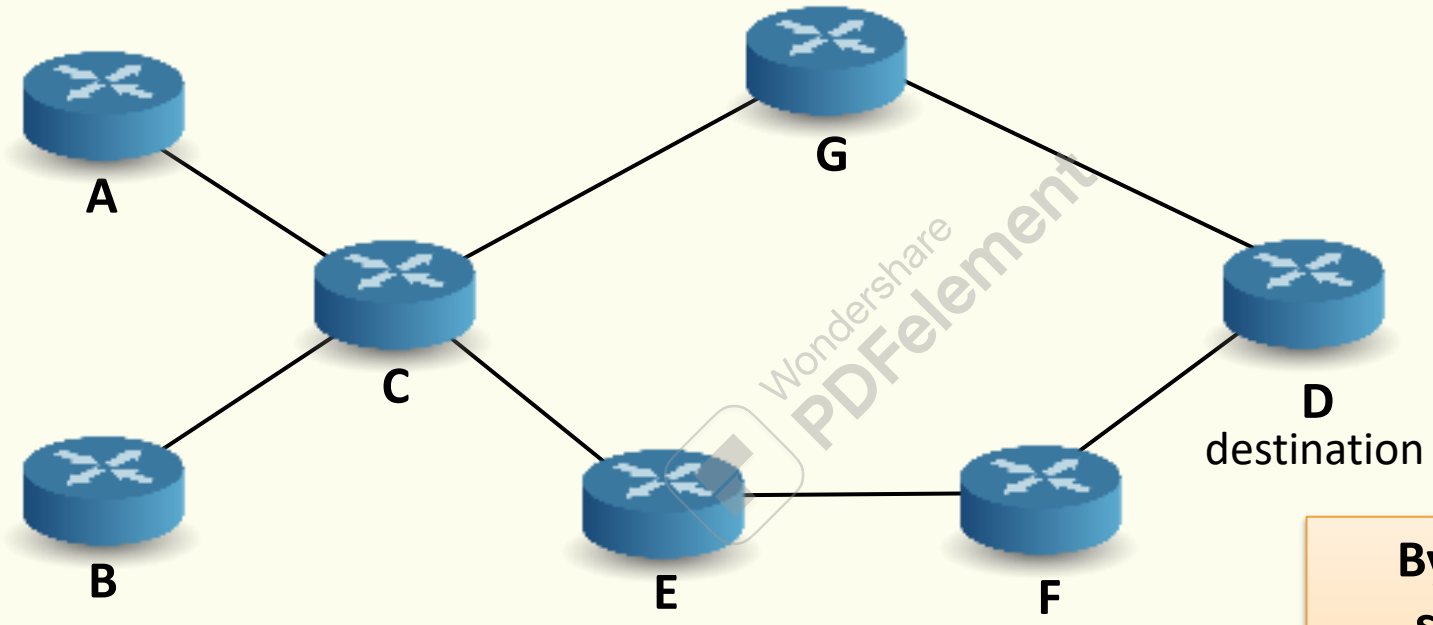
Requirements for TE

Remember: all the needed info is collected by OSPF



Computing paths that comply with a set of constraints

Enforcing traffic to be forwarded along these paths

By decoupling service from transport, MPLS is fundamental to support TE requirements

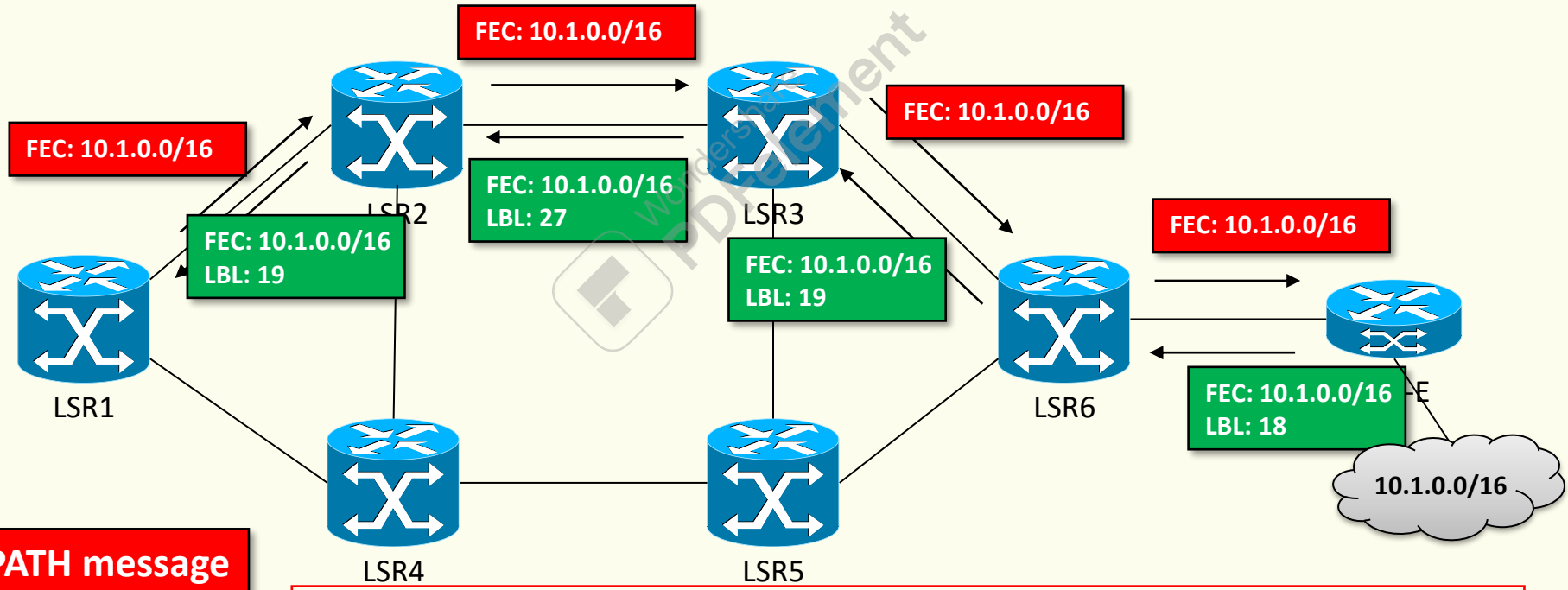
RSVP for label distribution

In the Control Plane

An IGP (Interior Gate Protocol) such as OPSF or IS-IS, with TE extension, is needed to collect data

- Ordered control with downstream on-demand

In order to fully establish LSBs, OSPF is not enough. It gets coupled with a new RSVP-TE protocol.



The head router that is setting up an LSB emits a "Path" message, and the tail router responds with a "Reservation" message (thus the RSVP name)

The PATH message carries in a field called Explicit Route Object(ERO) the info on the Path that an LSB has to follow.

By using the ERO, the FEC is not known by anyone other than the Head router

The Traffic Specification Object (Tspec) is a field in the PATH message that carries the needed constarints

RSVP-TE

- RSVP is extended to support LSP setup after CSPF computation
  - The LSP ***head-end*** has full (ordered) control of the setup
- Realized by the definition of new Objects carried by Path and Resv messages
  - *Path* objects: **Label Request Object** (LRO), **Explicit Route Object** (ERO), **Sender TSpec** (revisited)
  - *Resv* objects: **Label Object**
  - Common to both: **Record Route Object** (RRO)

A new RESV message is sent at every hop to the router's predecessor in the Path

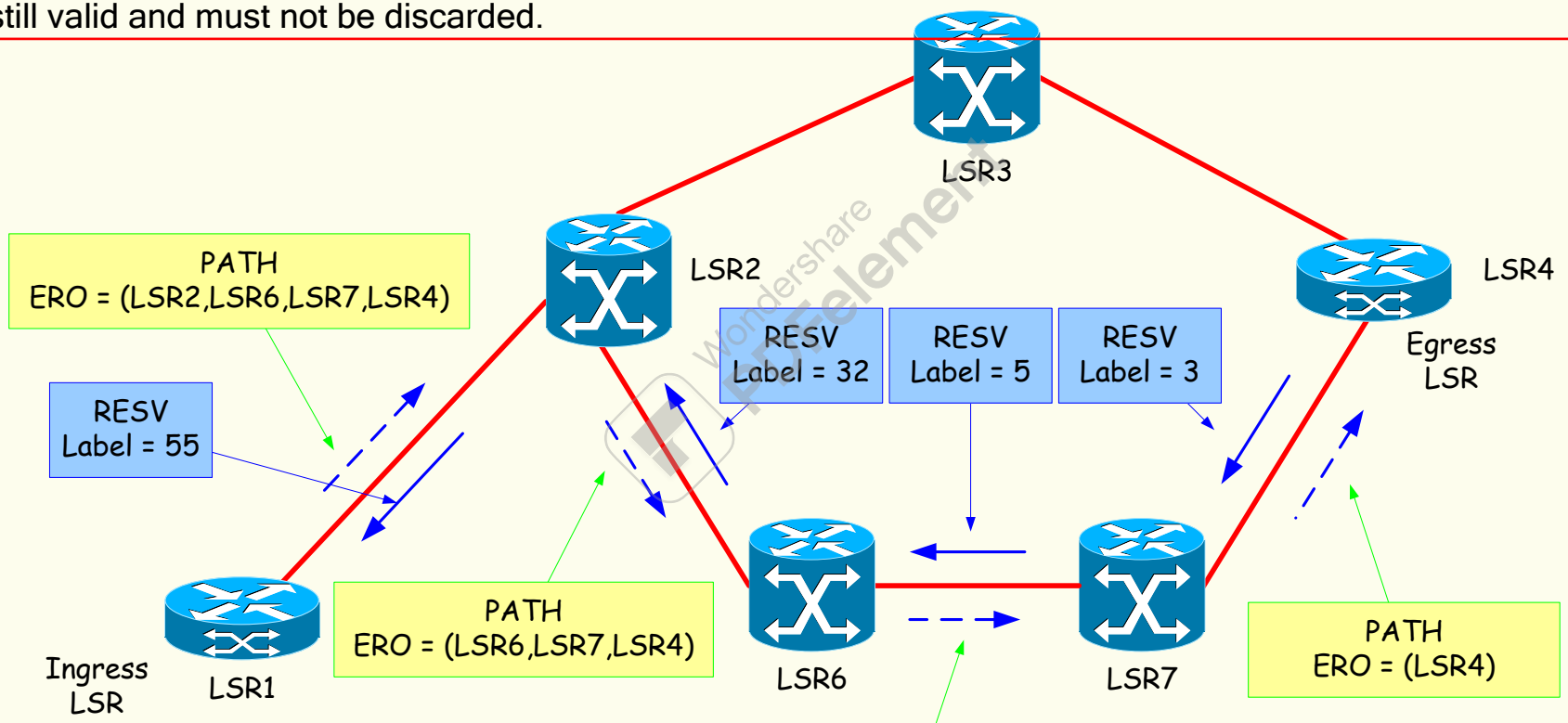
Constraint checking is done when the RSVP messages are received. Checking as late as possible is better, because things may change in the meantime. Constraint check is Link-based.

If a link on the Route does not meet constraints, a message is sent both Upstream and Downstream to cancel the Reservation and warn that the Path is invalid.

RSVP-TE

Labels are set up in each router when a successful constraints check on a RSVP message is made.

The introduction of all these constraints means that a lot of state management info is needed to keep each Label updated and functional. These are Soft States, meaning that they need periodic confirmation messages that they are still valid and must not be discarded.

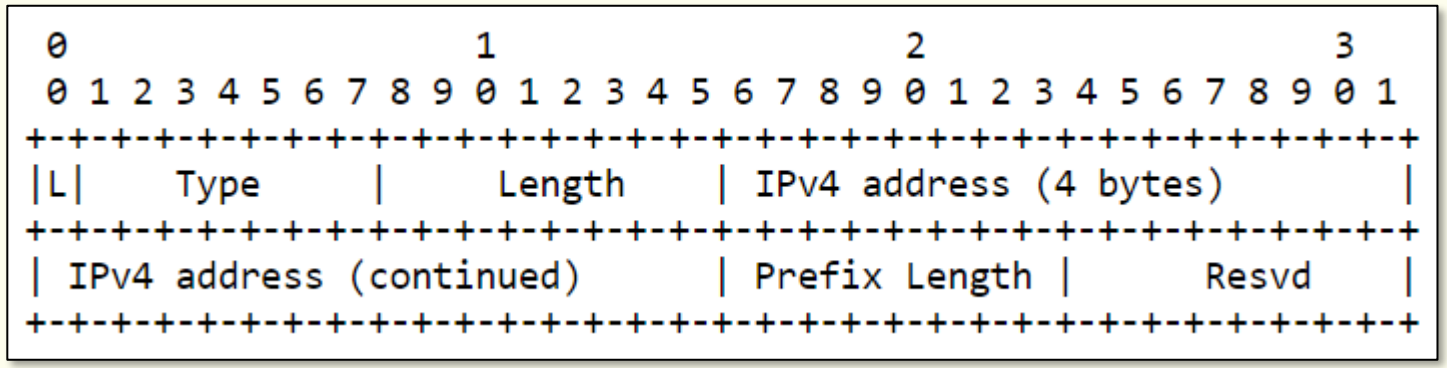


The ERO might specify only a Partial Path eg. different AS. The exact path taken must be recorded by each router. This is stored in the RRO, which is common to both messages.

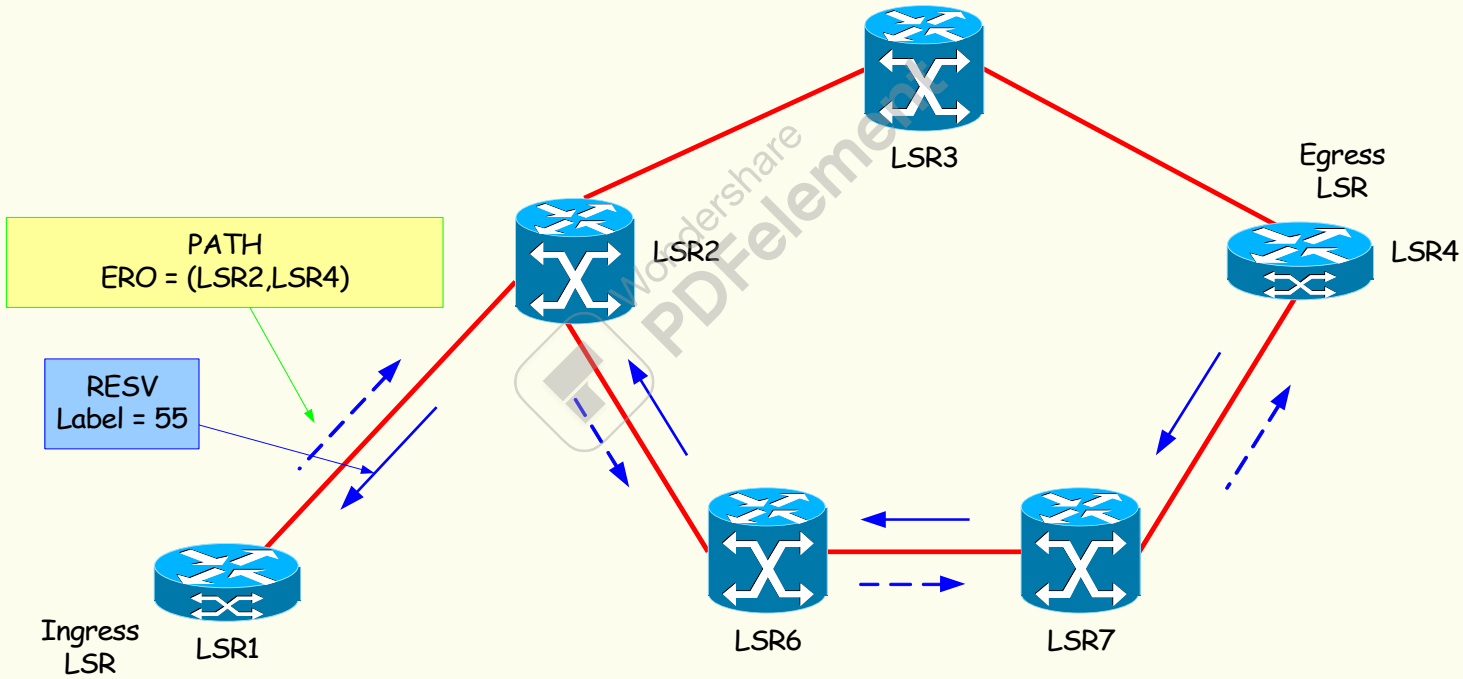
RSVP-TE

- Explicit Route Object
  - consists of a sequence of sub-objects, each representing an **abstract node**, i.e., **a group** of one or more routers
  - **Strict** vs. **loose** hops

A Strict hop lists the exact router, a Loose one is when the next hop is not fully known by the head router , and it has to be filled by the router at that link.



RSVP-TE


PATH
ERO = (LSR2,LSR4)
RESV
Label = 55
Ingress
LSR
LSR1
LSR2
LSR3
Egress
LSR
LSR4
LSR6
LSR7

RSVP-TE

- Admission control is **required** and performed at each hop
  1 – CSPF computation is not mandatory
  3 – Unreserved bandwidth on a link has changed after CSPF computation
  2 – The TED at the head-end is not accurate
- If LSP setup is successful, reservation updates are fed back to OSPF-TE
- **Bandwidth reservations are in the control plane only!**

There are 8 possible priorities for LSBs, where 0 is the most important and 7 the least.

In case of a failure, less important LSBs could be teared down in favour of more important ones

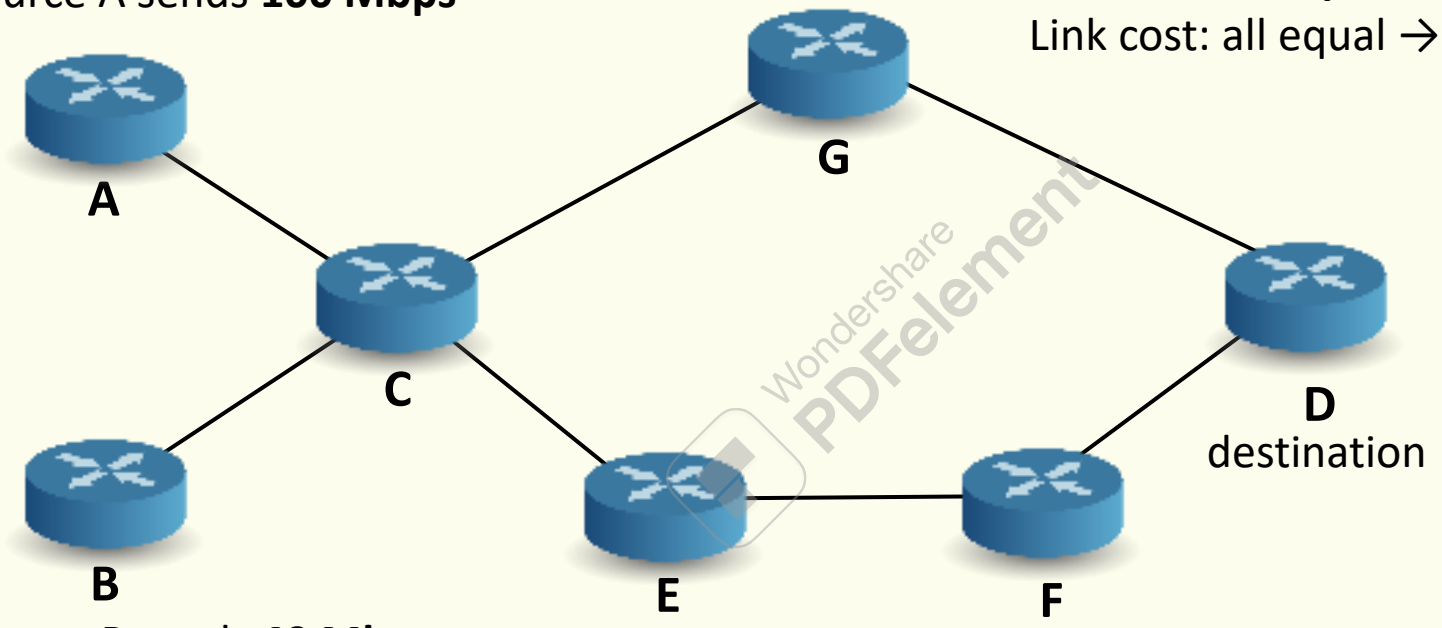
LSP priorities

- LSP have **priorities**, used to solve for resource contention
  - An important LSP is always established along the most optimal (shortest) path that fits the constraints, **regardless of existing reservations**
  - When LSPs need to reroute (e.g. after a link failure), important LSPs have a **better chance of finding an alternate path**
  - In the **absence of important LSPs**, resources can be reserved by less important LSPs

Application scenario [3]

source A sends **100 Mbps**

All link **150 Mbps**, but **E-F** (**50 Mbps**)
Link cost: all equal → Path cost = #hops



source B sends **40 Mbps**
Customer B buys a service with strict guarantees

1. **Shortest path → ok in normal conditions**
2. **What if link G-D fails?**

find paths between source/destination pairs that **comply with bandwidth constraints**, enforce the **priority of the path** sourced at B over that sourced at A

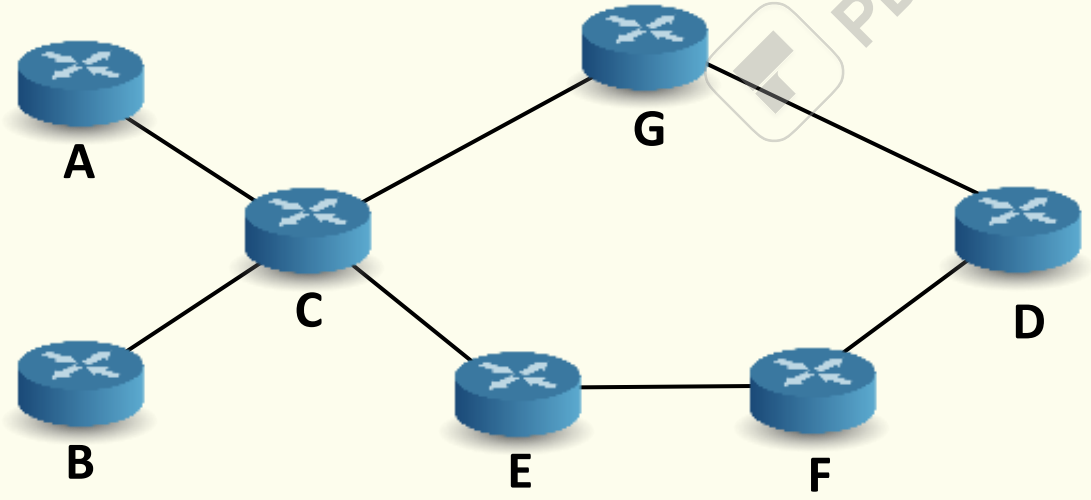
2 different priorities are used for Network Stability reasons. Some LSP might be less important, but they might need absolute stability once established.

LSP priorities

- Eight priority levels, **two** priorities per LSP
  - **Setup** priority (0 highest – 7 lowest): controls access to the resources when the LSP is established
  - **Hold** priority (0 highest – 7 lowest): controls access to the resources for an LSP that is already established
- When an LSP is set up, if not enough resources are available, the setup priority of the new LSP is compared to the hold priority of the LSPs using the resources in order to determine whether the new LSP **can preempt** any of the existing LSPs and take over their resources

LSP priorities

- Why distinct priorities?
  - Case 1: All LSPs have Hold Pri 0 & Setup Pri 7
  - Case 2: All LSPs have Hold Pri 7 & Setup Pri 0



1. a new LSP can never preempt an existing LSP and in turn can never be preempted

2. constant churn if two LSPs compete for the same resource

Reoptimization

- Due to dynamic changes, also the optimal solution for an LSP may change over time
- **Reoptimization** is the process of recomputing CSPF on each update
  - Trade-off between stability and optimization

- Without a full knowledge of **present** and **future** LSP requirements, **any algorithm is sub-optimal**

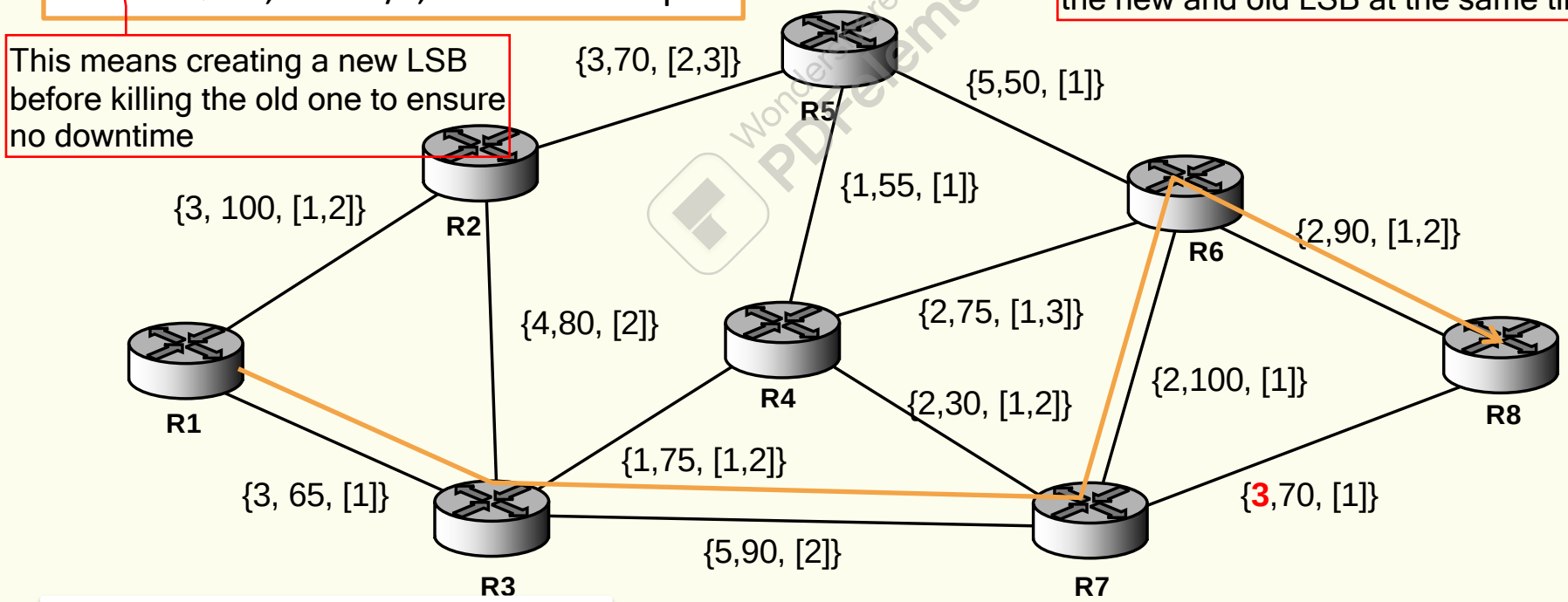
Reoptimization

- Re-routing an LSP without any traffic loss:
  make-before-break approach

LSP: R1→R8, 60Mb/s, exclude Group 3

This means creating a new LSB before killing the old one to ensure no downtime

The Make-before-break approach can only be used if the constraints (eg. Unused Bandwidth) are satisfied for both the new and old LSB at the same time.



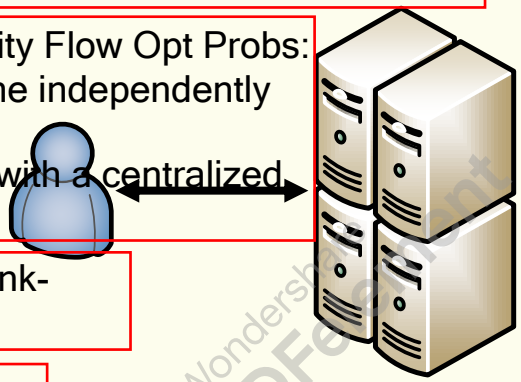
Shared Explicit (SE) reservation

Made at Provisioning/Set

Centralized (offline) TE

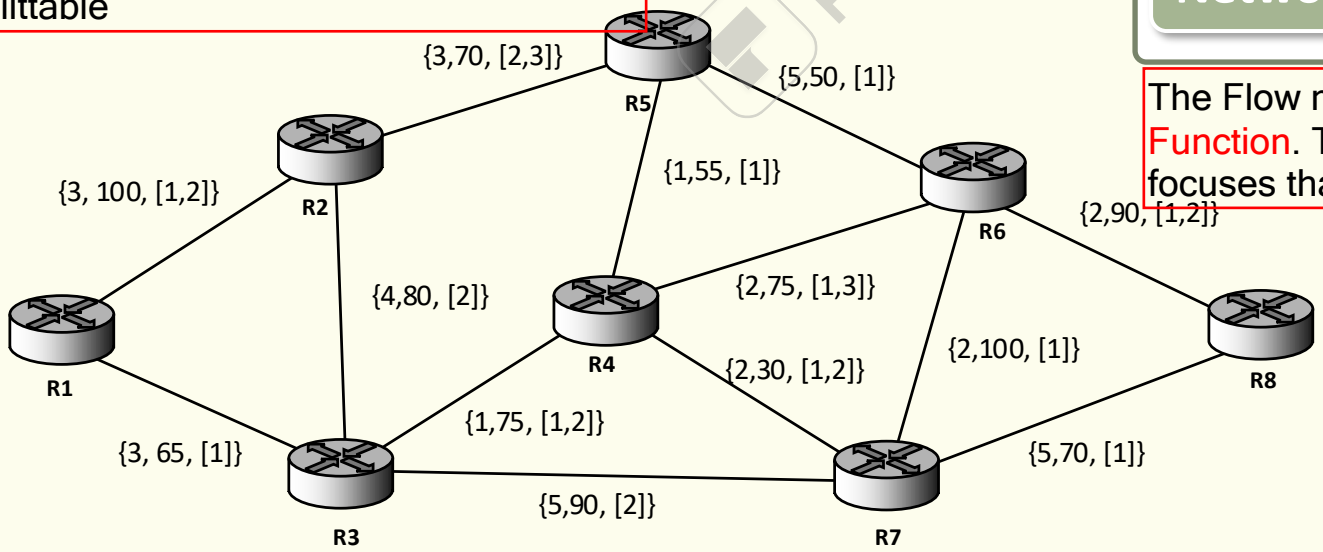
These system is usually modeled with Flow-based models. It is then optimized with a Commodity Flow Optimization Problem.

There exist mainly 2 types of Commodity Flow Opt Probs:
- Single-commodity(1 flow): can be done independently at each hop (eg. CSPF)
- Multi-commodity: can only be solved with a centralized approach



CSPF is a simplification because it is link-local

In Multi-commodity, Flows can be splittable



- Traffic demand estimation
- Topology and network state discovery
- Route computation
- Network configuration

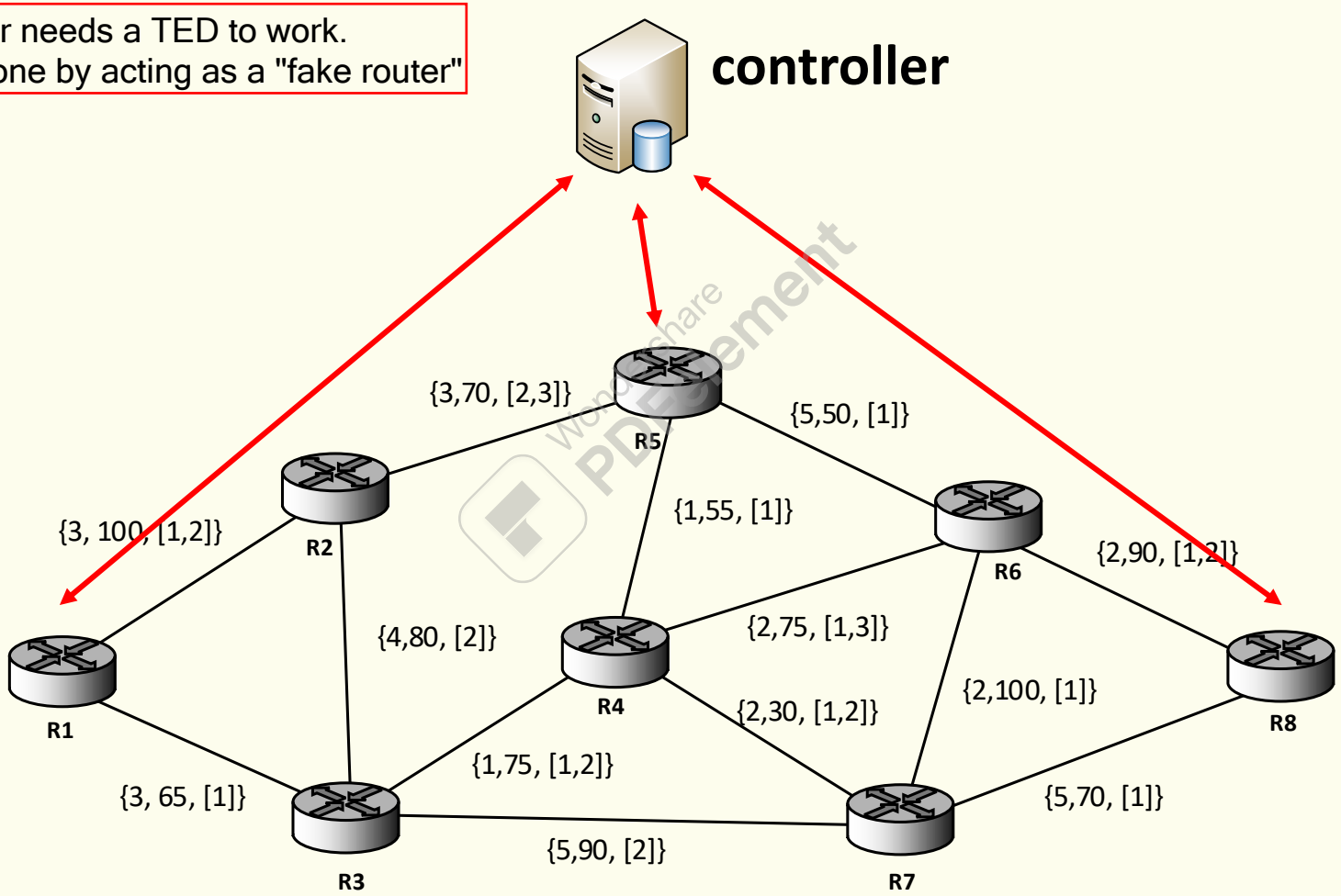
The Flow model needs an Objective Function. This can be one of several focuses that we want the network to have.

Some example Objective Functions are:
- Minimum cost
- Load balancing

Made at Runtime

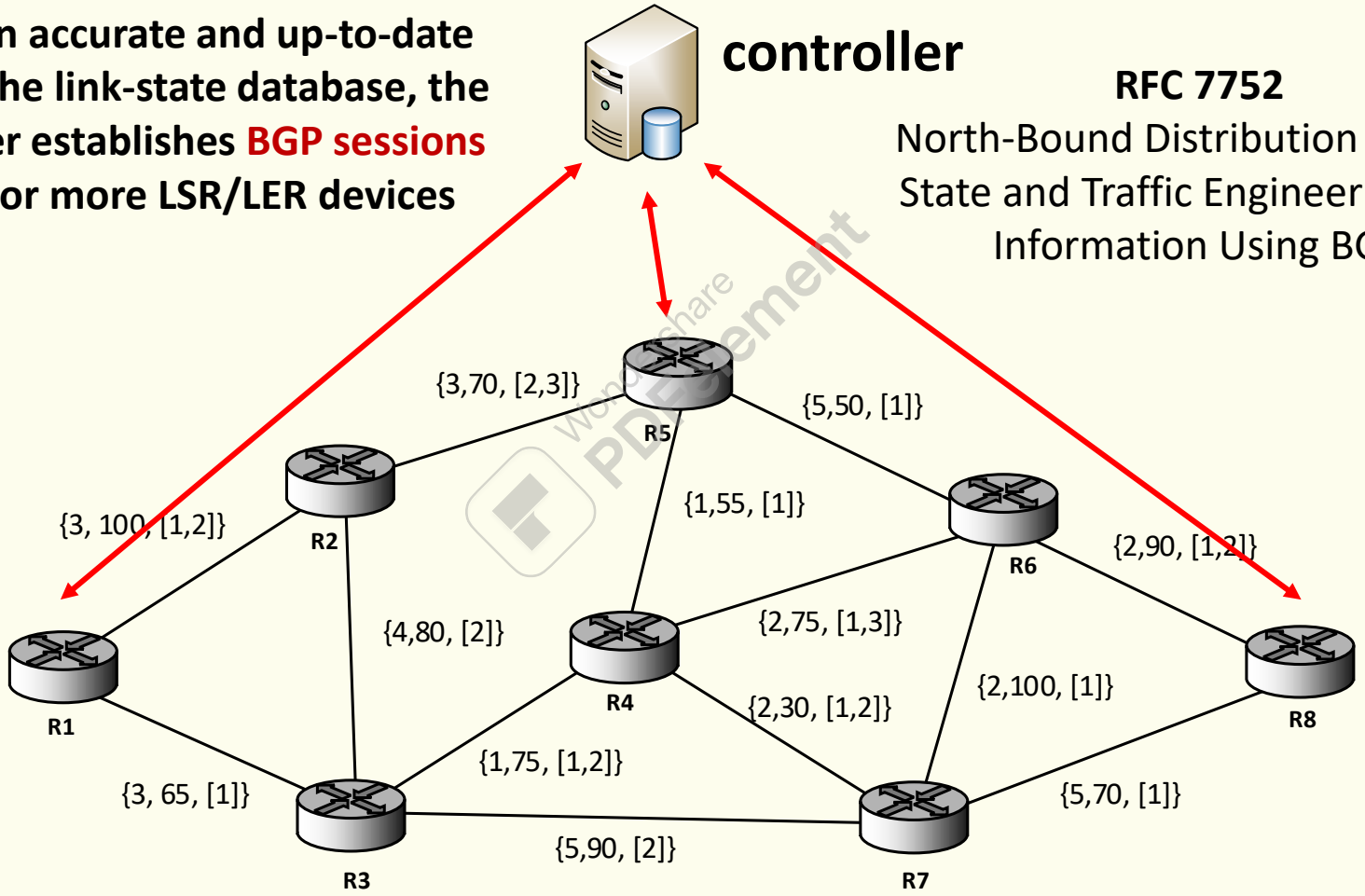
Centralized (online) TE

The Controller needs a TED to work.
It can obtain one by acting as a "fake router"



Centralized (online) TE

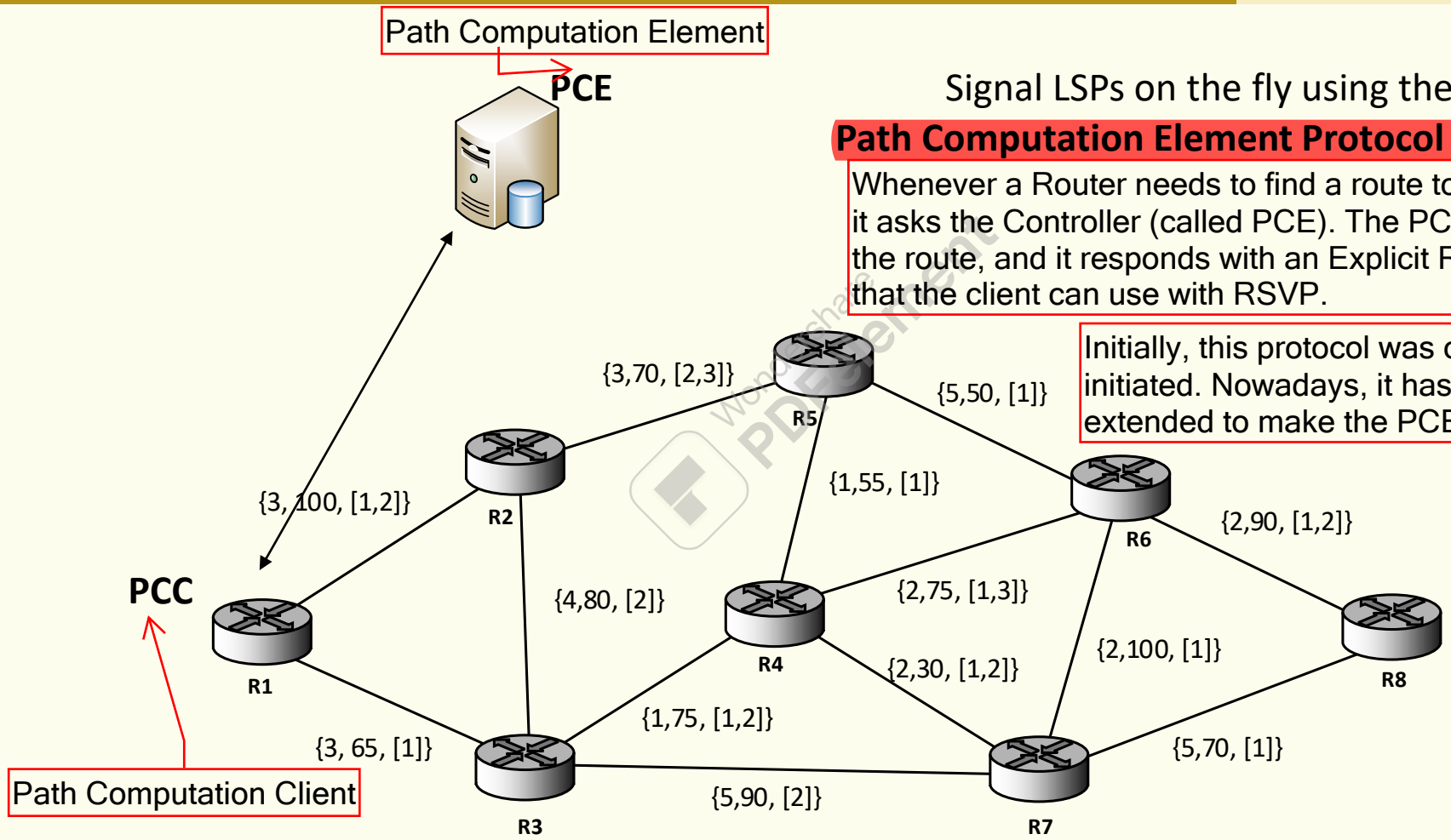
To get an accurate and up-to-date view of the link-state database, the controller establishes BGP sessions to one or more LSR/LER devices



RFC 7752

North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP

Path Computation Element



Signal LSPs on the fly using the

Path Computation Element Protocol (PCEP)

Whenever a Router needs to find a route to a target, it asks the Controller (called PCE). The PCE computes the route, and it responds with an Explicit Route Object that the client can use with RSVP.

Initially, this protocol was only client initiated. Nowadays, it has been extended to make the PCE update.

Protection and restoration

- Protection and restoration are mechanisms to handle failures
- It requires fast failure detection

A path keeps working even if one or more of its components fail

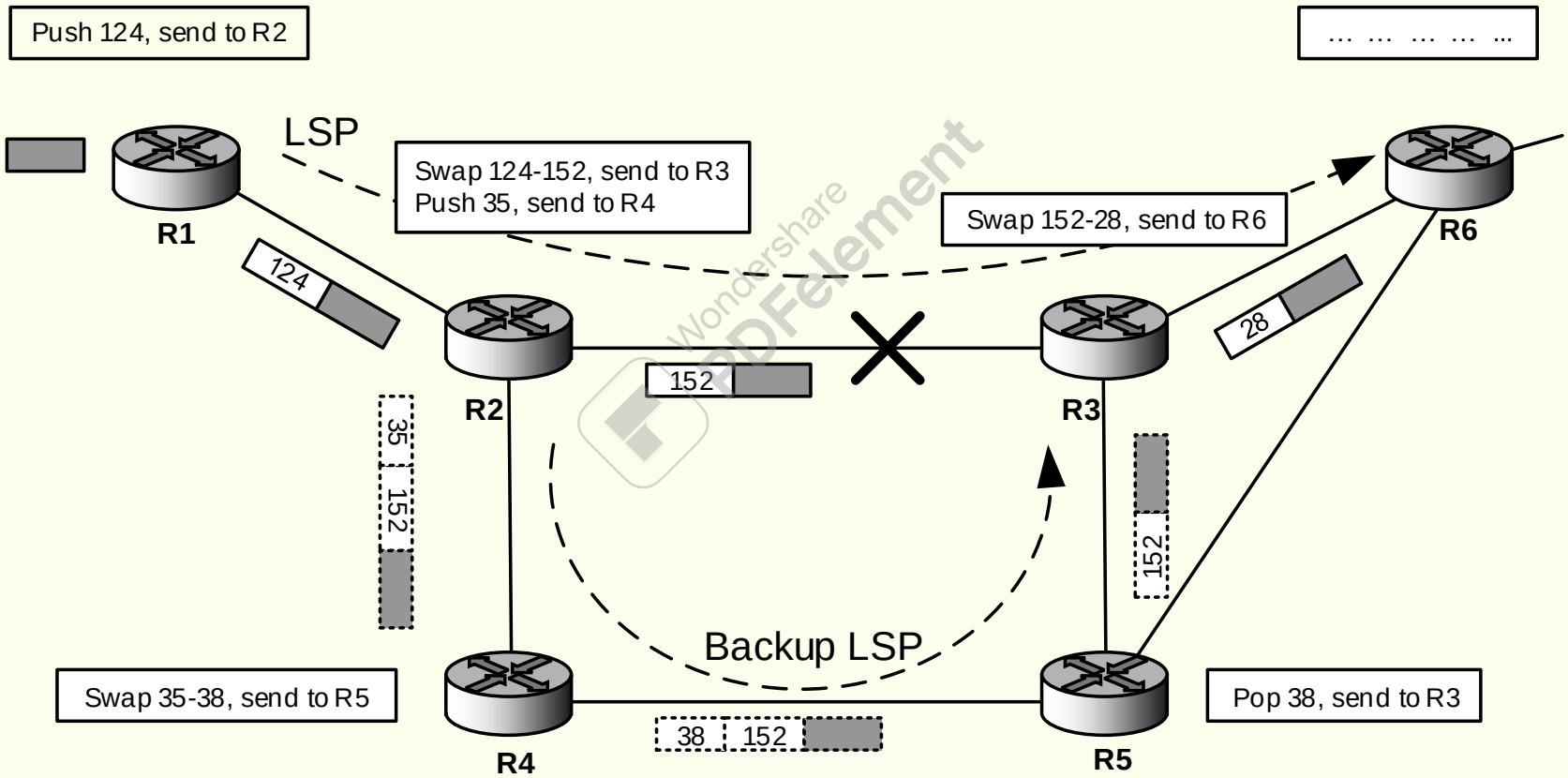
1. Path protection (end-to-end)
   - LSP protection is achieved using two LSPs: the *primary*, used under normal operation, and the *secondary*, used if there is a failure on the primary

Usually on standby

2. Local protection using **fast reroute**
   - Link vs. node protection
   - One vs. many LSP protected

Link protection, many LSPs

The first one that may detect a failure is R2. While the head router uses RSVP-TE to recalculate the LSP, R2 will start sending packets through an alternate route (R2-R4-R5-R3) to the next hop. It is a kind of a Virtual Link.



This mechanism is called Fast Rerouting, and the alternate path must be setup at provisioning time.

References

- I. Minei and J. Lucek, **MPLS-Enabled Applications: Emerging Developments and New Technologies**, 3rd Edition, Wiley, Dec. 2010
- L. Lenzini, E. Mingozzi, G. Stea, **Traffic Engineering** in *End-to-End Quality of Service Over Heterogeneous Networks*, 2008, Springer
- RFCs
  - **RFC3209**, RSVP-TE: Extensions to RSVP for LSP Tunnels, Sept. 2001
  - **RFC3630**, Traffic Engineering Extensions to OSPF, Sept. 2003
  - **RFC4090**, Fast Reroute Extensions to RSVP-TE for LSP Tunnels, May 2005
  - **RFC4655**, A Path Computation Element (PCE)-Based Architecture, Aug. 2006
  - **RFC5440**, Path Computation Element (PCE) Communication Protocol (PCEP), Mar. 2009