



**TECHNIQUES
DE L'INGÉNIEUR**

Réf. : **TE7594 V1**

Segment Routing

Date de publication :
10 février 2019

Cet article est issu de : **Technologies de l'information | Réseaux Télécommunications**

par **Bruno DECRAENE**

Mots-clés

Segment Routing | SR |
SR-MPLS | SRv6

Résumé Segment Routing est une solution de routage par la source dans les réseaux IP/MPLS. Il permet à un nœud du réseau, dit source, de spécifier un chemin particulier pour un paquet en indiquant ce chemin dans l'en-tête du paquet ouvrant ainsi de nouvelles possibilités en termes de programmation du routage, de protection de chemin, de protection rapide (Fast ReRoute) de lien ou nœud, de simplification et de réduction de la signalisation, de partage de charge ou de l'ingénierie de trafic. Segment Routing s'applique aux réseaux MPLS ou aux réseaux IPv6.

Keywords

Segment Routing | SR |
SR-MPLS | SRv6

Abstract Segment Routing is solution to provide source routing capabilities in IP/MPLS networks. Segment Routing (SR) allows a node, called the source, to enforce a specific path for a packet, by specify a list of waypoints in the packet header. SR open new opportunities in term of network programmability, path protection, fast protection again list or node failure (Fast ReRoute), simplification, states and signaling reduction in the network, load balancing and traffic engineering. Segment Routing is applicable to MPLS and IPv6 networks. In MPLS networks, Segment Routing MPLS (SR-MPLS) reuse as-is the existing MPLS architecture and data plane and only requires an extension in the control plane.

Pour toute question :

Service Relation clientèle
Techniques de l'Ingénieur
Immeuble Pleyad 1
39, boulevard Ornano
93288 Saint-Denis Cedex

Par mail :

infos.clients@teching.com

Par téléphone :

00 33 (0)1 53 35 20 20

Document téléchargé le : **06/01/2021**

Pour le compte : **7200038780 - ecole polytechnique // 129.104.72.250**

© Techniques de l'Ingénieur | tous droits réservés

Segment Routing

par **Bruno DECRAENE**
Ingénieur réseau

1. Rappels sur MPLS	TE 7 594 – 2
1.1 MPLS	— 2
1.2 LDP	— 3
1.3 RSVP-TE	— 3
2. Segment Routing MPLS (SR-MPLS)	— 3
2.1 Vue d'ensemble	— 3
2.2 Segments	— 4
2.2.1 Segments IGP	— 4
2.2.2 Segments BGP	— 7
2.2.3 Politique Segment Routing et Segment d'association	— 8
2.3 Fonctionnement	— 8
2.3.1 Plan de commande	— 8
2.3.2 Plan de commutation	— 9
2.4 Bénéfices de SR-MPLS	— 9
3. Segment Routing IPv6 (SRv6)	— 10
4. Applications	— 10
4.1 Réactions aux pannes	— 11
4.1.1 Fast ReRoute (FRR)	— 11
4.1.2 Loop Free Convergence	— 11
4.2 Ingénierie de trafic (TE)	— 11
4.2.1 Usage tactique	— 11
4.2.2 Réservation de bande passante	— 12
4.3 Opération du réseau	— 12
4.3.1 Monitoring du réseau	— 12
4.3.2 Monitoring des liens et tables de commutations	— 12
4.3.3 Simplification du réseau	— 12
5. Conclusion	— 13
6. Glossaire	— 13
7. Sigles, notations et symboles	— 13
Pour en savoir plus	Doc. TE 7 594

Segment Routing (SR) utilise le principe de routage par la source. Un nœud source dirige un paquet à travers une liste ordonnée d'instructions, appelées « segments ». Un segment peut représenter une instruction quelconque, telle qu'une instruction de type topologique (suivre un chemin) ou une instruction de type service (appliquer un service sur ce paquet). Segment Routing permet ainsi à un flux de suivre un chemin spécifique, sans avoir à créer et maintenir des états sur les équipements situés le long de ce chemin. Seul le routeur source maintient un état spécifique, sous la forme d'une liste de segments à ajouter sur le paquet. Cette liste de segments est appelée politique Segment Routing (SR policy). Un segment est identifié par un identifiant de segment (SID : Segment ID).

Segment Routing s'applique au plan de transfert MPLS (SR-MPLS) sans modification de l'architecture MPLS, ni du plan de transfert MPLS. Un segment est identifié par un label. Une liste de segments est encodée sous forme d'une pile

de labels. Le segment en cours de traitement correspond au label au sommet de la pile. Une fois le segment terminé, son label est retiré du haut de la pile.

Segment Routing s'applique au plan de transfert IPv6 (SRv6) par la définition d'un nouveau type d'en-tête de routage. Un segment est identifié par une adresse IPv6. Une liste ordonnée de segments est encodée par une liste ordonnée d'adresse IPv6 dans l'en-tête de routage IPv6. Le segment en cours de traitement est indiqué dans le champ adresse destination de l'en-tête IPv6. Le prochain segment est indiqué par un pointeur présent dans l'en-tête de routage.

Segment Routing supporte un plan de contrôle distribué ou centralisé.

Les segments génériques sont typiquement annoncés par les protocoles de routage IP distribués tels que les protocoles de routage interne à états de liens (OSPF et IS-IS) ou protocoles de routage inter-domaine (BGP).

Dans le mode distribué, la politique Segment Routing est calculée par le nœud souhaitant l'utiliser. Ce calcul utilise les informations topologiques et de segments distribuées par le protocole de routage.

Dans le mode centralisé, un contrôleur de réseau calcule la politique Segment Routing et décide quel nœud doit l'utiliser. Un serveur PCE (Path Computation Element) peut remplir ce rôle et dans ce cas utilise le protocole PCEP pour échanger ces informations avec les nœuds sources. Mais d'autres protocoles sont possibles tels que BGP ou NETCONF.

Segment Routing a de nombreuses applications que l'on peut regrouper en trois catégories : une meilleure réaction du réseau aux pannes (Fast ReRoute, convergence IGP sans micro-boucles) ; l'ingénierie de trafic (chemins disjoints, routage spécifique tel qu'une minimisation du délai, optimisation du routage des flux en fonction des capacités du réseau) ; des améliorations dans l'opération du réseau en termes de monitoring et de simplifications.

1. Rappels sur MPLS

1.1 MPLS

MPLS (Multi-Protocol Label Switching) est une technologie de commutation de paquet définie par l'IETF [RFC 3031]. À l'entrée du réseau MPLS, un paquet est encapsulé dans MPLS par ajout d'un petit entête de 32 bits (quatre octets) [RFC 3032]. La figure 1 illustre l'encapsulation d'un paquet IP dans un en-tête MPLS, mais MPLS peut encapsuler tout type de paquet, d'où son nom « Multi-Protocol ».

Cet en-tête MPLS contient quatre champs :

- « Label » contient un nombre qui indique le traitement à effectuer sur le paquet. Typiquement, il s'agit de l'envoyer vers une destination e.g. Brest. Cependant, dans le principe, toute signification est possible. L'association entre le label et sa signification est signalisée dans le plan de commande, par exemple à l'aide des protocoles LDP ou RSVP-TE ;
- « Traffic Class » (TC) qui correspond à une indication de classe de service similaire au champ « TC » d'IPv6 ou « Precedence » d'IPv4 ;
- « Time To Live » (TTL) indique une durée de vie maximale du paquet en cas de problème de routage créant une boucle. IPv4 et

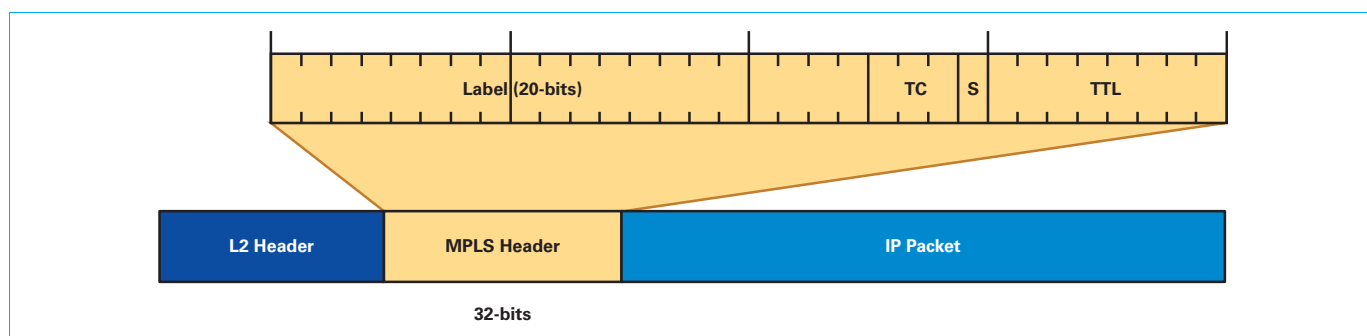


Figure 1 – En-tête MPLS

IPv6 ont le même champ. Elle est spécifiée en nombre de routeurs traversés ;

- « Bottom of Stack » (S) est un champ d'un seul bit très spécifique à MPLS. Il permet d'ajouter plusieurs en-têtes MPLS sur un paquet. S'il est à zéro, cet en-tête MPLS est suivi d'un autre en-tête MPLS. Ceci permet d'empiler un nombre quelconque de labels MPLS.

Lorsqu'un paquet contient plusieurs en-têtes MPLS, on parle d'une **pile de labels**. Chaque routeur achemine un paquet en consultant le label du haut de pile, sans consulter les autres labels. Arrivé à la destination indiquée par ce premier label, ce label de haut de pile est enlevé, puis le paquet est acheminé sur la base du label suivant, qui devient le nouveau label de haut de pile.

MPLS permet ainsi d'ajouter des instructions sur le paquet. Par exemple, en poussant deux labels on peut créer un service de réseau privé virtuel (VPN) :

- un premier label pour demander au réseau d'acheminer le paquet jusqu'à un routeur de périphérie de l'opérateur, par exemple situé à Brest ;
- un second label, qui ne sera visible que par le routeur de Brest, pour demander l'acheminement du paquet vers le client VPN 1.

Plus de précisions sur MPLS et les VPN BGP/MPLS sont disponibles dans [TE 7 535] et [TE 7 590].

Un label n'a pas de signification globale, ni de signification a priori. Il faut donc lui en donner une, et communiquer cette signification entre le(s) routeur(s) envoyant le paquet et le routeur le recevant. Bien que cette communication puisse être réalisée de manière statique par configuration, elle est habituellement réalisée par un protocole du plan de contrôle. L'architecture de MPLS n'impose pas un seul protocole de distribution de labels, et effectivement plusieurs protocoles peuvent distribuer les labels MPLS, par exemple LDP, RSVP-TE ou BGP. En général, c'est le routeur amené à recevoir le paquet qui choisit la valeur du label à utiliser.

1.2 LDP

« Label Distribution Protocol » (LDP) est un protocole de signalisation défini dans [RFC 5036]. LDP permet à deux routeurs voisins de signaler les labels utilisables et de leur associer une signification. Typiquement, cette signification correspond à une destination spécifiée sous la forme d'une ou plusieurs adresses IP.

Par **exemple**, un routeur utilise LDP pour annoncer à ses voisins qu'il associe le label « 53 » à la destination 192.0.2.0/32. Cela signifie que tout paquet qu'il reçoit avec le label « 53 » sera transmis vers la destination 192.0.2.0/32.

LDP ne fait pas de routage, mais s'appuie sur les protocoles de routages IP afin de déterminer la route vers la destination. Dans un réseau, les tunnels MPLS établis par LDP vont donc suivre l'arbre de plus court chemin vers la destination. Chaque tunnel est un arbre multipoint à point le long du plus court chemin. La figure 2 illustre un tunnel MPLS, établi par LDP à travers le réseau. LDP établit un tunnel multipoint à point vers chaque destination et donc N tunnels seront établis dans le réseau (un tunnel par destination).

1.3 RSVP-TE

« Resource Reservation Protocol extensions for Traffic Engineering » (RSVP-TE) est un protocole de signalisation défini dans [RFC 3209]. RSVP-TE permet d'établir un tunnel MPLS point à point à travers le réseau, selon un chemin particulier (quelconque). Il permet également à deux routeurs voisins de s'échanger le label MPLS utilisé pour ce tunnel. La figure 3 illustre l'ensemble des tunnels MPLS point à point vers D, établis par RSVP-TE. RSVP-TE établit un tunnel point à point entre chaque source et destination et donc N^2 tunnels seront établis dans le réseau (un tunnel par (source, destination)).

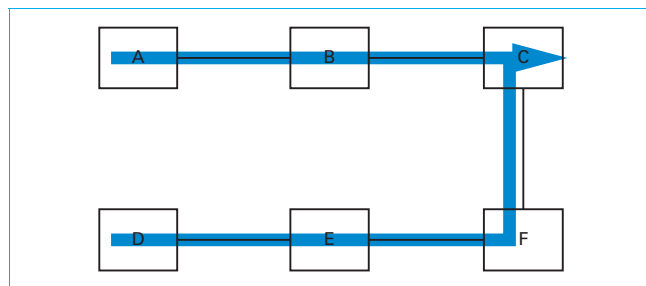


Figure 2 – Tunnel MPLS multipoint à point vers D, établi par LDP

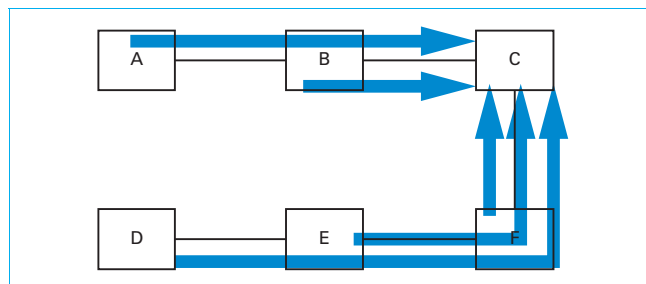


Figure 3 – Tunnels MPLS point à point vers D, établis par RSVP-TE

2. Segment Routing MPLS (SR-MPLS)

2.1 Vue d'ensemble

Segment Routing MPLS (SR-MPLS) est une évolution de MPLS :

- SR-MPLS ne change pas l'architecture de MPLS [RFC 3031] ;
- SR-MPLS ne change pas le plan de commutation MPLS donc ne nécessite pas de changement matériel ;
- SR-MPLS nécessite une évolution du plan de commande MPLS et donc nécessite une évolution logicielle des routeurs.

SR-MPLS repose sur deux principes :

1) Les labels MPLS ne sont plus annoncés dans un protocole de signalisation spécifique à MPLS mais sont annoncés dans le protocole de routage interne à états de liens, tels qu'IS-IS [ISO 10589] ou OSPF [IETF 2328].

2) Les tunnels nécessitant une route différente du plus court chemin ne sont pas signalisés par un protocole, ni installés dans la table de commutation des routeurs de transit, mais sont acheminés par imposition d'une pile de plusieurs labels préexistants. Cette liste de labels permet d'imposer une suite de nœuds ou d'interfaces à traverser successivement.

Le premier point permet à tous les routeurs de connaître tous les labels utilisés par chaque routeur. En effet, un protocole de routage interne à état de lien fonctionne par diffusion de l'information dans tout le réseau. Cela permet à tous les routeurs de connaître et d'utiliser les tunnels de tous les autres routeurs, contrairement à LDP où seuls les routeurs voisins d'un routeur R connaissent les labels utilisés par R et donc peuvent utiliser les tunnels traversant R.

Le second point permet à un routeur, dit source, d'envoyer des paquets le long de n'importe quel chemin à travers le réseau, sans besoin de signalisation ni de création de nouveaux états dans les tables de commutation. C'est une méthode de routage par la source. Ainsi, avec un petit nombre de tunnels génériques et mutualisés, appelés segments, un routeur source peut envoyer un

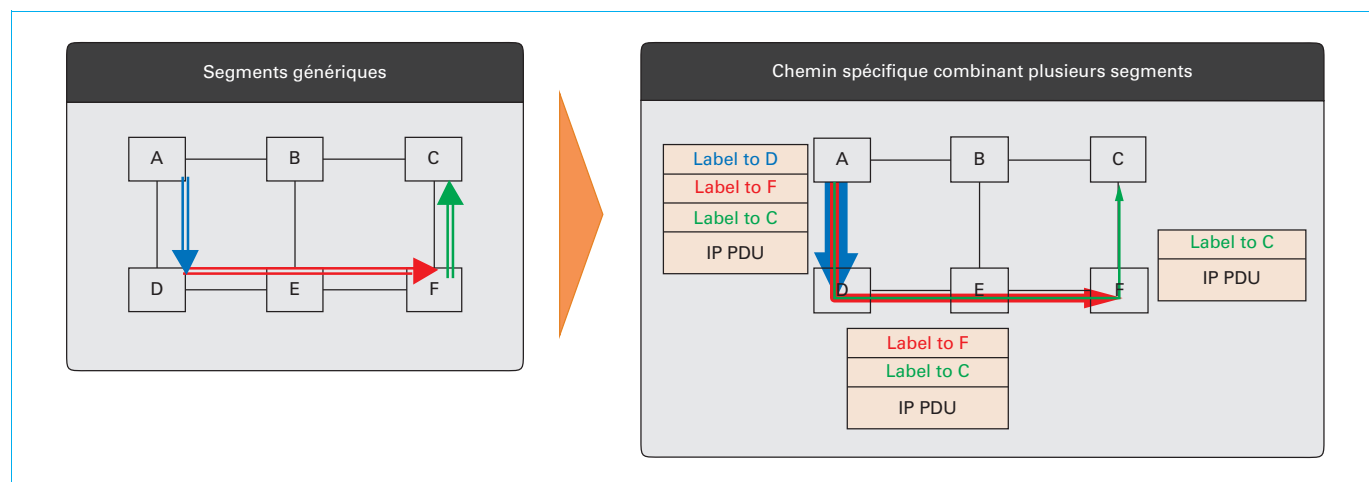


Figure 4 – Segment Routing combine plusieurs segments préexistants afin de suivre le chemin souhaité sans créer de nouveaux états dans le réseau

paquet le long d'un chemin quelconque. La figure 4 illustre l'utilisation de plusieurs segments existants, afin de suivre le chemin spécifique souhaité.

Les avantages de Segment Routing sont une très bonne scalabilité (résistance au facteur d'échelle) puisqu'aucun nouvel état n'est créé sur les routeurs de transits (contre une scalabilité en $O(N^2)$ pour RSVP-TE) ; et une utilisation instantanée du chemin car il ne nécessite ni signalisation, ni installation dans les tables de commutations (hors routeur source lui-même).

2.2 Segments

Un « segment » est une instruction qu'un nœud exécutera sur le paquet reçu.

Plusieurs types de segments sont définis, et d'autres pourront l'être à l'avenir. Des exemples de segments sont : acheminer le paquet selon le plus court chemin vers le routeur identifié par l'adresse IP 192.0.2.1 ; acheminer le paquet sur une interface de sortie spécifique ; appliquer un traitement X...

Un « segment ID (SID) » est un identifiant du segment. Pour SRv6, il s'agit d'une adresse IPv6. Pour SR-MPLS, il peut s'agir d'un label ou d'un index.

Un segment peut être local à un nœud ou global au domaine SR.

Un segment local est un segment qui n'est interprété que par un seul nœud.

Un segment global est un segment qui est interprété par tous les nœuds du domaine SR.

2.2.1 Segments IGP

Un « segment IGP » est un segment associé à une information annoncée dans un IGP à état de liens.

Un segment IGP peut être associé à un préfixe IP, et s'appellera un segment de préfixe. Un segment IGP peut être associé à une adjacence IGP, et s'appellera un segment d'adjacence.

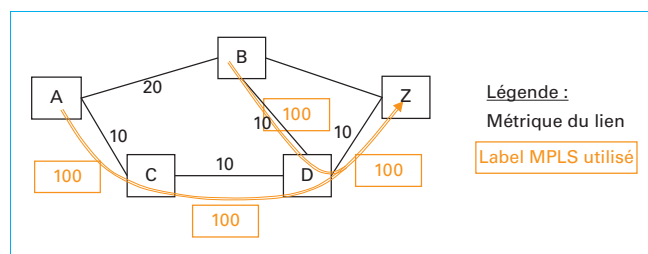


Figure 5 – Segment de préfixe IGP

2.2.1.1 Segment de préfixe (Prefix-SID)

Un « segment de préfixe IGP », ou « segment de préfixe » est un segment associé à un préfixe IGP, typiquement un préfixe IPv4 ou IPv6.

La terminologie anglaise est « IGP-Prefix-SID » ou « Prefix-SID ».

Un segment de préfixe est un segment global au sein de l'IGP, donc le segment sera installé dans la FIB de tous les routeurs de l'IGP. Il est routé de la même façon que le préfixe IGP. C'est un segment multipoint à point le long de l'arbre de plus court chemin IGP. S'il y a plusieurs chemins parallèles de même coût, le trafic sera réparti sur l'ensemble de ces chemins, comme en IP. Pour SR-MPLS, ce segment est identique à un LSP établi par le protocole LDP (figure 2). La figure 5 illustre un segment de préfixe IGP vers la destination Z identifiée par son adresse IP de loopback.

Pour des besoins plus avancés, Segment Routing permet de router le segment de préfixe selon un algorithme différent de celui du plus court chemin selon la métrique IGP. Pour cela, le segment est annoncé avec un numéro d'algorithme de routage à utiliser, et chaque routeur annonce les algorithmes de routage qu'il supporte. L'extension [Flex-Algo] permet même à l'opérateur du réseau de définir de nouveaux algorithmes de routage pour son réseau, en spécifiant lui-même l'algorithme de calcul (e.g. le plus court chemin), la métrique à minimiser (e.g. métrique IGP, métrique TE, délai) et les liens à utiliser/exclure (afin de restreindre la topologie à considérer). Ceci permet de définir des segments minimisant le délai, ou des segments restreints à une partie du réseau par exemple pour définir plusieurs plans de routages disjoints (e.g. pair, impair).

Le Segment ID (SID) est un nombre (label ou index) qui doit être provisionné par l'opérateur du réseau, de la même manière qu'une adresse IP. Une étape de configuration supplémentaire est donc nécessaire par rapport à l'utilisation de LDP. La figure 6 illustre un exemple de configuration d'un SID de segment de préfixe.

Le SID est annoncé dans l'IGP par un attribut du préfixe IP et encodé sous la forme d'un sous-TLV (Type Longueur Valeur), de la même façon que les autres attributs IGP (e.g. la métrique). La figure 7 détaille l'encodage du sous-TLV IS-IS encodant le SID du préfixe de segment. La figure 12 illustre un exemple de message LSP IS-IS annonçant un segment de préfixe attaché à un préfixe IPv4.

■ Segment de nœud (Node-SID)

Un segment de nœud est un segment de préfixe qui identifie un nœud de manière unique. La terminologie anglaise est « Node-SID ».

La seule contrainte est que ce préfixe ne doit pas être annoncé par plus d'un nœud. L'utilisation d'un segment de nœud est intéressante pour forcer un paquet à passer par un nœud particulier (nœud identifié dans la topologie de l'IGP link state), alors qu'un segment de préfixe pourrait être annoncé par plusieurs nœuds.

Un segment de nœud est annoncé comme un segment de préfixe avec le drapeau (flag) 'N' (N-Flag) positionné à 1.

■ Segment anycast (Anycast-SID)

Un segment anycast est un segment de préfixe annoncé par plusieurs nœuds.

Le fonctionnement est semblable à un préfixe IP anycast : le préfixe est routé sur le plus court chemin vers le nœud dont la métrique pour joindre le préfixe est la plus faible. SR impose que le SID annoncé soit le même pour tous les routeurs, de la même manière que le préfixe IP est le même.

■ Label identique sur tous les nœuds de transit ?

Avec Segment Routing, le SID n'est annoncé qu'une seule fois par le protocole de routage : c'est le nœud de destination qui annonce à la fois le préfixe IP et le SID SR. Lorsque le segment est global, ce segment est utilisable sur tous les routeurs et tous

```
interface Loopback0
  ipv4 address 25.0.1.17/32
!
router isis 1
  interface Loopback0
    address-family ipv4 unicast
      prefix-sid index 1
```

Figure 6 – Configuration d'un préfixe SID (Cisco IOS-XR)

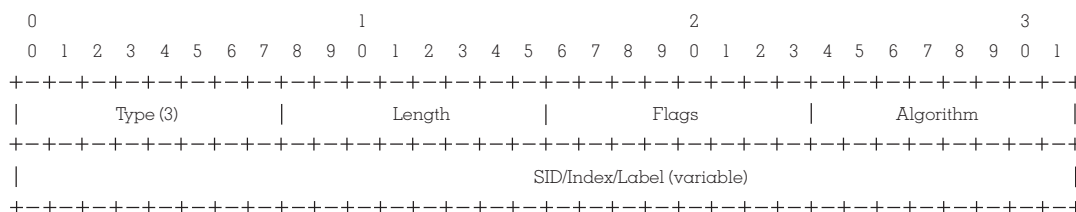


Figure 7 – Encodage dans IS-IS du SID de segment de préfixe

ceux-ci doivent installer le SID dans leur table de commutation. Pour SR-MPLS, si le SID est annoncé sous la forme d'un label MPLS, cela impose à tous les routeurs d'installer le même label dans leur table de commutation. Cela est plus aisé pour l'opérateur de réseau car le long du tunnel MPLS, le label commuté est toujours le même ce qui facilite l'identification des pannes dues à un label incorrect (alors qu'avec LDP ou RSVP-TE, le label change sur chaque routeur le long du tunnel). Mais cela est contraire à l'architecture MPLS qui définit que le label est local à chaque routeur et donc choisi librement par chaque routeur. Dans le cadre de SR-MPLS, il aurait été possible de changer ce point de l'architecture MPLS, mais lors de la conception de Segment Routing, il n'y a pas eu consensus au sein de l'industrie pour le faire, en particulier afin de faciliter l'introduction de SR-MPLS sur les équipements déjà déployés.

Ainsi, pour les segments globaux, SR-MPLS ne peut annoncer des SID sous forme d'un label. À la place, un mécanisme d'indirection supplémentaire a été ajouté :

- chaque routeur annonce un bloc de labels qu'il choisit librement. Le bloc sera utilisé pour y installer les segments globaux. Ainsi les labels MPLS restent localement choisis par chaque routeur. Ce bloc de labels est appelé SRGB (Segment Routing Global Block). Attention : le terme « Global » fait référence à segment global. Les labels de ce bloc sont par contre locaux au routeur qui annonce le bloc ;

- le SID est annoncé par le routeur destination sous forme d'un index. Chaque routeur de transit utilisera le label de son SRGB indexé par cet index.

La figure 8 illustre le fonctionnement de cette indirection :

- chaque routeur (de transit) annonce son SRGB. Par exemple, le nœud C annonce le bloc de label SRGB(C) := [1000 ; 2000] ;
- chaque routeur (destination) annonce son SID sous forme d'un index. Par exemple, le nœud Z annonce son SID index 100 ;
- il est possible de calculer le label (local) utilisé par un routeur quelconque (e.g. C) pour le segment global de Z par une simple addition :

Label utilisé par A pour joindre Z est

Label (A, Z) := SRGB (A) + Index (Z).

Dans notre exemple, le label est 1 000 + 100 = 1 100.

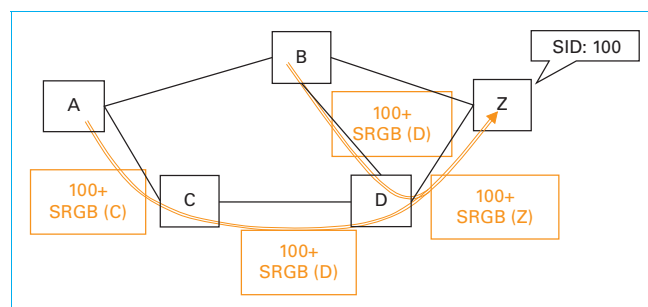


Figure 8 – SR-MPLS : fonctionnement du SRGB et des SID index

De manière intéressante, si l'opérateur configure le même SRGB sur chaque routeur, on retrouve le comportement d'origine avec, pour un segment global donné, le même label utilisé par tous les nœuds. C'est la configuration recommandée afin de simplifier le comportement du réseau. Certains types de routeurs peuvent ne pas supporter la configuration d'un SRGB quelconque, à cause de restriction matérielle ou logicielle. Dans ce cas, un SRGB différent est configuré sur ces routeurs. Même dans ce cas, en limitant le nombre de SRGB différents, et en les alignant sur les mêmes digits (par exemple les milliers), il reste aisé d'identifier le SID index qui restera constant sur les digits de rangs inférieurs.

Par **exemple** le label 16100 sur le routeur Z utilisant le SRGB [16000 ; 18000] et le label 19100 sur le routeur C utilisant le SRGB [19000 ; 21000] partagent les mêmes trois digits de poids faible indiquant directement le SID (100).

La figure 9 détaille l'encodage du sous-TLV IS-IS transportant le SRGB. Le SRGB est annoncé sous la forme du premier label utilisable (SID), puis de la taille du bloc en nombre de labels (Range). Si besoin, un SRGB peut être annoncé sous la forme de plusieurs blocs de labels, qui seront utilisables consécutivement comme un unique bloc virtuel. La figure 12 illustre un exemple de message LSP IS-IS annonçant un bloc de labels globaux (SRGB).

2.2.1.2 Segment d'adjacence (Adj-SID)

Un « segment d'adjacence IGP » ou « segment d'adjacence » est un segment unidirectionnel, routé le long de l'adjacence (lien) IGP indépendamment de la métrique de cette adjacence. La terminologie anglaise est « IGP-Adjacency-SID » ou « Adjacency-SID » ou « Adj-SID ».

Même si la norme permet plusieurs options, en pratique, un segment d'adjacence est un segment local, d'un seul bond (entre deux routeurs voisins) et annoncé dans l'IGP sous forme d'un label MPLS.

La figure 10 illustre un segment d'adjacence annoncé par le nœud B et routé vers le nœud Z. Son SID est le label 5000. Ce label est uniquement installé dans la table de commutation du nœud B avec la signification suivante :
LFIB (B) : label 5000 → pop du label, envoi du paquet sur l'interface vers Z.

Si le routeur A souhaite envoyer un paquet le long du chemin A-B-Z, il doit pousser deux SIDs : le segment de nœud pour rejoindre B, puis le segment d'adjacence pour utiliser l'adjacence B-Z. En effet, ce segment d'adjacence est local au routeur B et non connu des autres routeurs. Pour l'utiliser, il faut préalablement envoyer le paquet sur le routeur B.

Le SID est annoncé dans l'IGP par un attribut de l'adjacence IGP et encodé sous la forme d'un sous-TLV, de la même façon que les attributs IGP actuels (e.g. la métrique). La figure 11 détaille l'encodage du sous-TLV IS-IS encodant le SID du segment d'adjacence. La figure 12 illustre un exemple de message LSP IS-IS annonçant un segment d'adjacence attaché à une adjacence (un lien).

Un même SID peut être annoncé sur plusieurs adjacences différentes, typiquement des adjacences parallèles vers le même voisin. Dans ce cas, un partage de charge est réalisé sur l'ensemble des adjacences. Ce partage de charge peut être inégal si les valeurs de « Weight » sont annoncées comme différentes.

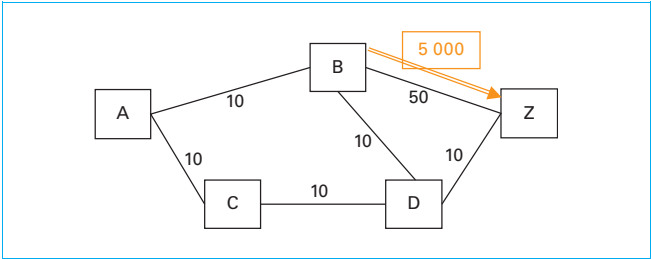


Figure 10 – Segment d'adjacence IGP

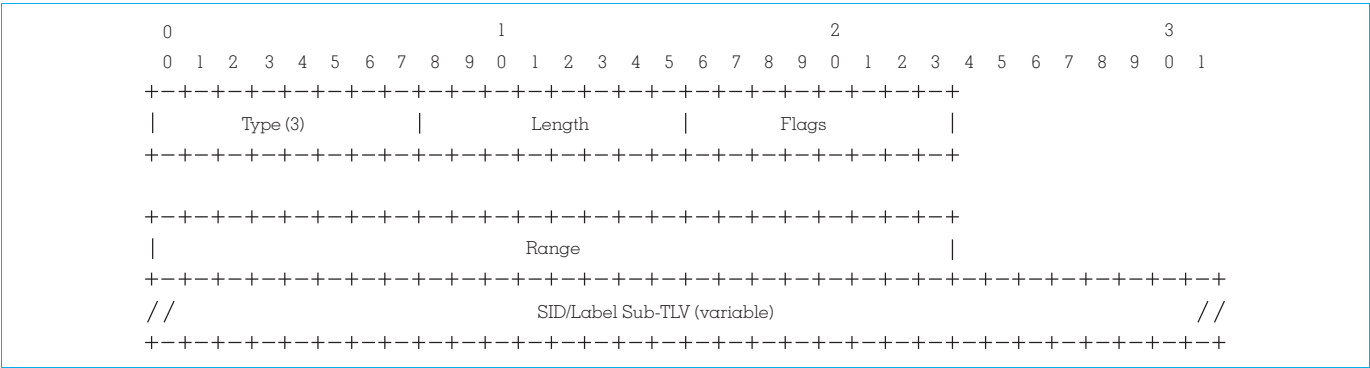


Figure 9 – Encodage dans IS-IS du bloc de labels (SRGB)



Figure 11 – Encodage dans IS-IS du SID de segment d'adjacence

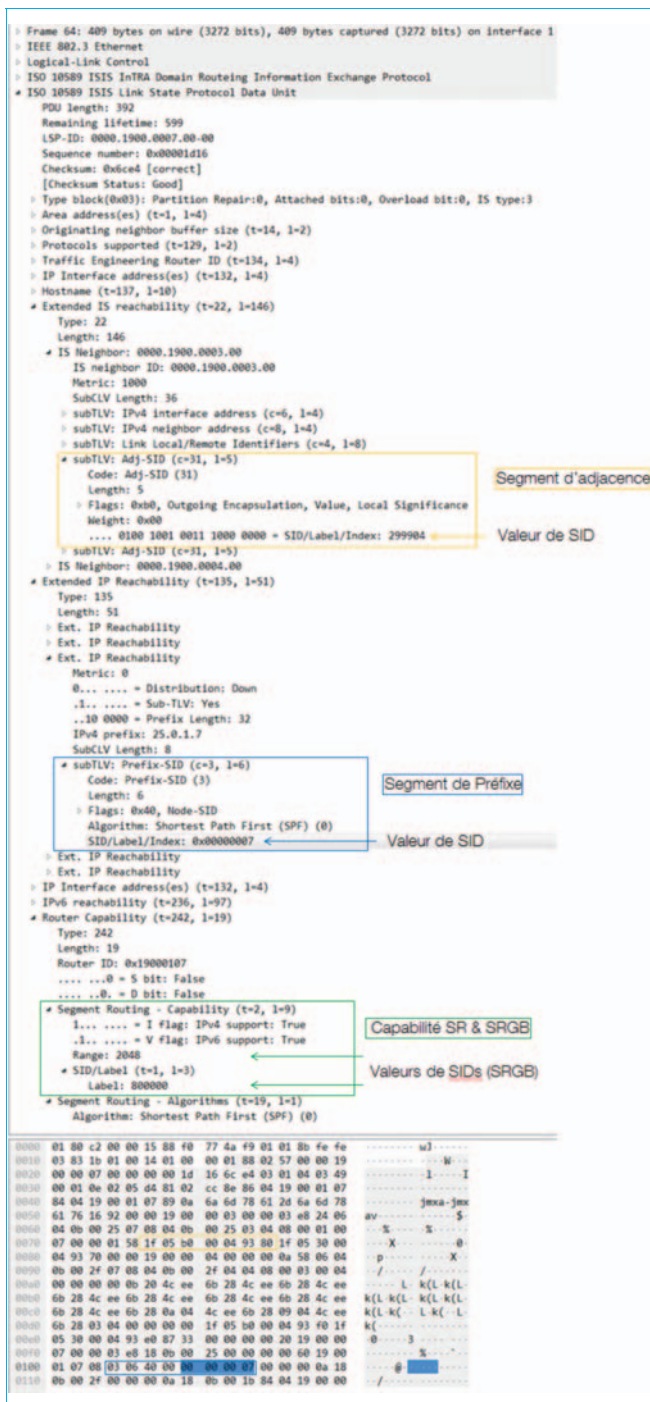


Figure 12 – Encodage des segments ID dans IS-IS

Plusieurs SID différents peuvent aussi être annoncés pour la même adjacence.

Par **exemple** un SID pour un ensemble d'adjacences parallèles – afin de pouvoir utiliser indifféremment l'ensemble des interfaces – et un SID pour chacune des interfaces – afin de pouvoir utiliser une interface spécifique.

Le SID/label annoncé est local au routeur. Il peut être choisi automatiquement et dynamiquement par le routeur. Il peut également être provisionné par l'opérateur, en particulier lorsqu'il est souhaitable que ce label soit persistant lors du redémarrage du routeur, voire de l'interface.

2.2.2 Segments BGP

Un « segment BGP » est un segment associé à une information annoncée dans BGP.

Un segment BGP peut être associé à un préfixe IP, et sera appelé un segment de préfixe BGP. Un segment BGP peut être associé à un peering EBGP, et sera appelé un segment de peering BGP.

2.2.2.1 Segments de préfixe BGP

Un segment de préfixe BGP est un segment BGP attaché à un préfixe BGP. C'est l'équivalent BGP d'un segment IGP de préfixe.

La terminologie anglaise est « BGP-Prefix segment ».

Le segment de préfixe BGP a été défini dans le contexte des data-center massifs utilisant le protocole BGP pour leur routage interne, tels que présentés dans [RFC 7938] pour les réseaux IP et dans le draft « BGP-Prefix Segment in large-scale data centers » [14] pour les réseaux Segment Routing. Néanmoins, ils peuvent être utilisés plus généralement dans les réseaux annonçant des routes MPLS [RFC 3107] [RFC 8277].

Le segment de préfixe BGP est annoncé en BGP sous forme de route IP labellisée [RFC 3107] [RFC 8277] à laquelle est ajouté un attribut BGP « BGP Prefix-SID » annonçant un index de label. Le fonctionnement est globalement identique aux routes BGP labellisées telles que définies dans [RFC 3107] [RFC 8277]. En particulier le label à utiliser pour l'émission d'un paquet MPLS est celui reçu dans la route labellisée. La nouveauté est que le label installé dans le label de commutation MPLS, et utilisé par les paquets reçus, n'est pas un label « au hasard » mais est déterminé par l'index de label (SRGB + index de label). Ainsi, sous réserve d'avoir le même SRGB, tous les routeurs sur le chemin vont allouer et utiliser le même label. Cela facilite la détection d'erreur pour l'opérateur du réseau, ainsi que le travail du contrôleur de réseau pour utiliser plusieurs segments consécutifs et ainsi choisir un chemin particulier.

2.2.2.2 Segments de peering BGP

Un segment de peering BGP est un segment BGP local attaché à un peering EBGP, c'est-à-dire à une interconnexion avec un autre système autonome.

La terminologie anglaise est « BGP Peering segment ».

Un segment de peering permet de diriger le trafic vers un point de sortie spécifique du système autonome, d'où leur nom anglais : « EPE Egress Peer Engineering ». La solution EPE est définie dans le draft « Segment Routing Centralized BGP Egress Peer Engineering » [15] et l'extension protocolaire permettant d'annoncer les segments dans BGP-LS est définie dans le draft « BGP-LS extensions for Segment Routing BGP Egress Peer Engineering » [16].

On peut distinguer trois types de segments de peering BGP :

- le segment de peering de nœud (PeerNode segment), pour diriger le trafic vers un routeur de bordure extérieur au système autonome ;
- le segment de peering d'adjacence (PeerAdj segment), pour diriger le trafic vers une interface externe ;
- le segment de peering de groupe (PeerSet segment), pour diriger le trafic vers un groupe de peer.

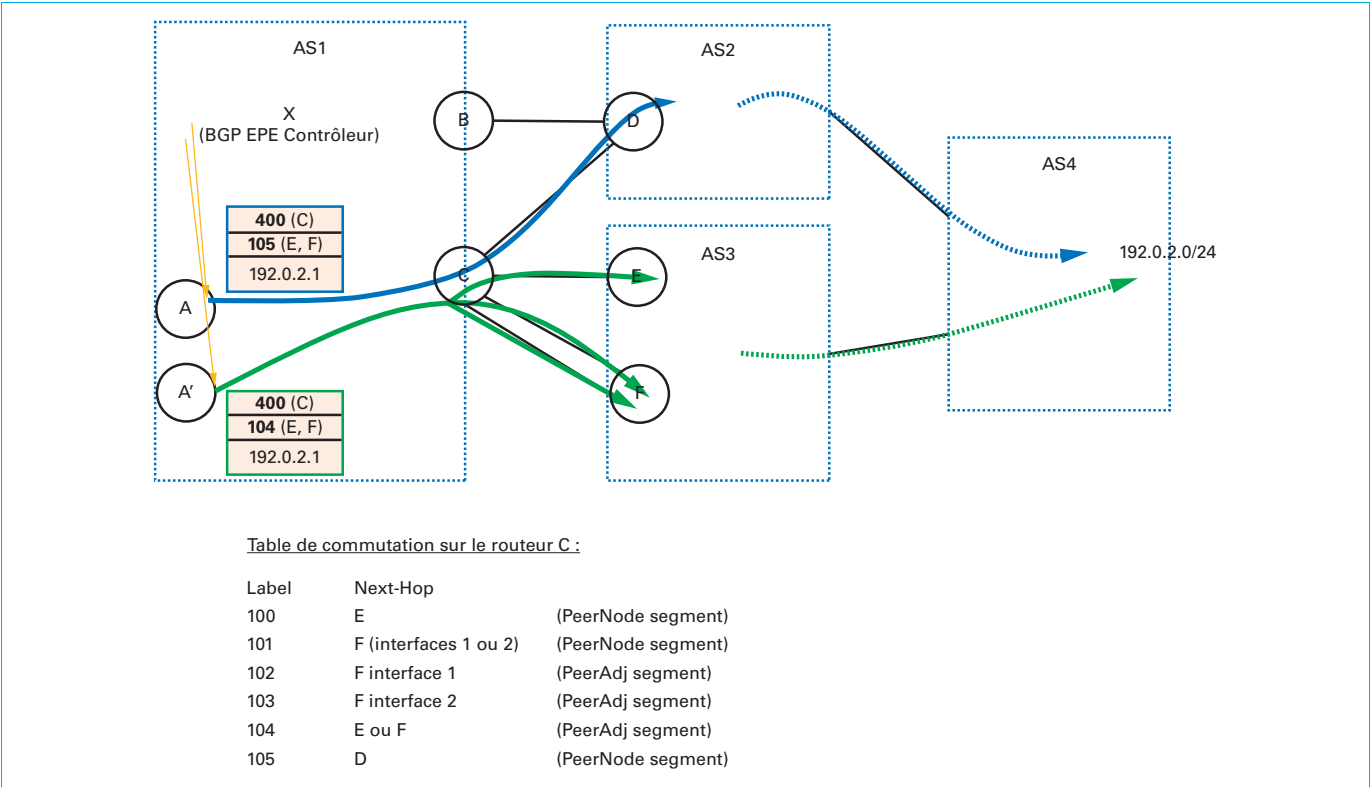


Figure 13 – Segments de peering BGP (EPE – Egress Peer Engineering)

La figure 13 illustre les segments de peering BGP et leurs usages. Un contrôleur peut ainsi choisir les chemins de sorties de son système autonome (AS) par flux et par routeur source :

- afin que le routeur A joigne la destination 192.0.2.0/24 via le routeur de sortie (ASBR) « C » puis via l'AS 2, le routeur A doit pousser un segment de nœud pour joindre C (label 400) puis un segment de peering de nœud (105) pour joindre le routeur D situé à la bordure de l'AS 2 ;
- afin que le routeur A' joigne la destination 192.0.2.0/24 via le routeur de sortie (ASBR) « C » puis via l'AS 3 à travers l'ensemble des interfaces de C vers l'AS 3, le routeur A' doit pousser un segment de nœud pour joindre C (label 400) puis un segment de peering de groupe (104) pour joindre le routeur E ou F situé à la bordure de l'AS 3.

Il faut noter que l'AS 1 ne peut contrôler que le chemin utilisé par son trafic sortant. En effet, son trafic entrant est contrôlé par l'AS envoyant le trafic, ici AS2 et AS3.

2.2.3 Politique Segment Routing et Segment d'association

Une politique Segment Routing définit la liste ordonnée de segments ajoutés par le nœud source.
La terminologie anglaise est « SR Policy ».

Les politiques SR sont définies dans le draft « Segment Routing Policies Architecture » [12] et leur annonce dans BGP spécifiée dans le draft « Advertising Segment Routing Policies in BGP » [13].

Une politique SR est identifiée par le tuple <source, couleur, destination>. Dans le contexte d'un routeur source particulier, elle peut être identifiée par le couple <couleur, destination>. La destination

est spécifiée sous la forme d'une adresse IP (v4 ou v6). La couleur est un identifiant sous la forme d'un nombre de 32 bits choisi par l'opérateur pour désigner le type de politique, par exemple « routage minimisant le délai ».

À une politique SR est associé un ou plusieurs chemins candidats. Le chemin candidat identifie la liste ordonnée de SID à utiliser, ou plusieurs listes en cas de partage de charge sur plusieurs chemins. Le chemin actif est celui qui a la meilleure préférence parmi les chemins candidats utilisables. La figure 14 illustre le modèle d'information d'une politique Segment Routing.

À chaque politique Segment Routing est associé un identifiant de segment d'association (Binding Segment). L'identifiant de segment d'association (Binding Segment ID : BSID) permet à tout nœud distant d'imposer cette politique sur un paquet, sans avoir besoin de connaître le détail de cette politique, ni être impacté par un changement de la liste de SIDs utilisée par la politique, en particulier lors d'un changement de topologie dans le réseau.

2.3 Fonctionnement

2.3.1 Plan de commande

Le plan de commande utilisé par Segment Routing est constitué de deux parties :

- la distribution des segments et leur identifiant associé. Ces informations sont relativement statiques et annoncées par les protocoles de routages IP distribués tels qu'OSPF, IS-IS et BGP (cf section 2.2) ;
- la communication avec un contrôleur centralisé, si utilisé :
 - le contrôleur peut obtenir les informations de topologie du réseau et d'identifiants de segments en participant passivement au(x) protocole(s) de routage ou via le protocole

```

SR policy POL1 <headend, color, endpoint>
Candidate-path CP1 <protocol-origin = 20, originator = 100:1.1.1.1, discriminator = 1>
Preference 200
Weight W1, SID-List1 <SID11...SID1i>
Weight W2, SID-List2 <SID21...SID2j>
Candidate-path CP2 <protocol-origin = 20, originator = 100:2.2.2.2, discriminator = 2>
Preference 200
Weight W3, SID-List3 <SID31...SID3i>
Weight W4, SID-List4 <SID41...SID4j>

```

Figure 14 – Modèle d'information d'une politique Segment Routing

BGP-LS [RFC 7752] (draft « BGP Link-State extensions for Segment Routing » [11]) BGP-LS redistribue ces informations de l'IGP sous une forme plus adaptée car indépendante du protocole de routage (IS-IS vs OSPF), à distance (sur TCP/IP) et sans risque d'interférence avec le protocole de routage (isolation),

- le contrôleur peut installer un chemin particulier sur un routeur source, en signalant une liste de segments à utiliser, via le protocole PCEP [5440] (draft PCEP Extensions for Segment Routing [10]) ou BGP (draft Advertising Segment Routing in BGP [13]).

2.3.2 Plan de commutation

Le plan de commutation de SR-MPLS utilise des paquets MPLS standards. Le paquet est routé par la source en utilisant une pile de labels MPLS. Chaque label identifie un segment. La pile de labels spécifie la liste des segments à utiliser consécutivement. Une fois que la source a ajouté la pile de labels, rien ne distingue ce paquet SR-MPLS d'un paquet MPLS :

- a) le réseau achemine le paquet le long du segment en cours par consultation du label en haut de pile ;
- b) arrivé à la fin du segment, le label est supprimé et le label suivant est utilisé (cf étape « a »).

La figure 4 illustre l'utilisation de plusieurs segments existants, afin de suivre le chemin spécifique souhaité. Le routeur A souhaite suivre le chemin A-D-E-F-C. Pour ce faire, il utilise consécutivement des segments préexistants. Ici il choisit d'utiliser trois segments :

- un segment de préfixe à destination de D (en bleu), qui permet de suivre le chemin A-D ;
- un segment de préfixe à destination de F (en rouge), qui permet de suivre le chemin D-E-F ;
- un segment de préfixe à destination de C (en vert), qui permet de suivre le chemin F-C.

D'autres choix de segments sont possibles, par exemple utiliser quatre segments d'adjacence (AD, DE, EF, FC).

Les opérations possibles sur le paquet sont identiques aux opérations MPLS. Leur nom a été changé afin de s'appliquer indifféremment au plan de transfert SR-MPLS ou SRv6 :

- PUSH est l'opération d'ajout d'un segment au sommet de la liste de segments. Elle est équivalente à l'opération PUSH en MPLS ;
- NEXT est l'opération effectuée à la fin d'un segment, afin de passer au segment suivant. Elle est équivalente à l'opération POP en MPLS ;
- CONTINUE est l'opération effectuée le long d'un segment. Elle est équivalente à l'opération SWAP en MPLS. Cependant, sous réserve d'utiliser le même SRGB sur tous les routeurs, SR-MPLS « remplace » (swap) le label reçu par le même label, alors qu'en MPLS classique, le label change à chaque saut.

2.4 Bénéfices de SR-MPLS

Par rapport à LDP, SR-MPLS permet une simplification en supprimant le protocole LDP et en particulier les difficultés d'interactions entre LDP et l'IGP tels que décrites dans [RFC 5443], [RFC 6138], [RFC 5919]. En effet, bien que LDP et IGP soient censés être indépendants – l'IGP déterminant le routage et LDP annonçant les labels –, en pratique l'utilisation par l'IGP d'une nouvelle interface pose problème tant que LDP n'a pas encore annoncé tous ses labels.

SR-MPLS permet également d'utiliser le même label MPLS sur tous les routeurs le long du segment. Cela simplifie en particulier la détection et la location des dysfonctionnements. Cela peut également améliorer la scalabilité de certains routeurs.

SR-MPLS permet d'utiliser n'importe quel chemin dans le réseau, alors que LDP ne permet que de suivre le plus court chemin. Ceci permet de nouvelles applications dans le réseau, telles que présentée dans la section 4. En particulier, cela permet une solution de Fast ReRoute capable de protéger toutes les pannes.

En contrepartie, SR-MPLS nécessite la configuration d'un identifiant de segment de préfixe/nœud pour chaque préfixe/nœud. Cette configuration est très similaire à la configuration d'une adresse IP sur le routeur.

Par rapport à RSVP-TE, SR-MPLS présente une meilleure scalabilité car sa scalabilité est en N contre kN^2 pour RSVP-TE (N étant le nombre de routeurs (destinations) dans le réseau, k étant le nombre de tunnels vers une destination). De plus, le chemin Segment Routing est utilisable immédiatement alors que RSVP-TE nécessite une signalisation de bout en bout ainsi que l'installation de nouveaux labels dans les tables de commutations. Enfin SR-MPLS permet d'utiliser les interfaces parallèles et plus généralement les chemins multiples de même coût, et ainsi utiliser l'ensemble de la bande passante du réseau sans devoir multiplier le nombre de tunnels.

En contrepartie, SR-MPLS nécessite d'ajouter plusieurs labels sur le paquet, ce qui peut être limitant sur le routeur source, mais également sur les routeurs de transit effectuant du partage de charge (car l'en-tête IP permettant d'identifier le flux est caché derrière un plus grand nombre de labels). L'utilisation de segments de préfixe plutôt que de segments d'adjacence permet de réduire ce nombre de labels, mais présente en contrepartie l'inconvénient de nécessiter plus de calculs dans le plan de commande du routeur source, et de ne pas permettre un contrôle aussi fin du chemin utilisé par le paquet en cas de changements de topologie. Ceci peut être gênant pour certaines applications spécifiques nécessitant un contrôle absolu du chemin.

SR-MPLS, utilisé seul, ne permet pas la réservation de bande passante le long du chemin. Cependant, d'une part cette fonction est très peu déployée, et d'autre part cette fonction peut être réalisée de façon plus précise par un serveur PCE stateful [RFC 5440] et [RFC 8231] (section 4.2.2).

3. Segment Routing IPv6 (SRv6)

SRv6 est la déclinaison de Segment Routing pour le plan de transfert IPv6.
SRv6 identifie un segment par une adresse IPv6.

Dans le plan de transfert, SRv6 est un nouveau type d'en-tête de routage IPv6 (Routing Extension Header) appelée en-tête Segment Routing (Segment Routing Header : SRH) représenté en figure 15. Le SRH est principalement constitué d'une liste d'adresses IPv6 identifiant les segments à traverser (Segment List), ainsi que d'un compteur indiquant le nombre de segments restant à parcourir (Segments Left) afin de pouvoir identifier le prochain segment à visiter. Il peut également contenir une liste de données au format TLV. Un champ « Tag » permet enfin d'identifier une classe de paquets.

- Trois types de nœuds peuvent être distingués :
- le nœud source crée le paquet IPv6 avec pour adresse IP destination l'adresse du premier segment. Si plus d'un segment est nécessaire, un en-tête SRH est ajouté. En corolaire, si un seul segment est nécessaire, le paquet envoyé est un paquet IPv6 natif ;
 - les routeurs de transit transportent le paquet vers la destination du segment en cours. Cette destination étant indiquée dans l'adresse de destination IPv6 habituelle, ces routeurs de transit sont des routeurs IPv6 standard. Ils n'ont pas besoin de supporter SRv6 ;

- les routeurs de terminaison de segments, qui ont annoncé les segments. Ces routeurs annonçant des segments doivent supporter l'extension SRv6 [6] qui peut ne pas être supportée par les routeurs précédemment déployés.

Sur le routeur de terminaison de segment, le traitement du SRH est le suivant :

- le champ « Segments Left » est décrémenté de 1 ;
- l'adresse IPv6 indiquée dans Segment List [Segments Left] est copiée dans l'adresse IPv6 de destination de l'en-tête IPv6.

À la date de cet article, SRv6 est moins mature que SR-MPLS car il a été standardisé plus tardivement et parce qu'il nécessite une évolution du plan de transfert des routeurs SRv6. À court terme, SRv6 paraît intéressant dans les parties de réseaux ne supportant pas MPLS, telles que les hôtes IPv6 (ordinateur, serveur, téléphone...), les réseaux de data center, les réseaux terminaux situés chez les utilisateurs résidentiels et entreprises. À court terme, il ne devrait pas remplacer massivement le protocole MPLS dans les réseaux d'opérateurs car ces routeurs à très haut débit peuvent ne pas (suffisamment) supporter l'extension SRv6 et parce que la majorité des applications actuelles est réalisable avec SR-MPLS. À plus long terme, SRv6 pourrait remplacer MPLS.

4. Applications

Segment Routing permet d'acheminer un paquet le long d'un chemin particulier, sans créer d'états dans le réseau de transit, ni nécessiter de signalisation. Cela permet de nouvelles applications

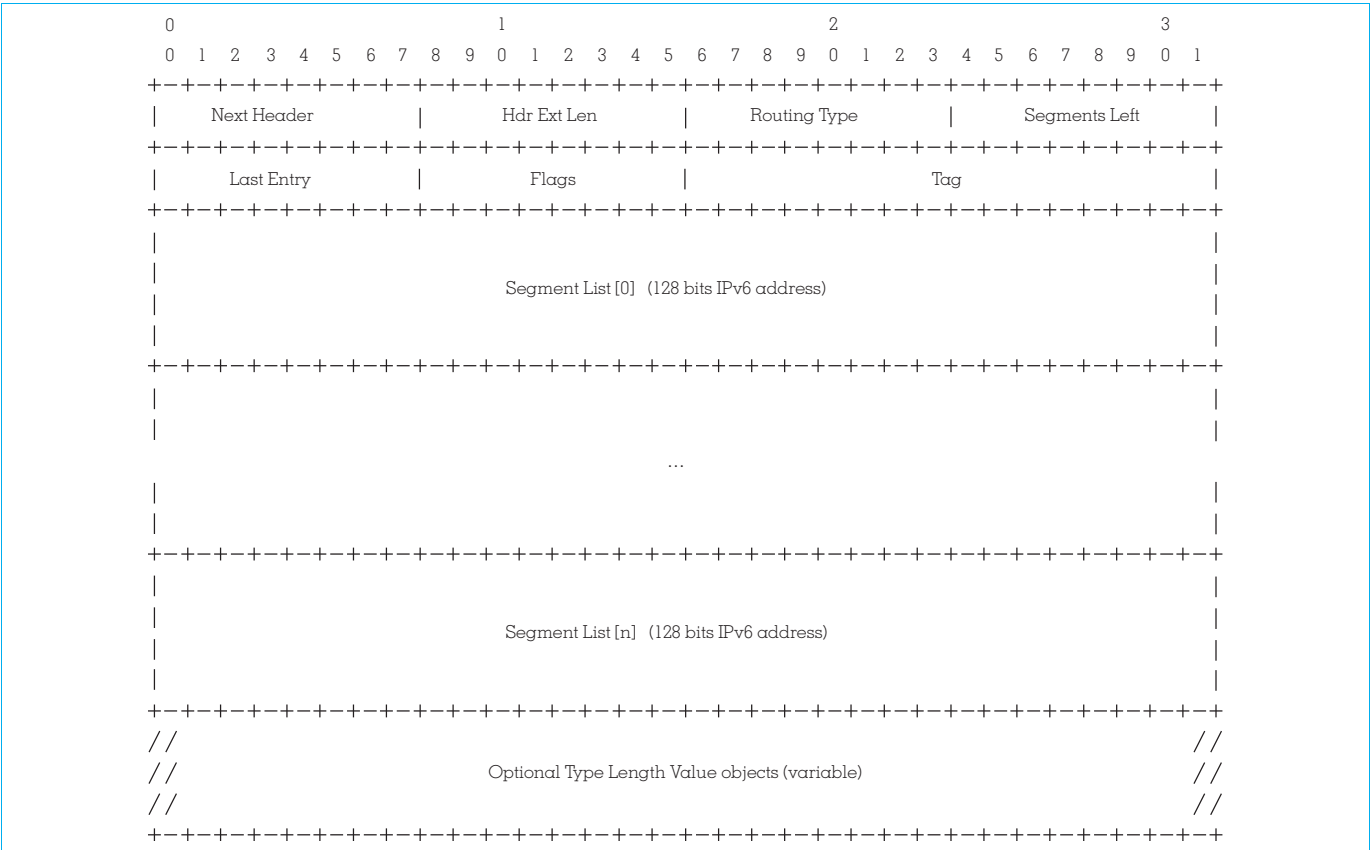


Figure 15 – SRv6 : en-tête SRH (Segment Routing Header)

dans les réseaux qui utilisaient précédemment IP ou MPLS/LDP. Pour les réseaux qui utilisaient MPLS/RSVP-TE, la majorité des applications étaient déjà possibles mais des changements sont apportés par Segment Routing.

Cette section explore trois familles d'applications :

- une meilleure réaction du réseau en cas de panne ;
- une meilleure adaptation du routage du trafic dans le réseau (ingénierie de trafic « TE ») ;
- une meilleure opération du réseau.

4.1 Réactions aux pannes

En cas de panne de lien ou de routeur dans le réseau, les paquets IP/MPLS, dont la route traverse l'élément en panne, sont détruits jusqu'à ce que le routage IP/MPLS s'adapte à la nouvelle topologie. Cette durée de re-routage est de l'ordre de 500 ms à 2 secondes dans les grands réseaux, pour les pannes impactant le protocole de routage interne. Compte tenu des hauts débits transportés dans un réseau, cela représente un nombre important de paquets (environ 10 millions de paquets pour la panne d'une interface à 100 Gbit/s).

4.1.1 Fast ReRoute (FRR)

Les solutions de Fast ReRoute permettent de réagir beaucoup plus rapidement, en 50 ms. Pour cela, elles se basent sur une réaction locale, pré-calculée avant la panne. Comme la réaction est locale, seul le nœud détectant la panne a connaissance de la panne. La difficulté est de pouvoir acheminer le paquet à travers le réseau, alors même que les tables de routage du réseau acheminent encore ce paquet vers l'élément en panne. Cela nécessite typiquement de pouvoir forcer ce paquet à suivre une route spécifique, différente de celle que le réseau utiliserait. Cela est difficile pour les réseaux IP et MPLS/LDP car leur plan de transfert ne permet que d'envoyer le paquet sur le plus court chemin installé dans le réseau. Segment Routing permet d'utiliser un chemin spécifique choisi par la source, et c'est donc l'outil nécessaire pour le Fast ReRoute.

En particulier la solution de Fast ReRoute TI-LFA (Topology Independent Loop Free Alternate) permet au réseau de protéger 100 % des pannes de liens ou de nœuds en 50 ms. De plus, durant la protection, le chemin utilisé est le plus court chemin entre le routeur détectant la panne et la destination du paquet, ce qui est une propriété novatrice. Pour plus d'information sur les solutions de Fast ReRoute et sur TI-LFA, le lecteur est invité à lire l'article [TE 7 592].

4.1.2 Loop Free Convergence

Même lorsqu'une solution de Fast ReRoute protège le trafic en 50 ms, l'opération de routage effectuée ensuite par le réseau peut provoquer des saturations temporaires d'interfaces entraînant des pertes de paquets et une augmentation des délais/de la gigue. Cela est dû au fait que le reroutage est une opération distribuée

sur l'ensemble des routeurs et qu'il n'est pas possible de garantir que tous les routeurs du réseau mettent à jour leur table de commutation exactement au même instant. De ce fait, une boucle de routage peut apparaître temporairement entre deux routeurs adjacents, comme illustré à la figure 16.

Segment Routing est un bon outil pour traiter le problème puisqu'il permet au routeur changeant son routage de forcer le routage du paquet selon le nouveau chemin, indépendamment des tables de routage utilisées par les autres routeurs du réseau. De plus, les protocoles de routage à états de liens tels qu'OSPF et IS-IS permettent au routeur source de calculer les tables de routage de tous les routeurs du réseau, à la fois dans l'ancienne topologie et dans la nouvelle topologie. Cela permet au routeur source de calculer les liens qui peuvent subir des micro-boucles et de limiter l'utilisation de Segment Routing à la traversée de ces liens.

4.2 Ingénierie de trafic (TE)

L'ingénierie de trafic consiste à adapter le routage du trafic afin d'optimiser l'usage du réseau.

4.2.1 Usage tactique

Un usage tactique limite l'adaptation du routage à un petit sous-ensemble du trafic, uniquement pour s'adapter à une situation particulière.

Les besoins peuvent être assez divers, tels que :

- le manque de bande passante sur un lien particulier, ce qui nécessite de rediriger une partie du trafic sur un autre chemin. En particulier, cela peut être utile dans certaines parties du monde où la bande passante peut être très chère ou tout simplement disponible en quantité insuffisante ;
- un besoin particulier exprimé par le client, tel que l'utilisation de deux routes disjointes à travers le réseau, ou le besoin de ne pas traverser un pays particulier suite à des risques de sécurité ;
- le besoin de minimiser le délai pour certaines applications ; en particulier lorsque les métriques de routage visent à maximiser la bande passante plutôt que minimiser le délai ; ou lorsque le routage fait du partage de charge entre le chemin nominal et le chemin alternatif plus long.

Pour toutes ces applications tactiques, la grande majorité du trafic utilise le plus court chemin IGP et donc n'a pas besoin de mécanisme supplémentaire. Pour le trafic devant suivre un chemin particulier, il suffit en général d'ajouter un petit nombre de segments afin de suivre le « détournement » souhaité.

Avec l'utilisation de segments BGP de peering, il est même possible de choisir l'interface de sortie du système autonome, ce qui est relativement peu aisé en l'absence de Segment Routing (cf. section 2.2.2.2 et figure 13).

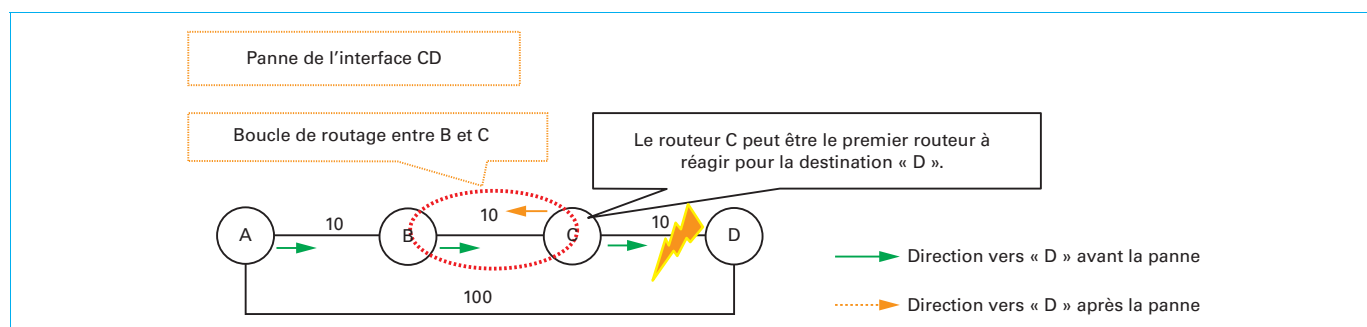


Figure 16 – Micro-boucle de routage durant l'opération de reroutage

Du fait de la forte scalabilité de Segment Routing, qui ne crée pas d'états dans le réseau de transit, il est envisageable de choisir un chemin application par application.

4.2.2 Réservection de bande passante

RSVP-TE permet de réserver de la bande passante le long du chemin utilisé. Segment Routing ne permet pas nativement de le faire car il n'y a pas de signalisation du chemin par la source.

- Cependant :
- très peu de déploiements réservent réellement la bande passante utilisée par le tunnel ;
 - la bande passante réservée est annoncée dans le protocole de routage interne relativement peu fréquemment afin de ne pas impacter le routage. De ce fait, en cas de changement de topologie/réervations significatives, les informations de bande passante réservées sont obsolètes donc peu utiles ;
 - SR permet de réserver de la bande passante de manière centralisée en utilisant le protocole PCEP stateful [IETF 5440] [IETF 8231]. Ce mode centralisé est beaucoup plus précis et à jour que le mode distribué de RSVP-TE. Néanmoins, le mode centralisé peut également poser des difficultés de scalabilité et de gestion de sa redondance.

4.3 Opération du réseau

4.3.1 Monitoring du réseau

En utilisant Segment Routing, une unique sonde peut envoyer des paquets le long de n'importe quel chemin du réseau. Ces paquets de sondes utilisent le vrai plan de commutation (hardware) utilisé par le réseau pour acheminer les paquets (contrairement à certains protocoles d'OAM). En testant de nombreux chemins, il est possible de monitorer chaque lien et nœud à partir d'une seule sonde placée où l'opérateur le souhaite. Il est ensuite possible d'estimer leurs taux de pertes de paquets et les délais qu'ils introduisent. La présentation [2] explique la solution et les résultats trouvés.

4.3.2 Monitoring des liens et tables de commutations

Dans un réseau MPLS, une mauvaise programmation des tables de commutations MPLS peut entraîner un mauvais acheminement des paquets. Ce type d'erreur est difficile à détecter

car les outils de management du routeur annoncent les labels que le routeur pense avoir programmé dans sa table de commutation, et non les labels réellement présents dans les tables de commutations hardware.

En utilisant Segment Routing, chaque nœud peut tester, de manière distribuée donc très scalable, ses propres liens et ses propres tables de commutations telles qu'elles sont utilisées par le trafic réel.

La figure 17 illustre un exemple où le routeur R2 teste sa table de commutation pour le label 16005 à destination de R5. Pour cela, il envoie un paquet MPLS LSP ping standard [RFC 8029] à son voisin amont R1, avec un premier label de segment d'adjacence (24012) pour forcer R1 à lui renvoyer le paquet après avoir supprimé le premier label. De ce fait, R2 reçoit un paquet de l'interface d'entrée et le segment à tester (node segment de R5). R2 achemine ce paquet en utilisant son plan de transfert habituel, celui que l'on souhaite effectivement tester. La suite de la procédure correspond au fonctionnement standard de LSP ping : du fait du TTL positionné à 3 sauts, le TTL du paquet va expirer sur le routeur aval (R3), qui va répondre au LSP Ping par une réponse permettant de vérifier que le chemin pris par le paquet est correct (ou incorrect).

Ainsi, R2 peut tester lui-même le fonctionnement de son plan de transfert réel, sans nécessiter de nouvelles extensions protocolaires sur ses voisins : le routeur amont a simplement acheminé le paquet SR-MPLS reçu et le routeur aval a répondu à la requête LSP ping habituelle.

4.3.3 Simplification du réseau

- SR apporte une simplification dans l'exploitation du réseau par :
- la suppression des sessions LDP ainsi que leurs interactions avec le routage interne. En particulier la suppression des sessions T-LDP avec des routeurs distants, qui peuvent être nécessaires pour le Fast ReRoute R-LFA ainsi que l'utilisation conjointe de RSVP-TE et de LDP ;
 - la suppression des sessions RSVP-TE utilisées pour des besoins tactiques dans les réseaux LDP ;
 - l'usage, pour un segment, du même label sur l'ensemble des routeurs du réseau. Cela facilite la recherche d'erreur de programmation des labels ;
 - l'utilisation d'une solution de Fast ReRoute qui fonctionne dans tous les cas sans besoin d'études/simulations, et utilisant un

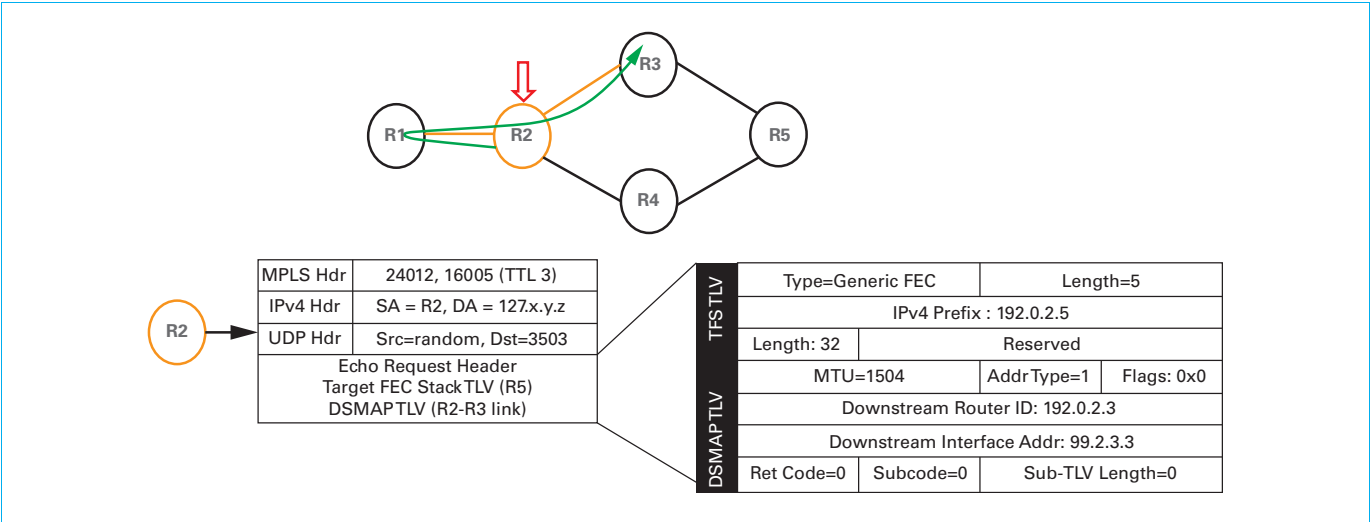


Figure 17 – Auto-vérification, par R2, de son plan de transit (Data plane monitoring)

chemin simple pour l'exploitation du réseau et son dimensionnement (à comparer aux algorithmes de chemins utilisés par LFA et RLFA) ;

– la possibilité de tester le fonctionnement de chemins avant de les mettre en production, ainsi que de tester des routes différentes lors de la recherche de panne (« et si le trafic passait par un autre chemin, aurais-je le même bug ? »).

5. Conclusion

SR-MPLS est une évolution de MPLS qui apporte beaucoup aux réseaux MPLS utilisant LDP. SR-MPLS ajoute en effet la capacité d'envoyer un paquet le long de n'importe quel chemin désiré. Les principales applications sont une meilleure réaction du réseau en cas de panne, l'ingénierie de trafic en particulier pour des besoins particuliers (tactiques) et une amélioration de l'opération des réseaux. Ces avantages ont permis une standardisation rapide de SR-MPLS, son implémentation par tous les constructeurs majeurs de routeurs et des déploiements variés.

SRv6 a été développé dans une seconde phase, en particulier du fait de l'évolution nécessaire du plan de transfert. Ce changement dans le plan de transfert nécessite plus de travaux en normalisation, une évolution du plan de transfert des routeurs, et une exigence matérielle accrue. Les bases de SRv6 sont désormais définies mais les travaux sur ses applications ne font que commencer. SRv6 a le potentiel de remplacer plusieurs technologies, en particulier dans le domaine des VPN ou réseau overlay dans les data center. Néanmoins, à ce jour, il doit encore faire ses preuves en termes de déploiements et d'adoption par l'industrie.

6. Glossaire

Segment ; *Segment*

Une instruction qu'un nœud exécutera sur le paquet reçu.

Un segment peut être une instruction topologique telle qu'acheminer le paquet selon le plus court chemin vers le routeur identifié par l'adresse IP 192.0.2.1 ; acheminer le paquet sur une interface de sortie spécifique.

Un segment peut également être une instruction de service consistant à appliquer un traitement sur le paquet.

Identifiant de segment ; *Segment ID (SID)*

Un nombre identifiant le segment et qui pourra être indiqué dans l'en-tête du paquet Segment Routing. Pour SRv6, il s'agit d'une adresse IPv6. Pour SR-MPLS il peut s'agir d'un label ou d'un index.

Segment de préfixe (IGP) ; *(IGP) Prefix Segment (Prefix-SID)*

Segment associé à un préfixe (IGP), typiquement un préfixe IPv4 ou IPv6.

Segment de nœud ; *Node Segment (Node-SID)*

Un segment de préfixe identifiant un nœud de manière unique.

Segment d'adjacence (IGP) ; *Adjacency Segment (Adjacency-SID/Adj-SID)*

Un segment unidirectionnel, routé le long de l'adjacence (lien) IGP indépendamment de la métrique de cette adjacence.

Politique Segment Routing *SR Policy*

Définit la liste ordonnée de segments ajoutés par le nœud source.

7. Sigles, notations et symboles

SR	Segment Routing
BGP	Border Gateway Protocol (protocole de routage inter-domaine, en particulier utilisé dans l'Internet)
FIB	Forwarding Information Base (Table de commutations IP)
FRR	Fast ReRoute
IGP	Internal Gateway Protocol par exemple OSPF ou IS-IS
IS-IS	Intermediate System to Intermediate System (IGP à état de liens (Link State))
LDP	Label Distribution Protocol (Protocole de signalisation permettant d'annoncer des labels MPLS et leur signification (destination) associée)
LFA	Loop Free Alternate (Solution de fast reroute)
LSP	Link State Packet (Message d'IGP à état de liens décrivant une partie de la topologie du réseau)
LSP	Label Switched Path (tunnel MPLS)
MPLS	Multi-Protocol Label Switching (Technologie de commutations de paquet basé sur des labels)
OAM	Operations, Administration and Maintenance
OSPF	Open Shortest Path First (IGP à état de liens (Link State))
PLR	Point of Local Repair (routeur activant le fast reroute)
RFC	Request For Comment (document produit par l'IETF, certains RFC sont des normes)
RLFA	Remote Loop Free Alternate (Solution de fast reroute)
SPF	Shortest Path First (Algorithme de calcul du SPT)
SPT	Shortest Path Tree (Arbre de plus court chemin)
SRLG	Shared Risk Link Group (groupe de liens partageant un même risque et donc pouvant tomber en panne en même temps)
TI-LFA	Topology Independent Loop Free Alternate (Solution de fast reroute)
VPN	Virtual Private Network (Réseau Privé Virtuel)
SRGB	Segment Routing Global Block (bloc de labels local à un routeur, utilisé pour les labels MPLS des segments globaux)
SID	Segment Identifier (identifiant de segment)
SR-MPLS	Segment Routing avec le plan de transfert MPLS
SRv6	Segment Routing avec le plan de transfert IPv6
TLV	Type Longueur Valeur/Type Length Value (Format d'encodage)

Segment Routing

par **Bruno DECRAENE**
Ingénieur réseau

Sources bibliographiques

- [1] FILSFILS (C.), MICHIELSEN (K.) et TALAULIKAR (K.). – *Segment Routing Part I*. CreateSpace Independent Publishing Platform.
- [2] GUILBAUD (N.) et CARLIDGE (R.). – *Google, "Localizing packet loss in a large complex network"*. NANOG 57. <https://www.nanog.org/meetings/nanog57/presentations/Tuesday/tues.general.GuilbaudCarlidge.Topology.7.pdf> (2013).
- [3] LISTE (J.). – *SR Data Plane Monitoring (SR-DPM) Demonstration*. Cisco <http://www.segment-routing.net/demos/2018-sr-data-plane-monitoring-demo/>.
- [4] IETF. – *Topology Independent Fast Reroute using Segment Routing*. draft-bashandy-rtgwg-segment-routing-ti-lfa, <https://tools.ietf.org/html/draft-bashandy-rtgwg-segment-routing-ti-lfa-00> (2018).
- [5] IETF. – *Segment Routing with MPLS data plane*. draft-ietf-spring-segment-routing-mpls, <https://tools.ietf.org/html/draft-ietf-spring-segment-routing-mpls-14> (2018).
- [6] IETF. – *IPv6 Segment Routing Header (SRH)*. draft-ietf-6man-segment-routing-header <https://tools.ietf.org/html/draft-ietf-6man-segment-routing-header-15> (2018).
- [7] IETF. – *SRv6 Network Programming*. draft-filsfils-spring-srv6-network-programming <https://tools.ietf.org/html/draft-filsfils-spring-srv6-network-programming-06> (2018).
- [8] IETF. – *IS-IS Extensions for Segment Routing*. draft-ietf-isis-segment-routing-extensions <https://tools.ietf.org/html/draft-ietf-isis-segment-routing-extensions-21> (2018).
- [9] IETF. – *OSPF Extensions for Segment Routing*. draft-ietf-ospf-segment-routing-extensions <https://tools.ietf.org/html/draft-ietf-ospf-segment-routing-extensions-27> (2018).
- [10] IETF. – *PCEP Extensions for Segment Routing*. draft-ietf-pce-segment-routing <https://tools.ietf.org/html/draft-ietf-pce-segment-routing-14> (2018).
- [11] IETF. – *BGP Link-State extensions for Segment Routing*. draft-ietf-idr-bgp-ls-segment-routing-ext <https://tools.ietf.org/html/draft-ietf-idr-bgp-ls-segment-routing-ext-11> (2018).
- [12] IETF. – *Segment Routing Policy Architecture*. draft-ietf-spring-segment-routing-policy <https://tools.ietf.org/html/draft-ietf-spring-segment-routing-policy-02> (2018).
- [13] IETF. – *Advertising Segment Routing Policies in BGP*. draft-ietf-idr-segment-routing-te-policy <https://tools.ietf.org/html/draft-ietf-idr-segment-routing-te-policy-05> (2018).
- [14] IETF. – *BGP-Prefix Segment in large-scale data centers*. draft-ietf-spring-segment-routing-msdc <https://tools.ietf.org/html/draft-ietf-spring-segment-routing-msdc-11> (2018).
- [15] IETF. – *Segment Routing Centralized BGP Egress Peer Engineering*. draft-ietf-spring-segment-routing-central-epe <https://tools.ietf.org/html/draft-ietf-spring-segment-routing-central-epe-10> (2017).
- [16] IETF. – *BGP-LS extensions for Segment Routing BGP Egress Peer Engineering*. draft-ietf-idr-bgp-ls-segment-routing-epe <https://tools.ietf.org/html/draft-ietf-idr-bgp-ls-segment-routing-epe-17> (2018).
- [17] IETF. – *IGP Flexible Algorithm*. draft-ietf-lsr-flex-algo <https://tools.ietf.org/html/draft-ietf-lsr-flex-algo-01> (2018).

À lire également dans nos bases

BONNIN (J.-M.). – *MPLS. Technologies de l'information, Réseaux et télécommunications* [TE 7 535] (2003).

NATAF (S.). – *VPN MPLS IPv6. Technologies de l'information, Réseaux et télécommunications* [TE 7 590] (2010).

DECRAENE (B.). – *IP Fast ReRoute (IP FRR). Technologies de l'information, Réseaux et télécommunications* [TE 7 592] (2019).

Sites Internet

Segment Routing
<http://www.segment-routing.net>

Normes et standards

RFC 2328	1998	OSPF Version 2. IETF.	RFC 5919	2010	Signaling LDP Label Advertisement Completion. IETF.
RFC 3031	2001	Multiprotocol Label Switching Architecture. IETF.	RFC 6138	2011	LDP IGP Synchronization for Broadcast Networks. IETF.
RFC 3032	2001	MPLS Label Stack Encoding. IETF.	RFC 7938	2016	Use of BGP for Routing in Large-Scale Data Centers. IETF.
RFC 3209	2001	RSVP-TE : Extensions to RSVP for LSP Tun- nels. IETF.	RFC 8231	2017	Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE. IETF.
RFC 5036	2007	LDP Specification. IETF.	RFC 3107	2001	Carrying Label Information in BGP-4. IETF.
RFC 5440	2009	Path Computation Element (PCE) Commu- nication Protocol (PCEP). IETF.			
RFC 5443	2009	LDP IGP Synchronization. IETF.			

SEGMENT ROUTING

RFC 8029	2017	Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures. IETF.	ISO 10589	2002	ISO, Intermediate System to Intermediate System Intra-Domain Routeing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473), International Standard 10589:2002, Second Edition [ISO 10589].
RFC 8277	2017	Using BGP to Bind MPLS Labels to Address Prefixes. IETF.			
RFC 8402	2018	Segment Routing Architecture. IETF.			

GAGNEZ DU TEMPS ET SÉCURISEZ VOS PROJETS EN UTILISANT UNE SOURCE ACTUALISÉE ET FIABLE

Techniques de l'Ingénieur propose la plus importante collection documentaire technique et scientifique en français !

Grâce à vos droits d'accès, retrouvez l'ensemble des **articles et fiches pratiques de votre offre**, **leurs compléments et mises à jour**, et bénéficiez des **services inclus**.



RÉDIGÉE ET VALIDÉE
PAR DES EXPERTS



MISE À JOUR
PERMANENTE



100 % COMPATIBLE
SUR TOUS SUPPORTS
NUMÉRIQUES



SERVICES INCLUS
DANS CHAQUE OFFRE

- + de 350 000 utilisateurs
- + de 10 000 articles de référence
- + de 80 offres
- 15 domaines d'expertise

- ☐ Automatique - Robotique
- ☐ Biomédical - Pharma
- ☐ Construction et travaux publics
- ☐ Électronique - Photonique
- ☐ Énergies
- ☐ Environnement - Sécurité
- ☐ Génie industriel
- ☐ Ingénierie des transports
- ☐ Innovation
- ☐ Matériaux
- ☐ Mécanique
- ☐ Mesures - Analyses
- ☐ Procédés chimie - Bio - Agro
- ☐ Sciences fondamentales
- ☐ Technologies de l'information

**Pour des offres toujours plus adaptées à votre métier,
découvrez les offres dédiées à votre secteur d'activité**

Depuis plus de 70 ans, Techniques de l'Ingénieur est la source d'informations de référence des bureaux d'études, de la R&D et de l'innovation.

www.techniques-ingenieur.fr

CONTACT : Tél. : + 33 (0)1 53 35 20 20 - Fax : +33 (0)1 53 26 79 18 - E-mail : infos.clients@teching.com

LES AVANTAGES ET SERVICES compris dans les offres Techniques de l'Ingénieur

ACCÈS



Accès illimité aux articles en HTML

Enrichis et mis à jour pendant toute la durée de la souscription



Téléchargement des articles au format PDF

Pour un usage en toute liberté



Consultation sur tous les supports numériques

Des contenus optimisés pour ordinateurs, tablettes et mobiles

SERVICES ET OUTILS PRATIQUES



Questions aux experts*

Les meilleurs experts techniques et scientifiques vous répondent



Articles Découverte

La possibilité de consulter des articles en dehors de votre offre



Dictionnaire technique multilingue

45 000 termes en français, anglais, espagnol et allemand



Archives

Technologies anciennes et versions antérieures des articles



Impression à la demande

Commandez les éditions papier de vos ressources documentaires



Alertes actualisations

Recevez par email toutes les nouveautés de vos ressources documentaires

*Questions aux experts est un service réservé aux entreprises, non proposé dans les offres écoles, universités ou pour tout autre organisme de formation.

ILS NOUS FONT CONFIANCE



www.techniques-ingenieur.fr

CONTACT : Tél. : + 33 (0)1 53 35 20 20 - Fax : +33 (0)1 53 26 79 18 - E-mail : infos.clients@teching.com