

*1. Servizi Multimediali e
Qualità del Servizio (QoS) su IP*
1.4. MPLS

Prof. Raffaele Bolla

R. Bolla Telematica 2, n. o.

MPLS introduzione

- MPLS (MultiProtocol Label Switching) è uno standard sviluppato dall'IETF con l'obiettivo di:
 - aumentare la velocità di trattamento del pacchetto all'interno dei nodi (relegando le funzioni più complesse ai bordi della rete);
 - agevolare l'allocazione di risorse, il *traffic engineering* e la realizzazione di particolari tipi di servizi.
- La specifica base è RFC 3031 (Gennaio del 2001).
- È il risultato dell'evoluzione di diverse tecnologie, tra le quali ricordiamo il *Tag Switching* di Cisco, l'*ARIS* di IBM, e il *Cell-Switched Router* di Toshiba.

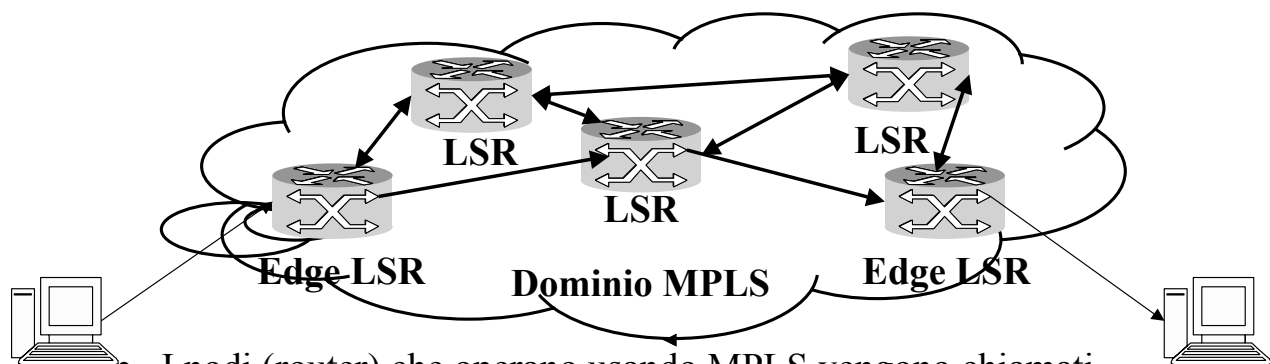
Funzionamento di massima

- MPLS architeturalmente si colloca in una sorta di livello intermedio (livello 2,5) fra il livello 3 di Rete (IP) ed il livello 2 di Linea.
- Si chiama “multi-protocol” perché, in linea di principio, è in grado di operare con qualunque protocollo di livello 3 (rete) anche se lo si applica tipicamente ad IP.
- Permette ai nodi che lo utilizzano di realizzare una commutazione su base “etichetta” e anche un instradamento tipo “Circuito Virtuale” su base flusso.

Lezione 1.4, v. 1.0

2.3

Domini e LSR



- I nodi (router) che operano usando MPLS vengono chiamati **Label Switching Router (LSR)**
- La parte di rete che questi nodi compongono viene chiamata **Dominio MPLS (MPLS Domain)**
- I nodi al confine del Dominio, ossia i nodi che ricevono/trasmettono traffico all'esterno del Dominio vengono chiamati

Lezione 1.4, v. 1.0

Edge Label Switching Router (ELSR)

2.4

Funzionamento di massima

- L'idea di base è che una certa tipologia di pacchetti che raggiungono un **ELSR** debbano venir trasportati all'interno del Dominio tramite MPLS ad un altro ELSR.
- In corrispondenza di un indirizzo di destinazione e di un tipo di trattamento richiesto (QoS) viene definita una specifica **Forwarding Equivalent Class (FEC)**.
- Una **FEC** individua quindi
 - Un aggregato di pacchetti diretto ad una stessa destinazione (intesa o come destinazione finale o come **ELSR**) che devono avere lo stesso trattamento

Funzionamento di massima

- Quanto un ELSR riceve un pacchetto IP :
 - compie una operazione di “Classificazione”, ossia in base a quanto contenuto nell'intestazione identifica l'eventuale FEC di appartenenza;
 - inserisce fra l'intestazione di livello 2 e il pacchetto IP una **Label**.
- Tale **Label**:
 - Identifica la FEC a cui il pacchetto appartiene
 - Ha una lunghezza costante e breve
 - Ha un significato “locale alla linea”

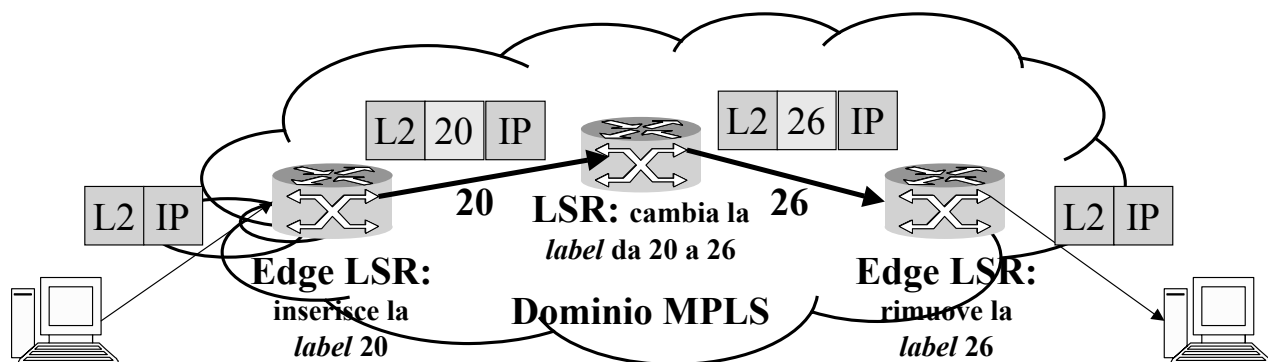
Funzionamento di massima

- Dall'**ELSR** di ingresso a quello di uscita tutte le operazioni di *forwarding* verranno effettuate utilizzando solo la **Label** e quindi l'intestazione del pacchetto IP non verrà più letta fino all'**ELSR** di destinazione.
- Gli **LSR** attraversati leggono la **Label**, trovano tramite essa in una tabella il FEC corrispondente ossia l'informazione sulla porta di uscita, la **Label** successiva ed il tipo di trattamento richiesto.
- L'**ELSR** di uscita (quindi l'ultimo nodo attraversato all'interno del Dominio MPLS) elimina la **Label** ed instrada il pacchetto nuovamente sulla base dell'indirizzo IP.

Lezione 1.4, v. 1.0

2.7

Un esempio



- Il percorso effettuato dai pacchetti appartenenti ad uno specifico **FEC** viene chiamato **Label-Switched Path (LSP)**

Lezione 1.4, v. 1.0

2.8

MPLS verso VC

- Due sono le differenze principali con un meccanismo a Circuito Virtuale tradizionale
 - Il livello di aggregazione; una FEC (e quindi una Label) aggrega in genere più flussi a partire dall'ELSR.
 - Non esiste una fase di *setup* del circuito attivata dalla sorgente, ma sono gli ELSR che attivano la procedura che introduce nelle tabelle di *forwarding* MPLS una nuova Label.

Vantaggi del MPLS 1/2

- Rispetto al *forwarding* IP tradizionale il modo di procedere di MPLS ha diversi vantaggi:
 - La procedura di *forwarding* richiede solo l'ispezione di una etichetta (Label) di dimensioni ridotte e l'esplorazione di una tabella relativamente semplice:
 - la ricerca in tabella non richiede un *longest prefix matching*.
 - La tabella di instradamento di MPLS può contenere molte informazioni sulle modalità di trattamento del pacchetto, mentre l'IP tradizionale può solo far uso delle informazioni contenute nell'intestazione del pacchetto.
 - » L'uso di queste informazioni, per altro, non rallenta molto il processo di *forwarding* perché esse non devono essere estratte dall'intestazione del pacchetto.

Vantaggi del MPLS 2/2

- Si posso definire percorsi che non siano decisi dinamicamente *hop-by-hop* ma scelti con opportune strategie di “Ingegneria del traffico”, questo permette di
 - » Ottenere un supporto efficace per la QoS.
 - » Migliorare l’utilizzo delle risorse di rete.
 - » Agevolare la realizzazione di alcuni servizi (VPN)

Struttura di MPLS

- MPLS prevede in sostanza tre passi fondamentali:
 - La definizione di una **FEC**
 - L’individuazione del percorso **LSP**.
 - La creazione (associazione al **FEC**) e distribuzione delle **Label** lungo il **LSP** (si osservi che questa operazione e la precedente si svolgono in modo parallelo e coordinato).
 - Il meccanismo di *forwarding* che comprende l’inserimento della **Label**, la commutazione sulla base di essa e la sua rimozione.

Trattamento dei pacchetti negli LSR

- Le operazioni che vengono effettuate sul pacchetto in transito nei LSR in relazione alle **Label** sono sostanzialmente tre:
 - *Pushing*, ossia l'inserimento della **Label**, che viene realizzata dall'**ELSR** di ingresso.
 - *Swapping*, ossia conversione dell'etichetta, realizzata contestualmente all'operazione di commutazione
 - *Popping*, ossia l'eliminazione di etichetta, effettuata dall'ultimo o dal penultimo LSR

Label e Label Binding

- Se R_s e R_d sono l'**LSR** sorgente e destinazione (intermedie) rispettivamente, possono accordarsi sul fatto che i pacchetti che R_s invierà a R_d con la **Label** L saranno legati al **FEC** F .
- Identificare una Label e assegnarla a una FEC significa realizzare un **Label Binding**
- L sarà quindi la **Label** di uscita per R_s e la **Label** di ingresso per R_d per quanto concerne il **FEC** F .
- Il valore di L sarà assegnato da R_d per evitare assegnamenti doppi (rispetto a R_s), ossia le Label vengono sempre assegnate dai router *downstream*.

Minimizzazione della dimensione delle tabelle MPLS

- Per mantenere il vantaggio di avere una commutazione molto veloce, si deve cercare di mantenere piccole le tabelle MPLS, ossia minimizzare il numero di **Label** usate per nodo.
- Ridurre il numero di **Label** significa pure ridurre il traffico di segnalazione
- La riduzione del numero di **Label** è ottenuta in tre modi
 - *Aggregation*
 - *Label Merging*
 - *Label stack*

Aggregation

- I **FEC** possono essere creati ogni volta che si vuole trattare un qualunque flusso di dati tramite MPLS
- I parametri che permettono di identificare nel **ELSR** i pacchetti appartenenti ad un **FEC** possono essere di vario tipo.
- I più immediati sono i *net-prefix* contenuti nel campo indirizzo di destinazione di IP.
- In ogni caso, ci si può trovare nelle condizioni di avere **FEC** diversi i cui flussi devono fare lo stesso percorso (quindi hanno lo stesso LSP).
- In questa situazione sarebbe opportuno non trattare questi flussi tramite **Label** diverse

Aggregation

- MPLS prevede la possibilità di associare più **FEC** ad una singola **Label** definendo una l'operazione di

Aggregation

che è realizzata derivando dai **FEC** aggregati un nuovo **FEC** “unione”.

- La decisione di chi deve realizzare la **Aggregation** e tra quali **FEC** è lasciata ai singoli LSR.

Label Merging

- Un altro caso che può presentarsi è l' avere pacchetti associati allo stesso **FEC** che arrivano ad un nodo con **Label** differenti (perché provenienti da **LSR** diversi).
- In questo caso è prevista la funzione di *Label Merging* tramite la quale un LSR può associare la stessa **Label** di uscita ai due flussi.
- Questa capacità non è imposta dallo standard a tutti i nodi.

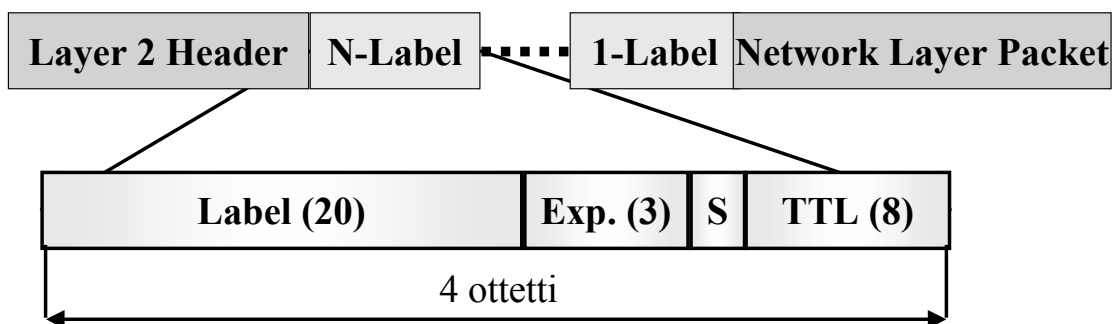
Label stack

- Le **Label** possono essere più di una per singolo pacchetto
- Una sequenza di **Label** viene indicata come **Label Stack**.
- Il **Label Stack** permette di realizzare una struttura gerarchica di instradamento/commutazione:
 - Il Dominio MPLS può essere suddiviso in sotto-domini ed i sotto-domini possono essere a loro volta suddivisi a più livelli.
 - Ogni sotto-dominio ha una **Label** specifica.
 - Il pacchetto entra nel Dominio con una singola **Label** e “guadagna” una nuova **Label** per ogni livello di sotto-dominio a cui accede.
 - All’uscita da ogni sottodominio viene eliminata la **Label** corrispondente.

Lezione 1.4, v. 1.0

2.19

Label



- **Label** (20 bits): l’etichetta vera e propria
- **Experimental** (3 bits): riservati per la sperimentazione
- **S** (1 bit): *Bottom of stack*, vale 1 se si tratta dell’ultima Label (1), vale 0 in tutti gli altri casi.
- **TTL** (8 bit): ci si copia al nodo di ingresso il TTL IP, il campo viene poi decrementato di 1 ogni LSR attraversato ed alla fine il valore viene inserito nel pacchetto IP.

Lezione 1.4, v. 1.0

2.20

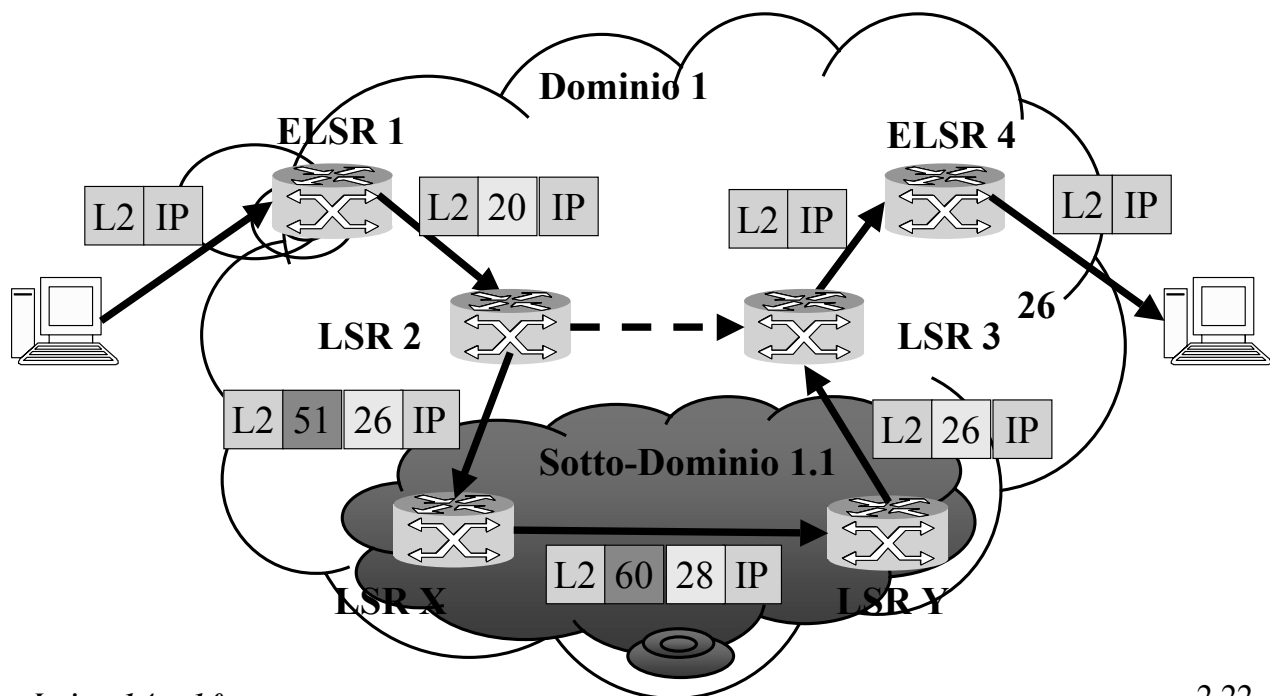
Label Popping

- Esiste un ulteriore modo per velocizzare le operazioni.
- Conviene fare il *popping* al penultimo LSR (***Penultimate Popping***), perché:
 - All'ultimo nodo il *forwarding* viene eseguito sulla base del pacchetto IP e quindi l'osservazione della Label è inutile
 - Lasciarlo significa costringere il nodo a cercare nella tabella MPLS per scoprire che deve eliminare la Label e quindi usare IP
- Non sempre si può fare il *Penultimate Popping* in quanto non è detto che l'LSR sia in grado di accorgersi di essere il penultimo del LSP.

Lezione 1.4, v. 1.0

2.21

Esempio 2



Lezione 1.4, v. 1.0

2.22

Selezione del percorso (LSP)

- La selezione del percorso avviene sfruttando le informazioni di instradamento fornite dai protocolli tradizionali (OSPF).
- Il metodo più immediato è quello di legare le FEC e i relativi LSP ai net-id o net-prefix IP.
- In pratica significa generare una FEC/Label e creare un corrispondente LSP per ogni linea della tabella d'instradamento IP del LSR.
- In questo caso l'aggregazione e la gerarchia potrebbe essere attivata seguendo l'aggregazione degli indirizzi IP.
- In alternativa la selezione del percorso può venir guidata in modo specifico.

Lezione 1.4, v. 1.0

2.23

Selezione del percorso (LSP)

- Ci sono due modalità per identificare il SLP
 - **Hop-by-Hop**
In questo caso ogni nodo associa ad una **FEC/Label** un *next-hop* in modo indipendente, basandosi per esempio sul protocollo di instradamento in uso (la sceglie sulla base del prefisso IP di destinazione collegato alla FEC).
 - **Explicit Routing**
E' in sostanza una forma di *source-routing* in cui l'ELSR di ingresso può stabilire a priori il percorso. La differenza sostanziale dal *source-routing* IP è che non richiede una complessa e lenta elaborazione dell'intestazione. Il percorso può essere scelto per configurazione o dinamicamente e può essere di tipo "loose" o "strict".

Lezione 1.4, v. 1.0

2.24

Controllo di un LSP

- Quando i **FEC** vengono fatti corrispondere a un prefisso di rete, si può procedere in due modi per l'attivazione delle **Label**:
 - ***Independent LSP Control***
 appena un nodo riconosce la presenza di un **FEC** (tramite la modifica indotta da un protocollo di instradamento della tabella IP) autonomamente realizza un **Label Binding** con i suoi *peer*.
 - ***Ordered LSP Control***
 il **Label Binding** può essere attivato solo dal **ELSR** di uscita (eventualmente “inizializzato” da quello di ingresso) o da un LSR che ha già ricevuto un *binding* dal suo *next-hop* per quel FEC.

Lezione 1.4, v. 1.0

2.25

Distribuzione delle Label

- La richiesta per un *binding* Label/FEC può partire sia da un LSR *upstream* (sorgente) che *downstream* (destinazione).
- L'assegnazione deve essere fatto in direzione *downstream* (a partire dalla destinazione)
- Un LSR *upstream* segnala al LSR *downstream* *binding* tramite uno specifico protocollo
- La segnalazione si realizza fra coppie (*peer*) di LSR.

Lezione 1.4, v. 1.0

2.26

Distribuzione delle Label

- Sono previsti quattro possibili protocolli per la distribuzione di Label
 - **Label Distribution Protocol (LDP)**: usato per le FEC create in corrispondenza di indirizzi IP unicast.
 - **RSVP, Constraint-based Routing LDP (CR-LDP)**: usati per assegnare Label per il Traffic Engineering e prenotazione di risorse.
 - **Protocol Independent Multicast (PIM)**
 - **BGP**: usato per Label esterne all'*autonomous system* (per esempio nelle VPN).

Distribuzione delle Label

- Le richieste di Label *binding* possono avvenire in due modi:
 - ***Downstream on demand***: su richiesta di un LSR al suo next-hop
 - ***Unsolicited Downstream***: un LSR può segnalare ai suoi predecessori un *binding* anche se questo non è stato esplicitamente richiesto

Instaurazione del LSP

- Due alternative:
 - ***Independent LSP Control***
 Ogni LSR, nel momento in cui riconosce la presenza di un nuovo **FEC** (per esempio in corrispondenza di un aggiornamento della tabella di instradamento IP da parte di un IGP) assegna una nuova **Label** e la distribuisce ai *peer*. Questo metodo non è attuabile con un *Explicit Routing*.
 - ***Ordered LSP Control***
 In questo caso l'**LSR** assegna l'etichetta solo se:
 - » è l'**ELSR** di uscita (finale) per quel **FEC**;
 - » se riceve per quel **FEC** un **LSR** in *downstream*.

Tunneling

- E' noto che in IP è possibile e a volte necessario realizzare un "*tunneling*".
- Tale *tunneling* avviene attraverso l'incapsulamento di un pacchetto IP in un altro pacchetto IP.
- MPLS permette di realizzare un *tunneling* IP in modo più efficiente, associando un FEC al traffico da trasportare e definendo il nodo di uscita come ELSR di destinazione.
- Sono previste due tipologie di *tunnels*
 - ***Explicitly routed***
 - ***Hop-by-hop***

Applicazioni di MPLS

- MPLS può essere usato per
 - gestire la QoS.
 - realizzare in modo più efficiente ed efficace una rete Diffserv
 - applicare tecniche di *Traffic Engineering* alle reti IP
 - realizzare in modo efficace un servizio di Virtual Private Network.

Traffic Engineering (TE)

- Il TE consiste nella ottimizzazione delle prestazioni di reti operative.
- In linea generale implica l'applicazione di tecnologie e principi scientifici concernenti la misura, la modellizzazione, la caratterizzazione e il controllo del traffico IP per raggiungere degli obiettivi di performance predefiniti.
- L'obiettivo principale è rendere più efficienti e affidabili le operazioni di rete ottimizzando contemporaneamente l'uso delle risorse e le prestazioni del traffico.
- MPLS fornisce degli strumenti per applicare tecniche di TE alle reti IP per quanto concerne la misura e il controllo.

Obiettivi del TE

- Si possono classificare in due insiemi principali:
 - Orientati al traffico
 - Si tratta principalmente di migliorare la QoS dei flussi di traffico
 - » Nel caso BE principalmente la perdita
 - » Nel caso di più classi di servizio anche il ritardo
 - Orientati alle risorse
 - Si tratta di ottimizzare l'uso delle risorse, in particolare la gestione efficiente della banda, evitando di trovarsi in condizioni in cui parte delle risorse sono sovraccariche e parte sotto-utilizzate.

Obiettivi del TE

- Dal punto di vista del traffico l'obiettivo complessivo potrebbe essere riassunto come:
 - minimizzazione della congestione “prolungata”**
- La congestione si presenta in genere in due casi:
 - Risorse inadeguate rispetto al carico offerto.
 - **Flussi di traffico distribuiti in modo inefficiente nella rete.**
- Il primo caso può essere affrontato usando due metodi (applicabili anche contemporaneamente):
 - Espansione della capacità
 - Meccanismi di controllo di congestione su traffici elastici (controllo dei tassi, controllo di flusso a finestra, gestione delle code,

Traffic Engineering (TE)

- Il secondo caso può essere affrontato dal TE.
- Il TE dovrebbe in pratica fornire le funzionalità di un meccanismo di controllo retroazionato adattativo composto da:
 - Un insieme di elementi di rete interconnessi
 - Un sistema di *monitoring* delle prestazioni della rete
 - Un insieme di strumenti di gestione delle configurazioni
- L'ingegnere del traffico stabilisce una politica di controllo, osserva il sistema tramite il *monitoring*, applica delle azioni di controllo per guidare la rete nello stato desiderato in accordo con la politica definita.
- Le azioni possono essere prese in risposta ad uno stato corrente o sulla base di previsioni ottenute tramite modelli previsionali.

Traffic Engineering (TE)

- Le azioni di controllo dovrebbero comprendere:
 - La modifica dei parametri di gestione del traffico (parametri di *scheduling*, di *queue management*, ...)
 - La modifica di parametri associati al routing (metriche, costi, percorsi)
 - La modifica di attributi e vincoli associati alle risorse.
- Il livello di intervento manuale dovrebbe essere ridotto al minimo.

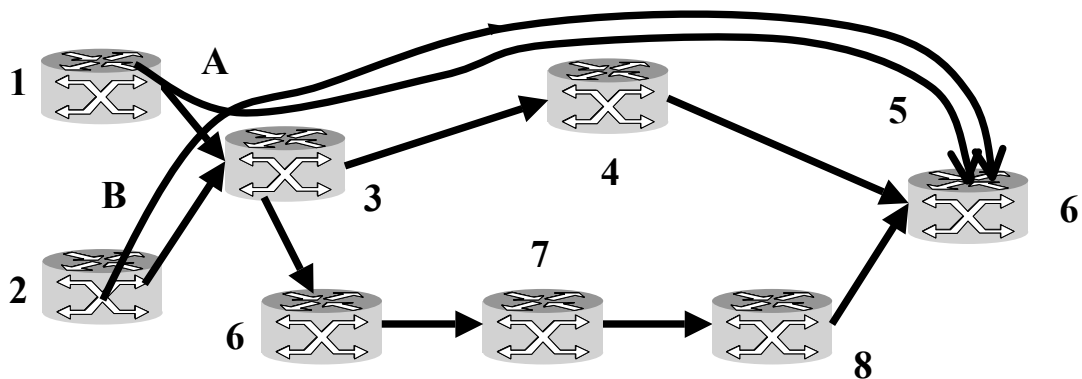
Perchè MPLS per il TE

- In linea di principio gli IGP (Interior Gateway Protocol) potrebbero essere lo strumento più naturale per l'applicazione delle tecniche di TE.
- Nella pratica non sono adatti a fornire un supporto adeguato ed anzi sono spesso fonte di inefficienze perché:
 - Sono basati su meccanismi *shortest path* che generano spesso condizioni di uso inefficiente delle risorse
 - Usano metriche semplici ed additive che non tengono conto della disponibilità di banda e delle caratteristiche del traffico
- E quindi il risultato è che generano condizioni di congestione quando
 - Gli *shortest path* di più flussi convergono su una singola linea/interfaccia
 - Un flusso viene instradato verso una linea/interfaccia che non ha sufficiente banda disponibile.

Lezione 1.4, v. 1.0

2.37

Perchè MPLS per il TE



- Sulla base degli algoritmi di instradamento tradizionali lo *shortest path* porta spesso ad una sotto-utilizzazione delle linee.

Lezione 1.4, v. 1.0

2.38

Perchè MPLS per il TE

- La soluzione “tradizionale” con la quale si è fino ad oggi operato è stato quello di utilizzare IP over ATM o IP over Frame-relay.
- In pratica si utilizzano i VP (*Virtual Circuit*) ATM per creare delle topologie virtuali fra router indipendenti dalla rete fisica effettivamente presente. Tali topologie virtuali sono considerate dagli IGP come reali (ossia un VP appare al router come una linea fisica).
- Questo modo di operare tramite un *overlay model* ha diversi vantaggi:
 - Permette di realizzare tramite VC un routing vincolato
 - Permette la configurazione di percorsi su base amministrativa
 - Permette la aggregazione di percorsi
 - Da un supporto per il CAC, *policing* e *shaping*.

Perchè MPLS per il TE

- In questo senso MPLS può essere applicato al posto di ATM e Frame-relay, con evidenti vantaggi quali ad esempio:
 - il livello di integrazione con IP
 - la possibilità di automatizzare gli interventi di TE.
 - Il minor impegno in termini di complessità e segnalazione introdotto.