Routing Algorithms and Traffic Engineering

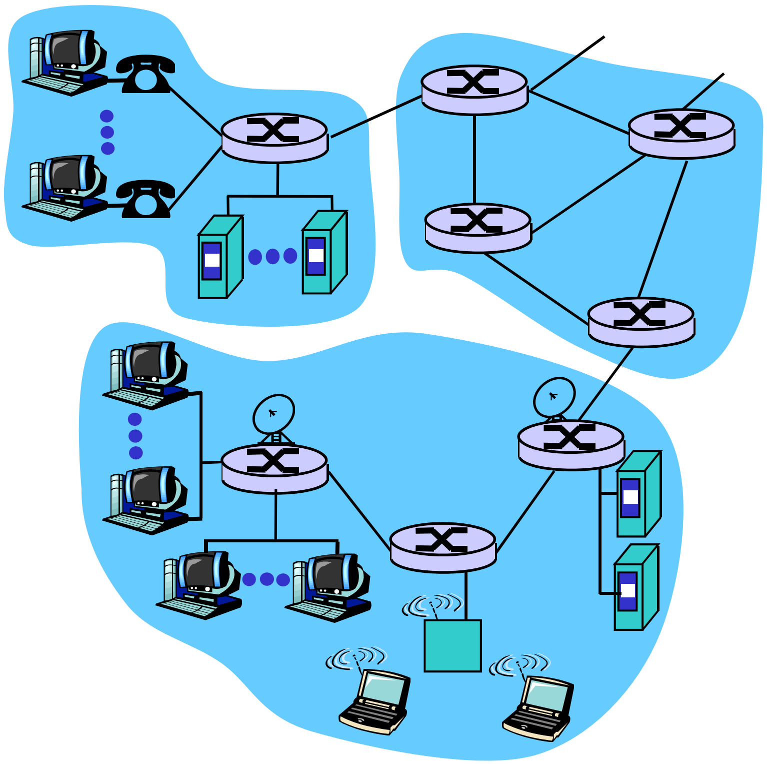
- MPLS and OSPF
- traffic engineering

- linear programming

Network layer functions

three important functions:

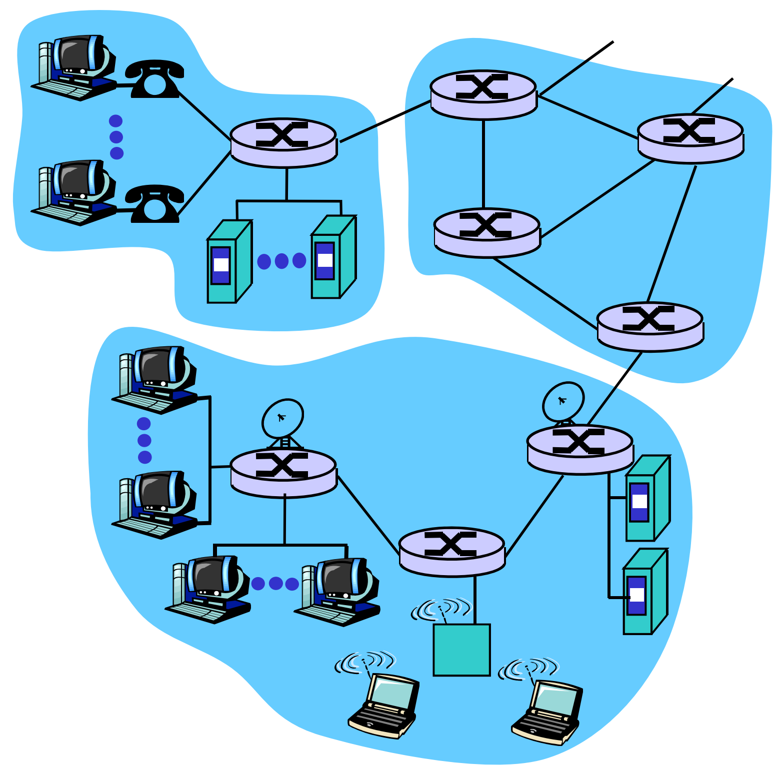
- *path determination:* route taken by packets from source to destination
- *forwarding:* move packets from router's input to appropriate router output
- *call setup:* some network architectures require router call setup along path before data flows



Network layer functions

three important functions:

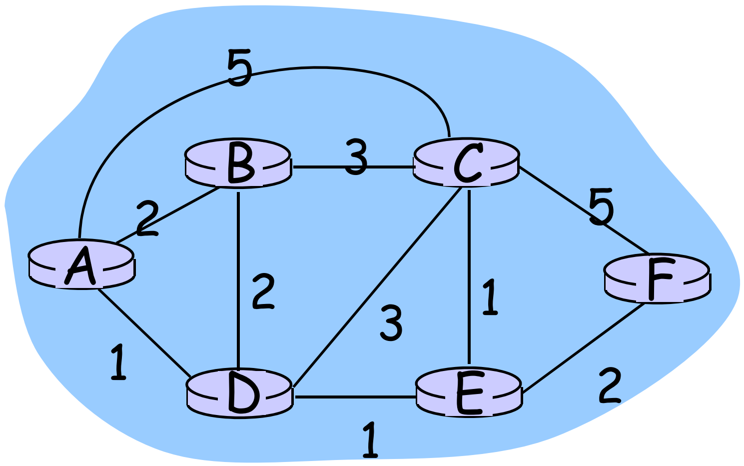
- ***path determination:*** **route taken by packets from source to destination**
  Routing algorithm
- *forwarding:* move packets from router's input to appropriate router output
- *call setup:* some network architectures require router call setup along path before data flows



Routing

Routing protocol

Goal: determine "good" path (sequence of routers) thru network from source to dest.



Graph abstraction for routing algorithms:

- graph nodes are routers
- graph edges are physical links
  - link cost: delay, $ cost, or congestion level

- "good" path:
  - typically means minimum cost path
  - other def's possible

Routing Algorithm classification

Global or decentralized information?

Global:

- all routers have complete topology, link cost info
- "link state" algorithms (Dijkstra)

Decentralized:

- router knows physically-connected neighbors, link costs to neighbors
- iterative process of computation, exchange of info with neighbors
- "distance vector" algorithms (Bellman Ford)

OSPF (Open Shortest Path First)

- link state protocol
  - link costs between 0 and 65,535
  - Cisco recommendation - link cost α 1/(link capacity)
  - rapid, loop-free convergence, scales well
  - topology map at each node, route computation using Dijkstra's algorithm
  - OSPF advertisement carries one entry per neighbor router, advertisements flooded to entire Autonomous System
- multiple equal-cost paths allowed: flow equally split on all outgoing links belonging to shortest paths
- IS – IS (intermediate system-intermediate system) similar

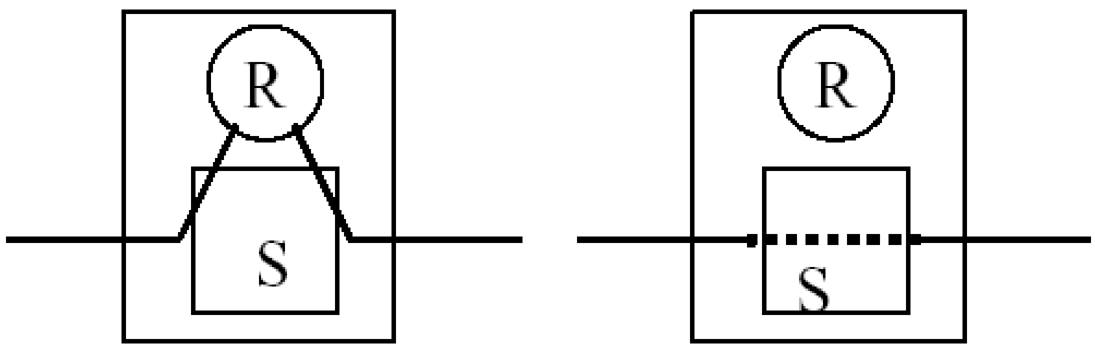
Routing vs Switching

- routing: based on address lookup, max prefix match
  - search operation
  - complexity O(logn) - O(n)
- switching: based on circuit numbers
  - indexing operation
  - complexity O(1)
  - scalable to large networks

$\Rightarrow$ MPLS

History: Ipsilon's IP Switching

- developed by Ipsilon
- routing software in every ATM switch
- initially, packets reassembled by routing software, forwarded to next hop
- long term flows transferred to separate VCs.



Ra

Ipsilon's IP Switching

ATM VCs set up when new IP "flows" seen, i.e., "data-driven" VC setup

- flow oriented traffic: ftp, ssh, http, multimedia
- short-lived traffic: DNS, SMTP, NTP
- Ipsilon claimed 90% of bytes flow-oriented
- runs as added software on ATM switch (12,000 lines of code)

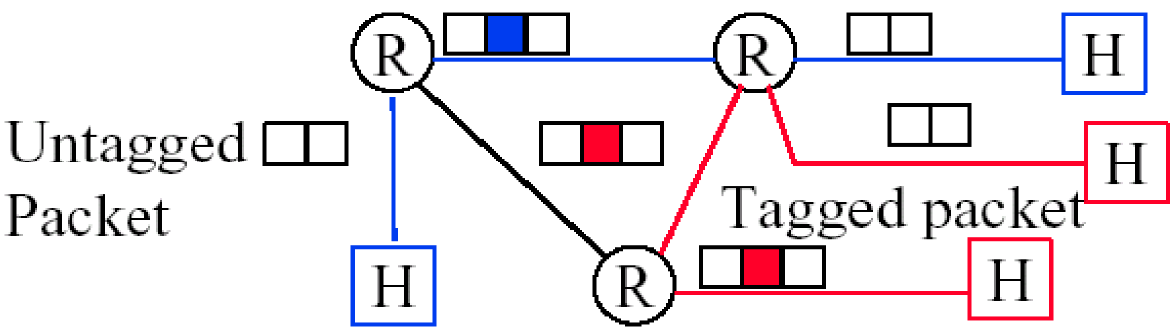
Issues with Ipsilon's IP switching

- VCI field used as ID
  VPI/VCI change at switch
  ⇒ must run on every ATM switch
  ⇒ non-IP switches not allowed between IP switches
  ⇒ subnets limited to one switch
- cannot support VLANs
- scalability: no. VCs > no. flows
  ⇒ VC explosion. 1000 setups/sec.
- QoS determined implicitly by flow class or RSVP
- ATM only

Tag Switching

- proposed by CISCO
- similar to VLAN tags
- tags explicit or implicit

- ingress router/host puts tag. exit router



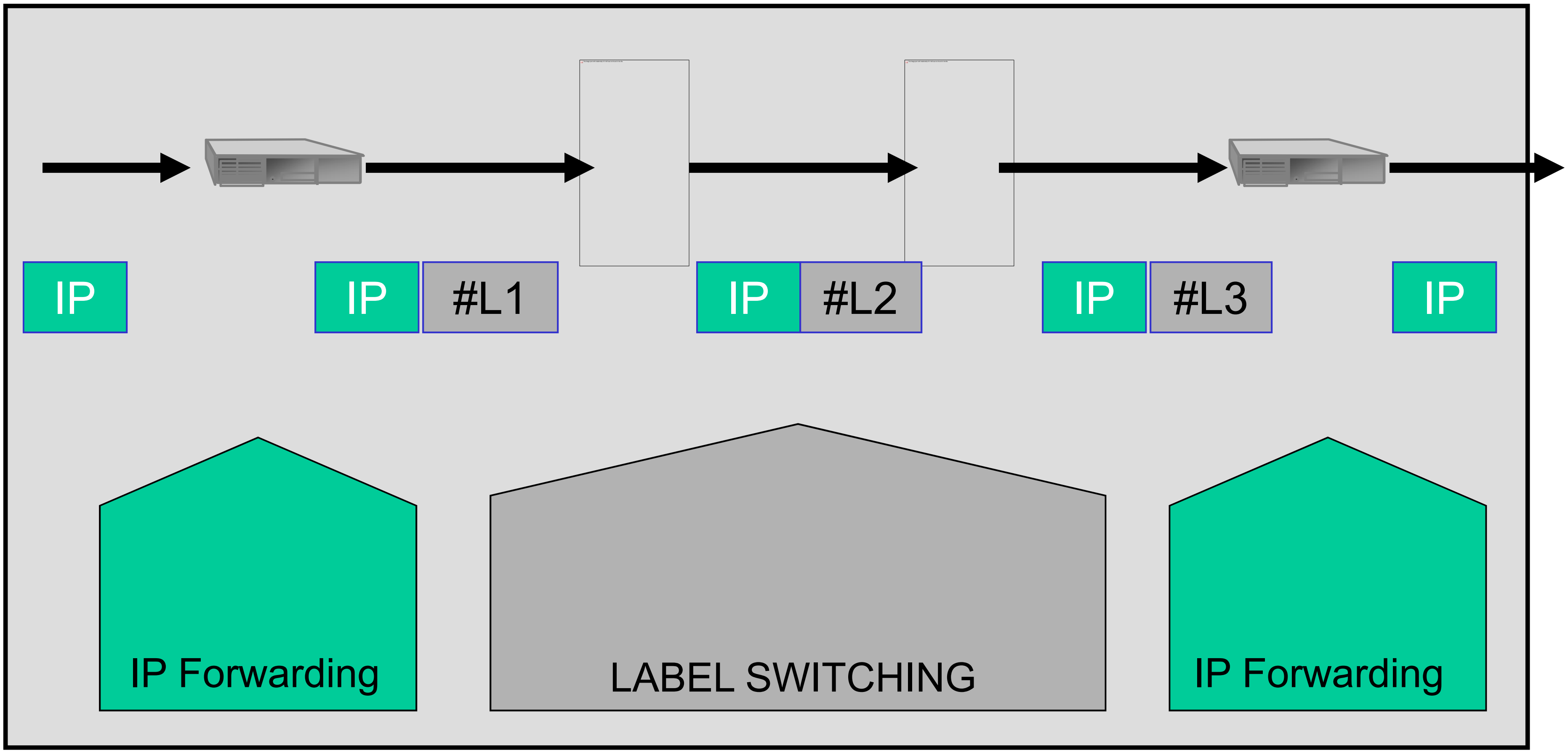
<u>Key difference:</u> tags set up in background using IP routing protocols (i.e. <u>control-driven</u> VC setup)

Alphabet Soup!

- CSR Cell Switched Router
- ISR Integrated Switch and Router
- LSR Label Switching Router
- TSR Tag Switching Router
- Multi layer switches
- Direct IP
- FastIP
- PowerIP

MPLS - IETF standard

MPLS concept: route at edge, switch in core



MPLS

- flows assigned labels, routing along Label Switched Path
- finer granularity for routing
  - can allow uneven traffic split
- not tied to any route computation algorithm

Traffic Engineering Framework

- knowledge of topology
- traffic matrix
  - K – set of origin destination flows
  - $k \in K$, d_k – demand, s_k – source, t_k – destination
- optimization criteria
  - minimize maximum utilization
  - keep utilizations below 60%

How does one set link weights?

Digression – linear programming

Linear program

$$\text{minimize} \quad \sum_{i=1}^{n} c_i x_i$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{i,j} x_j = b_i, \quad i = 1, \ldots, m$$

$$x_j \geq 0, \quad j = 1, \ldots, n$$

- polynomial time solution in n, m

Surplus variables

Slack variables

Free variables

Example: optimal routes

- topology $G = (V,E)$
- K - set of origin destination flows
  - $k \in K$, s_k - source, t_k - destination
- set of given link weights $\{w_{ij} : (i,j) \in E\}$
- X_{ij}^{k} fraction of flow k going over $(i,j) \in E$

Example

- decomposes into separate problems per flow k $\in$K

$$\min \sum_{(i,j)\in E} w_{ij} X_{ij}^{k}$$

$$s.t. \quad \sum_{j:(i,j)\in E} X_{ij}^{k} - \sum_{j:(j,i)\in E} X_{ji}^{k} = 0, \quad i \neq s_k, t_k$$

$$\sum_{j:(s_k,j)\in E} X_{ij}^{k} - \sum_{j:(j,s_k)\in E} X_{ji}^{k} = 1,$$

$$\sum_{j:(t_k,j)\in E} X_{ij}^{k} - \sum_{j:(j,t_k)\in E} X_{ji}^{k} = -1,$$

$$X_{ij}^{k} \geq 0$$

Interpretation

- let $\left\{\overline{X}_{ij}^{k}\right\}$ be optimal solutions
- if $\left\{\overline{X}_{ij}^{k}\right\}$ takes values 0 and 1, corresponds to shortest paths
- if $\left\{\overline{X}_{ij}^{k}\right\}$ takes other values, there exist multiple shortest paths.

Linear Program

$$\begin{aligned} \text{minimize} \quad & \sum_{\mathrm{i}} c_i x_j \\ \text{subject to} \quad & \sum_{\mathrm{j}} a_{i,j} x_j = b_i,\ i = 1,\ldots,m \\ & x_j \ge 0, \qquad j = 1,\ldots,n \end{aligned} \quad \rightarrow \quad \begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \ge 0, \end{aligned}$$

- x^0 is feasible if $Ax^0 = b$ and $x^0 \geq 0$

Basic solutions

Theorem of LP

Dual problem (symmetric form)

Primal		Dual	
minimize	$c^T x$	maximize	$y^T b$
subject to	$Ax \geq b$	subject to	$y^T A \leq c^T$
	$x \geq 0,$		$y \geq 0$

Dual problem (assymmetric form)

Primal

$$\text{minimize} \quad c^T x$$
$$\text{subject to} \quad Ax = b$$
$$x \geq 0,$$

Dual

$$\text{maximize} \quad y^T b$$
$$\text{subject to} \quad y^T A \leq c^T$$

Dual problem: properties.

- if x, y feasible, then $c^T x \geq y^T b$
- if x′, y′ feasible and if $c^T x' = y'^T b$, then x′ and y′ are optimal
- if either problem has finite solution, so does other, if either has unbounded solution, the other has no feasible solution.

Complementary slackness

Let x and y be feasible solutions. A necessary and sufficient condition for them to be optimal is that for all i

1. $x_i > 0 \Rightarrow y^T A_i = c_i$
2. $x_i = 0 \Leftarrow y^T A_i < c_i$

Here A_i is i-th column of A

Example: primal (P-SP)

$$\min \sum_{(i,j)\in E} w_{ij} X_{ij}^k$$

$$s.t. \quad \sum_{j:(i,j)\in E} X_{ij}^k - \sum_{j:(j,i)\in E} X_{ji}^k = 0, \quad i \neq s_k, t_k$$

$$\sum_{j:(s_k,j)\in E} X_{ij}^k - \sum_{j:(j,s_k)\in E} X_{ji}^k = 1,$$

$$\sum_{j:(t_k,j)\in E} X_{ij}^k - \sum_{j:(j,t_k)\in E} X_{ji}^k = -1,$$

$$X_{ij}^k \geq 0$$

Dual problem

- introduce dual variables $y_i^k, i \in V$
- dual problem, for all k∈K

$$\max \sum_{k\in K} y_{s_k}^k - y_{t_k}^k$$

$$s.t. \quad y_i^k - y_j^k \le w_{ij}, \ (i, j) \in E$$

Example: dual (D-SP)

- change of variables

$$U_i^k = y_{s_k}^k - y_i^k, \quad i \in V$$

- leads to

$$\max \sum_{k \in K} U_{t_k}^k, \quad k \in K$$
$$s.t. \quad U_j^k - U_i^k \leq w_{ij}, (i, j) \in E, k \in K$$
$$U_{s_k}^k = 0, \quad k \in K$$

Example

$$\max \sum_{k \in K} U_{t_k}^{k}, \quad k \in K$$

$$s.t. \quad U_j^k - U_i^k \le w_{ij}, (i,j) \in E, k \in K$$

$$U_{s_k}^k = 0, \quad k \in K$$

- $\{\overline{U}_i^k\}$ optimal solution to dual problem
- $\overline{X}_{ij}^k > 0 \Rightarrow \overline{U}_j^k - \overline{U}_i^k = w_{ij}, \quad \overline{U}_j^k$ length of shortest path from s_k to j
- $\overline{U}_{t_k}^k$ length of shortest path from s_k to t_k

How does one set link weights for OSPF?

Traffic engineering problem:
minimize maximum link utilization

- topology $G = (V,E)$
- c_{ij} - capacity of link $(i,j) \in E$
- K – set of origin destination flows
  - $k \in K$, d_k - demand, s_k - source, t_k - destination
- α – maximum link utilization

LP formulation

$$\text{minimize}\ \alpha$$

$$s.t. \sum_{j:(i,j)\in E} X_{ij}^{k} - \sum_{j:(j,i)\in E} X_{ji}^{k} = \begin{cases} 0, & i \neq s_k, t_k, k \in K \\ d_k, & i = s_k k \in K \\ -d_k, & i = t_k k \in K \end{cases}$$

$$\sum_{k\in K} X_{ij}^{k} \leq c_{ij}\alpha, \qquad (i,j) \in E$$

$$X_{ij}^{k} \geq 0$$

X_{ij}^{k} represents total data rate for k-th flow over link (i,j)

LP formulation

$$\text{minimize}\; \alpha + r\sum_{k\in K}\sum_{(i,j)\in E} X_{ij}^{k}$$

$$s.t. \sum_{j:(i,j)\in E} X_{ij}^{k} - \sum_{j:(j,i)\in E} X_{ji}^{k} = \begin{cases} 0, & i \neq s_k, t_k, k \in K \\ d_k, & i = s_k k \in K \\ -d_k, & i = t_k k \in K \end{cases}$$

$$\sum_{k\in K} X_{ij}^{k} \leq c_{ij}\alpha, \qquad (i,j) \in E$$

$$X_{ij}^{k} \geq 0$$

LP formulation

- can be many solutions with same α
- in case of tie, want solution with short paths

$\Rightarrow$ add term $r\sum_{k\in K}\sum_{(i,j)\in E} X_{ij}^{k}$
with small r to cost

- use standard LP algorithms (Simplex) to solve

Q: can we find link weights so that solution comes from shortest path problem?

Duality revisited

Primal		Dual	
minimize	$c^T x$	maximize	$y_1^T b_1 + y_2^T b_2$
subject to	$Ax = b_1$	subject to	$y_1^T A + y_2^T A' \le c^T$
	$A' x \ge b_2$		$y_2 \ge 0$
	$x \ge 0,$		

- free variables in primal $\Rightarrow$ equality constraints in dual

Dual formulation

- decision variables $\{U_i^k\}, \quad \{W_{ij}\}$

$$\max \sum_{k \in K} d_k U_{t_k}^k$$

$$s.t. \, U_j^k - U_i^k \leq W_{ij} + r, \quad k \in K, (i,j) \in E$$

$$\sum_{(i,j) \in E} c_{ij} W_{ij} = 1,$$

$$W_{ij} \geq 0, U_{s_k}^k = 0$$

Properties of primal-dual solutions

- optimal solution to primal problem $\{\overline{X}_{ij}^{k}\}, \overline{\alpha}$
  dual problem $\{\overline{U}_{i}^{k}\}, \{\overline{W}_{ij}\}$
- if $\overline{X}_{ij}^{k} > 0,$ **then** $\overline{U}_{j}^{k} - \overline{U}_{i}^{k} = \overline{W}_{ij} + r$
- can think of $\overline{U}_{j}^{k}$ as shortest path distance from s_k to j when link weights are $\{\overline{W}_{ij} + r\}$

Therefore: solution to TE problem is also solution to shortest path problem with

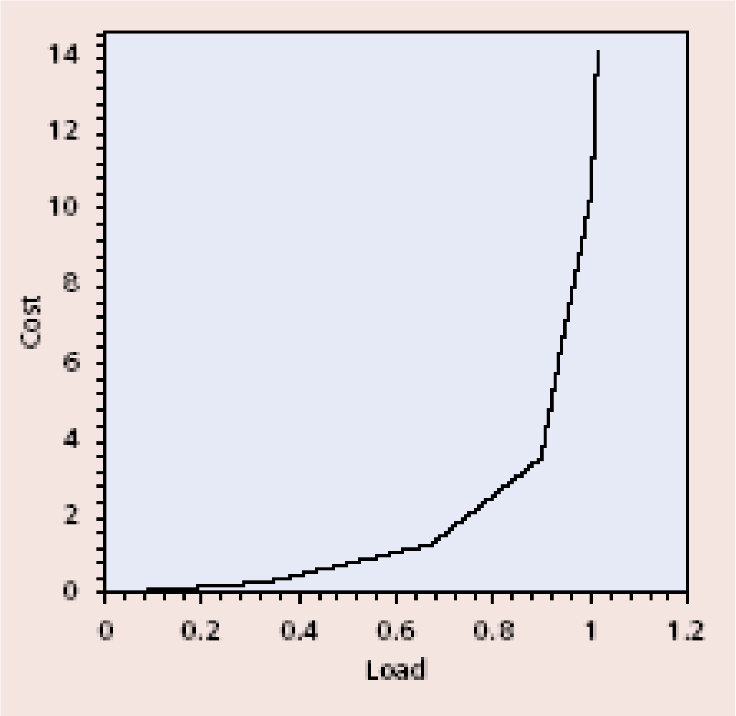
$$w_{ij} = \overline{W}_{ij} + r$$

Link weight assignment

- works for rich set of cost functions
- example:

$$\Phi = \sum_{(i,j)\in E} \Phi_{ij}\left(\sum_{k\in K} d_k X_{ij}^{k}\right)$$

- where Φ_{ij} are piecewise linear



Issues

- solutions are flow specific - need destination specific solutions
  - not a big deal, can reformulate to account for this
- solutions may not support equal split rule of OSPF
  - accounting for this yields NP-hard problem
    - see heuristics in FT paper
  - modify IP routing

One approach to overcome "splitting problem"

- current routing tables have thousands of routing prefixes
- instead of routing each prefix on all equal cost paths, selectively assign next hops to (each) prefix
  - i.e., remove some equal cost next hops assigned to prefixes
- goal: to approximate optimal link load

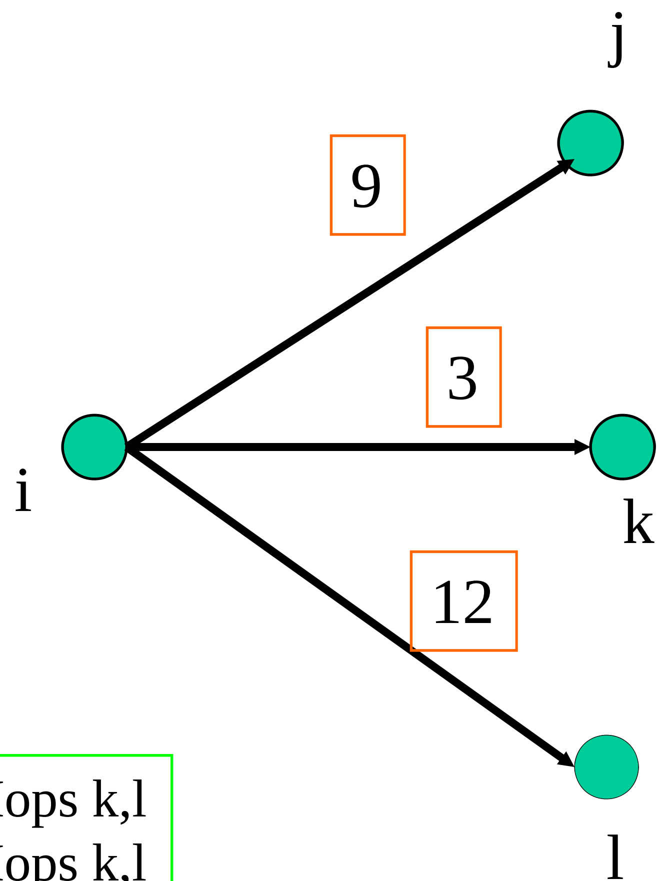
Example : EQUAL-SUBSET-SPLIT

Prefix A : 5

Prefix B : 1

Prefix C : 8

Prefix D : 10



Prefixes: C D

4 + 5 = 9

Prefixes: A B

2.5 + 0.5 = 3

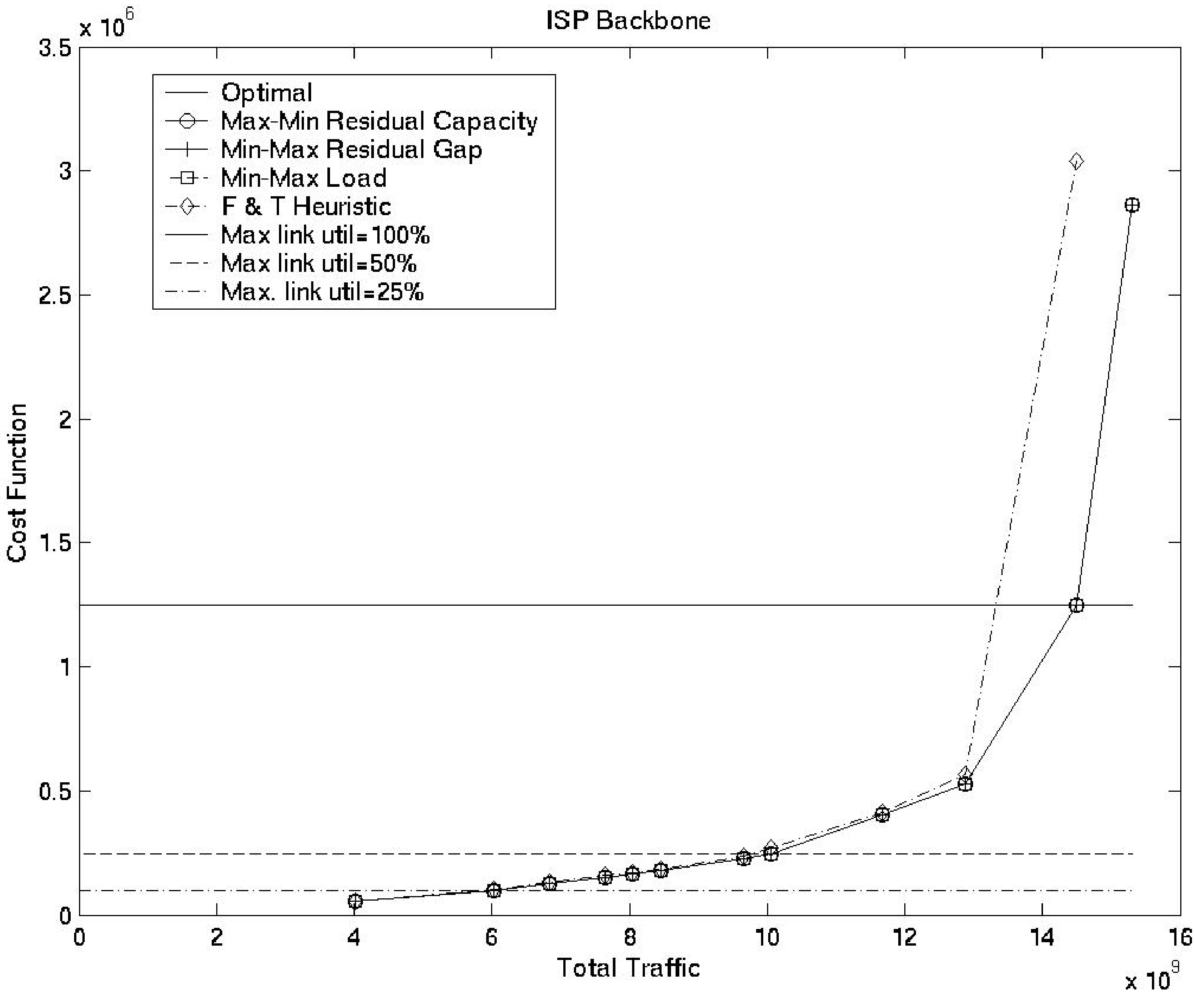
Prefixes: A B C D

2.5 + 0.5 + 4 + 5 = 12

Prefix A: Hops k,l
Prefix B : Hops k,l
Prefix C: Hops j,l
Prefix D: Hops j,l

Advantages

- requires no change in data path
- can leverage existing routing protocols
- current routers have 10,000s of routes in routing tables
  - provides large degree of flexibility in next hop allocation to match optimal allocation



ISP Backbone
x 10^6
3.5
3
2.5
2
1.5
1
0.5
0
Cost Function
Optimal
Max-Min Residual Capacity
Min-Max Residual Gap
Min-Max Load
F & T Heuristic
Max link util=100%
Max link util=50%
Max. link util=25%
0
2
4
6
8
10
12
14
16
Total Traffic
x 10^9

Summary

- can use OSPF/ISIS to support traffic engineering objectives
- performance objectives link weights
- equal splitting rule complicates problem
  - heuristics provide good performance
  - small changes to IP routing provide in better performance
- MPLS suffers none of these problems