

CIS 700/005 Networking Meets Databases

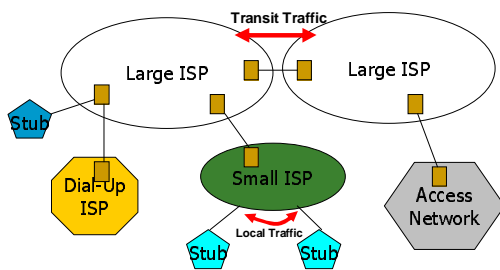
Boon Thau Loo
Spring 2007
Lecture 10

Several slides are courtesy of EE122 taught by Ion Stoica.

Outline

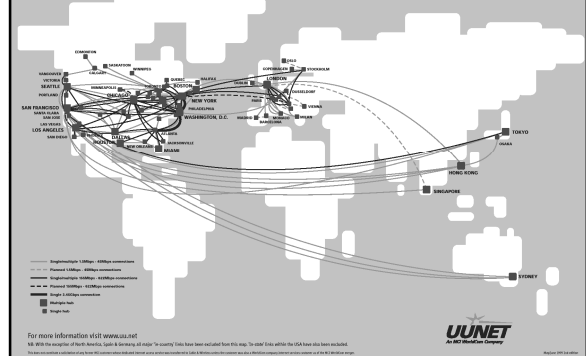
- Internet Architecture & Layering
- Intra-domain routing
- Inter-domain routing

Internet Structure



The Internet contains a large number of diverse networks

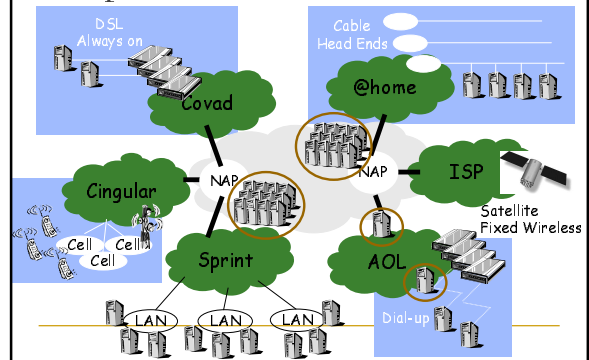
UUNET's Global Internet Backbone



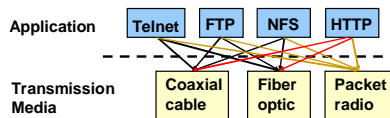
Tier 1 networks: The Global ISPs

- AT&T
- Global Crossing
- Level 3
- Verizon Business (formerly UUNET)
- NTT Communications
- Qwest
- SAVVIS
- Sprint Nextel Corporation

Computers Inside the Core



The Problem



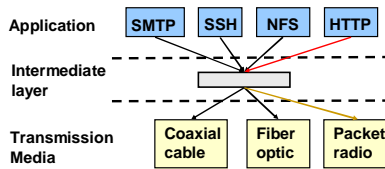
- Re-implement every application for every technology?
- No! But how does the Internet architecture avoid this?

Layering

- Layering is a particular form of modularization
- System is broken into a vertical hierarchy of logically distinct entities (layers)
- Service provided by one layer is based solely on the service provided by layer below
- Rigid structure:
 - (+) easy reuse
 - (-) performance suffers

Solution: Intermediate Layer

- Introduce an intermediate layer that provides a single abstraction for various network technologies
 - A new app/media implemented only once
 - Variation on "add another level of indirection"



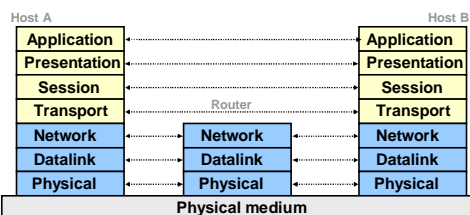
ISO OSI Reference Model for Layers

- Application
- Presentation
- Session
- Transport
- Network
- Datalink
- Physical

ISO: International Standards Organization
OSI: Open System Interface

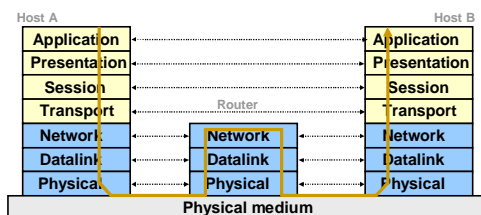
Who Does What?

- Seven layers
 - Lower three layers are implemented everywhere
 - Next four layers are implemented only at hosts



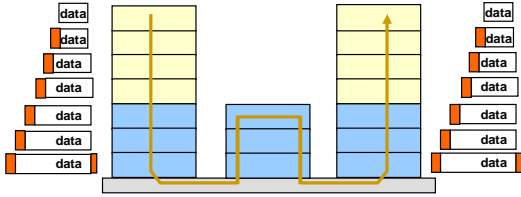
Physical Communication

- Communication goes down to physical network, then to peer, then up to relevant layer



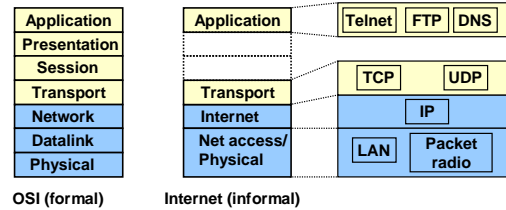
Encapsulation

- A layer can use only the service provided by the layer immediate below it
- Each layer may change and add a header to data packet

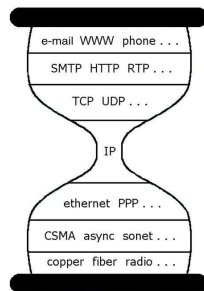


OSI vs. Internet

- OSI: conceptually define services, interfaces, protocols
- Internet: provide a successful implementation



Hourglass



Implications of Hourglass

Single Internet layer module:

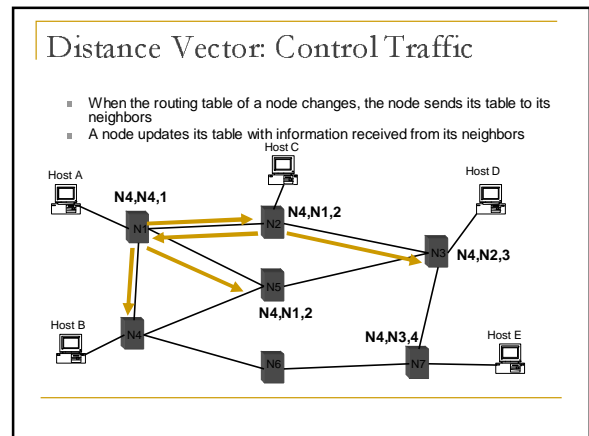
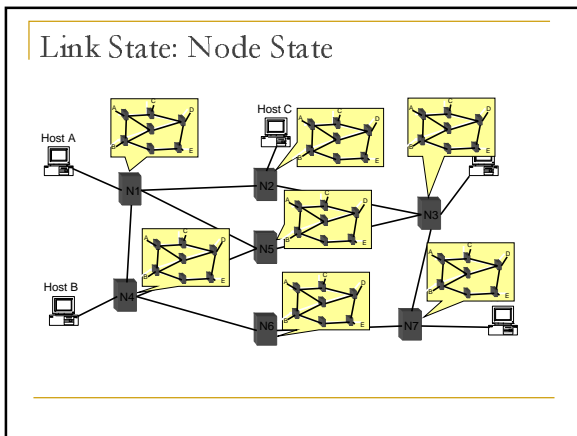
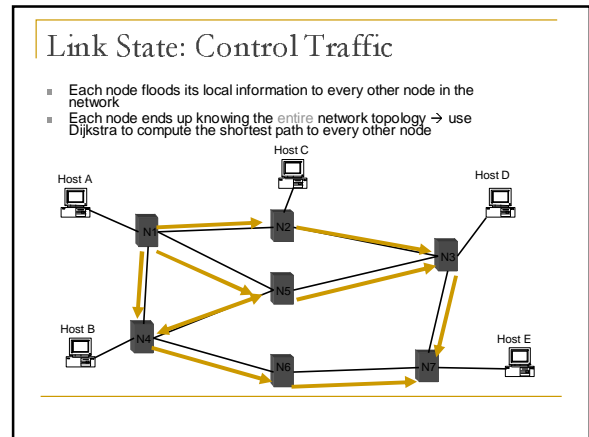
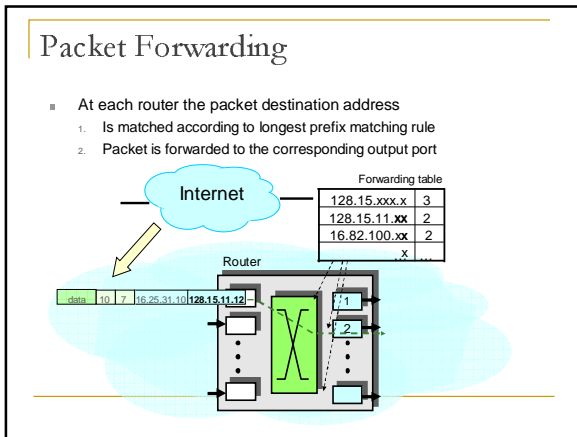
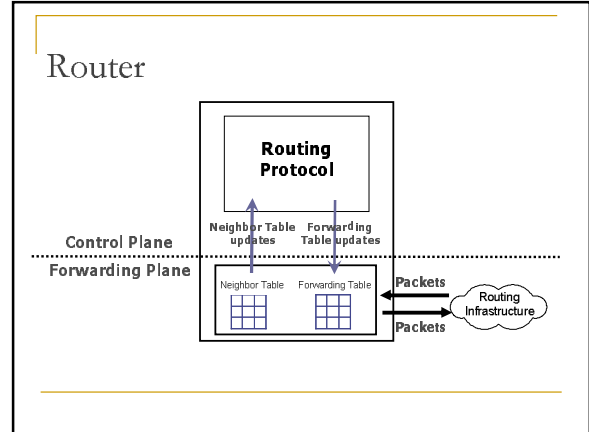
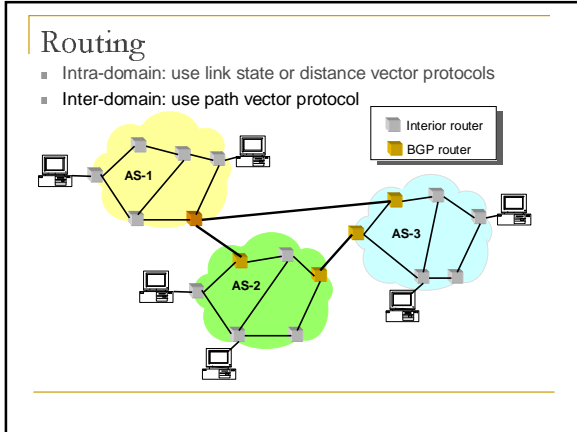
- Allows networks to interoperate
 - Any network technology that supports IP can exchange packets
- Allows applications to function on all networks
 - Applications that can run on IP can use any network
- Simultaneous developments above and below IP

End-to-end Argument

- Think twice before implementing functionality in the network
- If hosts can implement functionality correctly, implement it a lower layer *only* as a performance enhancement
- But do so only if it does not impose burden on applications that do not require that functionality

Outline

- Internet Architecture & Layering
- Intra-domain routing
- Inter-domain routing



Addressing

- ARP (Address Resolution Protocol):
 - MAC (Ethernet, 802.11) address <-> IP Address
- DHCP (Dynamic Host Configuration Protocol):
 - Request and obtain an IP address from DNS server
- DNS (Domain Name Service):
 - Hostname to IP address mapping

Support for Mobility

- What if the host moves around and changes its IP address?
- Use DHCP? What if host moves in the middle of a session?
- Solution:
 - Level of indirection
 - Home agent with permanent IP address
 - Stores information on mobile host
 - All nodes send packets to home agent, who forwards to mobile host
 - Foreign agent with permanent IP address
 - Tunnel traffic from home agent to foreign agent
 - Supports device in foreign network
- Issues: security and efficiency

IP Addressing

- Unicast – Send to one host
- Broadcast – Send to all possible destinations
- Multicast – Send to interested receivers associated with multicast address
- Anycast – Send to closest receiver associated with anycast address

IPv6

- Supports 2^{128} , as oppose to 2^{32} addresses
- IP Anycast
- Mobile IPv6
- IPSec
- Etc...

Outline

- Internet Architecture & Layering
- Intra-domain routing
- Inter-domain routing

Issues We Haven't Addressed

- Scaling
 - Router table size
- Structure
 - Autonomy
 - Policy

Scaling

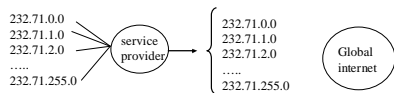
- Every router must be able to forward based on *any* destination IP address
 - Given address, it needs to know "next hop" (table)
- Naive: Have an entry for each address
 - There would be 10^8 entries!
- Better: Have an entry for a range of addresses
 - But can't do this if addresses are assigned randomly!

Original Addressing Scheme

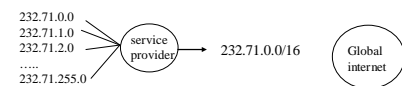
- Class-based addressing schemes:
 - 32 bits divided into 2 parts:
 - Class A $\left\{ \begin{array}{l} 0 \text{ network } 8 \text{ host} \\ \left. \begin{array}{l} 126 \text{ nets} \\ \sim 16\text{M hosts} \end{array} \right\}$
 - Class B $\left\{ \begin{array}{l} 0 \text{ network } 16 \text{ host} \\ 1 0 \text{ network } 16 \text{ host} \\ \left. \begin{array}{l} \sim 16\text{K nets} \\ \sim 65\text{K hosts} \end{array} \right\}$
 - Class C $\left\{ \begin{array}{l} 0 \text{ network } 24 \text{ host} \\ 1 1 0 \text{ network } 24 \text{ host} \\ \left. \begin{array}{l} \sim 2\text{M nets} \\ 254 \text{ hosts} \end{array} \right\}$
- Classless Inter-domain Routing (CIDR)
 - Variable prefix – prevents wastage
 - Prefix aggregation – lower overhead of Inter-domain routing
 - Routers match to longest prefix

Routing Table Size

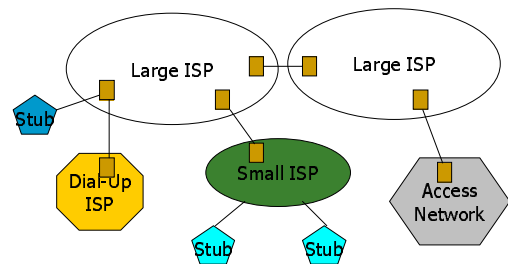
Without CIDR:



With CIDR:



Network Structure



The Internet contains a large number of diverse networks

Autonomous Systems (AS)

- Internet is not a single network!
- The Internet is a collection of networks, each controlled by different administrations
- An autonomous system (AS) is a network under a single administrative control

Implications

- ASs want to choose own local routing algorithm
 - Intra-domain routing algorithm, e.g., link state (OSPF), distance vector
- ASs want to choose own nonlocal routing policy
 - Inter-domain routing: BGP de facto standard

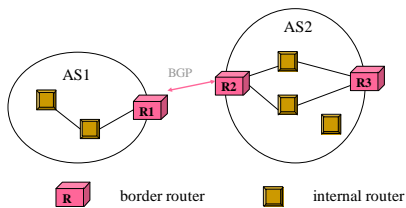
Interconnection

- IP unifies network technologies
 - Allows any network to communicate with another
- BGP unifies network organizations
 - Ties them into a global Internet

Border Gateway Protocol

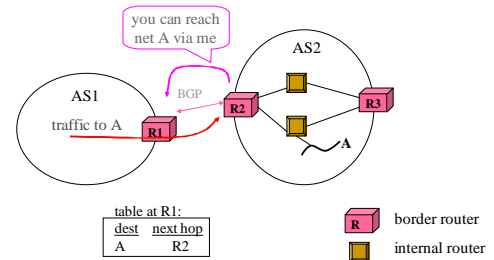
*ignore the details
pay attention to the "why"*

Who speaks BGP?



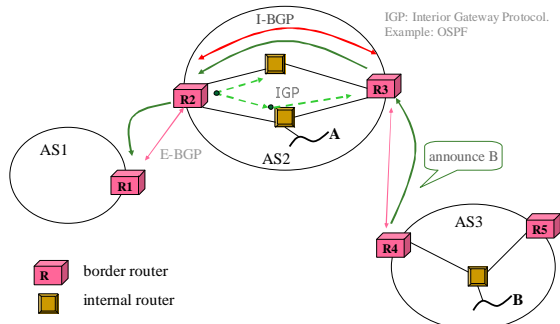
- Two types of routers
 - Border router (Edge), Internal router (Core)

Purpose of BGP



Share connectivity information across ASes

I-BGP and E-BGP



Issues

- What basic routing algorithm should BGP use?
- How are the routes advertised?
- How are routing policies implemented?
 - Policy routing: not always shortest path

Choice of Routing Algorithm

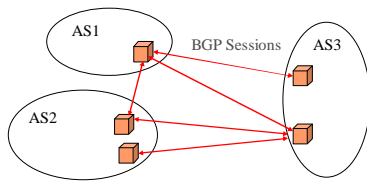
- Constraints:
 - Scaling
 - Autonomy (policy and privacy)
- Link-state?
 - Requires sharing of complete network information
 - Information exchanges doesn't scale
 - All policies exposed
- Distance Vector?
 - Scales and retains privacy
 - Can't implement policy
 - Can't avoid loops if shortest paths not taken

Path Vector Protocol

- Distance vector algorithm with extra information
 - For each route, store the complete path (ASs)
 - No extra computation, just extra storage
- Advantages:
 - Can make policy choices based on set of ASs in path
 - Can easily avoid loops

Advertising Routes

- One router can participate in many BGP sessions.
- *Initially* ... node advertises ALL routes it wants neighbor to know (could be > 50K routes)
- *Ongoing* ... only inform neighbor of changes
- TCP connections



Routes Have Attributes

- When a route is "advertised" it is described in terms of attributes:
 - next hop, AS-path, etc.
 - Other attributes: Origin, MED, Local Preference
- Steps in advertisements:
 - **Import policy:** Determine which route advertisements should be filtered
 - **Decision policy:** Select most desirable route
 - **Export policy:** Determine which neighbors to export routes to
- Example policies:
 - "Use provider B only to reach these addresses"
 - "Use the path that crosses the fewest number of AS"
 - "Use AS x in preference to AS y for route to destination z"

Is Reachability Guaranteed?

- In normal routing, if graph is connected then reachability is assured
- With policy routing, not always
- Issues:
 - Traffic engineering constraints
 - Optimal path is seldom achieved – good path that satisfies policies
 - "Hot potato" routing
 - Pass traffic to another AS as quickly as possible
 - Route to nearest exit point when there is more than one route to destination
 - High convergence delays

Roadmap

- Internet Indirection Infrastructure
 - DHT-based overlay to support flexible multicast, anycast, service composition and mobility
- Revisiting the IP waistline:
 - Active networks: Allow routers to run arbitrary code
 - Declarative networks: Allow routers to run database queries
- Routing on Flat Labels:
 - Clean-slate redesign – Get rid of location and route using DHTs
- Network Forensic Alliance
 - AS collaborate to locate origin of epidemic worm attacks

Announcements

- Feb 12 proposal deadline
- Volunteers to present twice
 - Improve presentation / class participation scores
 - Improve presentation skills
- In-class presentation:
 - Roughly 10 presentations, 15 minutes each
 - 19th Apr (Thursday) : 4:30pm – 7pm??
 - 20th Apr (Friday) : 1pm – 3:30pm??