



NETWORKERS 2004

DEPLOYING MPLS TRAFFIC ENGINEERING

SESSION RST-2603

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

1

Some Assumptions

Cisco.com

- You understand basic IP routing
- You understand MPLS concepts and operation
- You understand how a link-state protocol works
- Some knowledge of QoS is useful
- You will still be awake at the end of this

RST-2603
9866_05_2004_c2

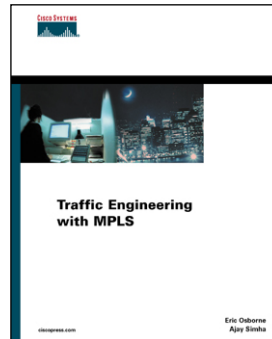
© 2004 Cisco Systems, Inc. All rights reserved.

2

A Blatant Plug

Cisco.com

- **Traffic Engineering with MPLS**
ISBN: 1-58705-031-5
- **Now available in Portuguese and Chinese!**



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

3

Agenda

Cisco.com

- **Traffic Engineering Overview**
- **Traffic Engineering Theory**
- **Configuration**
- **Protection**
- **Diffserv Traffic Engineering (DS-TE)**
- **Design and Scalability**
- **MPLS-VPN, Multicast and TE**
- **Summary**

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

4

TRAFFIC ENGINEERING OVERVIEW



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

5

Network vs. Traffic Engineering

Cisco.com

- **Network engineering**
Build your network to carry your predicted traffic
- **Traffic engineering**
Manipulate your traffic to fit your network
- **Traffic patterns are impossible to accurately predict**
- **Symmetric bandwidths/topologies, asymmetric load**
- **TE can be done with IGP costs, ATM/FR, or MPLS**

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

6

Motivation for Traffic Engineering

Cisco.com

- **Increase efficiency of bandwidth resources**
Prevent over-utilized (congested) links whilst other links are under-utilized
- **Ensure the most desirable/appropriate path for some/all traffic**
Override the shortest path selected by the IGP
- **Replace ATM/FR cores**
PVC-like traffic placement without IGP full mesh and associated $O(N^2)$ flooding
- **The ultimate goal is COST SAVING**
Service development also progressing

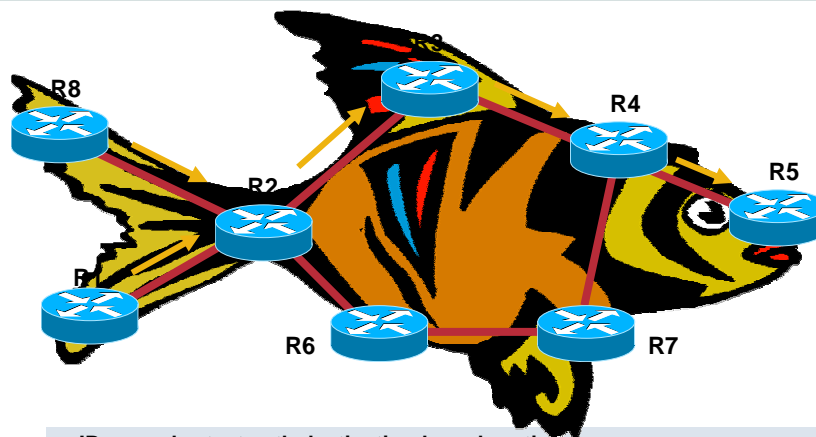
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

7

The “Fish” Problem (Shortest Path)

Cisco.com



- IP uses shortest path destination-based routing
- Shortest path may not be the only path
- Alternate paths may be under-utilized
- Whilst the shortest path is over-utilized

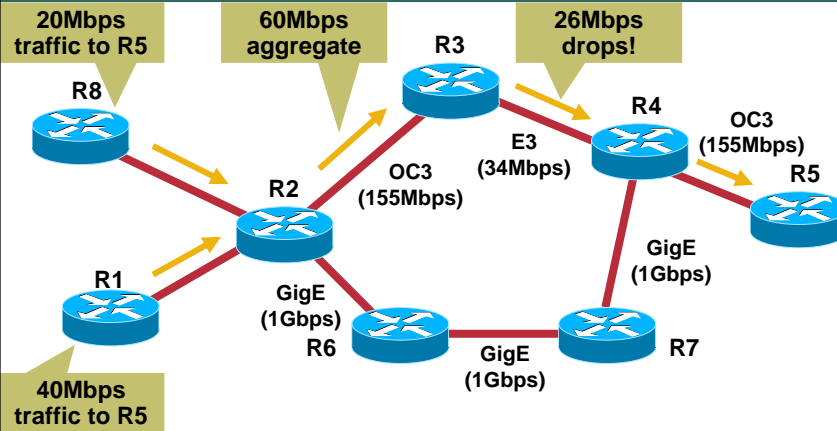
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

8

Shortest Path and Congestion

Cisco.com



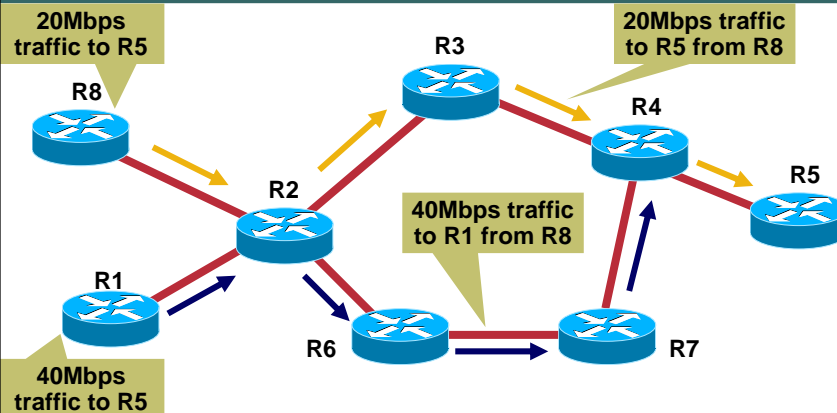
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

9

The TE Solution

Cisco.com



- MPLS Labels can be used to engineer explicit paths

- Tunnels are **UNI-DIRECTIONAL**

Normal path: R8 → R2 → R3 → R4 → R5

Tunnel path: R1 → R2 → R6 → R7 → R4

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

10

Terminology

Cisco.com

- **Constrained-Based Shortest Path First (CSPF)**

MPLS-TE uses CSPF to create a shortest path based on a series of constraints:

Bandwidth

Affinity/link attributes

...or an explicitly configured path

- **Tunnels are UNI-DIRECTIONAL!**



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

11

TRAFFIC ENGINEERING THEORY



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

12

Traffic Engineering Components

Cisco.com

- Information distribution
- Path selection/calculation
- Path setup
- Trunk admission control
- Forwarding traffic on to tunnel
- Path maintenance

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

13

Information Distribution

Cisco.com

- Need to flood TE information (Resource Attributes) across the network
 - Available bandwidth per priority level, a few other things
- IGP extensions flood this information
 - OSPF uses Type 10 (area-local) Opaque LSAs
 - ISIS uses new TLVs
- Basic IGP: {self, neighbors, cost to neighbors}
- TE extensions: {self, neighbors, cost to neighbors, available bandwidth to neighbors}
- TE bandwidth is a control-plane number only

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

14

Path Calculation

Cisco.com

- Once available bandwidth information and attributes are flooded, router may calculate a path from head to tail
 - Path may be explicitly configured by operator
- TE Headend does a “Constrained SPF” (CSPF) calculation to find the best path
- CSPF is just like regular IGP SPF, except
 - Takes required bandwidth and attributes into account
- Looks for best path from a head to a single tail (unlike OSPF)
- Minimal impact on CPU utilization using CSPF
- Path can also be explicitly configured

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

15

Path Setup

Cisco.com

- Once the path is calculated, it must be signaled across the network
 - Reserve any bandwidth to avoid “double booking” from other TE reservations
 - Priority can be used to pre-empt low priority existing tunnels
- RSVP used to set up TE LSP
 - PATH messages (from head to tail) carries LABEL_REQUEST
 - RESV messages (from tail to head) carries LABEL
- When RESV reaches headend, tunnel interface = UP
- RSVP messages exist for LSP teardown and error sig

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

16

Trunk Admission Control

Cisco.com

- **On receipt of PATH message**
 - Router will check there is bandwidth available to honour the reservation
 - If bandwidth available then RSVP accepted
- **On receipt of a RESV message**
 - Router actually reserves the bandwidth for the TE LSP
 - If pre-emption is required lower priority LSP are torn down
- **OSPF/ISIS updates are triggered**

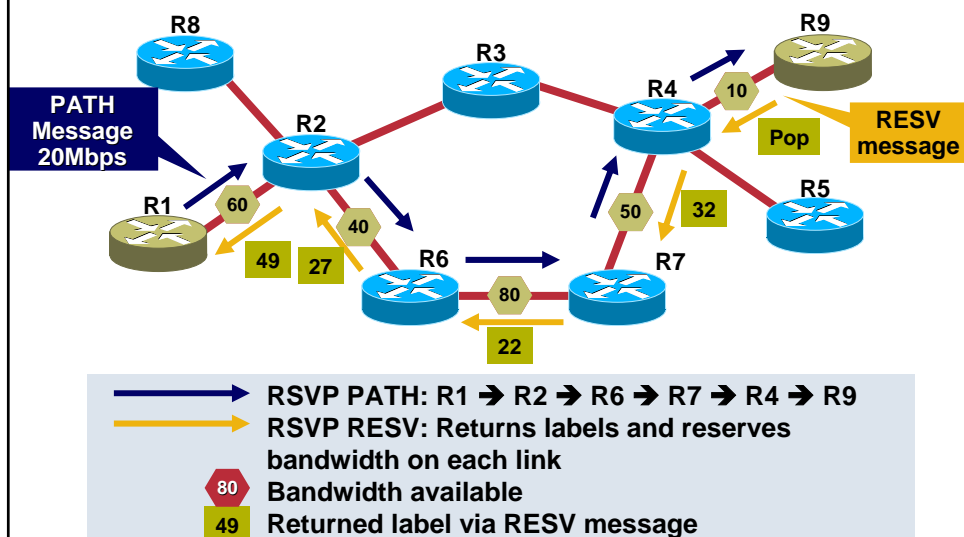
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

17

Path Setup Example

Cisco.com



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

18

Forwarding Traffic to a Tunnel

Cisco.com

- Static routing
- Policy routing
 - Global table only—not from VRF at present
- Autoroute
- Forwarding Adjacency

Static, autoroute, and forwarding adjacency get you unequal-cost load-balancing

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

19

Autoroute

Cisco.com

- Used to include TE LSP in SPF calculations
- IGP adjacency is **NOT** run over the tunnel!
- Tunnel is treated as a directly connected link to the tail
 - When tunnel tail is seen in PATH list during IGP SPF, replace outgoing physical interface with tunnel interface
 - Inherit tunnel to all downstream neighbors of said tail

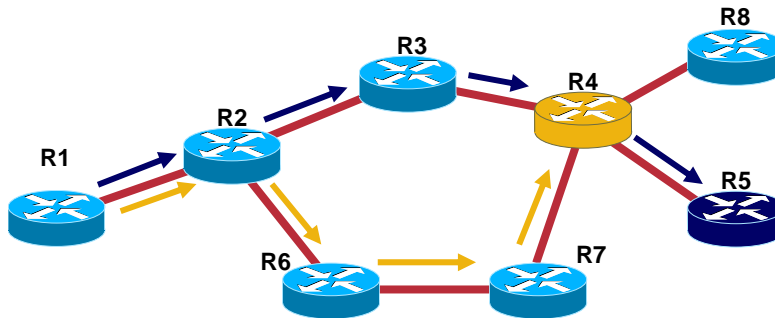
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

20

Autoroute Topology (OSPF and ISIS)

Cisco.com



Tunnel1: R1 → R2 → R3 → R4 → R5
 Tunnel2: R1 → R6 → R7 → R4

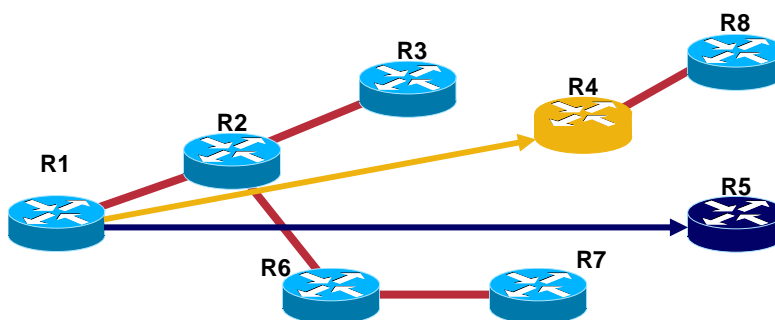
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

21

Autoroute Topology (OSPF and ISIS)

Cisco.com



From R1 router perspective:
 Next hop to R4 and R8 is Tunnel1
 Next hop to R5 is Tunnel2
 All nodes behind tunnel routed via tunnel

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

22

Forwarding Adjacency

Cisco.com

- Autoroute does not advertise the LSP into the IGP
- There may be a requirement to advertise the existence of TE tunnels to upstream routers
 - Like an ATM/FR PVC—attract traffic to a router regardless of the cost of the underlying physical network cost
- Useful as a drop-in replacement for ATM/FR (and during migration)
- Can get suboptimal forwarding (**NOT** loops) if you're not careful

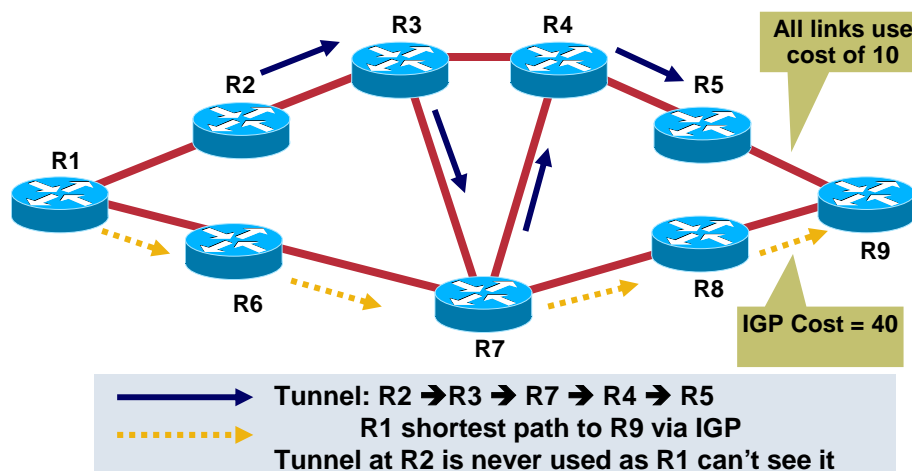
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

23

Forwarding Adjacency

Cisco.com



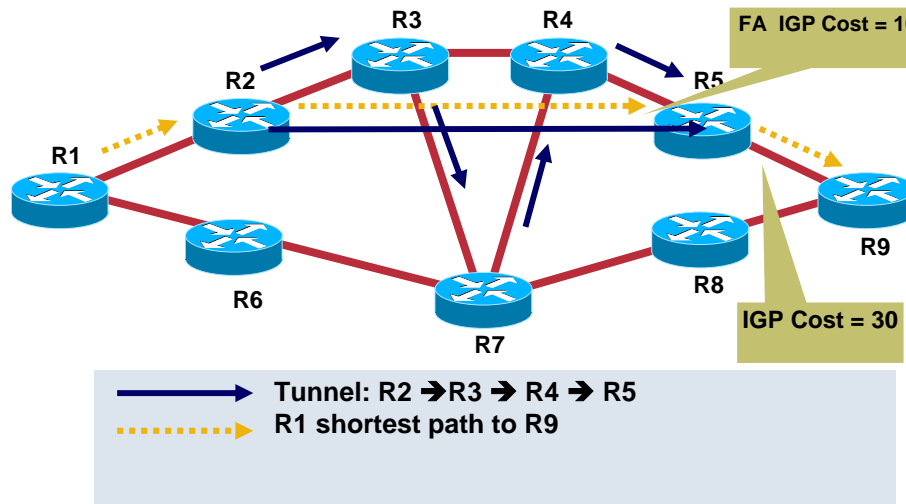
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

24

Advertise TE Links into IGP

Cisco.com



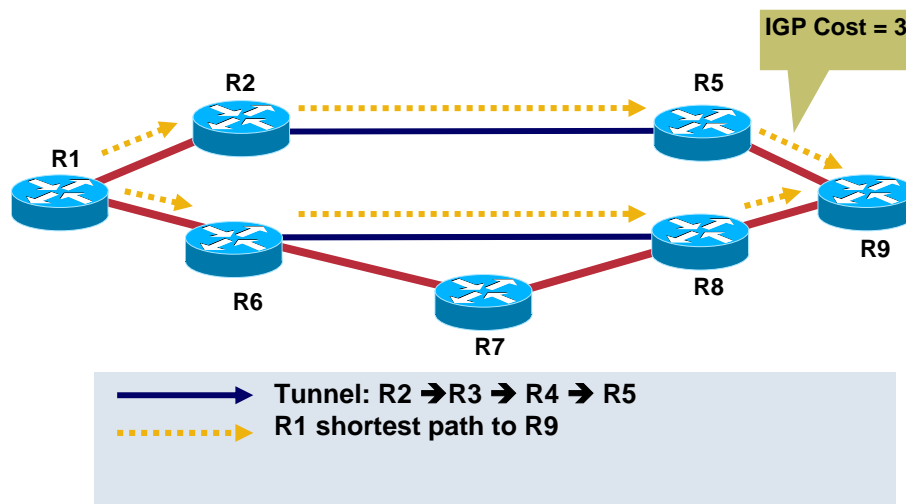
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

25

Load Balancing Across FA

Cisco.com



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

26

Unequal Cost Load Balancing

Cisco.com

- IP routing has equal-cost load balancing, but not unequal cost*
- Unequal cost load balancing difficult to do while guaranteeing a loop-free topology
- Since MPLS doesn't forward based on IP header, permanent routing loops don't happen
- 16 hash buckets for next-hop, shared in **rough** proportion to configured tunnel bandwidth or load-share value

*EIGRP Has 'Variance', but That's Not as Flexible

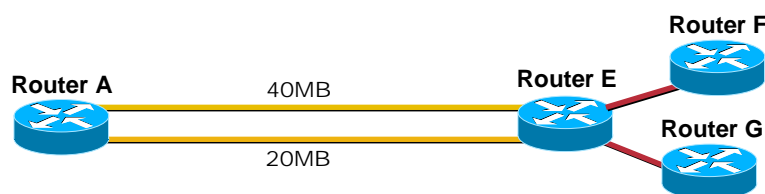
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

27

Unequal Cost: Example 1

Cisco.com



```
gsrl#show ip route 192.168.1.8
Routing entry for 192.168.1.8/32
  Known via "isis", distance 115, metric 83, type level-2
  Redistributing via isis
  Last update from 192.168.1.8 on Tunnel0, 00:00:21 ago
  Routing Descriptor Blocks:
    * 192.168.1.8, from 192.168.1.8, via Tunnel0
      Route metric is 83, traffic share count is 2
    192.168.1.8, from 192.168.1.8, via Tunnel1
      Route metric is 83, traffic share count is 1
```

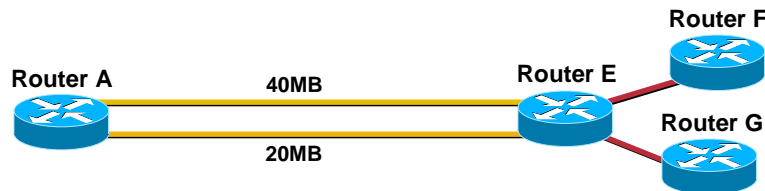
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

28

Unequal Cost: Example 1

Cisco.com



```
gsr1#sh ip cef 192.168.1.8 internal
```

```
.....
Load distribution: 0 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 (refcount 1)
```

Hash	OK	Interface	Address	Packets	Tags imposed
1	Y	Tunnel0	point2point	0	{23}
2	Y	Tunnel1	point2point	0	{34}

```
.....
```

**Note That the Load Distribution
Is 11:5—Very Close to 2:1, but Not Quite!**

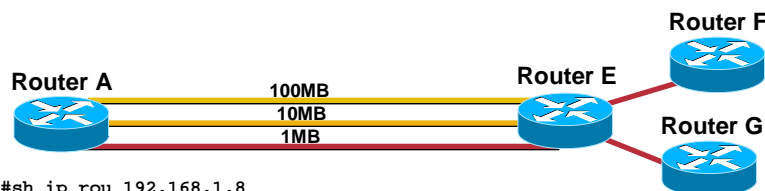
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

29

Unequal Cost: Example 2

Cisco.com



```
gsr1#sh ip rou 192.168.1.8
```

```
Routing entry for 192.168.1.8/32
```

```
Known via "isis", distance 115, metric 83, type level-2
```

```
Redistributing via isis
```

```
Last update from 192.168.1.8 on Tunnel2, 00:00:08 ago
```

```
Routing Descriptor Blocks:
```

```
* 192.168.1.8, from 192.168.1.8, via Tunnel0
  Route metric is 83, traffic share count is 100
  192.168.1.8, from 192.168.1.8, via Tunnel1
  Route metric is 83, traffic share count is 10
  192.168.1.8, from 192.168.1.8, via Tunnel2
  Route metric is 83, traffic share count is 1
```

Q: How Does 100:10:1 Fit Into a 16-Deep Hash?

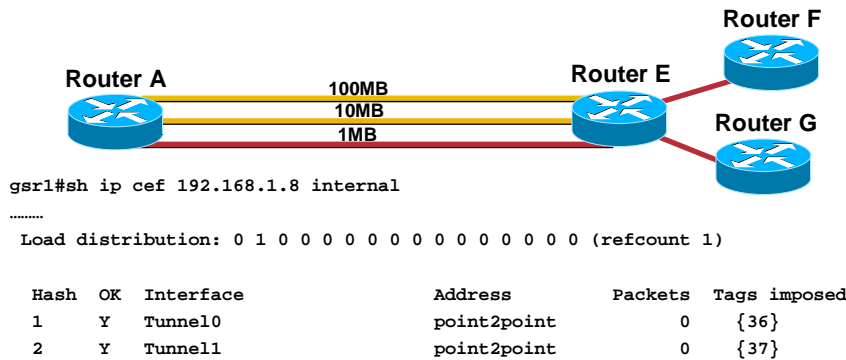
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

30

Unequal Cost: Example 2

Cisco.com



A: Any Way It Wants to! 15:1, 14:2, 13:2:1, it depends on the order the tunnels come up
Deployment Guideline: Don't use tunnel metrics that don't reduce to 16 buckets!

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

31

Path Maintenance

Cisco.com

- **Steady-state information load is low**
Especially with refresh reduction (RFC2961)
- **Path re-optimization**
Process where some traffic trunks are rerouted to new paths so as to improve the overall efficiency in bandwidth utilization
For example, traffic may be moved to secondary path during failure; when primary path is restored traffic moved back
- **Path restoration**
Comprised of two techniques; local protection (link and node) and path protection
Discussed later in protection section

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

32

CONFIGURATION



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

33

Prerequisite Configuration (Global)

Cisco.com

```
ip cef [distributed]  
mpls traffic-eng tunnels
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

34

Information Distribution

Cisco.com

- OSPF

```
mpls traffic-eng tunnels
mpls traffic-eng router-id loopback0
mpls traffic-eng area ospf-area
```

- ISIS

```
mpls traffic-eng tunnels
mpls traffic-eng router-id loopback0
mpls traffic-eng level-x
metric-style wide
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

35

Information Distribution

Cisco.com

- On each physical interface

```
interface pos0/0
  mpls traffic-eng tunnels
  ip rsvp bandwidth Kbps (Optional)
  mpls traffic-eng attribute-flags attributes (Opt)
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

36

Build a Tunnel Interface (Headend)

Cisco.com

```
interface Tunnel0
  ip unnumbered loopback0
  tunnel destination RID-of-tail
  tunnel mode mpls traffic-eng
  tunnel mpls traffic-eng bandwidth 10
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

37

Tunnel Attributes

Cisco.com

```
interface Tunnel0
  tunnel mpls traffic-eng bandwidth Kbps
  tunnel mpls traffic-eng priority pri [hold-pri]
  tunnel mpls traffic-eng affinity properties [mask]
  tunnel mpls traffic-eng autoroute announce
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

38

Path Calculation

Cisco.com

- Dynamic path calculation

```
int Tunnel0
  tunnel mpls traffic-eng path-option # dynamic
```

- Explicit path calculation

```
int Tunnel0
  tunnel mpls traffic path-opt # explicit name foo

ip explicit-path name foo
  next-address 1.2.3.4 [loose]
  next-address 1.2.3.8 [loose]
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

39

Multiple Path Calculations

Cisco.com

- A tunnel interface can have several path options, to be tried successively

```
tunnel mpls traffic-eng path-option 10 explicit name foo
tunnel mpls traffic-eng path-option 20 explicit name bar
tunnel mpls traffic-eng path-option 30 dynamic
```

- Path-options can each have their own bandwidth

```
tunnel mpls traffic-eng path-option 10 explicit name foo
  bandwidth 100
tunnel mpls traffic-eng path-option 20 explicit name bar
  bandwidth 50
tunnel mpls traffic-eng path-option 30 dynamic
  bandwidth 0
```

9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

40

LSP Attributes

Cisco.com

Configure on Tunnel:

```
tunnel mpls traffic-eng path-  
option 10 dynamic attributes  
foo
```

Attribute list 'foo' is defined at:

```
mpls traffic-eng lsp  
attributes foo  
  
bandwidth 25  
priority 2 2
```

• Attribute list options

- affinity
- auto-bw
- bandwidth
- lockdown
- priority
- protection
- record-route

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

41

Static and Policy Routing Down a Tunnel

Cisco.com

• Static routing

```
ip route prefix mask Tunnel0
```

• Policy routing (Global Table)

```
access-list 101 permit tcp any any eq www  
  
interface Serial0  
ip policy route-map foo  
  
route-map foo  
match ip address 101  
set interface Tunnel0
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

42

Autoroute and Forwarding Adjacency

Cisco.com

```
interface Tunnel0
  tunnel mpls traffic-eng autoroute announce
      OR
  tunnel mpls traffic-eng forwarding-adjacency
  isis metric x level-y (ISIS)
  ip ospf cost ospf-cost (OSPF)
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

43

Summary Configuration (1/2)

Cisco.com

```
ip cef {distributed}
mpls traffic-eng tunnels
interface Tunnel0
  tunnel mode mpls traffic-eng
  ip unnumbered Loopback0
  tunnel destination RID-of-tail
  tunnel mpls traffic-eng autoroute announce
  tunnel mpls traffic-eng path-option 10 dynamic
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

44

Summary Configuration (2/2)

Cisco.com

```
! Configure in IGP
mpls traffic-eng tunnels
mpls traffic-eng router-id Loopback0
mpls traffic-eng area ospf-area (OSPF)
mpls traffic-eng level-x (ISIS)
metric-style wide
!
! On Physical interface
interface POS0/0
    mpls traffic-eng tunnels
    ip rsvp bandwidth Kbps
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

45

You Want SHOW Commands?

Cisco.com

```
show mpls traffic-eng link-management admission-cont
show mpls traffic-eng link-management advertisements
show mpls traffic-eng link-management bandwidth-alloc
show mpls traffic-eng link-management igp-neighbors
show mpls traffic-eng link-management interfaces
show mpls traffic-eng link-management summary
show mpls traffic-eng forwarding-adjacency
show mpls traffic tunnel backup
show mpls traffic-eng fast-reroute database
show mpls traffic-eng tunnels
show mpls traffic-eng tunnels summary
show mpls traffic-eng what's-for-dinner
```

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

46

PROTECTION



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

47

Protection

Cisco.com

- Mechanism to minimize packet loss during a failure
- Pre-provisioned protection tunnels that carry traffic when a protected link or node goes down
- MPLS TE protection also known as **FAST REROUTE (FRR)**
- FRR protects against **LINK FAILURE**
For example, Fibre cut, Carrier Loss, ADM failure
- FRR protects against **NODE FAILURE**
For example, power failure, hardware crash, maintenance
- Real Soon Now: protection against **CONDUIT FAILURE (SRLG)**
Conduit may carry multiple fibres, don't want to protect C1:F1 with C1:F2

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

48

Categories of Fast Reroute Protection

Cisco.com

- **Local protection**
 - Link protection**
 - Node protection**
 - Protect a piece of the network (node or link)
 - 1:N scalability
 - Fast failure recovery due to local repair
- **Path protection**
 - Real soon now
 - Protects individual tunnels
 - 1:1 scalability

RST-2603
9866_05_2004_c2

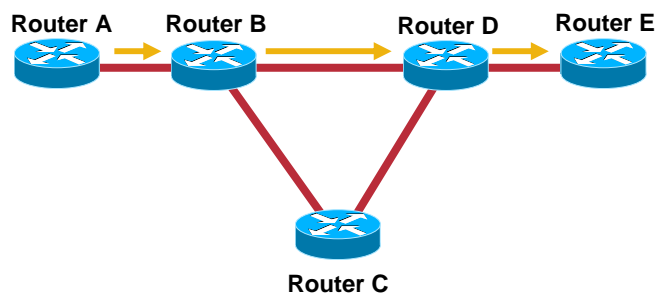
© 2004 Cisco Systems, Inc. All rights reserved.

49

Link Protection

Cisco.com

- TE Tunnel A → B → D → E →




RST-2603
9866_05_2004_c2

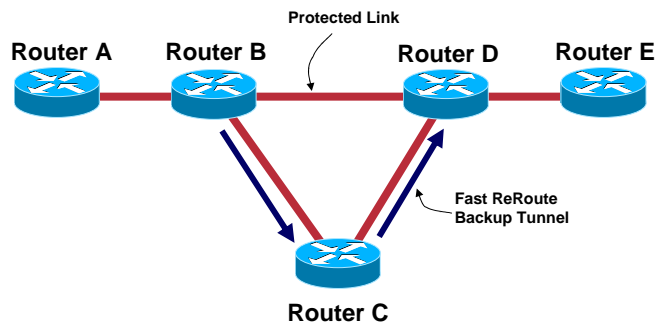
© 2004 Cisco Systems, Inc. All rights reserved.

50

Link Protection

Cisco.com

- B has a pre-provisioned backup tunnel to the other end of the protected link (Router D) $B \rightarrow C \rightarrow D$ 
- FRR relies on the fact that D is using global label space



RST-2603
9866_05_2004_c2

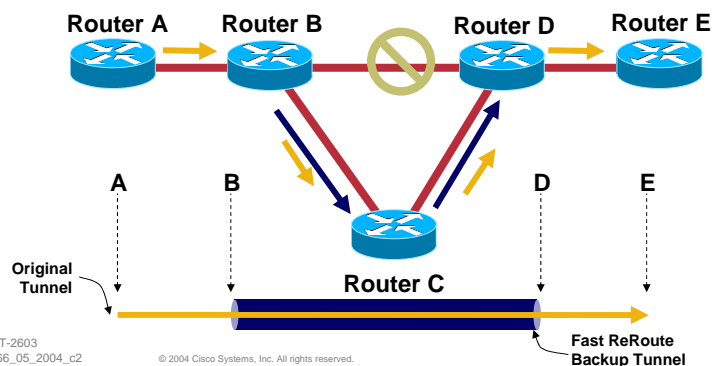
© 2004 Cisco Systems, Inc. All rights reserved.

51

Link Protection

Cisco.com

- When $B \rightarrow D$ link fails, $A \rightarrow E$ tunnel is encapsulated in $B \rightarrow D$ tunnel
- Backup tunnel is used until A can re-compute tunnel path as $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$ (~5-15 seconds or so)



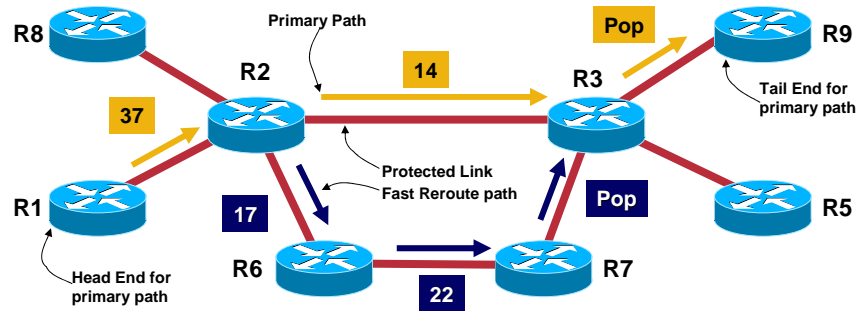
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

52

Link Protection Example

Cisco.com



Primary path: R1 → R2 → R3 → R9
Fast Reroute path: R2 → R6 → R7 → R3

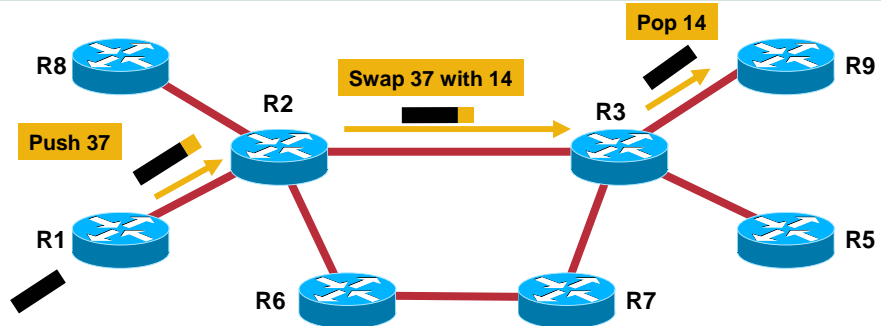
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

53

Normal TE Operation

Cisco.com



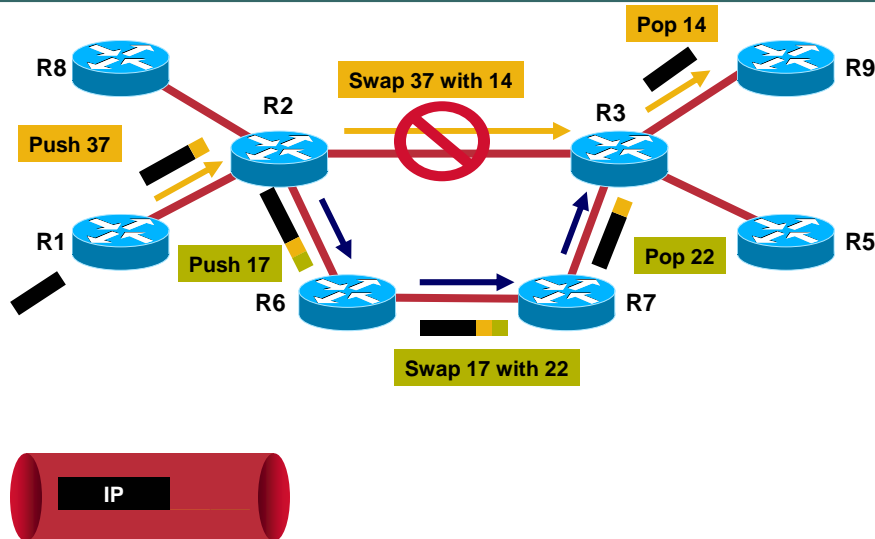
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

54

Fast Reroute Link Failure

Cisco.com



RST-2603
9866_05_2004_c2

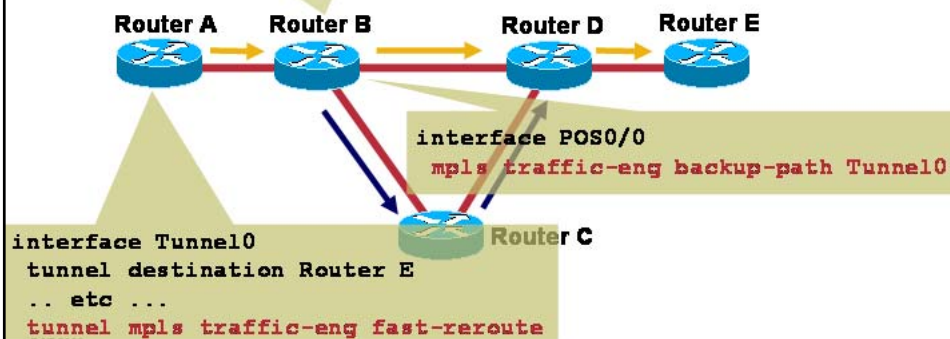
© 2004 Cisco Systems, Inc. All rights reserved.

55

Link Protection Configuration

Cisco.com

```
interface Tunnel0
  tunnel destination Router D
  ... explicit-path R2-R3-R4
  no tunnel mpls traffic-eng autoroute announce
```



RST-2603
9866_05_2004_c2

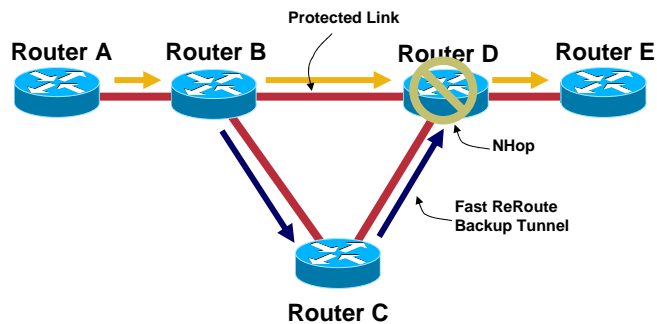
© 2004 Cisco Systems, Inc. All rights reserved.

56

Node Protection

Cisco.com

- What if Router D failed?
- Link protection would not help as the backup tunnel terminates on Router D (which is the NHop of the protected link)



RST-2603
9866_05_2004_c2

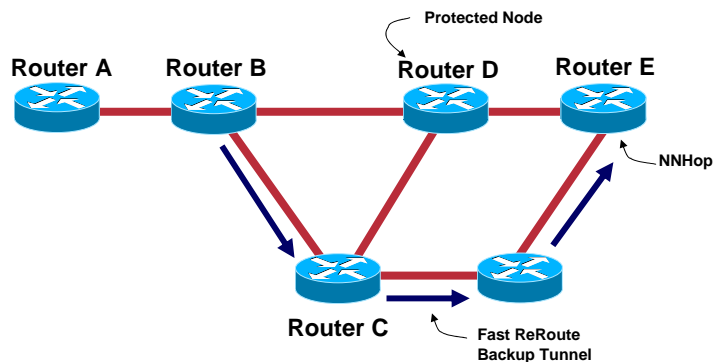
© 2004 Cisco Systems, Inc. All rights reserved.

57

Node Protection

Cisco.com

- **SOLUTION: NODE PROTECTION** (If network topology allows)
- Protect tunnel to the next hop **PAST** the protected link (NNhop)



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

58

Node Protection

Cisco.com

- **Node protection still has the same convergence properties as link protection**
- **Deciding where to place your backup tunnels is a much harder problem to solve on a large-scale**
- **For small-scale protection, link may be better**
- **Auto-tunnel and auto-mesh can help with this**
- **Configuration is identical to link protection, except where you terminate the backup tunnel (NNHop vs. NHop)**

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

59

Link and Node Protection Times

Cisco.com

- **Link and Node protection are very similar**
- **Protection times are commonly linear to number of protected items**
- **One nationwide provider gets ~35ms of loss**
- **New code on GSR E3 linecards gets a prefix-independent 2ms-4ms loss**

RST-2603
9866_05_2004_c2

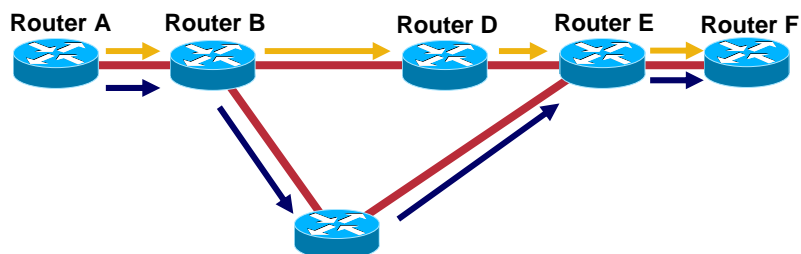
© 2004 Cisco Systems, Inc. All rights reserved.

60

Path Protection

Cisco.com

- **Path protection: Multiple tunnels from TE head to tail, across diverse paths**



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

61

Path Protection

Cisco.com

- **Least scalable, most resource-consuming, slowest convergence of all 3 protection schemes**
- **With no protection, worst-case packet loss is 3x path delay**
- **With path protection, worst-case packet loss is 1x path delay**
- **With link or node protection, packet loss is easily engineered to be subsecond (<100ms, <50ms, 4ms, all possible)**
- **Path protection is useful in a few places:**
 - Geographically constrained regions (e.g. Japan)
 - Only a few protected LSPs (one-off per-circuit AToM protection)

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

62

QoS/MPLS/MPLS-TE



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

63

No Need for TE or DiffServ

Cisco.com

Control
Plane



No contention anywhere, ever
So don't do any work!

Nothing

Forwarding
Plane

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

64

Need DiffServ, Don't Need TE

Cisco.com

Control
Plane

Temporary congestion can be
handled by differentiating
between packets at
forwarding time

“Managed Unfairness”

DiffServ

Forwarding
Plane

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

65

More Efficient Use of IP Bandwidth, but No Contention

Cisco.com

Control
Plane

Spread traffic around the
network with more flexibility
than just IP metrics

TE

Forwarding
Plane

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

66

Bandwidth Optimization and Congestion Management in Parallel

Cisco.com

Control Plane

Spread traffic around with more flexibility than the IGP offers

Managed unfairness during temporary congestion

TE + DiffServ

Forwarding Plane

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

67

Reserve Per-Class Bandwidth, Sort Of

Cisco.com

Control Plane

DS-TE

Control-plane reservation for multiple forwarding classes.

Still need DiffServ PHBs to provide actual service differentiation.

Forwarding Plane

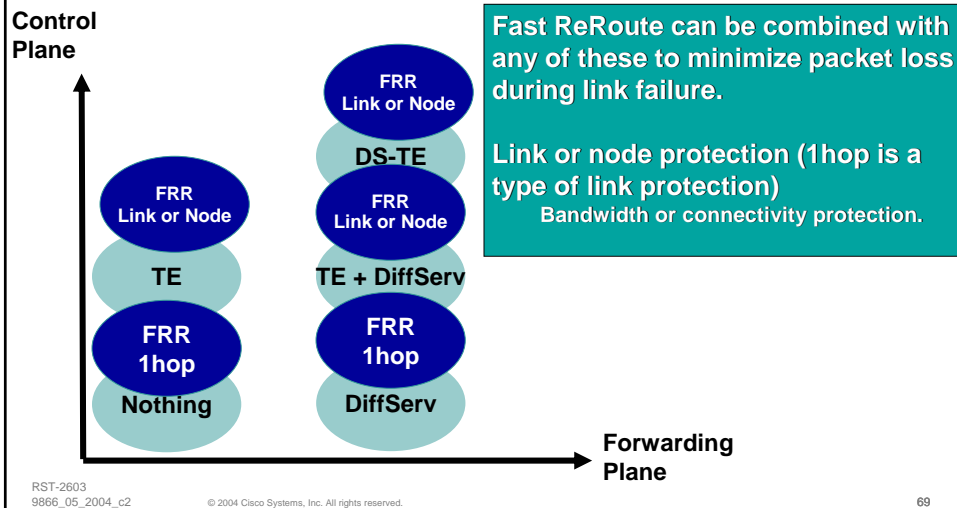
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

68

Minimize Packet Loss at Any Point

Cisco.com



DiffServ-Aware TE

Cisco.com

- Regular TE allows for one reservable bandwidth amount per link
- Regular (FIFO) queuing allows for one queue per link
- DiffServ queuing (e.g. LLQ) allows for more than one queue per link
- DS-TE allows for more than one reservable bandwidth amount per link
- Basic idea: connect PHB class bandwidth to DS-TE bandwidth sub-pool
- Still a **control-plane reservation only**

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

70

DESIGN AND SCALABILITY



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

71

Design Approach and Scalability

Cisco.com

Two Methods to Deploy MPLS-TE

- **Tactical**

As needed to clear up congestion

You only have tunnels when there is a problem (and you must remember to remove them)

- **Strategic**

Mesh of TE tunnels between a level of routers

Typically P to P but can be PE to PE in smaller networks

$N(N-1)$ LSPs (one in each direction)

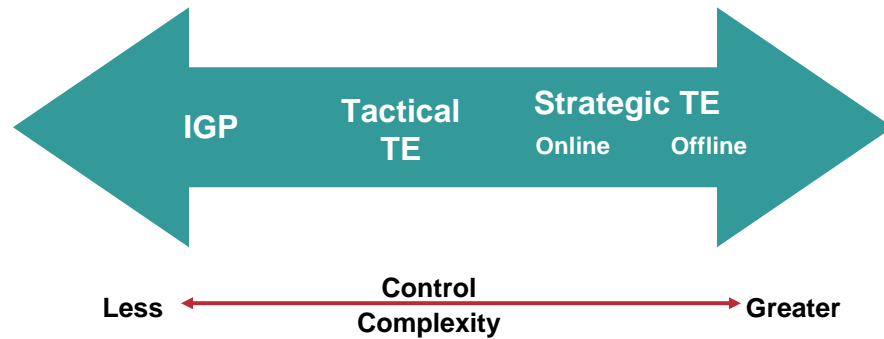
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

72

Design Spectrum

Cisco.com



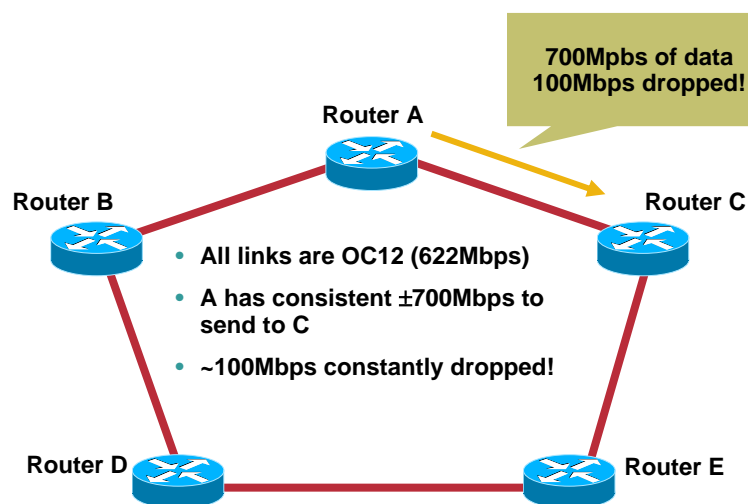
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

73

Tactical: Large ISP Case Study

Cisco.com



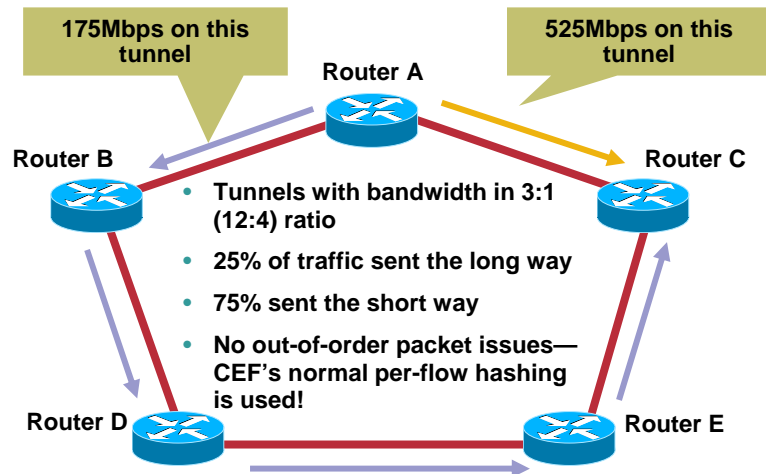
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

74

Multiple TE and Unequal Cost Load Balancing

Cisco.com



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

75

Tactical

Cisco.com

- As needed—Easy, quick, but hard to track over time
- Easy to forget why a tunnel is in place
- Inter-node BW requirements may change, tunnels may be working around issues that no longer exist

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

76

Strategic

Cisco.com

- Full mesh of TE tunnels between routers
- Initially deploy tunnels with 0 bandwidth
- Monitor tunnel interface statistics
 - ~Bandwidth used between router pairs
 - TE tunnels have interface MIBs
 - Make sure that $\Sigma \text{tunnel} \leq \Sigma \text{network BW}$
- As tunnel bandwidth is changed, tunnels will find the best path across the network

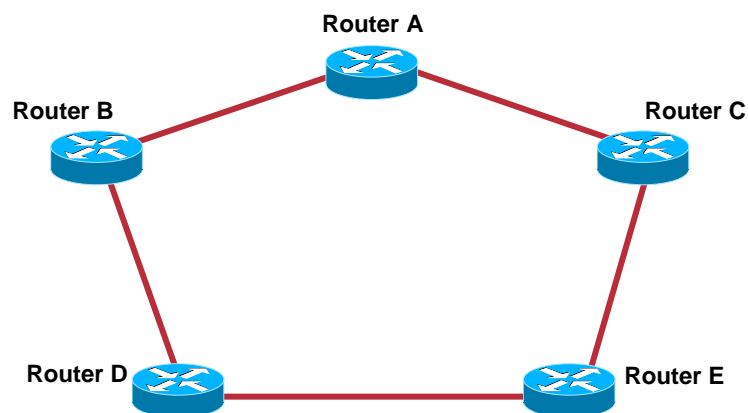
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

77

Strategic: Physical Topology

Cisco.com



RST-2603
9866_05_2004_c2

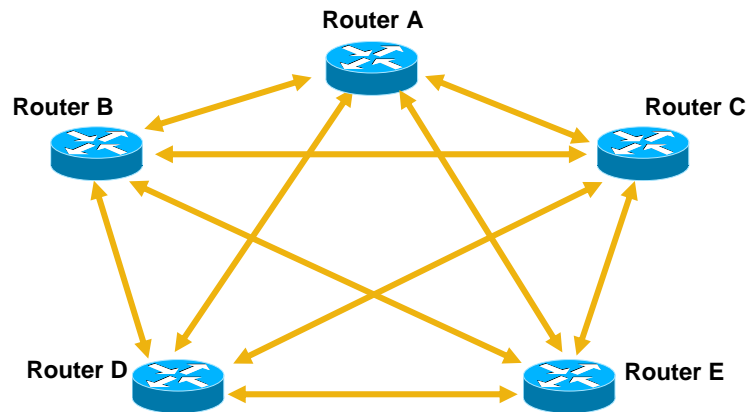
© 2004 Cisco Systems, Inc. All rights reserved.

78

Strategic: Logical Topology

Cisco.com

- Total of 20 tunnels in this network
- Each link is actually 2 unidirectional tunnels



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

79

Strategic

Cisco.com

- N routers, $N*(N-1)$ tunnels
- Routing protocols do not run over a TE tunnel
Unlike an ATM/FR full mesh!
- Tunnels are unidirectional
This is a good thing
Can have different bandwidth reservations in two different directions

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

80

Scalability

Cisco.com

Code	Number of Headend Tunnels	Number of Mid-Points	Number of Tails
12.0ST	600	10,000	5,000

- In late 2000/early 2001, we said, “how many tunnels come up in 5 minutes?”
- Answer as above
- With latest code, above converges in 2-3 minutes
- Haven’t need to formalize larger-number testing
Largest customer we know of has a network about ¼ this size
- Bottom line: MPLS-TE scalability is not the gating factor in scaling your network

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

81

Protection Scalability

Cisco.com

- N is number of nodes in the TE cloud (10-150)
- D is backbone degree of connectivity (4-6 avg, max 12-16)
- Primary full mesh: $O(N^2)$
- Link protection: additional $O(N \cdot D)$ tunnels
- Node protection: additional $O(N \cdot D^2)$ tunnels
- Path protection: additional $O(N^2)$ tunnels

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

82

Scalability

Cisco.com

<http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120limit/120st/120st14/scalable.htm>

- Search CCO for “Scalability Enhancements for MPLS Traffic Engineering”

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

83

MPLS VPN, MULTICAST, AND TE



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

84

TE and MPLS VPNs

Cisco.com

- **MPLS VPNs**
VPN label is carried across network by IGP label
- **MPLS TE label can serve as the IGP label**
TE tunnels must connect PE-PE...
- **Otherwise enable LDP/TDP on tunnel interface**
Using Targeted LDP session

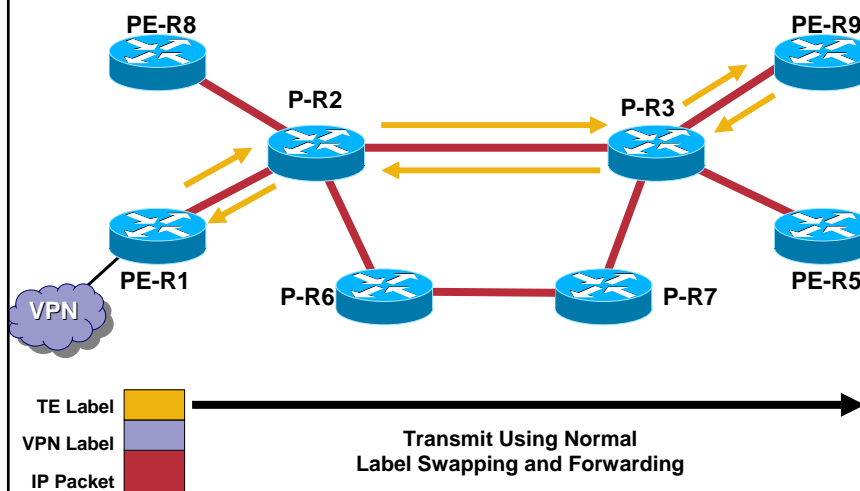
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

85

MPLS VPN over PE-PE TE Tunnel

Cisco.com



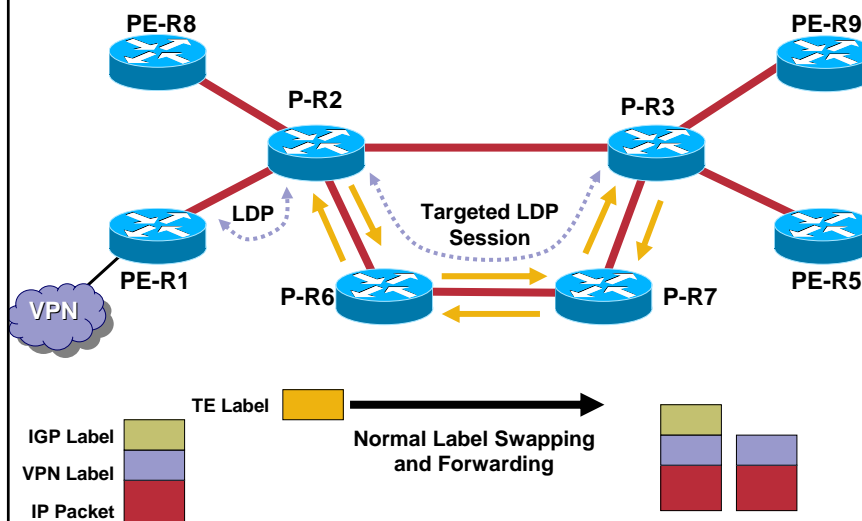
RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

86

MPLS VPN over P-P TE Tunnel

Cisco.com



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

87

TE and Multicast

Cisco.com

- Multicast requires RPF
- Multicast packets should arrive on interface that is shortest path to source
- Autoroute causes IGP to point to TE tunnel
- As packets don't exit tunnel interface, multicast breaks!
- Cisco IOS has a command to fix it (under IGP config)...

`mpls traffic-eng multicast intact`

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

88

SUMMARY



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

89

MPLS TE

Cisco.com

- **Helps optimize network utilisation (strategic)**
- **Assists in handling unexpected congestion (tactical)**
- **Provides fast reroute for link and node failures**
- **TE is only part of a method of guaranteeing bandwidth**

It is a control plane mechanism only

Must be used with traditional QoS mechanisms

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

90

Complete Your Online Session Evaluation!

Cisco.com

- WHAT:** Complete an online session evaluation and your name will be entered into a daily drawing
- WHY:** Win fabulous prizes! Give us your feedback!
- WHERE:** Go to the Internet stations located throughout the Convention Center
- HOW:** Winners will be posted on the onsite Networkers Website; four winners per day

RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

91



RST-2603
9866_05_2004_c2

© 2004 Cisco Systems, Inc. All rights reserved.

92