Traffic Engineering within MPLS Information Distribution

Sources:
MPLS Forum
E. Osborne and A. Simha, *Traffic Engineering with MPLS*, Cisco Press

MPLS Traffic Engineering – Information Distribution

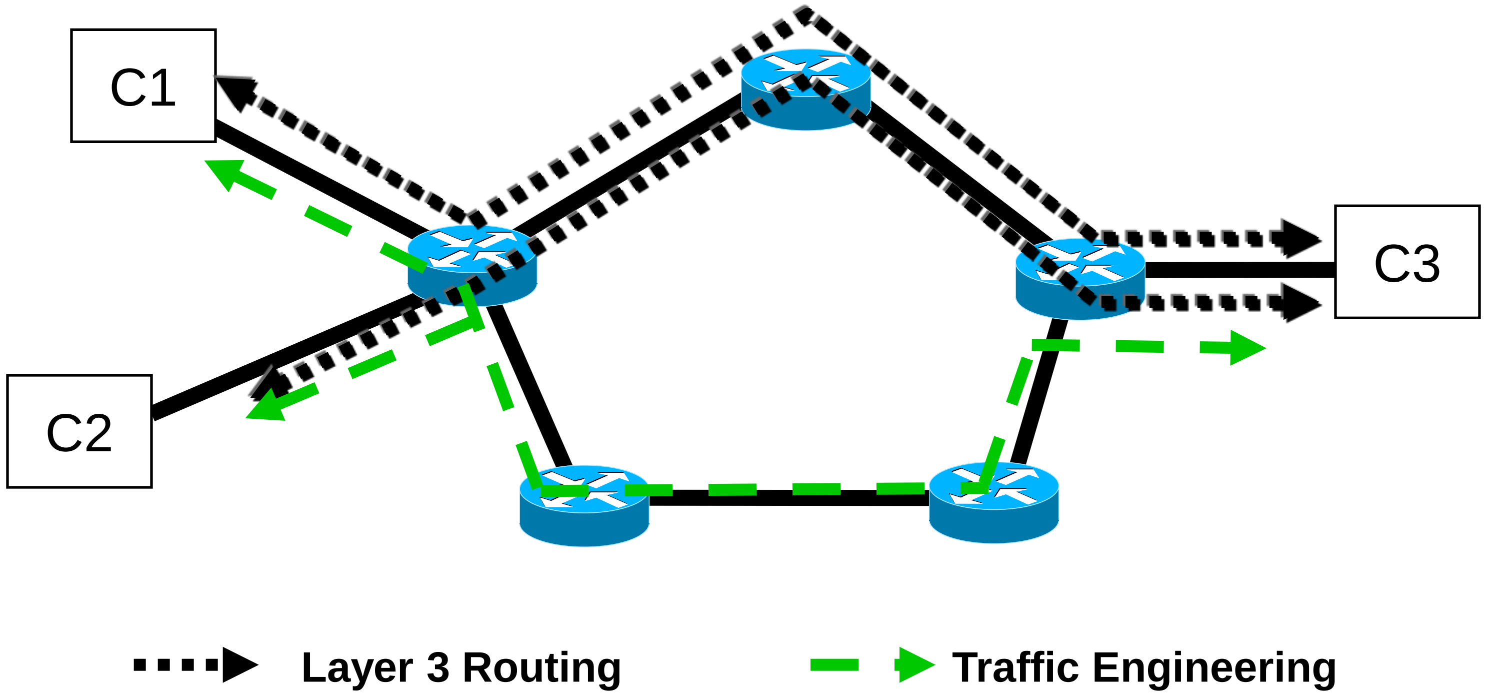
- Value added services enabled by MPLS Traffic Engineering
  - **Constraint-based routing**
  - QoS
  - Fast reroute
  - VPNs
  - …
- Need more information about constraint(s) than just network topology
  - Bandwidth, delay, etc.
- **What's involved in information distribution to support TE?**

Review Terminology...

- Network Engineering
  - ✓ "Put the ***bandwidth*** where the ***traffic*** is"
    - Physical cable deployment
    - Virtual connection provisioning

- Traffic Engineering
  - ✓ "Put the ***traffic*** where the ***bandwidth*** is"
    - Local or global control
    - On-line or off-line optimization of routes
    - Implies the ability to "**explicitly**" route traffic

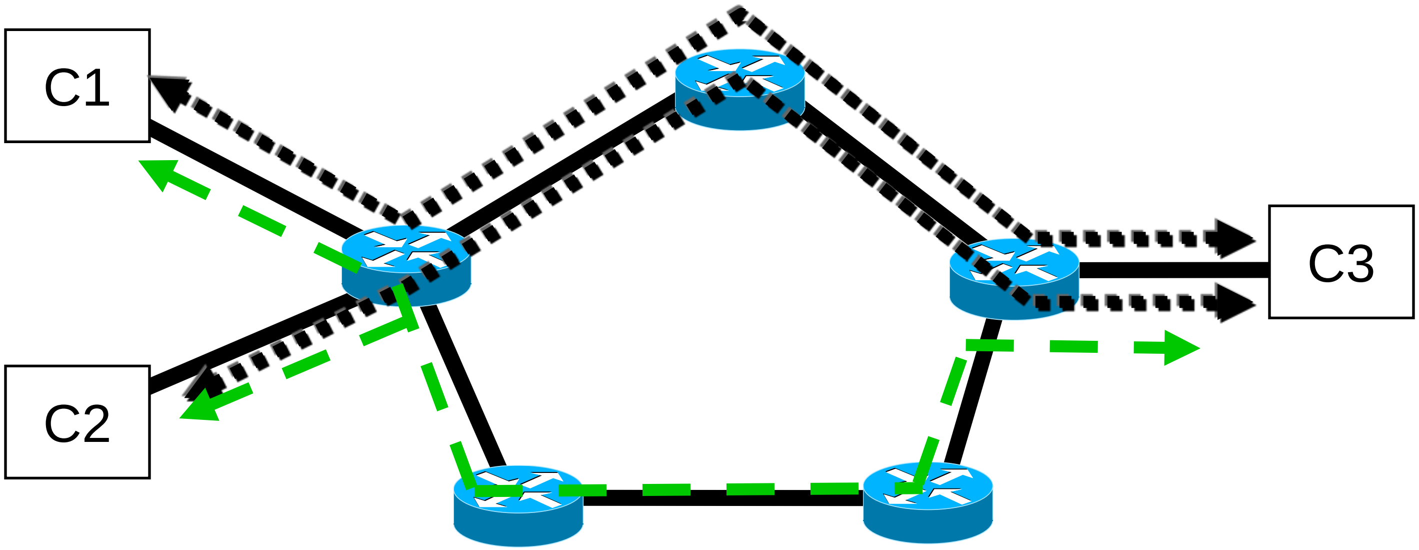
Traditional Traffic Engineering



- Move traffic from IGP path to less congested path

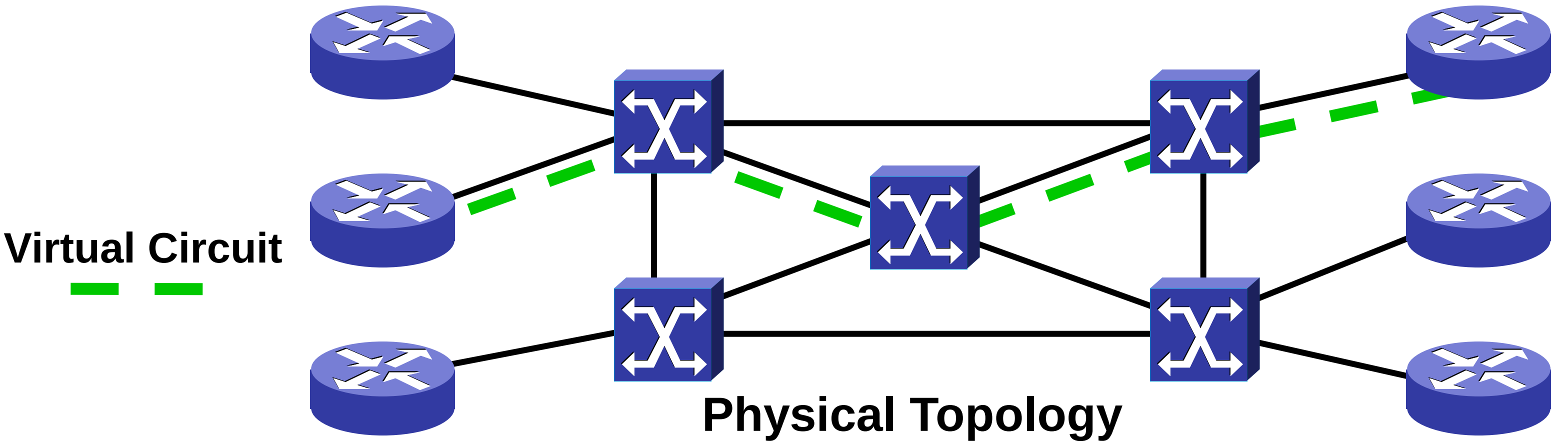
Traditional Traffic Engineering

Limitations



- TE Mechanisms
  - ✓ Over-provisioning
  - ✓ Metric manipulation
- Limitations
  - ✓ Some links become underutilized or overutilized
  - ✓ Trial-and-error approach

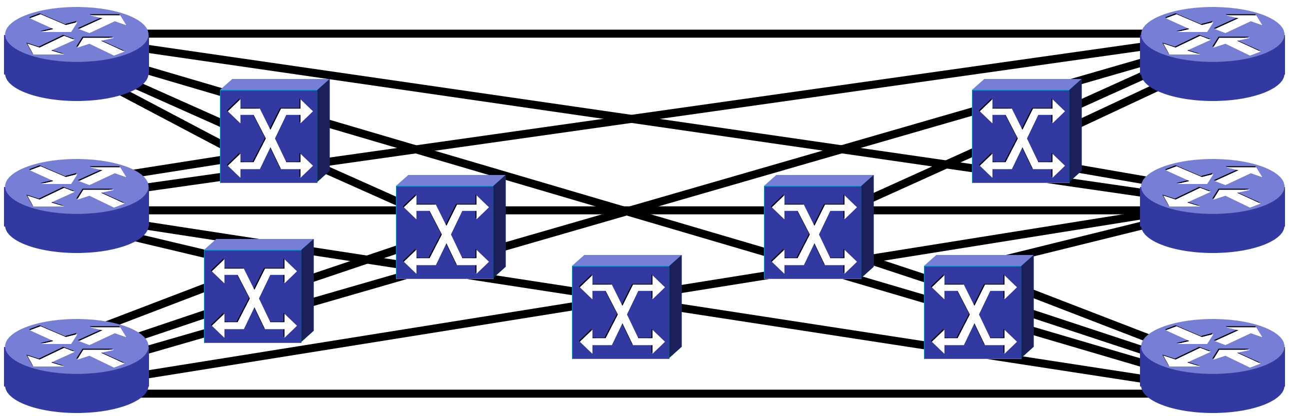
Traffic Engineering with ATM Core



- Infrastructure
  - ✓ Routed edge over ATM switched core
    - ➢ In 1990s, core routers were not fast enough
  - ✓ Introduced full Traffic Engineering (TE) ability
    - ➢ Dense VCs calculated offline
    - ➢ Overlay expensive and complex

Traffic Engineering with ATM Core
Limitations

Logical Topology



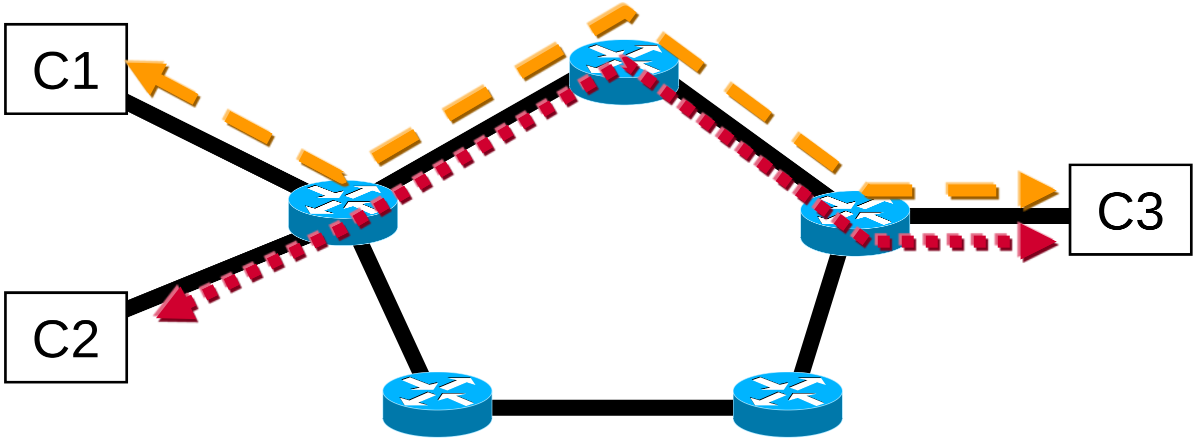
- TE Mechanisms
  - ✓ VC routing
  - ✓ Overlay network
- Benefits
  - ✓ Full traffic control
  - ✓ Per-circuit statistics
- Limitations
  - ✓ Overlay of IP and ATM
  - ✓ “N-squared” VCs
  - ✓ IGP Stress
  - ✓ Cell tax

MPLS Traffic Engineering

- Traditional TE controls traffic flows in a network
  - “The ability to move traffic away from the shortest path calculated by the IGP to a less congested path”
- MPLS Traffic Engineering
  - Allows **Explicit Routing** and set-up of LSP’s
  - Provides control over how LSP’s are recovered in the event of a failure
  - Enables Value Added Services
    - Virtual Private Networks – VPNs
    - Service Level Agreements - SLAs
    - Multi-media over IP solutions – MMoIP, VoIP
    - ATM over IP – easy and cheap for existing legacy networks

But, this is a simple example . . .

- Routing Protocols Create a "Shortest Path“ Route
- LSPs follow the "shortest path"



This mechanism does NOT give us Traffic Engineering

MPLS Traffic Engineering
Requires 3 main areas of extensions

- Enhancements to the *Routing Protocols*: ***Information Distribution***
  - ✓ OSPF ➔ OSPF-TE
  - ✓ ISIS ➔ ISIS-TE
- Enhancements to SPF to consider constraints: *Constraint-Based Routing (CSPF)*: ***Path Calculation***
  - ✓ Explicit route selection
  - ✓ Bandwidth parameters and recovery mechanisms defined
  - ✓ Connection Admission Controls (CAC) enforced
    - (policing, marking, metering, scheduling, queuing, etc)
- Enhancements to the *Signalling Protocols* to support explicit constraint-based routing: ***Path Creation***
  - ✓ LDP ➔ CR-LDP
  - ✓ RSVP ➔ RSVP-TE

What's involved in information distribution to support TE?

- Information distribution is broken down into three pieces:
  - ✓ ***What*** information is distributed and how to configure it
  - ✓ ***When*** information is distributed and how to control when flooding takes place
  - ✓ ***How*** information is distributed (protocol-specific details)

What information Is Distributed?

- The idea behind MPLS TE is to allow routers to build paths using information rather than the shortest IP path. But what information is distributed to allow the routers to make more *intelligent path calculations*?
  - ✓ Examples:
    - ➢ Path that has enough bandwidth, special attributes, low delay, …
    - ➢ Generally, information that has to do with ***TE objectives/requirements***
- MPLS TE works by using OSPF or IS-IS to distribute information about ***available resources***. Three main pieces of information are distributed *for* ***each link/interface***:
  - ✓ **Available bandwidth** information, broken down by **priority** to allow tunnels to preempt others
  - ✓ **Attribute flags**
  - ✓ **Administrative weight**

Available Bandwidth Information

- A key feature of MPLS TE is the capability to **reserve bandwidth** across the network
  - ✓ Every router needs to know available bandwidth for each interface
- How much bandwidth to allocate to the interface?
  - ✓ Also depends on oversubscription policies and the policy to enforce them
  - ✓ Reference: Cisco default is *75%* of the link bandwidth
- Main elements: interface, allocated, max, percentage. Example:
  - ✓ P04/2 233250K 466500K 50
- Need to keep track of currently allocated bandwidth to obtain currently **available or reservable bandwidth**
- Need both the ***per-interface*** and the ***per-tunnel*** *(TE LSP)* bandwidth
  - ✓ Why both?

Tunnel Priority

- Other information
- Some LSPs or tunnels are more important than others. For example, tunnels for voice traffic.
- Need capability to allow tunnels to **preempt** others.
  - ✓ **Each tunnels has a *priority***
  - ✓ Lower-priority tunnels are pushed out and are made to recalculate a path, and the resources are given to the higher-priority tunnel
  - ✓ 8 priority levels (0-7): lower value, higher priority
  - ✓ Destructive to other tunnels, use only necessary
  - ✓ In a real network, the preempted tunnel can have an alternative path for backup and the tunnel will come up
  - ✓ Example

Setup and Holding Priority

- Each tunnel actually has two priorities – a **Setup priority** and a **Hold priority** (RFC 3209)
- Setup priority to decide whether to admit the tunnel, Hold priority to compare priority if competition comes along for a new tunnel
  - ✓ Usually treated the same, but can be different
  - ✓ Application example: once the tunnel is setup, the Hold priority could be set to the highest, which means that it cannot be preempted by any other tunnels.
  - ✓ ***Hold priority must be >= Setup priority***, why?

Attribute Flags

- MPLS TE allows you to enable attribute flags.
- An attribute flag is a 32-bit bitmap on a link that can indicate the existence of up to 32 separate **properties** of that link.
  - ✓ ISPs have the freedom to manage these bits
  - ✓ Example:
    - ➢ Assuming 8-bit and a link that has attribute flags of 0x1 (0000 0001) means that the link is a satellite link.
    - ➢ If you want to build a tunnel that does not cross a satellite link, you need to make sure that any link the tunnel crosses has the satellite link bit set to 0
  - ✓ Need a mask

Administrative Weight or Metric

- For MPLS TE, two costs are associated with a link – the **TE cost** and the **IGP cost**.
  - ✓ Allow to present the TE path calculation with a different set of link costs than the regular IGP SPF sees.
  - ✓ Can change the cost advertised for the link, but only for traffic engineering. Why?
  - ✓ Useful in path calculation. Examples:
    - ➢ Networks that have both IP and MPLS TE traffic
    - ➢ Delay-sensitive link. Example: OC-3 land line and OC-3 satellite link have different delays, but with the same bandwidth.

When Information Is Distributed?

- IGP floods information about a link in three cases:
  - ✓ When a link goes up or down
  - ✓ When a link's configuration is changed (e.g., link cost)
  - ✓ When it's time to periodically flood the IGP information
- For MPLS TE, there is more to consider:
  - ✓ When link available bandwidth changes significantly
  - ✓ Link attribute(s) changed
- What is "significant"?

What is Significant?

- How to define significant?
  - ✓ Percentage of link bandwidth
  - ✓ Is it enough?
  - ✓ Rules are different for every network, situation, and link
  - ✓ Cisco uses default flooding **thresholds** (15, 30, 45, 60, 75, 80, 85, 90, 95, 96, 97, 98, 99, 100) on links. If the thresholds are crossed, link bandwidth is flooded.
- Flood insignificant changes periodically
  - ✓ If available bandwidth has changed and it hasn't been flooded, the changes will be flooded every 3 minutes (default value, but configurable), more frequently than IGP refresh interval
- If error, flood immediately
  - ✓ A path setup fails due to lack of bandwidth. Available bandwidth has been changed since the last time flooding occurred.
- Should be considered in TE methods.

How Information Is Distributed?

- MPLS TE in OSPF, hence OSPF-TE
- MPLS TE in IS-IS, hence ISIS- TE
- MPLS TE enhancements and IP-Extended TLVs are closely related.

0 1 2 3

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

Type (16 bits)	Length (16 bits)
Value (32 bits)	

  - ✓ Type 1: router address TLV: MPLS TE router ID
  - ✓ Type 2: link TLV: 9 sub-TLVs
    - Link type, link ID, local I/F IP addr, remote I/F IP addr, ***TE metric (cost, admin-weight)****, max link bw,* ***max reservable bw, unreserved bw (per priority), attribute flags***.

- Before you can do MPLS TE, support for wide metrics must be enabled.

OSPF TE Metric Extensions – Evolving Standards

- Last updated **Jan 9th, 2015**
- In certain networks, such as, but not limited to, financial information networks (e.g. stock market data providers), network performance criteria (e.g. **latency**) are becoming as critical to **data path selection** as other metrics. This document describes extensions to OSPF TE v.2 [RFC 3630] and v.3 [RFT 5329] to enable **network performance information** to be distributed in a scalable fashion. The information distributed using OSPF TE Metric Extensions can then be used to make path selection decisions based on network performance.
- This sub-TLV advertises the average **link delay (in micro-seconds, 24 bits)** between two directly connected OSPF neighbors. Example of TLV for delay:
-

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            TBD1               |               4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|A| RESERVED |                  Delay                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

OSPF Extension for MPLS Green TE

- **OSPF Extensions for MPLS Green Traffic Eng.**
  - ✓ draft-li-ospf-ext-green-te-01
  - ✓ Last updated Oct 14, 2013
  - ✓ The new TLV will be named as "Energy consumption on of Link TLV" (in Watts)
  - ✓ Used to calculate the path with the lowest energy consumption

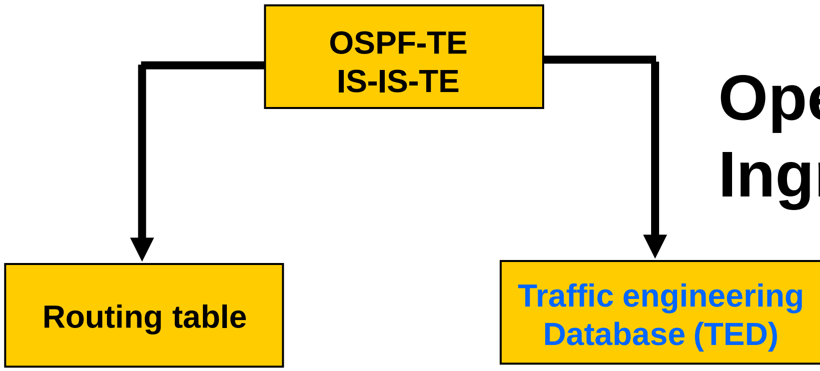
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type(TBD)            |          Length = 4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Energy consumption of link                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Constraint-Based Routing

Constraint-Based Routing

- Parameters over and above "best effort" are constraints
  - ✓ <u>Constraint</u> = order in which LSRs are reached
  - ✓ <u>Constraint</u> = description of traffic flow, bandwidth, delay, class, priority
  - ✓ <u>Constraint</u> = edge traffic conditioning functions such as marking, metering, policing, and shaping
  - ✓ <u>Constraint</u> = Recovery mechanism for "protection" of a working LSP
- Supports and enables QoS/CoS functions for;
  - ✓ IP DiffServ and IntServ

Constraint Route Signaling
Operational Model



Operations Performed by the Ingress LSR

1) Store information from IGP flooding

2) Store traffic engineering information

<u>OSPF and IS-IS - **TE** Extensions</u>

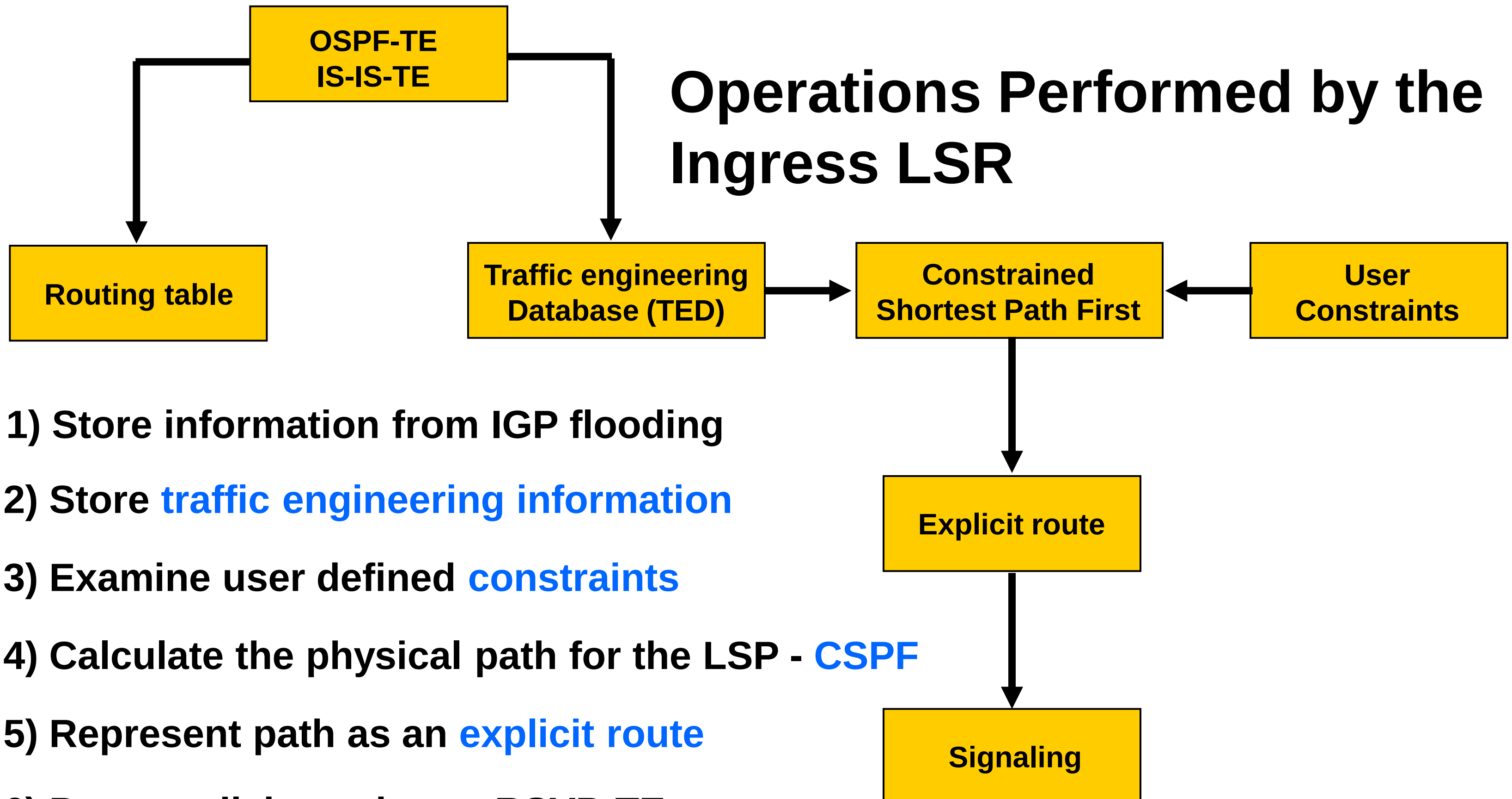
Distributed (piggybacked) on Opaque Link State Advertisements

Encoded as new Type Length Values (TLVs)

Metrics: Bandwidth, *Unreserved Bandwidth, Available Bandwidth, Delay, Delay-Jitter, Loss Probability, Administrative Weight, Economic Cost*

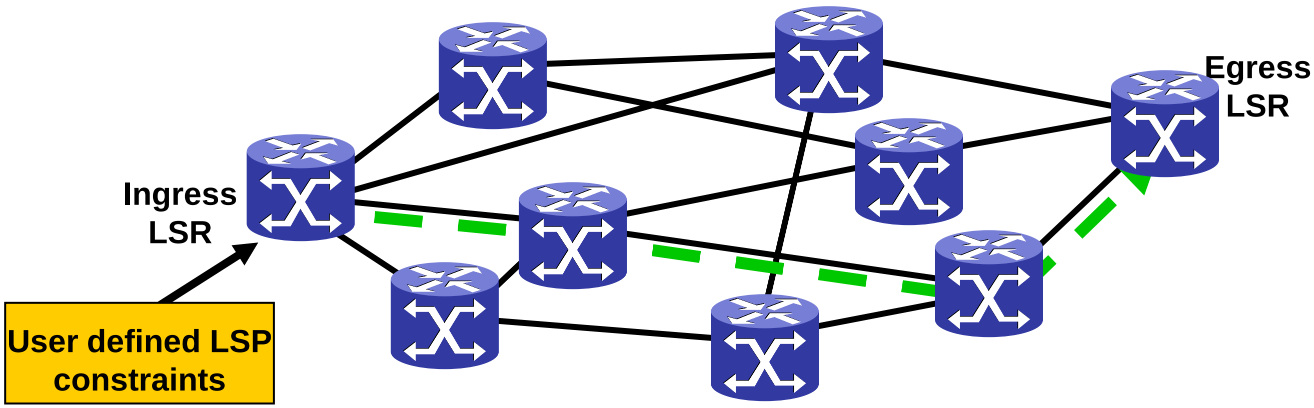
Constraint Route Signaling
Operational Model

Operations Performed by the Ingress LSR



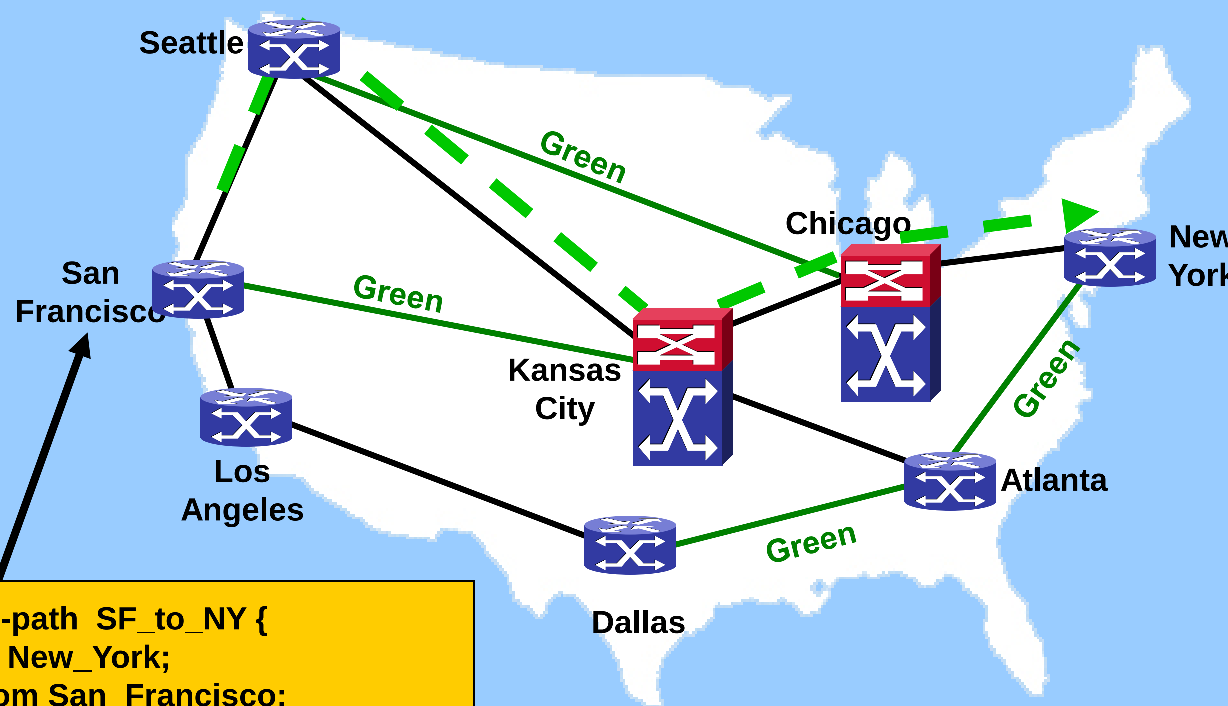
1) Store information from IGP flooding
2) Store traffic engineering information
3) Examine user defined constraints
4) Calculate the physical path for the LSP - CSPF
5) Represent path as an explicit route
6) Pass explicit routing to RSVP-TE or CR-LDP for signaling

Constraint Route Signaling



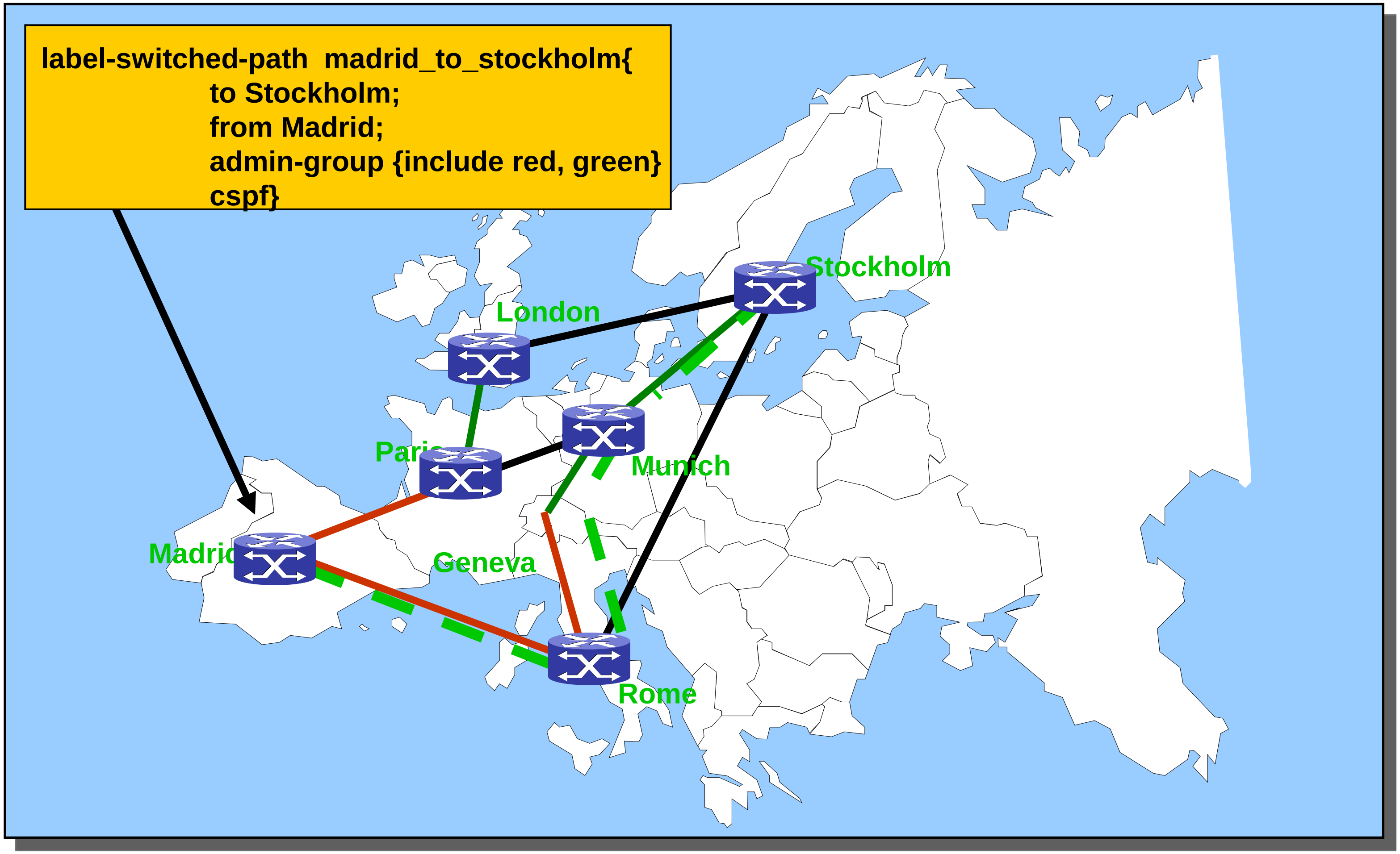
- Operator configures LSP constraints at ingress LSR
  - ✓ Bandwidth reservation
  - ✓ Include or exclude a specific link(s)
  - ✓ Include specific node traversal(s)
- Network actively participates in selecting an LSP path that meets the constraints

Constraint Route Signaling
Example




```
label-switched-path  SF_to_NY {
            to New_York;
            from San_Francisco;
            admin-group {exclude green}
            cspf}
```

Constraint Route Signaling
Example



Signaling Mechanisms

• LDP	Label Distribution Protocol
• CR-LDP	Constraint-Based Routing - Label Distribution Protocol
• RSVP-TE	Extensions to RSVP for Traffic Engineering
• BGP-4	Carrying Label Information in BGP- 4

Constraint-Based Routing

- CBR could be very challenging and complicated.
  - ✓ Example: need to deliver 60 bricks with only one bike.
  - ✓ Solution?

  - ✓ If > 1 constraint:
    NP-complete

The Challenge: A Practical Example

To support Green TE,

- Which one will be considered first?
  - ✓ Traditional TE metrics
  - ✓ Energy consumption

- How to calculate TE path?
  - ✓ Distributed
  - ✓ Centralized: e.g., SDN