QoS routing:

Given two or more performance metrics—e.g., delay and bandwidth—find path with delay less than target delay $D$ (e.g., 100 ms) and bandwidth greater than target bandwidth $B$ (e.g., 1.5 Mbps)

$\rightarrow$ from shortest path to best QoS path

$\rightarrow$ multi-dimensional QoS metric

$\rightarrow$ other: delay, hop count, etc.

How to find best QoS path that satisfies all requirements?

Brute-force

• Enumerate all possible paths

• Rank them
How many paths are there:

- If there are $n$ nodes, there can be up to

$$\frac{n(n - 1)}{2}$$

undirected links

- Hence, from source $S$ there can be up to

$$(n - 1) (n - 2) \cdots 3 \ 2 \ 1 = (n - 1)!$$

paths

- By Stirling’s formula

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

→ superexponential

→ too many for brute-force
Is there a more clever or better algorithm?

→ as of Apr. 13, 2004: unknown

→ specifically: QoS routing is NP-complete

→ strong evidence there may not exist good algorithm

In networking: several problems turn out to be NP-complete

→ e.g., scheduling, control, ... 

→ “P = NP” problem

→ one of the hardest problems in science ever

Doesn’t matter too much for QoS routing

→ little demand for very good algorithm
Policy routing:

→ “policy” is not precisely defined

→ anything goes

Routing criteria include

• Performance
  → e.g., shortest path

• Trust
  → what in the world is it?

• Economics
  → pricing
  → flexibility through multiple providers

• Politics, social issues, etc.
Implementation

Major Internet routing protocols:

- RIP (v1 and v2): intra-domain, Bellman-Ford
  → also called “distance vector”
  → metric: hop count
  → UDP
  → nearest neighbor advertisement
  → popular in small intra-domain networks

- OSPF (v1 and v2): intra-domain, Dijkstra
  → also called “link state”
  → metric: average delay
  → directly over IP: protocol number 89
  → broadcasting via flooding
  → popular in larger intra-domain networks
• IS-IS: intra-domain, Dijkstra
  → “link state”
  → directly over link layer (e.g., Ethernet)
  → more recently: also available over IP
  → flooding
  → popular in larger intra-domain networks

• Source routing: packet specifies path
  → implemented in various link layer protocols
  → ATM call set-up: circuit-switching
  → IPv4/v6: option field
  → mostly disabled
  → large ISPs: sometimes used internally for diagnosis
BGP (Border Gateway Protocol):

- Inter-domain routing
  → border routers vs. backbone routers

→ peering between two AS’s
→ exchanges: peering between multiple AS’s
• CIDR addressing
  → i.e., \textit{a.b.c.d}/x

• Routing table look-up: maximum prefix matching
  → e.g., route aggregation

• Metric: policy
  → e.g., shortest-path, trust, pricing
  → meaning of “shortest”
  → mechanism: path vector routing
  → BPG update message
AS H → AS F → AS C → AS B → AS A
AS H → AS G → AS D → AS E

AS F → C → B → A; a.b.c.d/x
AS F → G → D → B → A; a.b.c.d/x
AS C → B → A; a.b.c.d/x
AS D → B → A; a.b.c.d/x

→ AS-PATH (path vector)
BGP-update procedure:

Upon receiving BGP update message from neighbor to target $A$

1. Store AS-PATH reachability info for target $A$
2. Determine if new path to $A$ should be adopted
   → policy
   → path should be unique
   → BPG table & IP routing table update
3. Determine who to advertise reachability for target $A$
   → selective advertisement

Note: if shortest-path then same as Dijkstra in-reverse
   → global advertisement advertisement
BGP-withdrawal:

1. Use BGP keep-alive message to sense/prompt neighbor
2. If keep-alive does not arrive within certain time, assume node is down
3. Send BGP withdraw message for neighbor who is deemed down
   → may trigger further updates

Other BGP features:

- BGP runs over TCP
  → port number 179
  → i.e., “application layer” protocol
- BPG-4 (1995); secure BGP
Performance

Route update frequency:

→ routing table stability vs. responsiveness
→ rule: not too frequently
→ 30 seconds
→ stability wins
→ hard lesson learned from the past (sub-second)
→ e.g., TTL

Other factors for route instability:

→ selfishness (e.g., fluttering)
→ BGP’s vector path routing
→ inherently unstable: chain reaction
→ more frequent: slow convergence
→ target of denial-of-service (DoS) attack
Route amplification:

→ shortest AS path ≠ shortest router path
→ e.g., may be several router hops longer
→ AS graph vs. router graph
→ inter- vs. intra-domain routing: separate subsystems

Route asymmetry:

→ routes are not symmetric
→ estimate: > 50%
→ mainly artifact of inter-domain policy routing
→ also intra-domain: e.g., hot potato
→ various performance implications
Black holes:

→ persistent unreachable destination prefixes
→ BGP routing problems
→ further aggrevated by DNS
→ purely application layer: end system problem

Topology: