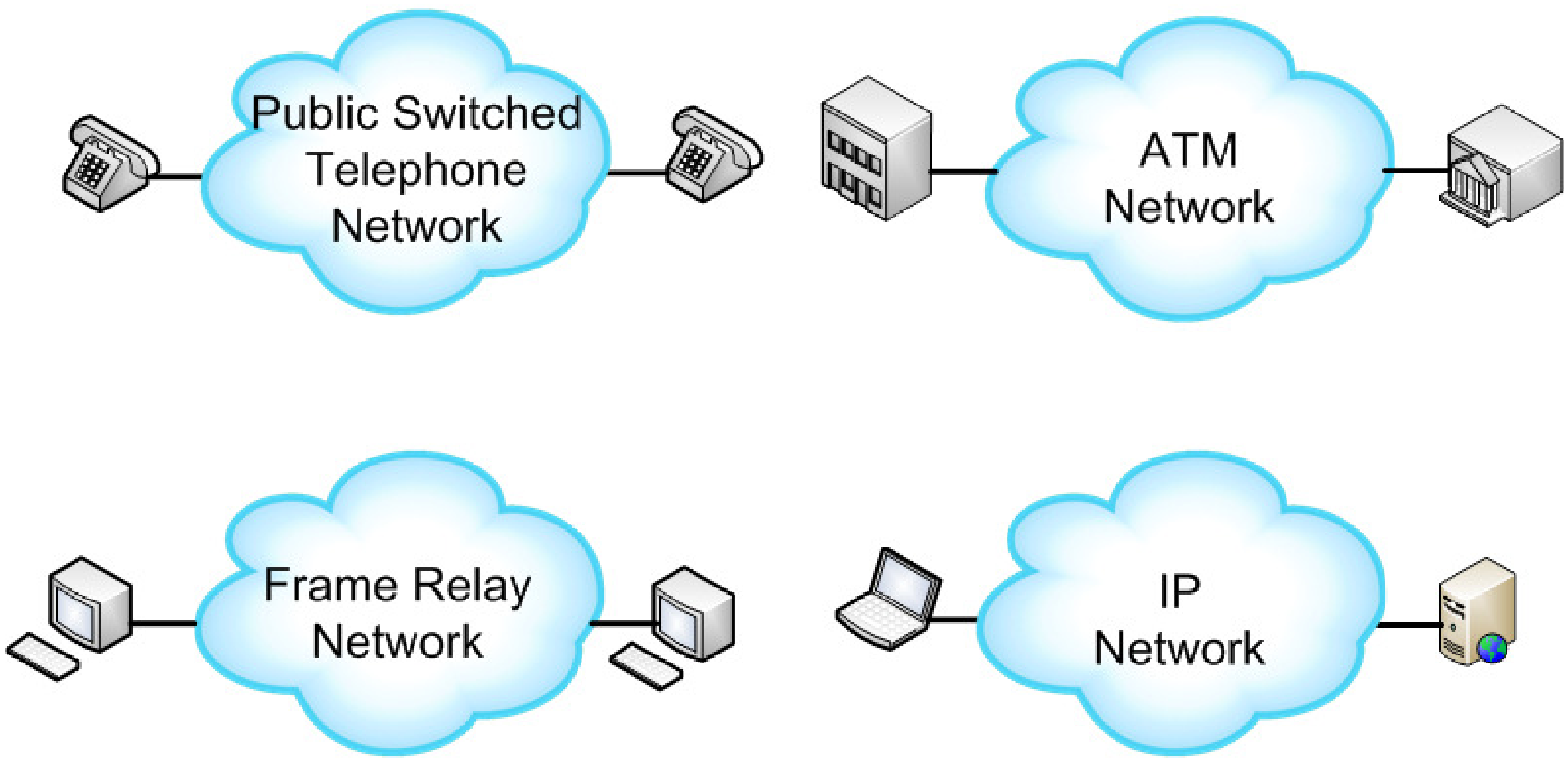
# MPLS Traffic Engineering

- An offshoot of MPLS was that it enabled Traffic Engineering capabilities via the label-switching approach
  - » Constraint-based routing could now account for:
    - Bandwidth
    - Link colors
    - Loose or Strict source routes
    - Etc.



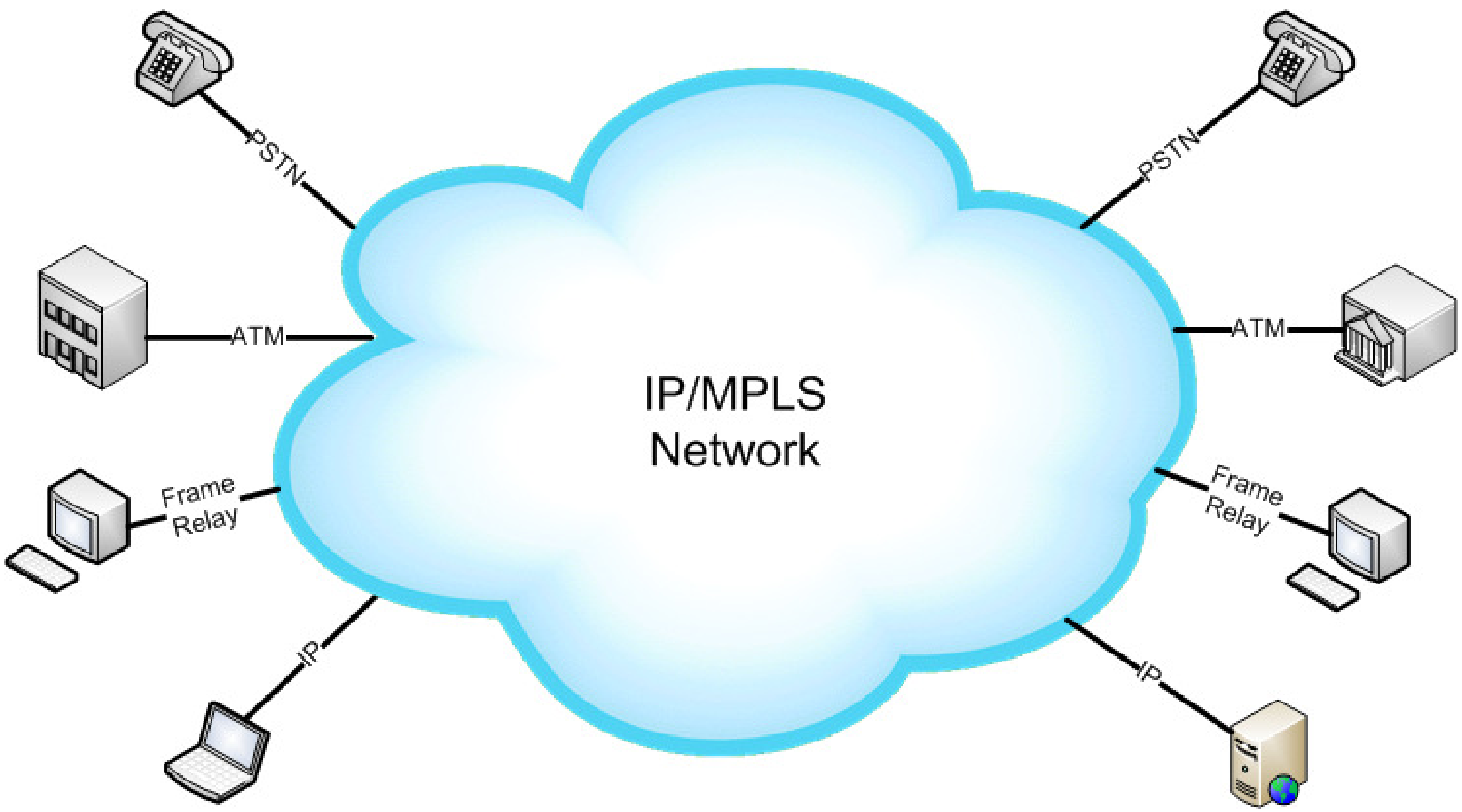
# MPLS Virtual Private Networks

- Before MPLS, Providers typically had to support multiple disparate networks to support different types of customers



# MPLS Virtual Private Networks

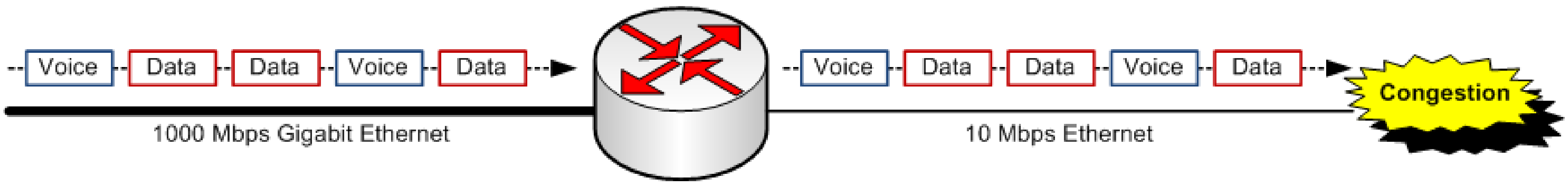
- By using the labels inherent in MPLS, Providers could overlay other types of networks onto their IP/MPLS core network, realizing significant cost savings by collapsing many networks into one
  - » Labels are used to stitch endpoints, i.e. Label A for ATM, label B for Frame Relay



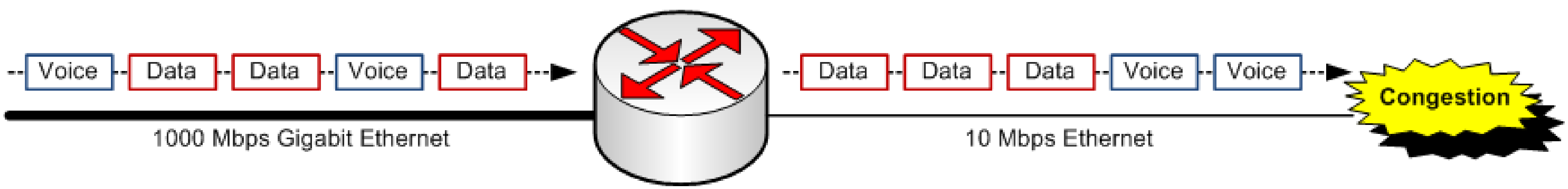

ShortestPathFirst Consulting Services
www.shortestpathfirst.net

# Quality of Service

- Quality of Service (QoS) is the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow
- QoS guarantees are important if the network capacity is insufficient, especially for real-time streaming multimedia applications such as VoIP, online games and IP-TV, since these often require fixed bit rates and are delay sensitive
- A network experiencing congestion that is not utilizing QoS will affect all traffic

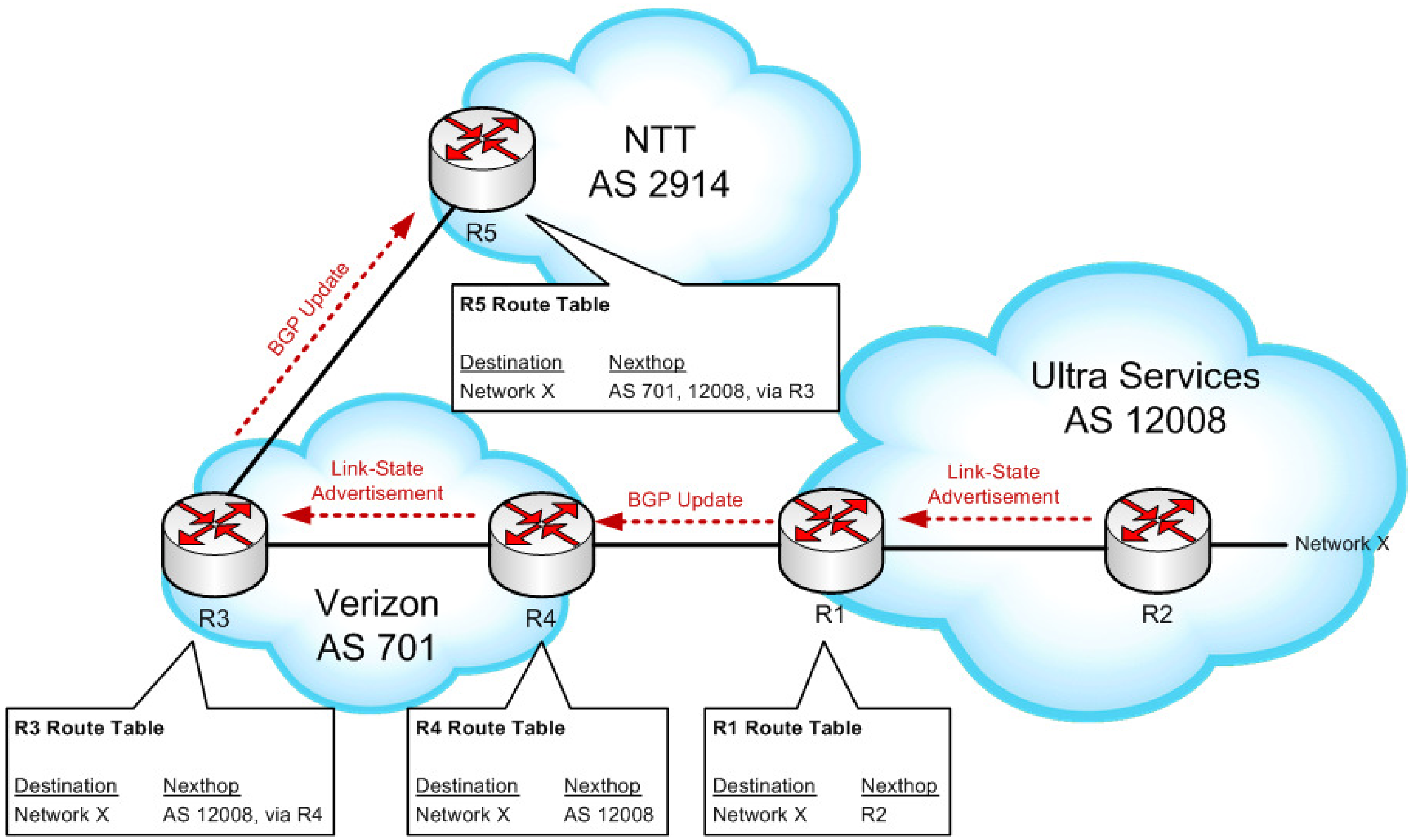



- Networks employing the use of QoS mechanisms can prioritize certain types of traffic to ensure that high priority traffic will be serviced above that of Best Effort (BE) traffic



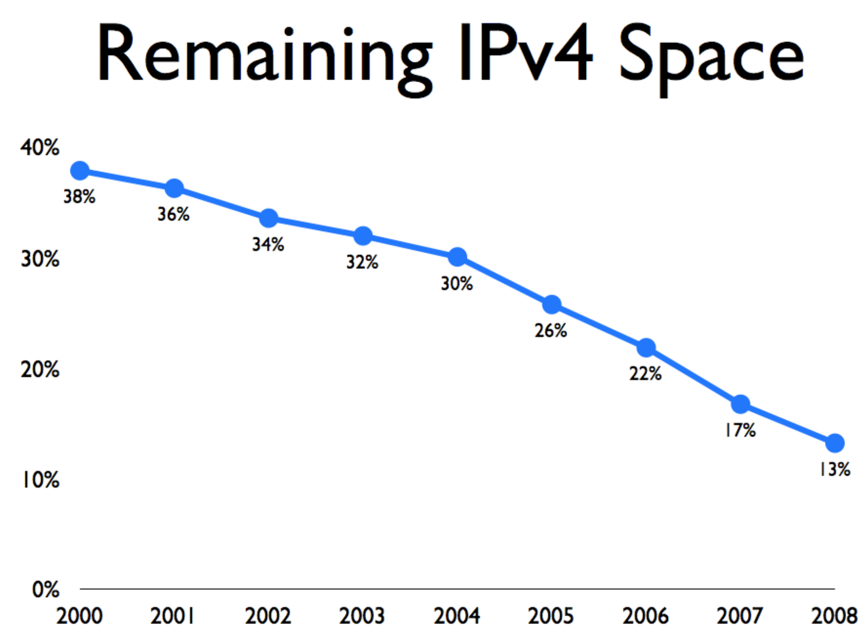
# Tying it all Together

- Through this complex interaction of IGPs and EGPs working together, routers build routing tables which allow them to forward packets towards their intended destination



# The Future of Internet Routing

- The current version of IP, IP Version 4 utilizes a 32-bit based addressing scheme
  - » 2^32 = ~4 Billion hosts
  - » Exacerbating the situation is reserved address space, wasteful address assignments and poor address aggregation schemes




- A New Way Forward: IPv6
  - » 128 bit addressing scheme
  - » Trillions upon trillions of hosts
  - » An IP address for your Refrigerator!
  - » Most current routing protocols can support this evolution (i.e. IS-IS, OSPFv3, RIPng, BGP4 w/ IPv6 NLRI

# Conclusion

- Packet switching evolved from a desire to move away from fixed-path circuit-switching paradigms
- Dynamic routing protocols evolved due to a need to dynamically exchange routing information in large-scale networks, allowing for automatic advertisement of network prefixes and seamless failover in the event of failures
- Distance-Vector protocols have largely fallen out of favor due to their slow convergence time, as well as their susceptibility to form routing loops
- BGP is the defacto Exterior Gateway Protocol in common use today
- Although IP is based on packet-switching and is not nearly as reliable as circuit-switching, a number of initiatives are under way to allow IP networks to behave more like that of traditional circuit-switched networks
- Separate networks can be virtually collapsed over a common IP core using MPLS in order to realize huge CapEx and OpEx savings
- More services will continue to move to the Internet in order to reach its large customer base as well as to realize the significant cost savings of using IP transport

# Questions?