

Une brève introduction à MPLS

Chaput Emmanuel



ENSEEIHT

2024-2025



Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

1 / 139

Plan I

- 1 Introduction
- 2 Le plan de données
- 3 La signalisation : le protocole LDP
- 4 La signalisation : le protocole RSVP
- 5 L'ingénierie de trafic avec MPLS
- 6 Les VPNs fondés sur MPLS/BGP
- 7 Références bibliographiques

Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

2 / 139

Introduction

- 1 Introduction
 - L'origine de MPLS
 - Le succès de MPLS
 - Les principes de base de MPLS
- 2 Le plan de données
- 3 La signalisation : le protocole LDP
- 4 La signalisation : le protocole RSVP
- 5 L'ingénierie de trafic avec MPLS
- 6 Les VPNs fondés sur MPLS/BGP

Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

3 / 139

Qu'est-ce que MPLS ?

- MultiProtocol Label Switching
- Protocole permettant d'acheminer des données dans un réseau
- Fondé sur la notion de commutation de *label*
- Capable de véhiculer presque tout type de données
 - *e.g.* IP, ATM, Ethernet, ...
- Capable de fonctionner sur tout type de réseau
 - *e.g.* IP, ATM, Ethernet, ...
- Introduit pour profiter du meilleur des deux mondes
 - Le mode *datagramme* de IP
 - Le mode *circuit virtuel* d'ATM
- MPLS s'est révélé un puissant outil d'ingénierie de trafic

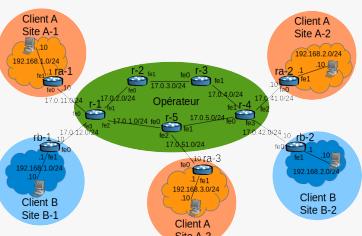
Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

4 / 139

Quel est le problème avec IP ?



- Comment acheminer le trafic vers chaque site de chaque client ?
- Comment réservrer des ressources pour garantir un service ?
- Comment router avec un tel plan d'adressage ?

Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

5 / 139

L'origine de MPLS

- 1 Introduction
 - L'origine de MPLS
 - Le succès de MPLS
 - Les principes de base de MPLS

Chaput Emmanuel

Une breve introductionàMPLS

2024-2025

6 / 139

L'origine de MPLS

- Quel déploiement de IP sur ATM ?
 - LAN Emulation [15]
 - Interconnexion de réseaux locaux au travers d'ATM
 - Continuité de service, transparence
 - Maintient de la diffusion (via le BUS), de l'adressage (via le LES)
 - Difficultés de passage à l'échelle (LANE 1.0), pas de QoS (introduite dans LANE 2.0 [12, 16])
 - Classical IP [19, 25, 20]
 - Gestion des sous-réseaux (LIS)
 - Résolution d'adresse (ATMARP)
 - Encapsulation LLC/SNAP sur AAL5
 - Inconvénient dans CLIP
 - Le routage entre LIS interdit les connexions ATM de bout en bout
 - Les communications passent par le routeur, ...
 - Introduction de NHRP [21]
 - NBMA Next Hop Resolution Protocol
 - Permet de connaître l'adresse (de "niveau 2") du prochain *hop* au travers d'un NBMA
 - Autorise les connexions ATM entre membres de plusieurs LIS sans routage



Introduction de MPOA

- *MultiProtocol Over ATM* [13, 14]
- Intégration de NHRP et LANE
- Principes
 - Les clients MPOA (MPC) établissent des connexions ATM directes entre eux
 - L'adresse ATM du MPC cible est déterminée en interrogeant le serveur MPOA (le MPC)
 - La demande d'adresse est routée et se propage au travers de NHRP
- Les données circulent ensuite au travers de la connexion via une encapsulation LLC/SNAP (et éventuellement une étiquette)

Routage ou commutation ?

- Routage (IP, Ethernet, ...)
 - Chercher dans une table de routage
 - Flexibilité (multicast, anycast, ...)
 - Dynamique
- Commutation (X.25, ATM, Frame Relay, ...)
 - Consulter une table de commutation
 - Établissement d'un circuit virtuel
- MPLS a pour objectif initial de profiter des avantages de chacun [30]
 - Souplesse du routage IP
 - Efficacité de la commutation ATM

Le succès de MPLS

1 Introduction

- L'origine de MPLS
- Le succès de MPLS
- Les principes de base de MPLS

Le succès de MPLS

- MPLS ne doit pas son succès à ses objectifs initiaux
 - Le routage IP est très efficace
 - ATM n'a pas eu le déploiement espéré
- C'est en revanche un excellent outil d'ingénierie de trafic [6]
 - Définition de classes de trafic
 - Classification en entrée du réseau puis commutation simple dans le cœur
 - Chaque classe peut être traitée (routée, ...) spécifiquement
 - Sans remettre en cause IP
- Il peut encapsuler divers protocoles
 - Et permettre ainsi la mise en place de VPN
- Il peut être étendu pour gérer des technologies très variées
 - Generalized MPLS [22]

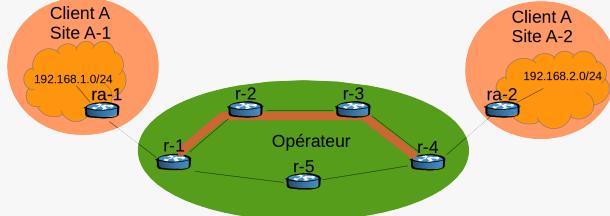
Les principes de base de MPLS

1 Introduction

- L'origine de MPLS
- Le succès de MPLS
- Les principes de base de MPLS

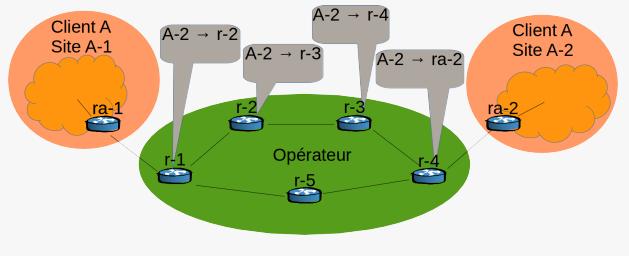
Les principes de base de MPLS

- Utiliser les avantages des deux mondes
 - Profiter de la souplesse du routage (par exemple de IP)
 - Utiliser la performance de la commutation (par exemple de ATM)
- Plutôt que router (puis commuter) sur chaque routeur chaque paquet émis depuis Site A-1 vers Site A-2
 - Routons-les en r-1 dans un "conduit"
 - Communions les en fonction de ce conduit sur r-2, r-3 et r-4
 - Routons-les vers Site A-2 en sortie du conduit



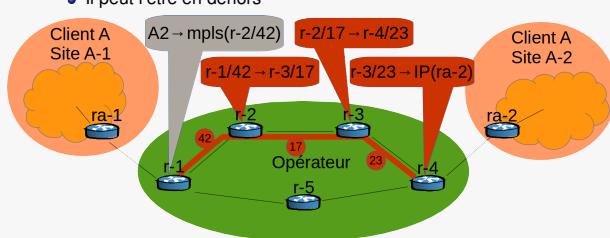
Un exemple simple (1/2)

- IPv4 sans MPLS
 - Le paquet est routé sur chaque entité IP



Un exemple simple (2/2)

- IPv4 avec MPLS
 - Le paquet n'est plus routé à l'intérieur du "conduit"
 - Il peut l'être en dehors



Un peu de vocabulaire

- Un chemin MPLS, ou LSP (*Label Switched Path*)
 - Constitue un "conduit" depuis un point vers un autre
 - Il est unidirectionnel
- Les routeurs MPLS ou LSR (*Label Switching Router*)
 - Routeur d'entrée (I-LER : *ingress LER* ou *Label Edge Router*)
 - Routeurs de cœur
 - Routeur de sortie (E-LER : *egress LER*)
- Une classe d'équivalence ou FEC (*Forwarding Equivalence Class*)
 - Regroupe un ensemble de paquets devant être acheminés dans le même LSP
 - Par exemple en fonction de leur adresse destination
- Un label
 - Est un identifiant dans un paquet MPLS permettant de désigner sa FEC
 - Il peut être ajouté dans un en-tête spécifique ou utiliser un champ existant (eg les VPI/VCI d'ATM)

Introduction

1 Introduction

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

3 La signalisation : le protocole LDP

4 La signalisation : le protocole RSVP

5 Les VPNs fondés sur MPLS/BGP

7 Références bibliographiques

Les labels

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

Les labels MPLS

- L'information en fonction de laquelle un "paquet MPLS" est communiqué est un **label**
 - Négocié entre deux LSR (ou LER)
 - Portée locale
 - À comparer aux identifiants de canaux logiques x.25 par exemple
- Peut utiliser des mécanismes existants (VPI/VCI, DLCI, ...)
- Shim label entre L2 et L3 :
 - **Label** l'identifiant à proprement parler
 - **Exp** pour un usage expérimental (QoS [1], ...)
 - **S** pour l'empilement des labels (voir plus loin)
 - **TTL** *Time To Live*
- Certaines valeurs sont réservées



Exemple

Wireshark screenshot showing an MPLS label switched packet (0x8847). The packet structure is highlighted with a red box, showing fields like L2 (Ethernet II), Label (0x8847), Exp (0x00), S (0x01), TTL (253), and L3 (IP).

L'empilement des labels

- 2 Le plan de données
- Les labels
 - L'empilement des labels
 - Fonctionnement d'un LSR
 - Gestion du TTL
 - Le Penultimate Hop Popping (PHP)
 - Les tables de correspondances
 - Illustration sur un exemple
 - Annexe : la configuration des équipements

L'empilement des labels

- Les LSP MPLS peuvent "s'emboîter" les uns dans les autres [29]
 - Conduits de différentes portées et différentes granularités
- Les labels sont alors empilés
 - Une pile de labels est en fait une séquence de ces labels
 - Le premier label de la séquence est le sommet de la pile
 - Le dernier label est identifié par un bit **S = 1**
- Un paquet MPLS est traité en fonction de son label de sommet de pile
 - Cela peut être une commutation classique, ou SWAP
 - Un dépilement (suppression du label de sommet) ou POP
 - Un empilement (ajout d'un label) ou PUSH

Fonctionnement d'un LSR

- 2 Le plan de données
- Les labels
 - L'empilement des labels
 - Fonctionnement d'un LSR
 - Gestion du TTL
 - Le Penultimate Hop Popping (PHP)
 - Les tables de correspondances
 - Illustration sur un exemple
 - Annexe : la configuration des équipements

Fonctionnement d'un LSR

- Le **ingress LER**
 - Doit "insérer" les paquets dans un LSP
 - Par exemple suite à une décision de routage
 - Utilise le label négocié avec le LSR aval
 - Le label est éventuellement empilé (PUSH)
- Un LSR de cœur
 - Commute les paquets en fonction de leur label
 - Les place sur la voie de sortie correspondante vers le *NextHop*
 - Change la valeur du label (SWAP)
- Le **egress LER**
 - Doit "extraire" les paquets du LSP
 - Ils seront ensuite injectés dans le système correspondant (eg routés)
 - Le label est supprimé (POP)

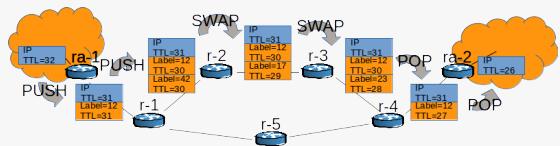
Gestion du TTL

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

Gestion du TTL

- Gestion cohérente avec celle de IP
- Seul le label en haut de pile est traité
- En cas de TTL à 0, le message ICMP est envoyé *dans le sens des données*!



(Ici on suppose que le client fait lui aussi du MPLS)

Le Penultimate Hop popping (PHP)

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

Le Penultimate Hop Popping (PHP)

- Un LSR peut dépiler le label situé en sommet de pile
 - Avant de faire passer le paquet au LER
 - Cela permet de décharger ce dernier
 - Signalisation par un numéro de label spécifique
 - Label 3 (dit *implicit NULL*)
 - Annoncé par le LER
- Utile en particulier dans les VPN MPLS
 - Un label supplémentaire est utilisé pour identifier les VLAN
 - Le LER n'a donc plus que ce label à gérer
- Inconvénient : le champ EXP est perdu
- Utilisation d'un label spécifique
 - Label 0 (dit *explicit NULL*)
 - Le LER le supprime, mais ne fait aucune recherche

Les tables de correspondances

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

Les tables de correspondances

- Elles sont décrites dans les standards [30]
- Elles permettent de spécifier
 - Comment acheminer un paquet doté d'un label (c'est la NHLFE)
 - Comment traiter un paquet reçu avec un label (c'est la ILM)
 - Comment insérer un paquet dans un LSP en fonction de sa FEC (c'est la FTN)
- Leur mise en œuvre effective peut varier d'une implantation à l'autre
- En lien étroit avec les outils de routage du système (en particulier sur les LER)

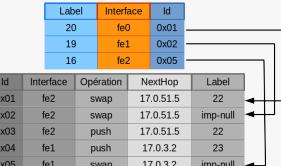
La Next Hop Label Forwarding Entry (NHLFE)

- Elle permet au LSR de déterminer comment traiter le paquet
- Elle contient les éléments suivants
 - L'interface (et la méthode d'encapsulation et d'encodage des labels) vers le prochain LSR
 - L'action à mener sur les labels
 - Remplacer le label au sommet
 - Remplacer le label au sommet et en empiler d'autres
 - Dépiler des labels

Id	Interface	Opération	NextHop	Label
0x01	fe2	swap	17.0.51.5	22
0x02	fe2	swap	17.0.51.5	imp-null
0x03	fe2	push	17.0.51.5	22
0x04	fe1	push	17.0.3.2	23
0x05	fe1	swap	17.0.3.2	imp-null

La Incoming Label Map (ILM)

- Utilisée pour les paquets entrants avec un label
 - Table de commutation
- Elle donne une correspondance entre un label d'entrée et une NHLFE
 - Sinon le paquet est routé (sur le e-LER)



La FEC-to-NHLFE map (FTN)

- Établit une correspondance entre une FEC et une NHLFE
 - Routage d'une classe de trafic vers un LSP MPLS
 - Cela peut être plus qu'une "simple" route
 - En fonction du protocole, de la source, ...

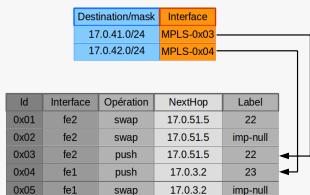
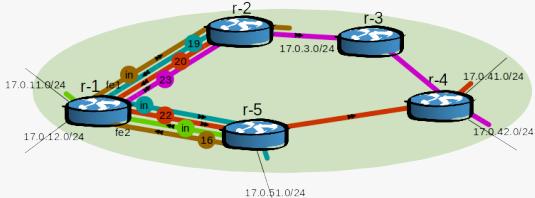


Illustration sur un exemple

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple
- Annexe : la configuration des équipements

Illustration sur un exemple



(Observons quelques LSP sur r-1)

Consultation (partielle) sur Cisco (1/2)

```
r-1#sh mpls ip binding detail
17.0.3.0/24, rev 4
  in label: 16
    Advertised to:
      17.0.3.2:0          17.0.51.5:0
      out label: 16        lsr: 17.0.51.5:0
      out label: imp-null  lsr: 17.0.3.2:0      inuse
17.0.11.0/24, rev 12
  in label: 1
    Advertised to:
      17.0.3.2:0          17.0.51.5:0
      out label: 19        lsr: 17.0.3.2:0
      out label: 20        lsr: 17.0.51.5:0
17.0.41.0/24, rev 18
  in label: 20
    Advertised to:
      17.0.3.2:0          17.0.51.5:0
      out label: 22        lsr: 17.0.3.2:0
      out label: 22        lsr: 17.0.51.5:0      inuse
17.0.42.0/24, rev 20
  in label: 21
    Advertised to:
      17.0.3.2:0          17.0.51.5:0
      out label: 23        lsr: 17.0.3.2:0
      out label: 23        lsr: 17.0.51.5:0      inuse
17.0.51.0/24, rev 16
  in label: 19
    Advertised to:
      17.0.3.2:0          17.0.51.5:0
      out label: 21        lsr: 17.0.3.2:0
      out label: imp-null lsr: 17.0.51.5:0      inuse
```

Consultation (partielle) sur Cisco (2/2)

```
r1#show mpls forwarding-table
Local Outgoing Prefix Bytes tag Outgoing Next Hop
tag tag or VC or Tunnel Id switched interface
16 Pop tag 17.0.3.0/24 4560 Fa1/0 17.0.2.2
17 Pop tag 17.0.5.0/24 0 Fa2/0 17.0.1.5
18 17.0.4.0/24 0 Fa1/0 17.0.2.2
19 17.0.4.0/24 1376 Fa2/0 17.0.1.5
20 Pop tag 17.0.41.0/24 0 Fa2/0 17.0.1.5
21 17.0.41.0/24 7198 Fa1/0 17.0.1.5
22 17.0.42.0/24 4592 Fa2/0 17.0.1.5
```

Chaput Emmanuel

Une brève introduction à MPLS

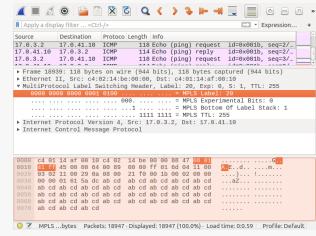
2024-2025

37 / 139

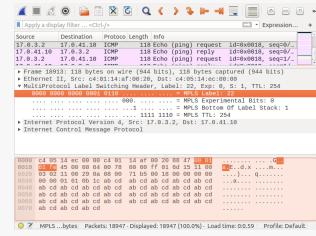
Observation des messages (1/2)

Observez un paquet venant de 17.0.3.0/24 et allant vers 17.0.41.0/24

- Entre r-2 et r-1



- Entre r-1 et r-5



Observation des messages (2/2)

Observez un paquet venant de 17.0.41.0/24 et allant vers 17.0.3.0/24

- Entre r-5 et r-1

Annexe : la configuration des équipements

2 Le plan de données

- Les labels
- L'empilement des labels
- Fonctionnement d'un LSR
- Gestion du TTL
- Le Penultimate Hop Popping (PHP)
- Les tables de correspondances
- Illustration sur un exemple

Annexe : la configuration des équipements

- Les outils Linux
- Les outils Cisco

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

40 / 139

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

39 / 139

Dans le noyau Linux

Le noyau Linux intègre progressivement les fonctionnalités MPLS

- Depuis le noyau 4.1
- LER, LSR
- ...
- Uniquement plan de données
- Configuré par la commande `sysctl`
 - Classique sous Linux
- Manipulé au travers de la commande `ip`
 - Pas de nouvelle commande spécifique
 - Bien intégré avec le reste du réseau
- Plan de contrôle
 - LDP, RSVP, ...
 - Dans quagga, un jour peut-être ?
 - Dans FRR, le protocole LDP est disponible

2 Le plan de données

- Annexe : la configuration des équipements
 - Les outils Linux
 - Les outils Cisco

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

41 / 139

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

42 / 139

Configuration du noyau

- Il faut charger les modules (un pour le routage, l'autre pour le *forwarding*, le troisième pour l'*offloading*)

```
# modprobe mpls_router
# modprobe mpls_gso
# modprobe mpls_iptunnel
```
- Activation d'une interface (sinon, tout paquet MPLS reçu par cette interface est détruit)

```
# sysctl -w net.mpls.conf.eth0.input=1
```
- Définir le nombre de labels du système (par défaut 0, donc pas de label !)

```
# sysctl -w net.mpls.platform_labels=1048575
```

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

43 / 139

Configuration du noyau

Attention au *Reverse Path Filtering* [7]

- Principe : rejeter un paquet dont l'interface d'arrivée ne mène pas vers son adresse source
 - Technique de protection contre le *spoofing*
 - Un paquet "sortant" d'un LSP MPLS ne peut alors pas être "ré-injecté" sur le LER
 - Il peut cependant être transmis à une autre machine (réalisant ainsi du PHP)
 - Pour éviter cette difficulté, on peut déverrouiller le filtrage sur l'interface de *loopback*
- ```
sysctl -w net.ipv4.conf.all.rp_filter=0
sysctl -w net.ipv4.conf.lo.rp_filter=0
```

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

44 / 139

## Router du trafic dans un LSP

### Configuration d'un LER *ingress*

- ```
# ip route add 192.168.10.0/24 encap mpls 100 via inet
192.168.20.12
```
- Route le trafic destiné au réseau cible
 - Ici 192.168.10.0/24
 - L'encapsule dans du MPLS avec le label précisé
 - Ici 100
 - Envie le paquet au *Next Hop* fourni
 - Ici en IPv4 à l'adresse 192.168.20.12

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

45 / 139

Router du trafic MPLS

Configuration d'un LER *egress*

- ```
ip -f mpls route add 200 via inet 192.168.30.2 dev eth0
```
- On commute le trafic MPLS entrant avec le label indiqué
    - Ici 200
  - On envoie le trafic désencapsulé vers le *Next Hop* spécifié
    - Ici en IPv4 à l'adresse 192.168.30.2
  - Attention si c'est une adresse locale
    - Désactiver le *Reverse Path Filtering*!

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

46 / 139

## Commuter du trafic MPLS

### Configuration d'un LSR

- ```
# ip -f mpls route add 200 as 300 via inet 192.168.30.2
dev eth0
```
- On commute le trafic entrant avec le label indiqué
 - Ici 200
 - On l'envoie vers l'interface définie
 - Ici en IPv4 vers 192.168.30.2
 - En remplaçant le label par la valeur fournie
 - Ici le label 300

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

47 / 139

Un exemple



- On souhaite acheminer du trafic entre les sites de chacun des clients
 - ... dans les deux sens !
- Au travers de 2 LSPs MPLS construits à la main
- Puis observer le trafic sur les routeurs de l'opérateur
- Télécharger et installer les fichiers de configuration
 - <http://www.enseeih.fr/~chaput/exercices/mpls-linux-2.bz2>
- Utilisable sous une version récente de Linux (nécessite Linux 4.1)
 - Attention également à la version de *iproute2*

Fichiers de configuration



Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

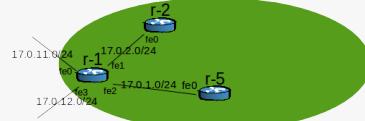
48 / 139

Les outils Cisco

2 Le plan de données

- Annexe : la configuration des équipements
 - Les outils Linux
 - Les outils Cisco

Configuration des équipements



Activation de MPLS

```
r-1(config)# mpls ip
```

Activation de MPLS sur une interface

```
r-1(config)#interface fastEthernet 1/0
r-1(config-if)#mpls ip
```

Observation de la configuration

```
r-1#show mpls interfaces
Interface          IP      Tunnel   Operational
FastEthernet1/0    Yes (ldp)  No       Yes
FastEthernet2/0    Yes (ldp)  No       Yes
r-1#
```

La signalisation : le protocole LDP

1 Introduction

2 Le plan de données

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- Les sessions ciblées
- Annexe : la configuration des équipements

4 La signalisation : le protocole RSVP

Le protocole LDP

Label Distribution Protocol [2, 3]

- Conçu dans le cadre du groupe de travail MPLS de l'IETF
- Son objectif est de distribuer des labels le long d'un LSP
- Se fonde sur un protocole de routage existant

Facilement extensible

- Utilise la notion de TLV (Type/Length/Value)

Plusieurs étapes principales

- Découverte des pairs (*peers*)
 - Par un protocole *Hello*
- Établissement et maintien de sessions entre pairs
 - Échange de labels
 - Échange de FECs

5 Les VPNS fondés sur MPLS/BGP

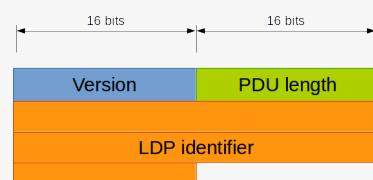
6 Références bibliographiques

Les messages LDP

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- Les sessions ciblées
- Annexe : la configuration des équipements

Les messages LDP

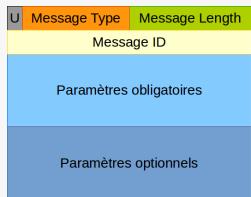


Entête commun suivi d'un ou plusieurs messages

LDP Identifier l'espace des labels

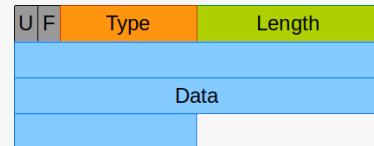
- Les quatre premiers octets identifient le pair et doivent être uniques
- Les deux derniers identifient l'espace de nommage sur le pair (0 si espace de nommage global au système)

Les messages LDP



- Le bit **U** signifie qu'un message inconnu peut être ignoré en silence (à 0 il signifie qu'il faut alerter l'émetteur)
- Les paramètres (obligatoires comme optionnels) dépendent de la nature du message

Les TLV



- Le bit **U** signifie qu'un TLV inconnu peut être ignoré en silence (à 0 il signifie qu'il faut alerter l'émetteur et détruire l'intégralité du message)
- Le bit **F** (significatif si **U = 1**) stipule qu'un TLV inconnu doit être propagé

Découverte des pairs

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- Les sessions ciblées
- Annexe : la configuration des équipements

Découverte des pairs

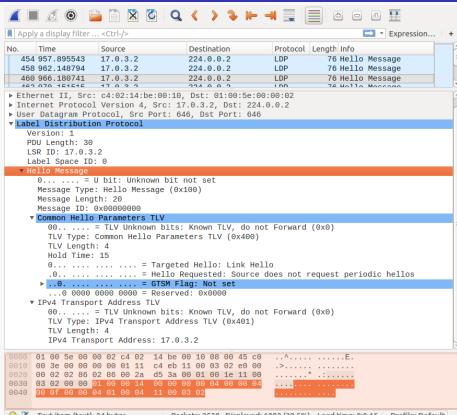
Deux techniques de découverte des pairs

- Mécanisme de base (*Basic Discovery Mechanism*)
 - Pour des pairs en lien direct
 - Émission périodique de messages *Hello* en UDP vers une adresse multicast et un port définis
- Mécanisme étendu (*Extended Discovery Mechanism*)
 - Pour des pairs non directement reliés
 - Émission périodique de messages *Hello* en UDP vers l'adresse unicast du paire à un port défini

Échange de messages *Hello*

- Délai de validité (*Hold Time*), adresse à utiliser, ...

Découverte des pairs



Établissement de la session

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- Les sessions ciblées
- Annexe : la configuration des équipements

Établissement de la session

- Une session est établie entre deux pairs
 - Une pour chaque espace de nommage de labels
 - Le protocole de transport utilisé est TCP
 - Établissement de la connexion TCP
 - Utilisation d'adresses annoncées dans le protocole Hello (ou pas)
 - Détermination de qui joue le rôle actif/passif (comparaison des adresses)
 - Ouverture de la connexion par l'actif
 - Initialisation de la session LDP avec échange de paramètres
 - Version du protocole
 - Méthode de distribution des labels
 - Timers
 - Plages de VPI/VCI ou DLCI ou autres
 -

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 61 / 139

Une brève introduction à MPLS

2024-2025

61 / 139

Établissement de la session

```
Apply a display filter: <Ctrl+>
No. Time Source Destination Protocol Length Info
664 1250.996389 17.0.3.2 17.0.4.3 LDP 98 Initialization Mes...
665 1251.001389 17.0.3.2 17.0.4.3 TCP 60 64962 - 646 [ACK]
666 1251.001409 17.0.3.2 17.0.4.3 UDP 60 64962 - 646 [ACK]
667 1251.020857 17.0.3.2 17.0.4.3 TCP 60 646 - 64962 [ACK]
668 1251.020873 17.0.3.2 17.0.4.3 TCP 1 no 356 Address Message Is...
[1 rows]
View: List
PDU Length: 40
RSU ID: 17.0.3.2
Label Space: 14
Initialization Message
  0... .... 5 = bit: Unknown bit not set
  Message Type: Common Session Initialization Message (0x208)
  Message Length: 22
  Message ID: 0x000000014
Common Session Parameters TLV
  00... 1 = TLV Unknown bits: Known TLV, do not Forward (0x0)
  TLV Type: Common Session Parameters TLV (0x500)
  TLV Length: 14
  + Parameters:
    Session Protocol Version: 1
    Session Keepalive Timer: 180
    0... 1 = Session Label Advertisement Discipline: Downstream Unsolicited proposed
    0... .... 5 = Session Loop Detection: Loop Detection Disabled
    Session Path Vector Limit: 0
    Session Max Hops: 0
    Session Receiver LSR Identifier: 17.0.4.3
    Session Receiver Label Space Identifier: 0
  * Keep Alive Message
    0... .... 5 = bit: Unknown bit not set
    Message Type: Keep Alive Message (0x201)
    Message Length: 20
    Message ID: 0x000000015
[1 rows]
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025

Une brève introduction à MPLS

2024-2025

62 / 139

Spécification des FEC

- Une FEC est définie par un ensemble de *FEC elements*
 - Un *FEC element* est une définition de paquets à inclure dans la FEC
 - Par exemple un *Prefix FEC element* permettant de définir une plage d'adresse
 - Mais aussi *Pseudo Wire identifier FEC element* [23]



Chaput Emmanuel Une brève introduction à MPLS 2024-2025 63 / 139

Une brève introduction à AMPL

2024-2025

63 / 139

Spécification des FEC

Chaput Emmanuel | Une brève introduction à l'AMPL | 2024-2025 | 65 / 139

Une brève introduction à MBSR

2024-2025

65 / 139

Échange des labels

- 3 La signalisation : le protocole LDP
 - Les messages LDP
 - Découverte des pairs
 - Établissement de la session
 - Spécification des FEC
 - **Échange des labels**
 - Les sessions ciblées
 - Annexe : la configuration des é

Chaput Emmanuel | Une brève introduction à MELIS | 2024-2025

Une brève introduction à MPI 5

2024-2025

66 / 139

Deux modes de distribution

- **Downstream on Demand label distribution**
 - Demande par le LSR amont d'un mapping entre une FEC et un label
- **Downstream Unsolicited label distribution**
 - Le LSR aval peut également envoyer une telle correspondance "spontanément"
- MPLS prévoit les deux techniques, sans préférence
 - Les deux peuvent être implantées
 - Y compris simultanément
 - Les deux LSR doivent donc se mettre d'accord sur la technique à utiliser

Les sessions ciblées

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- **Les sessions ciblées**
- Annexe : la configuration des équipements

Les sessions ciblées

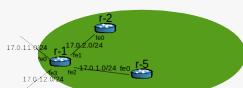
- Les sessions LDP sont établies entre voisins directs
 - Suite à l'utilisation de messages `hello` multicastés sur les liens qui les relient
- Comment établir une session entre des routeurs distants ?
 - On parle de *Targeted LDP Session*
 - Peuvent également être mises en place entre voisins directs
- Fondées sur des messages `hello` transmis en *unicast*
 - L'adresse du destinataire doit être spécifiée explicitement
 - Permet une meilleure résilience

Annexe : la configuration des équipements

3 La signalisation : le protocole LDP

- Les messages LDP
- Découverte des pairs
- Établissement de la session
- Spécification des FEC
- Échange des labels
- **Les sessions ciblées**
- Annexe : la configuration des équipements

Configuration des équipements (Cisco)



● Utilisation de LDP pour la distribution des labels

```
r-1# mpls label protocol ldp
```

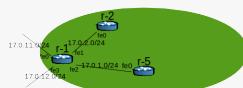
● Observation des découvertes LDP

```
r-1# show mpls ldp discovery
Local LDP Identifier: 17.0.12.1:0
Discovery Sources:
  Interfaces:
    FastEthernet1/0 (ldp): xmit
    FastEthernet2/0 (ldp): xmit
```

● Attention, pour établir une session LDP avec un voisin, l'adresse qui identifie celui-ci doit être joignable. Si besoin :

```
r-1(config-if)# mpls ldp discovery transport-address 17.0.12.1
```

Configuration des équipements (Cisco)



● Observation du voisinage LDP

```
r-1#sh mpls ldp neighbor
Peer LDP Ident: 17.0.3.2:0; Local LDP Ident 17.0.12.1:0
  TCP connection: 17.0.3.2.646 -> 17.0.12.1.15002
  State: Oper; Msgs sent/rcvd: 13/13; Downstream
  Up time: 00:00:15
  LDP discovery sources:
    FastEthernet1/0, Src IP addr: 17.0.2.2
    Addresses bound to peer LDP Ident:
      17.0.2.2          17.0.3.2
```

● Observation des sessions LDP

```
r-1#sh mpls ldp bindings
  tib entry: 17.0.1.0/24, rev 2
    local binding: tag: imp-null
    remote binding: tsr: 17.0.3.2:0, tag: 16
  tib entry: 17.0.2.0/24, rev 6
    local binding: tag: imp-null
    remote binding: tsr: 17.0.3.2:0, tag: imp-null
```

Configuration de LDP (avec FRR sur Linux)

• Dialogue avec le démon LDPD

```
$ telnet 127.0.0.1 2612
Password:
r> enable
r# configure terminal
r(config) #
```

• Activation de ldp

```
r(config)# mpls ldp
r-1(config-ldp)# router-id 1.2.3.4
```

• Choix de l'adresse utilisée lors des échanges (ici en IPv4)

```
r-1(config-ldp)# address-family ipv4
r-1(config-ldp-af)# discovery transport-address 5.6.7.8
```

• Ajout d'une interface (ici en IPv4)

```
r-1(config-ldp)# address-family ipv4
r-1(config-ldp-af)# interface eth0
r-1(config-ldp-af-if)# end
```

Configuration de LDP (avec FRR sur Linux)

• Observation des voisins connus

```
r-1# show mpls ldp neighbor
AF ID State Remote Address Uptime
ipv4 17.0.2.2 OPERATIONAL 17.0.2.2 00:16:31
```

• Observation des labels

```
r-1# sh mpls ldp binding
AF Interface Nexthop Local Label Remote Label In Use
ipv4 17.0.1.0/24 17.0.2.2 imp-null 20 no
ipv4 17.0.1.0/24 17.0.5.5 imp-null imp-null no
ipv4 17.0.41.0/24 17.0.2.2 20 18 no
ipv4 17.0.41.0/24 17.0.5.5 20 21 yes
ipv4 17.0.51.0/24 17.0.2.2 17 23 no
ipv4 17.0.51.0/24 17.0.5.5 17 imp-null yes
...
```

• Observation des interfaces

```
r-1# sh mpls ldp interface
AF Interface State Uptime Hello Timers ac
ipv4 fe1 ACTIVE 00:20:00 5/15 1
ipv4 fe2 ACTIVE 00:20:00 5/15 1
```

La signalisation : le protocole RSVP

1 Introduction

2 Le plan de données

3 La signalisation : le protocole LDP

4 La signalisation : le protocole RSVP

- Introduction
- La protection

5 L'ingénierie de trafic avec MPLS

6 Les VPNs fondés sur MPLS/BGP

7 Références bibliographiques

Introduction

4 La signalisation : le protocole RSVP

- Introduction
- La protection

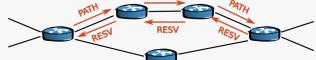
Introduction

- Le protocole RSVP a été conçu pour la qualité de service
 - Resource Reservation Protocol [10]
- Parfaitement adapté à MPLS
 - Discussion de proche en proche entre routeurs
 - N'impose pas le choix de la route
 - Permet de mobiliser des ressources, ou de définir des états
- Permet de construire des LSP
 - Unidirectionnels
 - Point à point
 - Initiés par le routeur d'entrée (le I-LER)

Pourquoi RSVP ?

- Mode d'attribution des labels *Downstream On Demand*
 - Requête transmise dans le sens des données
 - Labels attribués dans le sens inverse
- RSVP est le candidat idéal
 - Offre la signalisation permettant de réserver des ressources sur un chemin
 - Définit deux types de messages lors de l'établissement du chemin
 - Le message PATH part de la source et permet aux routeurs sur le chemin de découvrir le trafic
 - Le message RESV part de la destination et permet de mettre en place le chemin
 - Permet le rerouting, la détection de boucle, ...
- On utilise donc ces messages pour transmettre
 - La demande depuis le LER *ingress*
 - Les labels depuis le LER *egress*

Utilisation de RSVP



- Utilisation des deux messages de base de RSVP
- Les messages PATH émis dans le sens des données
 - Permettent de définir la FEC
- Les messages RESV émis dans le sens inverse
 - Permettent de mettre en place le tunnel
 - Définissent les labels à utiliser de proche en proche

La protection

4 La signalisation : le protocole RSVP

• Introduction

• La protection

- Le "One-to-one backup"
- Le "Facility backup"

Problématique

- Que faire en cas de défaillance sur un LSP ?
 - Le trafic n'est pas routé
 - Aucune incidence de la réaction des protocoles de routage
- Re-routage d'un LSP ?
 - Processus lourd (voir la signalisation de mise en place)
 - Conséquences sur l'ordre des paquets transmis
- Le problème n'est pas anodin
 - L'ingénierie de trafic des opérateurs est notamment fondée sur MPLS
 - Besoin de mécanismes automatiques efficaces
 - MPLS est également utilisé pour des applications exigeantes (eg voix sur IP)
 - Besoin de réactions rapides (dizaines de ms)

Le re-routage rapide par RSVP

• Principes [24]

- Établissement anticipé de LSP de secours
- Réparation "locale"
 - Un LSR décide localement de basculer sur un LSP de secours
 - Rapide (pas de mise en place ni même signalisation)
- Deux méthodes proposées
 - "One-to-one backup"
 - Pour chaque LSP à protéger, un LSP de secours est construit
 - "Facility backup"
 - Un même LSP peut être utilisé pour protéger plusieurs LSP
 - La protection peut se faire sur une partie ou l'intégralité du chemin

Le "One-to-one backup"

4 La signalisation : le protocole RSVP

• La protection

- Le "One-to-one backup"
- Le "Facility backup"

Le "One-to-one backup"

• Pour un LSP à protéger

- Des LSP de "détour" peuvent être mis en place
 - Sur chaque LSR (sauf le E-LER !)
- Le point d'entrée dans un LSP de protection est appelé PLR
 - Point of Local Repair
 - C'est lui qui redirige le trafic en cas de panne
- Le trafic est ré-injecté dans le LSP d'origine dès que possible

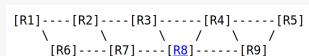


Figure issue de [24]

Le "Facility backup"

- ④ La signalisation : le protocole RSVP
 - La protection
 - Le "One-to-one backup"
 - Le "Facility backup"

Le "Facility backup"

- Pour un ensemble de LSP à protéger
 - Des LSP de "détour" peuvent être mis en place
 - Chaque LSP de secours peut protéger plusieurs LSP
 - Meilleur passage à l'échelle
- Considérations similaires au "One-to-one backup"
- Utilisation du *label stacking* de MPLS

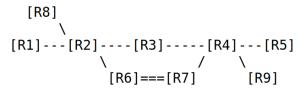


Figure issue de [24]

L'ingénierie de trafic avec MPLS

- ① Introduction
- ② Le plan de données
- ③ La signalisation : le protocole LDP
- ④ La signalisation : le protocole RSVP
- ⑤ L'ingénierie de trafic avec MPLS
 - Introduction
 - Le routage
 - La signalisation
 - Annexe : configuration des équipements

⑥ Les VPNs fondés sur MPLS/BGP

⑦ Références bibliographiques

Introduction

⑤ L'ingénierie de trafic avec MPLS

- Introduction
- Le routage
- La signalisation
- Annexe : configuration des équipements

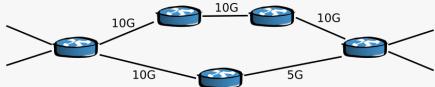
Problème de l'ingénierie de trafic

- Quel est l'objectif de l'ingénierie de trafic ?
 - L'évaluation et l'optimisation des performances du réseau [5]
 - Assurer une qualité de service satisfaisante à l'utilisateur
 - Utiliser au mieux les ressources du réseau
- Tâche complexe
 - Trafics nombreux et variés
 - Réseaux complexes, interconnexion
 - Objectifs antagonistes
 - Assurer un niveau de qualité de service de bout en bout aux trafics utilisateurs
 - Utiliser de façon économique et fiable les ressources du réseau

Introduction

- Quelle ingénierie de trafic dans IP ?
 - Routage de proche en proche
 - Pas de réservation de ressource
 - Routage fondé uniquement sur l'adresse destination
- Que permet MPLS ?
 - Mise en place de LSP sur un chemin, on les appelle des TE-LSP
 - Les LSP sont connus des routeurs qu'ils traversent
 - Le trafic peut y être injecté en fonction de différents paramètres
- MPLS-TE est une solution complète d'ingénierie de trafic
 - OSPF-TE ou un autre protocole de routage intégrant des fonctions de *Traffic Engineering*
 - RSVP-TE ou un autre protocole de signalisation, ...

Le problème



- Comment choisir la meilleure route ?
 - Les protocoles classiques vont choisir la plus courte
 - Possibilité de modifier les paramètres
- Comment partager la charge ?
 - Les outils classiques type ECMP ne savent pas gérer la diversité [32, 17]
 - On peut saturer un chemin et sous-utiliser l'autre
 - Que faire en cas de panne ?
- Idée : utiliser des LSP pour cela

La solution MPLS-TE

L'architecture MPLS-TE est fondée sur plusieurs éléments [6]

- Calculer des chemins
 - Routage avec contraintes
 - Besoin d'étendre les protocoles existants
 - Calcul de route intégrant ces contraintes
 - Besoin de changer les algorithmes de routage
- Signaler la mise en place des LSP
 - Échanger des informations sur les LSP
 - Label, FEC, ...
 - Caractéristiques du trafic
 - Le protocole LDP n'y suffit pas en l'état
 - Une extension a été proposée nommée CR-LDP
 - Abandonnée au profit de RSVP-TE
- Réserver des ressources associées aux LSP
 - Vérifier que les routeurs sont capables d'assurer le service
 - Possibilité d'*overbooking*

Le routage

5 L'ingénierie de trafic avec MPLS

- Introduction
- Le routage
 - Introduction
 - Extension à OSPF
- La signalisation
- Annexe : configuration des équipements

Introduction

5 L'ingénierie de trafic avec MPLS

- Le routage
 - Introduction
 - Extension à OSPF

Introduction

- Inadéquation des protocoles et algorithmes actuels
 - Les protocoles ne véhiculent pas d'information sur l'utilisation des liens
 - Les algorithmes ne prennent pas en compte de telles considérations
 - RIP ne calcule pas un chemin complet, il ne peut même pas être étendu
- Nécessité de proposer des extensions
 - Ajout de champs spécifiques dans les messages protocolaires
 - Capacité des liens
 - Ressources utilisées par les LSP déjà établis
 - Définition d'un nouvel algorithme de calcul de route
 - Prévoir la disponibilité des ressources nécessaires à chaque LSP
 - Application de cet algorithme à la demande

L'algorithme de routage

- Calcul d'un chemin
 - En utilisant l'algorithme de base de l'IGP
 - En tenant compte des contraintes du tunnel à établir
 - Sur la base des ressources encore disponibles
 - Lorsque le tunnel doit être mis en place
- On parle de PCALC (*Path Calculation*) ou CSPF (*Constraint Shortest Path First*)
- Il s'agit donc d'une modification relativement simple de l'algorithme de base
 - Celui-ci peut être également exécuté sans contrainte
 - Mises à jour de OSPF, ISIS, ...

Extension à OSPF

5 L'ingénierie de trafic avec MPLS

- Le routage
 - Introduction
 - Extension à OSPF

Extension à OSPF

Utilisation des LSA opaques [11, 9]

- Introduits pour permettre un déploiement de nouvelles fonctionnalités
 - Permettent de véhiculer des informations, éventuellement de façon opaque
 - Un TE LSA est défini [18]
- Définition de TE TLVs [18]
 - Type Length Value
- Au plus haut niveau
 - Router Address
 - Une adresse stable du routeur
 - link
 - Description d'un lien au moyen de sous-TLV

Les caractéristiques des liens pour le TE

9 sous-TLV sont définis [18]

- Link type
 - Est-ce un lien point à point ou multipoint ?
- Link ID
 - Identifiant sur 4 octets (adresse du vis-à-vis ou du DR)
- Local interface IP address
 - Liste des adresses du link
- Remote interface IP address
 - Adresse(s) du système distant ou 0.0.0.0

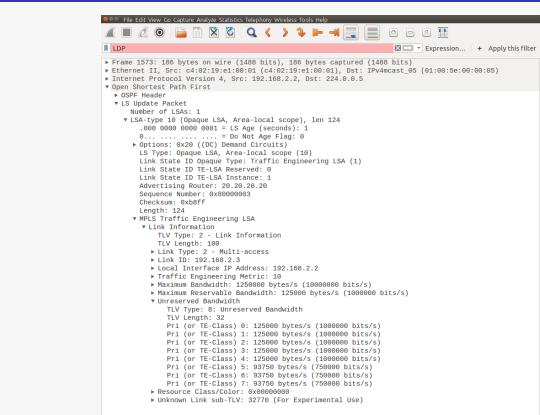
Les caractéristiques des liens pour le TE

- Traffic engineering metric
 - Métrique affectée par l'administrateur
- Maximum bandwidth
 - Débit maximum sur ce lien, dans ce sens, en octets par seconde
- Maximum reservable bandwidth
 - Débit réservable (peut être supérieur au débit max)
- Unreserved bandwidth
 - Débit non réservé
- Administrative group
 - Groupes assignés par l'administrateur

Principes de base

- Toute modification de lien doit être rapidement propagée
 - Pas nécessairement immédiatement
 - Des seuils peuvent être définis
- Le fonctionnement de base de OSPF perdure
 - Propagation des LSA en cas de changement de la topologie
 - Pas plus fréquent que MinLSInterval
- Toute nouvelle information ou mise-à-jour reçue doit être insérée dans la base
 - Aucun calcul nécessaire à ce moment

Un exemple



La signalisation

5 L'ingénierie de trafic avec MPLS

- Introduction
- Le routage
- La signalisation
 - Extension de RSVP
- Annexe : configuration des équipements

Extension de RSVP

5 L'ingénierie de trafic avec MPLS

- La signalisation
- Extension de RSVP

Extension de RSVP



- Définition de RSVP-TE [4]
- Utilisation des deux messages de base de RSVP
- Les messages PATH émis dans le sens des données
 - Permettent de définir les besoins du tunnel
 - Déterminent la disponibilité des ressources
- Les messages RESV émis dans le sens inverse
 - Permettent de mettre en place le tunnel
 - Enclanchent la réservation des ressources
 - Définissent les labels à utiliser de proche en proche

Un exemple

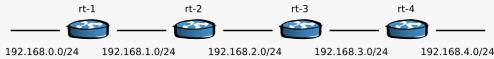
Un exemple

Annexe : configuration des équipements

5 L'ingénierie de trafic avec MPLS

- Introduction
- Le routage
- La signalisation
- Annexe : configuration des équipements

Activation de l'ingénierie de trafic



- Activation sur le système

```
rt-1(config)# ip cef
rt-1(config)# mpls traffic-eng tunnels
rt-1(config)# interface Loopback0
rt-1(config-if)# ip address 10.10.10.10 255.255.255.255
```

- Activation d'une interface (réservons 1 Mbit/s)

```
rt-1(config-if)# mpls ip
rt-1(config-if)# mpls traffic-eng tunnels
rt-1(config-if)# ip rsvp bandwidth 1000
```

- Configuration d'OSPF-TE (le routage étant configuré)

```
rt-1(config)# router ospf 1
rt-1(config-router)# mpls traffic-eng area 0
rt-1(config-router)# mpls traffic-eng router-id Loopback0
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 109 / 139

Création d'un tunnel

- Création de l'interface sans adresse (car unidirectionnelle)

```
rt-1(config)# interface Tunnel1
rt-1(config-if)# ip unnumbered Loopback0
```

- Définition de la destination et du mode (MPLS-TE)

```
rt-1(config-if)# tunnel destination 40.40.40.40
rt-1(config-if)# tunnel mode mpls traffic-eng
```

- On lui affecte une priorité de création et une de maintien

```
rt-1(config-if)# tunnel mpls traffic-eng priority 5 5
```

- Calcul dynamique du chemin (option 1, donc la préférée)

```
rt-1(config-if)# tunnel mpls traffic-eng path-option 1 dynamic
```

- On lui affecte un débit de 250 kbit/s, pas de routage

```
rt-1(config-if)# tunnel mpls traffic-eng bandwidth 250
rt-1(config-if)# no routing dynamic
```

- Utilisons-le pour une route

```
rt-1(config)# ip route 192.168.5.0 255.255.255.0 Tunnel1
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 110 / 139

Observation des données de TE

- Sur rt-1

```
rt-1# show mpls forwarding-table
Local Outgoing Prefix Bytes tag Outgoing Next Hop
tag tag or Tunnel Id switched interface
14 Untagged 192.168.2.0/24 0 Fa0/1 192.168.1.2
17 Untagged 192.168.3.0/24 0 Fa0/1 192.168.1.2
18 Untagged[T] 192.168.5.0/24 0 null point2point

[T] Forwarding through a TSP tunnel.
View additional tagging info with the 'detail' option
```

- Sur rt-2

```
rt-2# show mpls forwarding-table
Local Outgoing Prefix Bytes tag Outgoing Next Hop
tag tag or WC or Tunnel Id switched interface
16 Untagged 192.168.2.0/24 0 Fa0/1 192.168.2.3
17 17 10.10.10.10 1 [1] 590 Fa0/1 192.168.2.3
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 111 / 139

Observation des données de TE

```
rt-1# show ip ospf database opaque-area
OSPF Router with ID (10.10.10.10) (Process ID 1)
Type-10 Opaque Link Area Link States (Area 0)
LS age: 1316
Options: (No TOS-capability, DC)
Link ID: 192.168.1.1
Link State ID: 1.0.0.0
Opaque Type: 1
Opaque ID: 0
Advertising Router: 10.10.10.10
LS Seq Number: 80000004
Checksum: 0x218F
Length: 132
Forwarding number: 0
MPLS Router ID: 10.10.10.10
Link connected to Broadcast network
Link ID: 192.168.1.1
Interface Address: 192.168.1.1
Admin Metric: 10
Maximum bandwidth: 1250000
Maximum reservable bandwidth: 125000
Number of Priority: 8
Priority 0 : 125000 Priority 1 : 125000
Priority 2 : 125000 Priority 3 : 125000
Priority 4 : 125000 Priority 5 : 93750
Priority 6 : 93750 Priority 7 : 93750
Affinity Bit : 0x0
IGP Metric: 10
Number of Links: 1
...
...
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 112 / 139

Observation des tunnels

```
rt-1# show mpls traffic-eng tunnels
Name: rt-1_1 (Tunnel1) Destination: 40.40.40.40
Status: Admin: up Oper: up Path: valid Signalling: connected
Config Parameters:
  Bandwidth: 250 kbps (Global) Priority: 5 5 Affinity: 0x0/0xFFFF
  Metric Type: TE (default)
  AutoRoute: disabled Lockdown: disabled Loadshare: 250 bw-based
  auto-bw: disabled
InLabel: -
OutLabel: FastEthernet0/1, 17
 RSVP Signalling Info:
  Src 10.10.10.10, Dst 40.40.40.40, Tun_Id 1, Tun_Instance 1
  RCV Path Info:
    My Address: 192.168.1.1
    Explicit Route: 192.168.1.2 192.168.2.2 192.168.2.3 192.168.3.3
    Record Route: NONE
    Speci: ave rate=250 kbytes, burst=1000 bytes, peak rate=250 kbytes
  RSVP Resv Info:
    Record Route: NONE
    Speci: ave rate=250 kbytes, burst=1000 bytes, peak rate=250 kbytes
History:
Tunnel1:
  Time since created: 56 minutes, 25 seconds
  Time since path change: 55 minutes, 11 seconds
Current LSP:
  Uptime: 55 minutes, 12 seconds
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 113 / 139

Observation de la topologie

```
rt-3# show mpls traffic-eng topology
My_System_id: 30.30.30.30, Globl Link Generation 18
Signalling error holddown: 10 sec
IGP Id: 10.10.10.10, MPLS TE Id:10.10.10.10 Router Node
link[0 ]:DR Intf Address:192.168.1.1, gen:18
  frag_id 0, Intf Address:192.168.1.1
  TE metric:10, IGP metric:10, attribute_flags:0x0
  physical_bw: 10000 (kbps), max_reservable_bw_global: 1000 (kbps)
  max_reservable_bw_sub: 0 (kbps)

  Global Pool Sub Pool
  Total Allocated Reservable Reservable
  BW (kbps) BW (kbps) BW (kbps)
  -----
bw[0]: 0 1000 0
bw[1]: 0 1000 0
bw[2]: 0 1000 0
bw[3]: 0 1000 0
bw[4]: 0 1000 0
bw[5]: 250 750 0
bw[6]: 0 750 0
bw[7]: 0 750 0
...
...
```

Chaput Emmanuel Une brève introduction à MPLS 2024-2025 114 / 139

La famille d'adresses VPN-IPv4

- MBGP introduit la notion de famille d'adresses [8]
 - Capacité à router autre chose que de l'IPv4
 - Par exemple IPv6
- L'utilisation de BGP pour les VPN se fonde sur une nouvelle famille
 - La famille d'adresses VPN-IPv4 [28]
 - 12 octets
 - Permet d'identifier de façon unique des réseaux IPv4
 - Pas utilisé dans le routage

La famille d'adresses VPN-IPv4

Nombre sur 12 octets

- Champ Type (2 octets)
- Champ Value (6 octets)
 - Sous champ Administrator
 - Identifiant globalement unique
 - ASN (Type = 0)
 - Adresse IPv4 (Type = 1)
 - ASN 32 bits (Type = 2)
 - Sous champ Assigned Number
 - Identifiant localement unique
- Adresse IPv4

Les 8 premiers octets sont qualifiés de *Route Distinguisher* [28].

Le routage

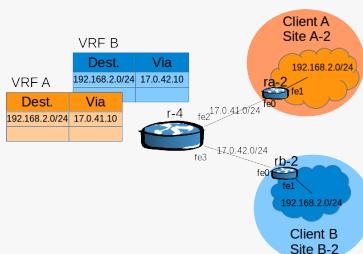
6 Les VPNs fondés sur MPLS/BGP

- Introduction
- L'adressage
- Le routage
- Un exemple
- Annexe : configuration des équipements

Le routage

- Chaque PE dispose d'un ensemble de VRF
 - VPN Routing and Forwarding
 - Une par VPN et une par défaut
- Chaque paquet est routé en fonction de la VRF correspondante
 - Fonction de l'AC par lequel il arrive
 - Éventuellement de son adresse source
 - Par défaut si pas de VRF associée
- Construction des tables de routages
 - Entre PE et CE
 - Entre PEs
 - Statique, dynamique, ...

Le routage



- Les deux VRF du routeur r-4 liées à ses clients
- Contenu échangé avec le routeur du client correspondant (ra-2 et rb-2 ici)
- Indépendantes entre elles et avec la table de routage classique

L'attribut *Route Target* de BGP

- Route Target Community* [31]
- Attribut étendu transitif
- Permet de désigner un ensemble de routeurs
- Routeurs susceptibles d'utiliser une route ainsi qualifiée
- Fonction équivalente à l'attribut *Community*
- Permet également de répartir la charge entre plusieurs ensembles de *Route Reflectors*
 - Chacun gérant les routes d'un ensemble de *Route Target Community*

Annonce de routes entre CE et PE

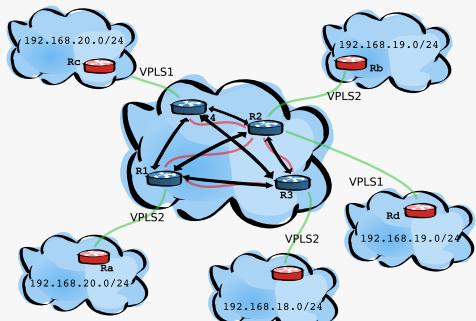
- Différentes possibilités en fonction de la nature du réseau du client
 - Réseau de transit
 - "Stub AS"
- Techniques classiques envisageables
 - Routage statique
 - RIP
 - OSPF (SLAS véhiculés grâce aux attributs étendus de BGP4) [27]
 - BGP, *a priori* solution la plus souple
- Routes annoncées par PE au CE
 - Route par défaut souvent suffisante

Un exemple

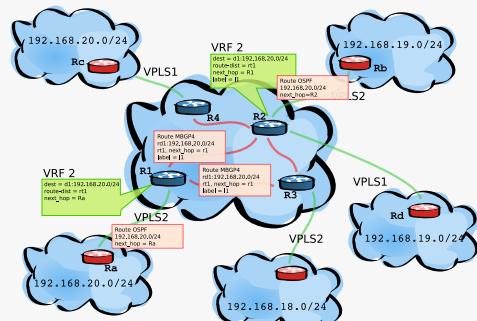
6 Les VPNs fondés sur MPLS/BGP

- Introduction
- L'adressage
- Le routage
- Un exemple**
- Annexe : configuration des équipements

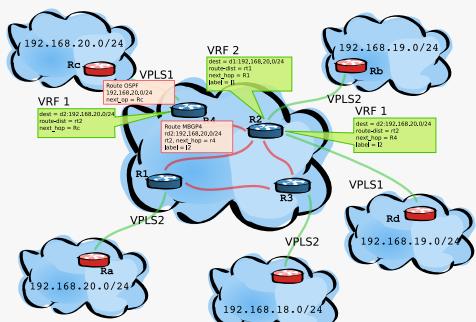
Annonce de routes entre CE et PE : un exemple (1)



Annonce de routes entre CE et PE : un exemple (2)



Annonce de routes entre CE et PE : un exemple (3)



Acheminement du trafic

- Réception d'un paquet par un PE
 - Utilisation du VRF associé au lien de réception
- Encapsulation dans un paquet MPLS
 - Label associé à la route
- Acheminement au travers du backbone
 - Par un chemin MPLS en empilant un label
 - Par un tunnel IP [33]
- Sur le routeur PE de sortie
 - Désencapsulation du paquet
 - Émission sur le lien en fonction du label
 - Et éventuellement de la table de routage

Annexe : configuration des équipements

6 Les VPNs fondés sur MPLS/BGP

- Introduction
- L'adressage
- Le routage
- Un exemple
- Annexe : configuration des équipements

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

133 / 139

Gestion des VPN (1/3)

FRR permet de gérer les VPN MPLS BGP par le biais des éléments suivants

- BGP qui va échanger les informations sur ces VPN
 - Mais IPv4 n'est plus la seule famille d'adresses
 - Il faut configurer les voisins BGP avec lesquels on échange
- MPLS qui doit être configuré par ailleurs mais dont la distribution des labels peut être réalisée par FRR via le protocole LDP

```
router bgp 65
no bgp default ipv4-unicast
neighbor 10.0.65.1 remote-as 65
neighbor 10.0.65.1 update-source eth0
address-family ipv4 vpn
neighbor 10.0.65.1 activate
```

Chaput Emmanuel
Une brève introduction à MPLS
2024-2025
134 / 139

Gestion des VPN (2/3)

FRR permet de gérer les VPN par le biais des éléments suivants

- Les familles d'adresses supportées par BGP
 - Par exemple la famille `ipv4 vpn`
 - Il faut activer les voisins BGP avec lesquels on les échange
 - Les VRF au sein desquelles on peut échanger via BGP des routes IPv4 obtenues par OSPF par exemple
- ```
router bgp 4832
address-family ipv4 vpn
neighbor 192.168.10.12 activate
```
- ```
router bgp 42 vrf toto
address-family ipv4 unicast
redistribute ospf
```

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

135 / 139

Gestion des VPN (2/2)

FRR permet de gérer les VPN par le biais des éléments suivants

- MPLS dont il faut demander à BGP d'exporter les labels de la façon suivante
 - Les *Route Distinguisher* associés à cette famille d'adresses et *Route Target* utilisés par BGP
- ```
router bgp 24 vrf titi
address-family ipv4 unicast
label vpn export auto
```
- ```
router bgp 33 vrf brd
address-family ipv4 unicast
rd vpn export 10.0.10.33:1
rt vpn both 1:33
import vpn
export vpn
```

Chaput Emmanuel
Une brève introduction à MPLS
2024-2025
136 / 139

Configuration d'une VRF (ici sur Cisco)

- Création d'une VRF
- Sélection du *route distinguisher*
- Ajout d'une *route target* importée
- Ajout d'une *route target* exportée
- Les deux à la fois
- Sélection de la VRF associée à une interface

```
rt-1(config)# ip vrf test
```

```
rt-1(config-vrf)# rd 1:100
```

```
rt-1(config-vrf)# route import 1:1
```

```
rt-1(config-vrf)# route export 1:2
```

```
rt-1(config-vrf)# route both 1:3
```

```
rt-1(config-vrf)# ip vrf forwarding test
```

Chaput Emmanuel

Une brève introduction à MPLS

2024-2025

137 / 139

Exemple d'un client utilisant RIP

(voir la configuration dans fish-topo)

- Sur le PE, on crée la VRF
- On place l'interface dans la VRF
- On active une instance de RIP dans cette VRF
- Sur le client, on ajoute le réseau à RIP

```
ip vrf client1
rd 1:1
route-target export 1:1
route-target import 1:1
```

```
interface FastEthernet1/0
ip vrf forwarding client1
ip address 192.168.11.1 255.255.255.0
```

```
router rip
!
address-family ipv4 vrf client1
network 192.168.11.0
neighbor 192.168.11.10
redistribute bgp 1 metric transparent
```

```
router rip
!
address-family ipv4 vrf client1
network 192.168.11.0
network 192.168.11.0
neighbor 192.168.11.10
```

Chaput Emmanuel
Une brève introduction à MPLS
2024-2025
138 / 139

Configuration de BGP

- On configure les voisins BGP (on se présente avec une adresse routable)

```
router bgp 1
neighbor 17.0.2.1 remote-as 1
neighbor 17.0.2.1 update-source FastEthernet0/0
```

- On y active IPv4 (si c'est utile par ailleurs)

```
address-family ipv4
neighbor 17.0.2.1 activate
```

- On active les VPN !

```
address-family vpng4
neighbor 17.0.2.1 activate
neighbor 17.0.2.1 send-community extended
```

- On redistribue dans la VRF de chaque client

```
address-family ipv4 vrf client1
redistribute rip metric 1
```

- [1] L. Andersson and R. Asati.
Multiprotocol Label Switching (MPLS) Label Stack Entry : "EXP" Field Renamed to "Traffic Class" Field.
Technical Report 5462, February 2009.
- [2] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas.
LDP Specification.
Technical Report 3036, Internet Engineering Task Force, January 2001.
Obsoleted by RFC 5036.
- [3] L. Andersson, I. Minei, and B. Thomas.
LDP Specification.
Technical Report 5036, October 2007.
Updated by RFCs 6720, 6790.
- [4] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow.
RSVP-TE : Extensions to RSVP for LSP Tunnels.

Technical Report 3209, December 2001.
Updated by RFCs 3936, 4420, 4874, 5151, 5420, 5711.

- [5] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao.
Overview and Principles of Internet Traffic Engineering.
RFC 3272 (Informational), May 2002.
Updated by RFC 5462.
- [6] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus.
Requirements for Traffic Engineering Over MPLS.
Technical Report 2702, September 1999.
- [7] F. Baker and P. Savola.
Ingress Filtering for Multihomed Networks.
RFC 3704 (Best Current Practice), March 2004.
- [8] T. Bates, Y. Rekhter, R. Chandra, and D. Katz.
Multiprotocol Extensions for BGP-4.
Technical Report 2858, Internet Engineering Task Force, June 2000.
Obsoleted by RFC 4760.

- [9] L. Berger, I. Bryskin, A. Zinin, and R. Coltn.
The OSPF Opaque LSA Option.
Technical Report 5250, July 2008.
- [10] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin.
RFC 2205 : Resource reservation protocol (rsvp) – version 1 functional specification.
Standards track, IETF, September 1997.
- [11] R. Coltn.
The OSPF Opaque LSA Option.
Technical Report 2370, July 1998.
Obsoleted by RFC 5250, updated by RFC 3630.
- [12] The ATM Forum Technical Committee.
Lan emulation over atm, verion 2 luni specifications.
Technical Report Version 2, ATM Forum, July 1997.
- [13] The ATM Forum Technical Committee.
Multi-protocol over atm.
Technical Report Version 1.0, ATM Forum, July 1997.

- [14] The ATM Forum Technical Committee.
Multi-protocol over atm.
Technical Report Version 1.1, ATM Forum, May 1999.
- [15] The ATM Forum Technical Committee.
Lan emulation over atm version 1.0.
Technical Report af-lane-0021.000, ATM Forum, January 1995.
- [16] The ATM Forum Technical Committee.
Lan emulation over atm version 2 - Inni specification.
Technical Report AF-LANE-0112.000, ATM Forum, February 1999.
- [17] C. Hopps.
Analysis of an equal-cost multi-path algorithm.
Technical report, Internet Engineering Task Force, November 2000.
- [18] D. Katz, K. Komppella, and D. Yeung.
Traffic Engineering (TE) Extensions to OSPF Version 2.
RFC 3630 (Proposed Standard), September 2003.

- [19] M. Laubach.
Classical IP and ARP over ATM.
RFC 1577, IETF, January 1994.
- [20] M. Laubach and J. Halpern.
Classical IP and ARP over ATM.
Technical Report 2225, IETF, April 1998.
Updated by RFC 5494.
- [21] J. Luciani, D. Katz, D. Piscitello, B. Cole, and N. Doraswamy.
NBMA Next Hop Resolution Protocol (NHRP).
Technical Report 2332, IETF, April 1998.
- [22] E. Mannie.
Generalized Multi-Protocol Label Switching (GMPLS) Architecture.
Technical Report 3945, October 2004.
Updated by RFC 6002.

- [23] L. Martini, E. Rosen, N. El-Aawar, T. Smith, and G. Heron.
Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP).
Technical Report 4447, "Internet Engineering Task Force", April 2006.
- [24] P. Pan, G. Swallow, and A. Atlas.
Fast Reroute Extensions to RSVP-TE for LSP Tunnels.
Technical Report 4090, May 2005.
- [25] M. Perez, F. Liaw, A. Mankin, E. Hoffman, D. Grossman, and A. Malis.
ATM Signaling Support for IP over ATM.
Technical Report 1755, February 1995.
- [26] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear.
RFC 1918 : Address allocation for private internets.
Best current practice, Internet Engineering Task Force, February 1996.
- [27] E. Rosen, P. Psenak, and P. Pillay-Esnault.
OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs).
Technical Report 4577, "Internet Engineering Task Force", June 2006.
- [28] E. Rosen and Y. Rekhter.
BGP/MPLS IP Virtual Private Networks (VPNs).
Technical Report 4364, "Internet Engineering Task Force", February 2006.
Updated by RFCs 4577, 4684.
- [29] E. Rosen, D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li, and A. Conta.
MPLS Label Stack Encoding.
Technical Report 3032, Internet Engineering Task Force, January 2001.
Updated by RFCs 3443, 4182.
- [30] E. Rosen, A. Viswanathan, and R. Callon.

- RFC 3031 : Multiprotocol label switching architecture.
Standards track, IETF, January 2001.
- [31] S. Sangli, D. Tappan, and Y. Rekhter.
BGP Extended Communities Attribute.
Technical Report 4360, "Internet Engineering Task Force", February 2006.
- [32] D. Thaler and C. Hopps.
Multipath Issues in Unicast and Multicast Next-Hop Selection.
RFC 2991 (Informational), November 2000.
- [33] T. Worster, Y. Rekhter, and E. Rosen.
Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE).
Technical Report 4023, "Internet Engineering Task Force", March 2005.