Traffic Engineering for ISP Networks

Jennifer Rexford

Internet and Networking Systems

AT&T Labs - Research; Florham Park, NJ

http://www.research.att.com/~jrex

Joint work with Anja Feldmann, Albert Greenberg, Carsten Lund, Nick Reingold, and Fred True, and AT&T IP Services

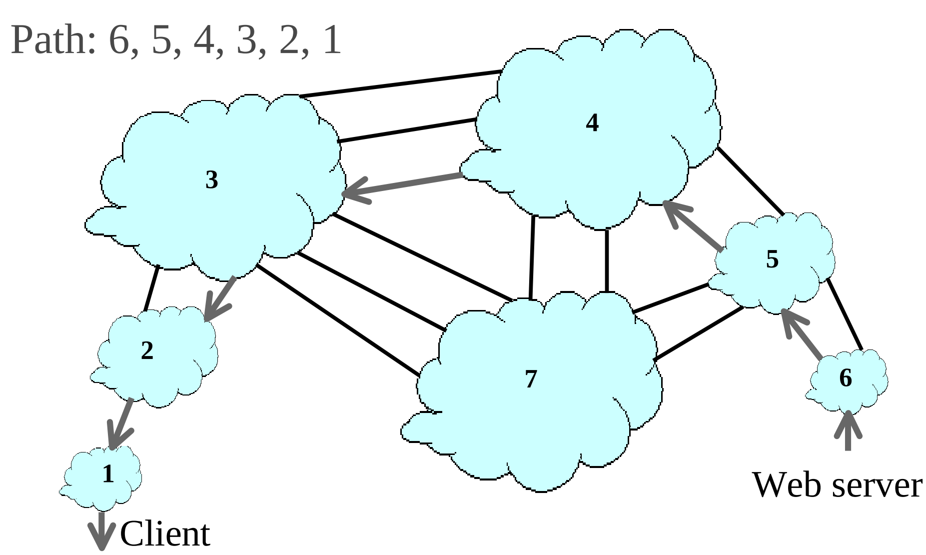
Outline

- Background
  - – Internet architecture
  - – Interdomain and intradomain routing
  - – Internet service provider backbone
- Traffic engineering
  - – Optimizing network configuration to prevailing traffic
  - – Requirements for topology, routing, and traffic info
- Traffic demands
  - – Volume of load between edges of the network
  - – Measurement methodology
  - – Analysis of the demands on AT&T's IP Backbone

Internet Architecture

- Divided into Autonomous Systems
  - Distinct regions of administrative control (~6000-7000)
  - Set of routers and links managed by a single institution
  - Service provider, company, university, …
- Hierarchy of Autonomous Systems
  - Large, tier-1 provider with a nationwide backbone
  - Medium-sized regional provider with smaller backbone
  - Small network run by a single company or university
- Interaction between Autonomous Systems
  - Internal topology is not shared between ASes
  - … but, neighboring ASes interact to coordinate routing

Autonomous Systems (ASes)



Characteristics of the Internet

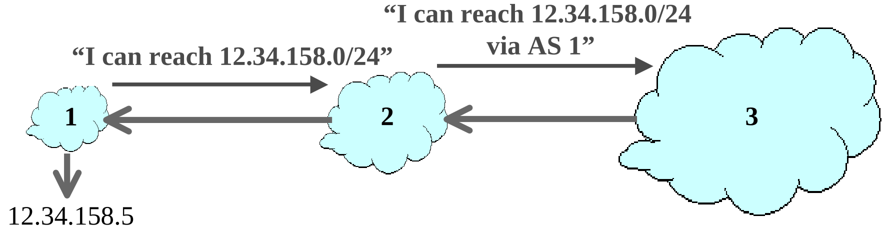
- The Internet is
  - Decentralized (loose confederation of peers)
  - Self-configuring (no global registry of topology)
  - Stateless (limited information in the routers)
  - Connectionless (no fixed connection between hosts)
- These attributes contribute
  - To the success of Internet
  - To the rapid growth of the Internet
  - … and the difficulty of controlling the Internet!

Internet – Interconnection of Networks

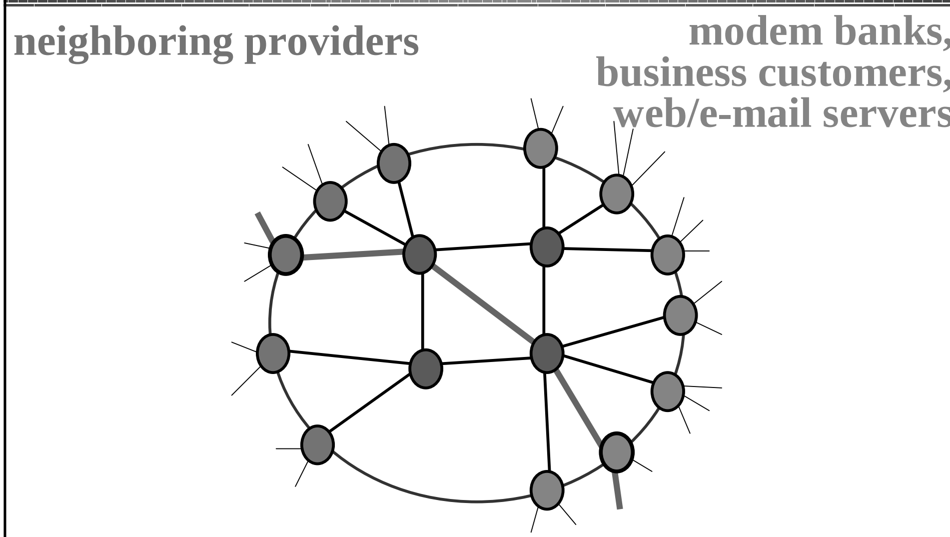
- Internet Protocol
  - Transmit a single packet from one host to another
  - Packet includes the IP address of the sender and receiver
  - Packets may be lost, delayed, or delivered out of order
  - Hosts perform retransmission and reordering of packets
- IP address
  - 32-bit IP addresses divided into octets (12.34.158.5)
  - Allocated to institutions in contiguous blocks or prefixes
  - 12.34.158.0/24 is a 24-bit prefix with 2^8 IP addresses
  - Routing of IP packets in the Internet is based on prefixes

Interdomain Routing (Between ASes)

- ASes exchange info about who they can reach
- Local policies for path selection (which to use?)
- Local policies for route propagation (who to tell?)
- Policies configured by the AS's network operator



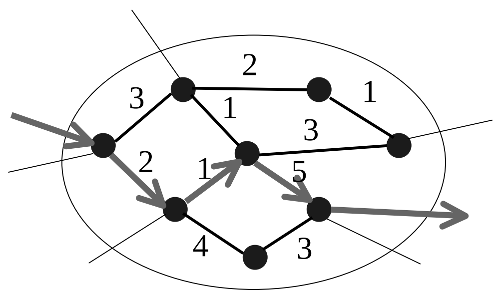
Internet Service Provider Backbone



How should traffic be routed through the ISP backbone?

Intradomain Routing (Within an AS)

- Routers exchange information to learn the topology
- Routers determine "next hop" to reach other routers
- Path selection based on link weights (shortest path)
- Link weights configured by AS's network operator
- … to engineer the flow of traffic



Traffic Engineering in an ISP Backbone

- Topology of the ISP backbone
  - Connectivity and capacity of routers and links
- Traffic demands
  - Expected/offered load between points in the network
- Routing configuration
  - Tunable rules for selecting a path for each traffic flow
- Performance objective
  - Balanced load, low latency, service level agreements …
- Question: Given the *topology* and traffic *demands* in an IP network, which *routes* should be used?

State-of-the-Art in IP Networks

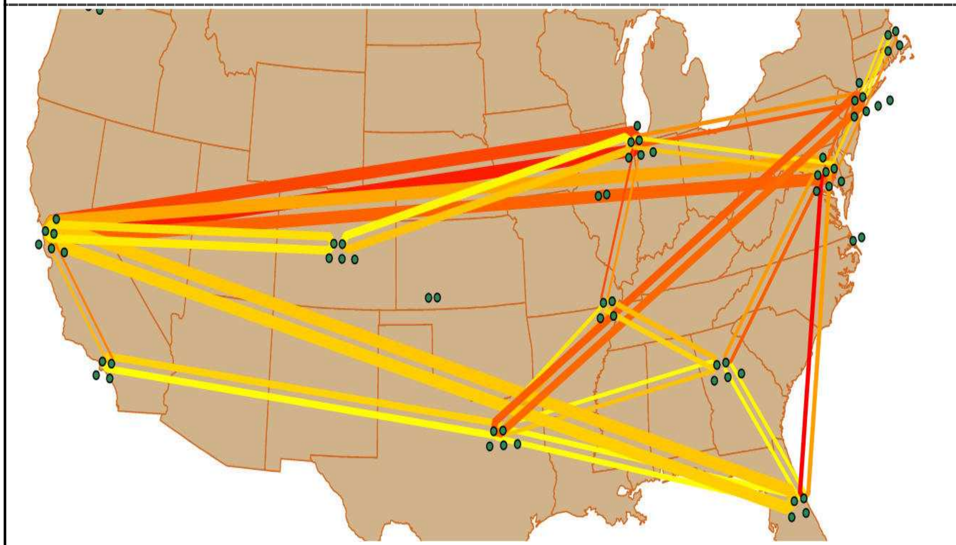
- Missing input information
  - The topology and traffic demands are often unknown
  - Traffic fluctuates over time (user behavior, new appls)
  - Topology changes over time (failures, growth, reconfig)
- Primitive control over routing
  - The network does not adapt the routes to the load
  - The static routes are not optimized to the traffic
  - Routing parameters are changed manually by operators

(But, other than that, everything is under control…)

Example: Congested Link

- Detecting that a link is congested
  - Utilization statistics reported every five minutes
  - Sample probe traffic suffers degraded performance
  - Customers complain (via the telephone network?)
- Reasons why the link might be congested
  - Increase in demand between some set of src-dest pairs
  - Failed router/link within the AS causes routing change
  - Failure/reconfiguration in another AS changes routes
- How to determine *why* the link is congested???
  - Need to know the *cause*, not just the *manifestations*!
- How to alleviate the congestion on the link???

Link Utilization



Utilization: link color (high to low)

Requirements for Traffic Engineering

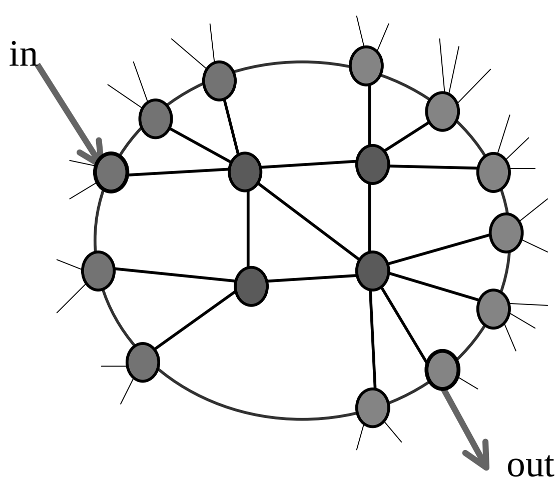
- Models
  - **Traffic demands**
  - Network topology/configuration
  - Internet routing algorithms
- Techniques for populating the models
  - **Measuring/computing the traffic demands**
  - Determining the network topology/configuration
  - Optimizing the routing parameters
- Analysis of the traffic demands
  - **Knowing how the demands fluctuates over time**
  - **Understanding the traffic engineering implications**

Modeling Traffic Demands

- Volume of traffic V(s,d,t)
  - From a particular *source* s
  - To a particular *destination* d
  - Over a particular *time* period t
- Time period
  - Performance debugging -- minutes or tens of minutes
  - Time-of-day traffic engineering -- hours
  - Network design -- days to weeks
- Sources and destinations
  - Individual hosts -- interesting, but huge!
  - Individual prefixes -- still big; not seen by any one AS!
  - Individual edge links in an ISP backbone -- hmmm….

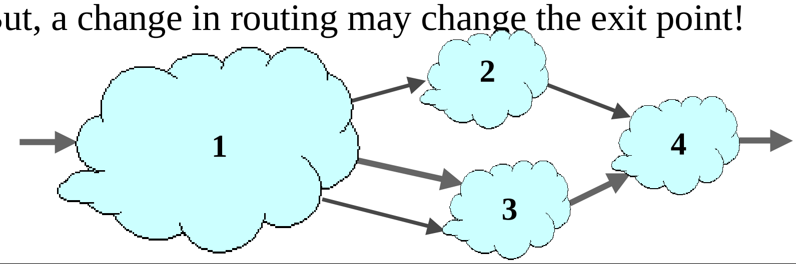
Traffic Matrix

Traffic matrix: V(in,out,t) for all pairs (in,out)



Problem: Multiple Exit Points

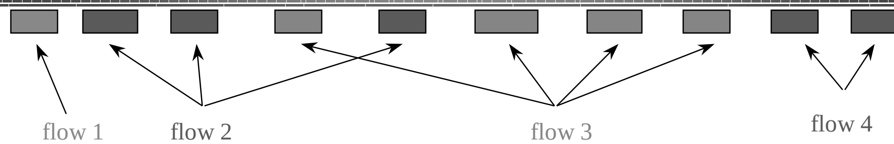
- ISP backbone is in the *middle* of the Internet
  - Multiple connections to other autonomous systems
  - Destination is reachable through multiple exit points
  - Selection of exit point depends on intradomain routes
- Problem with traditional point-to-point models
  - Want to predict impact of changing intradomain routing
  - But, a change in routing may change the exit point!



Traffic Demand

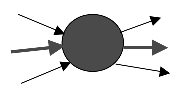
- Definition: V(in, {out}, t)
  - Entry link (in)
  - Set of possible exit links ({out})
  - Time period (t)
  - Volume of traffic (V(in,{out},t))
- Computing the traffic demands
  - Measure the traffic where it *enters* the ISP backbone
  - Identify the set of *exit* links where traffic could leave
  - Associate traffic with the entry link and set of exit links
  - Aggregate over all traffic with same in, {out}, and t

Measuring Flows Rather Than Packets



- IP flow abstraction
  - Set of packets with “same” src and dest IP addresses
  - Packets that are “close” together in time (a few seconds)
  - The closest thing to a “call” in the Internet
- Cisco NetFlow
  - Measure all flows between AT&T and neighbors
  - Extremely large amount of data (~100 GB/day)

NetFlow Data

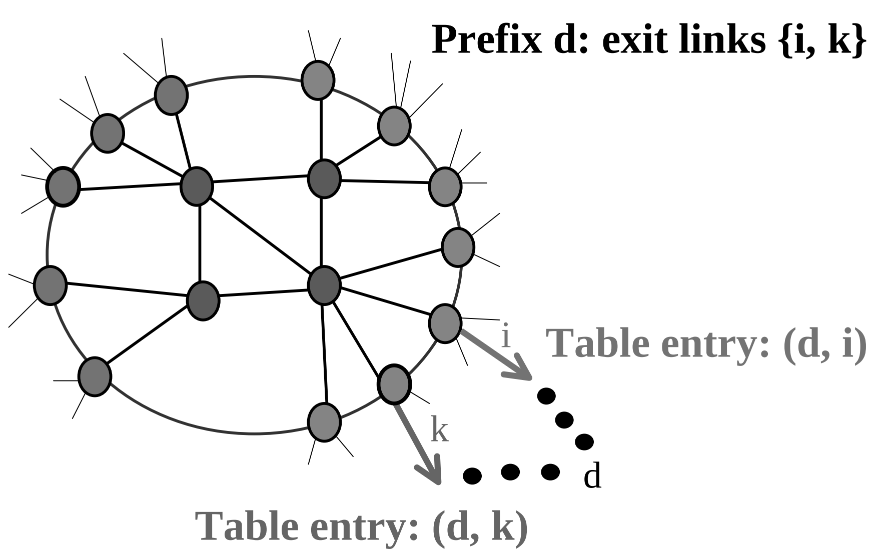


- Source and destination information
  - Source and **destination** IP addresses (hosts)
  - Source and destination port numbers (application)
  - Source and destination AS numbers
- Routing information
  - Source and destination IP prefix (network address)
  - **Input** and output links at this router
- Traffic information
  - **Start** and **finish** time of flow (in seconds)
  - Total number of **bytes** and packets in the flow

Identifying Where the Traffic Can Leave

- Traffic flows
  - Each flow has a dest IP address (e.g., 12.34.156.5)
  - Each address belongs to a prefix (e.g., 12.34.156.0/24)
- Forwarding tables
  - Each router has a table to forward a packet to "next hop"
  - Forwarding table maps a prefix to a "next hop" link
- Process
  - Dump the forwarding table from each router
  - Identify the entries where the "next hop" is an exit link
  - Identify the set of exit links associated with each prefix
  - Associate a flow's dest address with the set of exit links

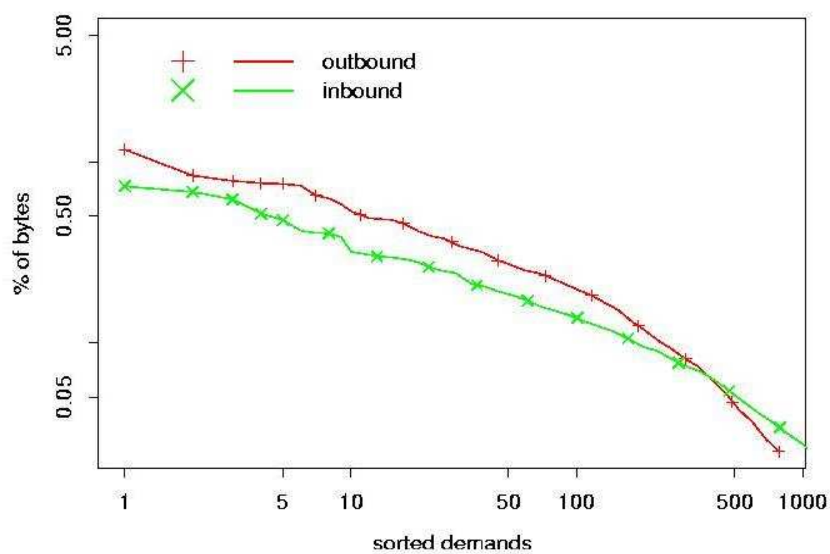
Locating the Set of Exit Links for Prefix d



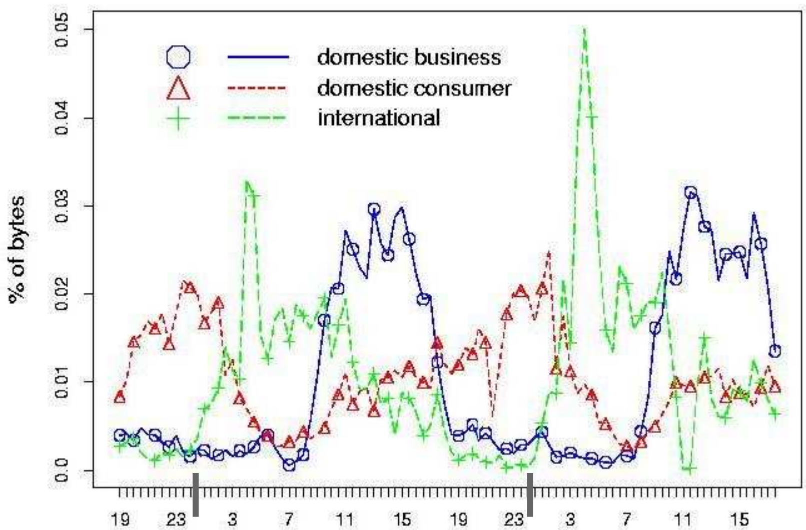
Experimental Results: AT&T IP Backbone

- Measurement for four days in November 1999
  - Netflow data
  - Forwarding tables
  - Topology and routing configuration
- Traffic analysis
  - Distribution of traffic volume across demands
    - » Small % of demands are responsible for most traffic
  - Time-of-day fluctuations in traffic volumes
    - » U.S. business, U.S. residential, International
  - Stability of traffic demands across hours and days
    - » Initial results suggest some stability, but need to study

Proportion of Traffic in Top Demands (Log Scale)



Time-of-Day Effects (San Francisco)

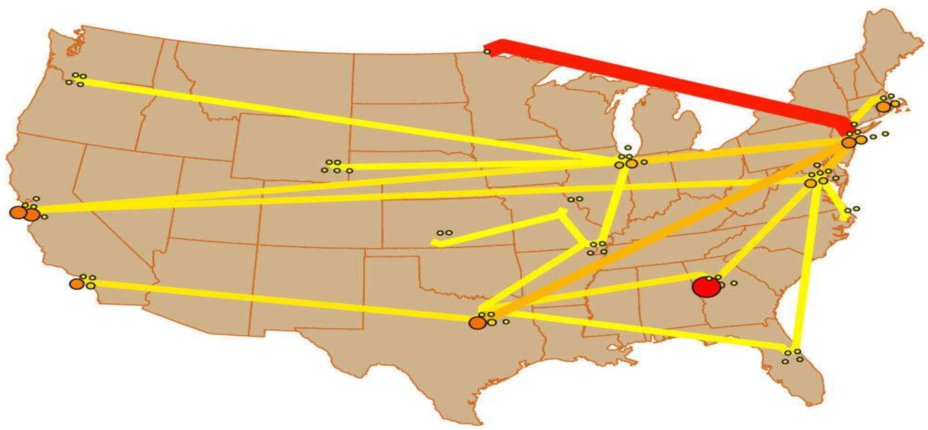


Traffic-Engineering Implications

- Small number of demands contribute most traffic
  - – Can optimize routes for just the heavy hitters
  - – Can measuring a small fraction of the traffic
  - – Must watch out for changes in load and set of exit links
- Time-of-day fluctuations
  - – Reoptimize routes a few times a day (three?)
- Traffic (in)stability
  - – Select routes that are good for different demand sets
  - – Reoptimize routes after sudden changes in load

Traffic Flow Through AT&T's IP Backbone

Source node: public peering link (NAP) in New York
Destination nodes: AT&T access routers



Color/size of node: proportional to traffic to this router (high to low)
Color/size of link: proportional to traffic carried (high to low)

Conclusions

- Internet traffic engineering is hard
  - Decentralized (over 6000 Autonomous Systems)
  - Connectionless (traffic sent as individual packets)
  - Changing (topological changes, traffic fluctuations)
- Traffic engineering requires knowing the demands
  - Interdomain traffic has multiple possible exit points
  - Demand as the load from entry to set of exit points
  - Not available from traditional measurement techniques
- Measurement of traffic demands
  - Derivable from flow-level measurements at *entry* points
  - … and "next hop" forwarding info from *exit* points

Ongoing Work

- Detailed analysis of traffic demands
  - Statistical properties (how to study stability?)
  - Implications for traffic engineering
- Online computation of traffic demands
  - Distributed flow-measurement infrastructure
  - Online aggregation of flow data into demands
- Network operations ("operations" research?)
  - Efficiently detecting sudden changes in traffic or routing
  - Optimizing routes based on topology and demands
  - Planning the design of the network over time
  - Getting the network to run itself… ☺

Interesting Problems

- Inferring the traffic demands from less information
  - sampling, active probes, inference from utilization
- Optimizing routes subject to fluctuating demands
  - optimal routes per demand set vs. good for all sets
- Techniques for analyzing stability of demand sets
  - multidimensional data (in, {out}, time)
- Detecting shifts in the distribution of load
  - random changes vs. change in underlying distribution
- Joint route optimization across multiple ASes
  - optimizing routes without divulging topology & traffic