Last Time/Today

- OSPF, Inter-domain routing
- Today: Discussion of Peering/transit paper
- Today: Discussion of Paxon's E2E routing behavior paper
- Today (maybe): Finish some IP related stuff (Tunneling, Mobile IP, VPNs, Ad hoc routing ...)

Parallel – Telephone Network

- Local Exchange Carriers (LECs), and Inter Exchange Carriers (IXCs)
- LECs provide IXCs access to customers (by law)
- A call goes across multiple phone company “wires”
  - Caller gets bill
  - Money is split up among carriers
  - Other models also used (e.g., first carrier gets all; pay per use ...)
  - How the money is split is called “settlement”

Inter-domain Routing

- OSPF: link state; uses hierarchy to allow scalability
  - do we lose anything?
- Autonomous systems: scale routing and allow distributed management
- Inter-domain routing between AS’s
- BGP: distance vector with paths; designed to support policy driven routing
- Path inflation and other anomalies due to (1) hierarchy; and (2) policy
Peering Example

- East Net Purchases Transit from USA Net
  - It receives access to all the entries in USA Net's routing tables
  - It is advertised to all USA Net's peers
  - Effectively, it becomes a client

Internet

- No regulations
  - ISPs don't have to talk to each other
- No standard settlement model; tough to come up with one (why?)
- Works due to “shared goodwill”
  - Pay for connectivity not per packet

Peering vs. Transit

- Peering: Two ISPs provide connectivity for each other's customers
  - Usually for free
  - Not a transitive relationship
- Transit: One ISP provides another with connectivity to every place it knows
  - $$$
### Tier-1 ISP

- Tier-1 ISP: an ISP that reaches everywhere
  - Useless definition, everyone claims to be Tier 1
- More useful definition, ISP big enough not to need to purchase transit
- Tier 2 buys transit from Tier 1
- Example: Sprint, UUNET, Genuity, MCI

### Transit

- What is it that gets bought?
  - Not paying for use, but for connectivity
  - Purchasing Routing Table entries!
- It's really hard to be a small ISP
  - Big ISPs have more value to offer to small ISPs than vice versa
- Peering and Transit are two ends of a continuum
  - Which routes do you give away and which do you sell? to whom? and under what conditions?

### How is Peering Accomplished?

- Private Peering: Create a point to point link between peering ISPs
- Public (or Exchange) Peering
  - Use a shared exchange point for the connection
  - Exchange is a building where ISP’s can interconnect
  - Anyone can use it if they pay
  - Peering still negotiated bilaterally (i.e., not everyone using the exchange has to be connected to everyone else)
  - “Long” article has excellent discussion of exchanges/NAPs
Why Peer?

- Transit is expensive!! Remember, you are buying the connectivity, not just the wires
- Peering with other ISPs can reduce transit expenses
  - Can also reduce latency (see Paxon’s routing paper)
- Communication patterns not uniform
  - Traffic tends to be localized (e.g., geographically)
  - Peer with the ISPs you communicate most heavily with

Why not Peer?

- Traffic is not symmetric
  - Peering is a “zero-sum” agreement
  - But one ISP may be generating more traffic than it is receiving
- Big hassle to negotiate peering
- Big boys not interested in helping small ISPs become more competitive
- Hard to deal with problems if there is no dollar exchange

Who to Peer With?

- Technical Issues
  - Where does my traffic go?
- Business Issues
  - Other partnerships/deals
  - Peering Policies (becoming more open, see pointers in paper)
  - Cost of exchanging traffic (leased line/exchange fees)

Model for Tier-2 ISPs

1. Buy transit from Tier-1’s
2. Peer at public exchange points to reduce transit cost
3. Private peering with key ISPs
4. When you are big enough negotiate peering with transit providers
   - What determines big enough?
Summary/Discussion

- Interdomain routing
  - Exchange Reachability information + attributes/hints
  - Local Policy to decide which route to use

- Traffic Exchange policies are a big issue – $$$
  - Complicated because of lack of accepted economic model (who creates value?)
  - Hard to be a small ISP

- Business issues can have serious implications on technical/operational efficiency of the Internet
  - Explanation for the pathologies in Paxon’s paper?

- Implicit trust?
  - what if an ISP drops “peering” traffic not destined to its clients?
  - Are there concerns for transit relationships?

Routing Pathologies Discovered

- Persistent loops: loop unresolved by end of traceroute: 0.13-0.16% (0.21 // 0.44%)
- Temporary loops: loop resolved by end of traceroute: 0.055-0.078% (0.06% // 0.22%)
- Outage greater than 30 sec: repeatedly dropped probe packets (0.96% // 2.2%)
- Mid-stream change: wrong path: 0.004% (1 instance)
- Infrastructure failure: route terminates inside network
- Total: (1.5 // 3.4%)

Methodology

- Traced routes between 37 hosts across the world (actually, USA and Europe)
- Periodic, exponentially distributed (means of hours and days)
- An Internet that looks significantly different than today’s (at least in scale)

- Study provided detailed analysis of trace data (which is available if you want to do your own additional analysis)

- Are there concerns for transit relationships?
- What if an ISP drops “peering” traffic and not?
- Implict trust?

- Explanation for the pathologies in Paxon’s papers? An example of applying measurement to study the current commercial structure of the Internet (hope you read/enjoyed the CAIDA paper)
- An example of applying measurement to study the Internet as seen by end hosts
- Explore “wide-area” routing behavior in the Internet

Paxon’s E2E Routing Behavior Paper

- How would you react if an ISP decided to charge customers for peering traffic?
- How would you react if an ISP decided to charge customers for peering traffic?
- How would you react if an ISP decided to charge customers for peering traffic?
- How would you react if an ISP decided to charge customers for peering traffic?

- Traffic Exchange policies are a big issue – $$$
- Local Policy to decide which route to use
- Exchange aggregated/summary information + attributes/attributes
- Interdomain routing

SUNY-BINGHAMTON – CS528 Fall ’07 Lecture #13
SUNY-BINGHAMTON – CS528 Fall ’07 Lecture #13
SUNY-BINGHAMTON – CS528 Fall ’07 Lecture #13
Routing Asymmetry

- How often do the routes A to B and B to A differ?
  - Routes differing in cities traversed: 49% !!
- Reflection of policy routing
  - If shortest path, asymmetry should be rare
  - What do you think could be the explanation?
- Asymmetry implications?
  - Consider sliding window with asymmetric forward/return path
  - Even with stable routes, HTTP request delay may be shorter than response delay – why?

Discussion

- Routing Loops within a single AS
  - BGP works well
- Pathologies occur most during the times when the network is congested (surprise surprise)
  - Reason to be hopeful?

Routing Stability

- How often do routes between two end-points change?
- Prevalence: How likely am I to observe the same route in the future?
- Host median is 82% (strongly dominated by a single route)
- Persistence: How long before a route will change?
  - Routing changes occur over a range of time scales
    - 10s of minutes: 9%
    - 6+ hours: 23%
    - Days: 68%
    - Changes occur mostly within a network (AS)
- Importance of route stability?

Discussion

- If the experiments were repeated today, do you think the results would be different?
  - Would have been a nice project for someone!
- Paper shows the behavior, but does not provide explanation
  - What is the culprit? protocols? policy? wrong implementations? bad hardware?
- Is this behavior likely to get better? Does it matter?
- Next time: Internet Path Inflation
IP in IP (Tunneling)

What does tunnelling do?
- Allows a packet to be encapsulated and sent to another host
- Allows packet indirection using the same IP infrastructure

Aside:
- How is it different from source routing?
- How is this different from IP spoofing?
- How is it different from NAT?

Why do it?
- possible applications?

How is it done?

The IP Packet is encapsulated by another IP header with the tunnel address
- Protocol field is set to IP-in-IP protocol
- When tunnel destination receives the tunneled packet, it strips the header and sends the original packet on
- For efficiency, an abbreviated IP header is defined for tunnelling

Virtual Private Networks

- Corporations would like private networks to connect their different campuses
- Companies cannot afford dedicated physical networks (or leased lines)
- Solution – use tunneling
  - Tunneling creates a virtual circuit
  - Encrypt payload if security is needed

Routing for Mobile Networks – Last Hop

- Mobile hosts can move to another network – Problem:
  - How to keep existing connections/IP address
- One solution – Mobile IP. Idea is, when a host moves:
  - Foreign Agent (FA) assigns a care-of-address
  - FA informs Home Agent (HA) of care-of-address
  - HA tunnels packets destined to the mobile to the care-of-address
  - FA strips tunnel header and delivers packet
- Any problems?
  - What about hosts on the home network?
  - How efficient? Can it be optimized?
Discussion

• Which of the two approaches is better?
• Are there security issues in either?

Triangulation

• Consider a source node (S) talking to a mobile node M
  • Packets from S have to be forwarded through home agent
  • Packets from M can be sent directly to S
    • Will this work with ingress (and other) filtering?
  • Route looks like a triangle – inefficient
    • IETF draft proposes solution to this problem
    • Bad: requires changes to S’s TCP/IP stack

Generalizing the Idea: Overlay Networks

• Deploy an Overlay Network over multiple Internet Sites
  • Applications configure themselves and route traffic among participating overlay nodes
  • Many applications can use such a model to implement features not available in a vanilla IP network
  • End to End argument in action?
  • Do we have to use tunneling?
• E.g., Peer-to-Peer networks, Application Level Multicast, Secure Overlays, VPNs, Content Delivery systems, even your second class project
• Important and emerging research area. Many research issues. We will revisit later.
Routing – Ad Hoc Networks

• Recall – ad hoc networks are networks made up completely of mobile devices
  – Scenario – handheld devices on infantry in enemy territory; search and rescue operations; conference; cars on a highway

• Problem: computer moves, network really dynamic
• Can we just use traditional routing algorithms such as RIP or OSPF?
  – RIP is too slow to converge
  – Change is an exception in wired protocol, it is the norm in ad hoc networks

Routing in Ad Hoc Networks

• Two flavors of routing algorithms

• Table-driven, or proactive:
  – similar to traditional algorithms – construct a table of paths to all nodes in the network
  – High Overhead, but low startup time

• On-demand, or reactive:
  – when you need a path to a node, flood the network to search for it
  – Lower overhead, higher startup time

Problems with DSDV/Table Driven

• Control message overhead grows with \( O(n^2) \)
  – A lot of overhead incurred proactively when a node moves in or out of range
  – What if no one is communicating with it?

• Each node must have complete table

• Alternative – Reactive routing
  – Compute route only when needed
  – Low routing overhead, but high latency
  – Big flood of the network when communication is needed
On-demand (or reactive) Routing

- When you need to find a path to a destination, a route discovery process is initiated
  - Flood the network with a route request packet
  - When a node receives a route request looking for it, it reverses the path and sends a route reply back
  - Intermediate nodes can set their routing table entries
- Once a reply is received, the packet can be sent
- Examples, Dynamic Source Routing (DSR, Johnson et al), Ad Hoc On Demand Distance Vector (AODV, Perkins et al)
- Any concerns?
  - Should we use a flood on every packet? Too expensive
  - What if the route disappears while in use?
  - It is too expensive to detect a route breaking (several MAC layer timeouts)

Example – DSR

- Source Routing protocol; intermediate nodes do not have to keep information
- Route request
  - Each node receives the route request, adds itself as a hop and forwards it along
  - When destination receives a route request, it reverses the path and sends a reply
- Source uses path with smallest number of hops
- When a link breaks, the node along the link that failed sends a route error packet to source
  - Source looks for a new route

Optimizations

- Route flood too expensive
  - Flood returns multiple paths — retain in a cache
  - Localize the flood
  - Based on location (Location Aided Routing, Vaidya et al)
  - Based on previous path (Query Localization, Das et al)
  - Reply from cache rather than continuing flood?
  - Cache inconsistency
- MAC level broadcast unreliable