Groupe de travail Réseau
Request for Comments : 5286
Catégorie : Sur la voie de la normalisation
Traduction Claude Brière de L'Isle

A. Atlas, éditeur, BT
A. Zinin, éditeur, Alcatel-Lucent
septembre 2008

Spécification de base pour réacheminement rapide sur IP : solutions de remplacement sans boucle

Résumé
Le présent document décrit l'utilisation de remplaçants sans boucle (LFA, *Loop-Free Alternate*) pour assurer la protection locale du trafic en envoi individuel dans les purs réseaux IP et MPLS/LDP dans l'éventualité d'une seule défaillance, de liaison, de nœud, ou de groupe de liaisons à risques partagés (SRLG, *Shared Risk Link Group*). Le but de cette technologie est de réduire la perte de paquets qui survient lorsque les routeurs convergent après un changement de topologie dû à une défaillance. La réparation rapide de la défaillance est réalisée par l'utilisation de prochain bonds de sauvegarde pré-calculés qui sont sans boucle et d'utilisation sûre jusqu'à l'achèvement du processus de convergence de réseau répartie. Cette approche simple n'exige pas de prise en charge de la part d'autres routeurs. La mesure dans laquelle ce but peut être atteint par cette spécification dépend de la topologie du réseau.

Table des matières

1. Introduction

Les applications pour les services multimédia interactifs tels que la voix sur IP (VoIP, *Voice over IP*) et les pseudo filaires peuvent être très sensibles aux pertes de trafic, comme il s'en produit quand une liaison ou un routeur est défaillant dans le réseau. Le temps de convergence d'un routeur est généralement de l'ordre de la centaine de millisecondes ; le trafic d'application peut être sensible à des pertes supérieures à des dizaines de millisecondes.

Comme exposé dans la [RFC5714], minimiser les pertes de trafic exige un mécanisme pour que le routeur adjacent à une défaillance invoque rapidement un chemin de réparation, qui est minimalement affecté par une reconvergence suivante. La présente spécification décrit un tel mécanisme qui permet à un routeur dont la liaison locale est défaillante de transmettre le trafic à une solution de remplacement pré-calculée jusqu'à ce que le routeur installe les nouveaux prochains bonds principaux sur la base de la nouvelle topologie de réseau. La terminologie utilisée dans la présente spécification est donnée dans la [RFC5714]. Le mécanisme décrit suppose que l'acheminement dans le réseau est effectué en utilisant un protocole d'acheminement d'état de liaison -- OSPF [RFC2328] [RFC2740] [RFC5340] ou IS-IS [RFC1195] [RFC2966] (pour IPv4 ou IPv6). Le mécanisme suppose aussi que le chemin principal et le chemin de remplacement sont tous deux dans la même zone d'acheminement.

Quand une liaison locale est défaillante, un routeur doit actuellement signaler l'événement à ses voisins via l'IGP, recalculer de nouveaux prochains bonds principaux pour tous les préfixes affectés, et seulement ensuite installer ces nouveaux prochains bonds principaux dans le plan de transmission. Jusqu'à ce que les nouveaux prochains bonds principaux soient installés, le trafic dirigé sur les préfixes affectés est éliminé. Ce processus peut prendre des centaines de millisecondes.

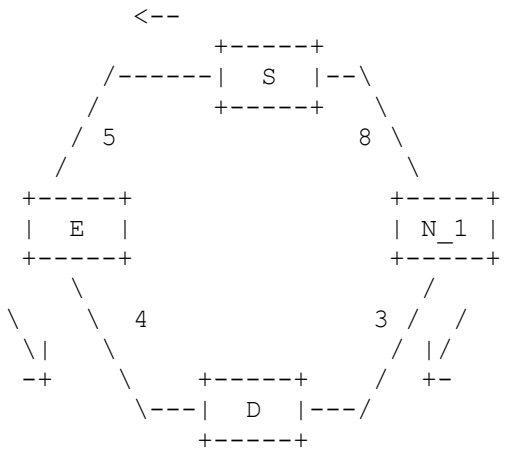


Figure 1 : Topologie de base

Le but du réacheminement rapide IP (IPFRR, *IP Fast ReRoute*) est de réduire le temps de réaction à la défaillance à des dixièmes de milliseconde en utilisant un prochain bond de remplacement pré calculé, dans l'éventualité d'une défaillance du prochain bond principal actuellement choisi, afin que le remplaçant puisse être rapidement utilisé quand la défaillance est détectée. Un réseau qui a cette caractéristique subit moins de pertes de trafic et moins de micro boucles de paquets qu'un réseau sans IPFRR. Il y a des cas où les pertes de trafic sont quand même possibles car la couverture de l'IPFRR varie, mais dans la pire des situations possibles, un réseau avec IPFRR est équivalent à l'égard de la convergence de trafic avec un réseau sans IPFRR.

Pour préciser le comportement du réacheminement rapide IP, considérons la simple topologie de la Figure 1. Quand le routeur S calcule son plus court chemin au routeur D, il détermine d'utiliser la liaison au routeur E comme son prochain bond principal. Sans le réacheminement rapide IP, cette liaison est le seul prochain bond que le routeur S calcule pour atteindre D. Avec le réacheminement rapide IP, S cherche aussi un prochain bond de remplacement à utiliser. Dans cet exemple, S va déterminer qu'il pourrait envoyer du trafic destiné à D en utilisant la liaison au routeur N_1 et donc S va installer la liaison à N_1 comme son prochain bond de remplacement. Un peu plus tard, la liaison entre le routeur S et le routeur E pourrait avoir une défaillance. Quand cette liaison est défaillante, S et E vont être les premiers à le détecter. À la détection de la défaillance, S va cesser d'envoyer du trafic destiné à D par E via la liaison défaillante, et va plutôt envoyer le trafic au prochain bond de remplacement pré calculé de S, qui est la liaison à N_1, jusqu'à ce qu'un nouveau SPF fonctionne et que ses résultats soient installés. Comme avec le prochain bond principal, un prochain bond de remplacement est calculé pour chaque destination. Le processus de calcul d'un prochain bond de remplacement n'altère pas le prochain

bond principal calculé via un SPF standard.

Si dans l'exemple de la Figure 1, le coût de liaison de N_1 à D a augmenté de 3 à 30, alors N_1 ne va pas être un remplacement sans boucle, parce que le coût du chemin de N_1 à D via S va être 17 alors que le coût de N_1 directement à D serait de 30. Dans les réseaux réels, on peut souvent rencontrer cette situation. L'existence d'un prochain bond de remplacement sans boucle convenable dépend de la topologie et de la nature de la défaillance pour laquelle le remplacement est calculé.

La présente spécification utilise la terminologie introduite dans la [RFC5714]. En particulier, elle utilise Distance_opt(X,Y), abrégée en D_opt(X,Y) pour indiquer la plus courte distance de X à Y. S est utilisé pour indiquer le routeur de calcul. N_i est un voisin de S ; N est utilisé comme abréviation quand un seul voisin est discuté. D est la destination considérée.

Un voisin N peut fournir un remplacement sans boucle (LFA, *Loop-Free Alternate*) si et seulement si

Distance_opt(N, D) < Distance_opt(N, S) + Distance_opt(S, D)

Inégalité 1 : Critère d'absence de boucle

Un sous ensemble de remplacement sans boucle est un chemin vers l'aval qui doit satisfaire une condition plus restrictive qui est applicable à des scénarios de défaillance plus complexes :

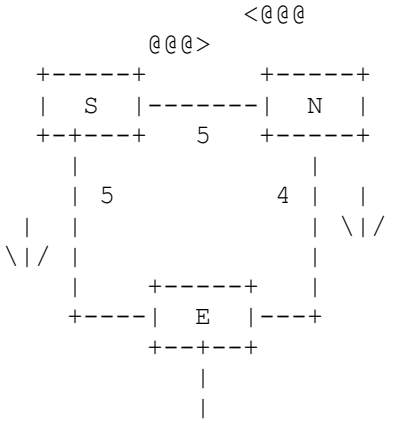
Distance_opt(N, D) < Distance_opt(S, D)

Inégalité 2 : Critère de chemin vers l'aval

1.1. Scénarios d'échec

Le prochain bond de remplacement peut protéger contre la défaillance d'une seule liaison, la défaillance d'un seul nœud, la défaillance d'une ou plusieurs liaisons au sein d'un groupe de liaisons à risque partagé, ou une combinaison d'elles. Chaque fois que se produit une défaillance qui est plus étendue que ce que le remplacement était destiné à protéger, il y a la possibilité d'un trafic temporairement en boucle (noter encore qu'une telle boucle ne durerait que jusqu'au prochain calcul complet de SPF). L'exemple où un nœud a une défaillance quand le remplacement fournissait seulement la protection de la liaison est illustré ci-dessous. Si des défaillances simultanées inattendues se produisent, alors des micro-boucles peuvent survenir car les remplacements ne sont pas pré-calculés pour éviter l'ensemble de liaisons défaillantes.

Si seule la protection de liaison est fournie et si le nœud est défaillant, il est possible que le trafic qui utilise les remplacements subisse des micro-boucles. Ce problème est illustré à la Figure 2. Si Link(S->E) échoue, alors le remplacement qui protège la liaison via N va fonctionner correctement. Cependant, si le routeur E échoue, alors S et N vont tous deux détecter une défaillance et passer à leurs remplacements. Dans cet exemple, cela causerait la redirection par S du trafic sur N et par N du trafic à S et causant donc une boucle de transmission. Un tel scénario peut se produire à cause de l'hypothèse clé que tous les autres routeurs dans le réseau transmettent sur la base du plus court chemin, qui est violée à cause d'une seconde défaillance simultanée corrélée – une autre liaison connectée au même voisin principal. Si il n'y a pas d'autres mécanismes de protection pour traiter la défaillance du nœud, une défaillance du nœud est alors un problème quand on utilise seulement des LFA de protection de liaison.



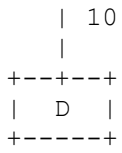


Figure 2 : Remplacements de protection de liaison causant une boucle lors d'une défaillance de nœud

L'apparition de micro-boucles de trafic via les remplacements causées quand une défaillance plus étendue que prévu se produit peut être empêchée en choisissant seulement des chemins vers l'aval comme remplacements. Une micro-boucle due à l'utilisation de remplaçants peut être évitée en utilisant des chemins vers l'aval parce que chaque routeur successif sur le chemin vers la destination doit être plus proche de la destination que son prédécesseur (selon la topologie avant la défaillance). Bien que l'utilisation de chemins vers l'aval assure que des micro-boucles ne se produisent pas via les remplaçants, une telle restriction peut limiter sévèrement la couverture des remplaçants. Dans la Figure 2, S serait capable d'utiliser N comme remplaçant vers l'aval, mais N ne pourrait pas utiliser S ; donc, N n'aurait pas de remplaçant et éliminerait le trafic, évitant donc la micro-boucle.

Comme on le montre ci-dessus, l'utilisation d'un LFA de protection de nœud (décrite au paragraphe 3.2) ou d'un chemin vers l'aval, fournit une protection contre des micro-boucles dans le cas d'une défaillance de nœud. Il y a des topologies où il peut y avoir un LFA de protection de nœud, un chemin vers l'aval, les deux, ou aucun des deux. Un nœud peut choisir un LFA de protection de nœud ou un chemin vers l'aval sans risquer de causer des micro-boucles en cas de défaillance du nœud voisin. Alors qu'un LFA de protection de liaison et de nœud garantit la protection contre la défaillance de la liaison ou du nœud, un chemin vers l'aval fournit seulement la protection contre la défaillance de la liaison et peut ou non fournir la protection contre une défaillance du nœud selon la protection disponible au nœud aval, mais il ne peut pas causer de micro-boucle. Par exemple, dans la Figure 2, si S utilise N comme chemin vers l'aval, bien qu'aucune boucle ne puisse se produire, le trafic ne va pas être protégé en cas de défaillance du nœud E parce que N n'a pas de chemin de réparation viable, et va simplement éliminer le paquet. Cependant, si N a un LFA de protection de liaison et de nœud ou un chemin vers l'aval via quelque autre chemin (non montré) alors la réparation peut réussir.

Comme la fonction de LFA de protection de liaison et de nœud est supérieure à celle des chemins de protection de liaison en aval, un routeur DEVRAIT choisir un LFA de protection de liaison et de nœud plutôt qu'un chemin de protection de liaison en aval. Si il y a des destinations pour lesquelles un LFA de protection de liaison et de nœud n'est pas disponible, alors par définition le chemin pour toutes ces destinations à partir de tout voisin du routeur (S) qui fait le calcul doit être à travers le nœud (E) à protéger (autrement il y aurait un LFA de protection du nœud pour cette destination). Par conséquent, si il existe un chemin vers l'aval pour le nœud protégé comme destination, ce chemin vers l'aval peut alors être utilisé pour toutes les destinations pour lesquelles un LFA de protection de liaison et de nœud n'est pas disponible ; l'existence d'un chemin vers l'aval peut être déterminée par une seule vérification de la condition Distance_opt(N, E) < Distance_opt(S, E).

Il peut être désirable de trouver un remplaçant qui puisse protéger contre d'autres défaillances corrélées (dont la défaillance de nœud est une instance spécifique). En général, elles sont traitées par des groupes de liaisons à risques partagés (SRLG, *Shared Risk Link Group*) où toute liaison dans le réseau peut appartenir au SRLG. Les SRLG généraux peuvent ajouter une complexité inacceptable aux calculs de recherche d'un remplaçant sans boucle.

Cependant, une sous catégorie de SRLG est intéressante et peut être appliquée seulement durant le choix d'un remplaçant acceptable. Cette sous catégorie est pour exprimer les défaillances corrélées des liaisons qui sont connectées au même routeur, par exemple, si il y a plusieurs sous interfaces logiques sur la même interface physique, comme des VLAN sur une interface Ethernet, si plusieurs interfaces utilisent le même accès physique à cause de la mise en canaux, ou si plusieurs interfaces partagent une défaillance corrélée parce qu'elles sont sur la même carte de ligne. Cette sous catégorie de SRLG va être appelée un SRLG local. Un SRLG local a toutes ses liaisons membres qui ont une extrémité connectée au même routeur. Donc, le routeur S pourrait choisir un remplaçant sans boucle qui n'utilise pas une liaison dans le même SRLG local comme prochain bond principal. On peut se protéger de la défaillance de SRLG locaux appartenant à E via la protection de nœud, c'est-à-dire, en prenant un remplaçant de protection de nœud sans boucle.

Lorsque la protection de SRLG est fournie, c'est dans le contexte de la zone OSPF ou IS-IS particulière, dont la topologie est utilisée dans le calcul de SPF pour les remplaçants sans boucle. Si un SRLG contient des liaisons dans plusieurs zones, des remplaçants séparés de protection de SRLG vont être nécessaires dans chaque zone traversée par le trafic affecté.

1.2 Langage des exigences

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS",

"RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

2. Applicabilité des mécanismes décrits

Les mécanismes de réacheminement rapide IP décrits dans le présent mémoire couvrent seulement l'acheminement intra domaine, avec OSPF [RFC2328] [RFC2740] [RFC5340] ou IS-IS [RFC1195] [RFC2966] comme IGP. Précisément, le réacheminement rapide pour l'acheminement BGP inter-domaines ne fait pas partie de cette spécification.

Certains aspects du comportement d'acheminement inter-zones OSPF expliqués au paragraphe 6.3 et à l'Appendice A impactent la capacité du routeur de calculer les prochains bonds de sauvegarde pour assurer les trajectoires du trafic. Afin d'éviter des micro-boucles et assurer la couverture requise, certaines contraintes sont appliquées aux réseaux OSPF multi-zones :

a. Les remplaçants sans boucle ne devraient pas être utilisés dans la zone cœur si il y a des liaisons virtuelles configurées sauf si, pour chaque zone de transit, il y a un maillage complet des liaisons virtuelles entre tous les routeurs de bordure de zone (ABR, *Area Border Router*) dans cette zone. Les remplaçants sans boucle peuvent être utilisés dans des zones non de cœur sans considération de si il y a des liaisons virtuelles configurées.

b. Les remplaçants sans boucle ne devraient pas être utilisés pour des chemins inter-zones dans une zone qui contient plus d'un ABR de remplacement [RFC3509].

c. Les remplaçants sans boucle ne devraient pas être utilisés pour les chemins d'AS externe ou de routeur bordure de système autonome (ASBR, *Autonomous System Border Router*) dans une zone non cœur d'un réseau où il existe un ABR qui est annoncé comme ASBR dans plusieurs zones non de cœur et où il existe un autre ABR qui est dans au moins deux des mêmes zones non de cœur.

d. Les remplaçants sans boucle ne devraient pas être utilisés dans une zone non de cœur d'un réseau pour des chemins d'AS externe où un préfixe d'AS externe est annoncé avec le même type de métrique externe par plusieurs ASBR, qui sont dans des zones non de cœur différentes, avec une adresse de transmission de 0.0.0.0 ou par un ou plusieurs ASBR avec des adresses de transmission dans plusieurs zones non de cœur quand un ABR existe simultanément dans deux ou plus de ces zones non de cœur.

3. Calcul du prochain bond de remplacement

En plus de l'ensemble de prochains bonds principaux obtenus par un calcul d'arborescence de plus court chemin (SPT, *Shortest Path Tree*) qui fait partie de la fonction standard d'acheminement d'état de liaison, les routeurs qui prennent en charge le réacheminement rapide IP calculent aussi un ensemble de prochains bonds de sauvegarde qui sont engagés quand une défaillance locale se produit. Ces prochains bonds de sauvegarde sont calculés pour fournir le type de protection requis (c'est-à-dire, protection de liaison et/ou protection de nœud) et pour garantir que quand la défaillance attendue se produit, le trafic de transmission à travers eux ne va pas résulter en une boucle. De tels prochains bonds sont appelés des remplaçants sans boucle (LFA, *Loop-Free Alternate*) dans la présente spécification.

En général, pour être capable de calculer l'ensemble des LFA pour une destination D spécifique, un routeur a besoin de connaître les éléments d'informations de base suivants :

- La distance du plus court chemin du routeur calculant à la destination (Distance_opt(S, D))

- La distance du plus court chemin des voisins IGP du routeur à la destination (Distance_opt(N, D))

- La distance du plus court chemin des voisins IGP du routeur à lui-même (Distance_opt(N, S))

- Distance_opt(S, D) est normalement disponible à partir du calcul régulier de SPF effectué par les protocoles d'acheminement d'état de liaison. Distance_opt(N, D) et Distance_opt(N, S) peuvent être obtenues en effectuant des calculs de SPF supplémentaires du point de vue de chaque voisin IGP (c'est-à-dire, en considérant le sommet du voisin comme la racine de la SPT-- appelée ci-après SPT(N) – plutôt que de calculer celle du routeur, appelée SPT(S)).

La présente spécification définit une forme de protection de SRLG limitée aux SRLG qui incluent une liaison à laquelle le routeur calculant est directement connecté. Ce seul ensemble de SRLG pourrait causer une défaillance locale ; le routeur calculant ne calcule les remplaçants que pour traiter une défaillance locale. Les informations sur l'appartenance de la liaison locale au SRLG sont configurées manuellement. Les informations sur l'appartenance d'une liaison distance au SRLG peuvent être obtenues dynamiquement en utilisant la [RFC4205] ou la [RFC4203]. Définir SRLG_local(S) comme étant l'ensemble des SRLG qui incluent une liaison à laquelle le routeur calculant S est directement connecté. Seul SRLG_local(S) est intéressant pour ce calcul, mais le routeur calculant doit traiter correctement les changements à SRLG_local(S) déclenchés par des changements de l'appartenance au SRLG de liaison locale.

Afin de choisir parmi tous les LFA disponibles qui fournissent la protection requise de SRLG pour une certaine destination, le routeur calculant a besoin de retracer l'ensemble des SRLG dans SRLG_local(S) qu'implique le chemin à travers un voisin IGP spécifique. Pour ce faire, chaque nœud D dans la topologie du réseau est associé à SRLG_set(N, D) qui est l'ensemble des SRLG qui seraient traversés si le trafic pour D était transmis à travers N. Pour calculer cet ensemble, le routeur initialise SRLG_set(N, N) pour chacun de ses voisins IGP comme étant vide. Durant le calcul de SPT(N), quand un nouveau sommet V est ajouté à la SPT, son SRLG_set(N, V) est réglé à l'union des ensembles de SRLG associés à ses parents, et les ensembles de SRLG dans SRLG_local(S) qui sont associés aux liaisons provenant des parents de V à V. L'union de l'ensemble des SRLG associés à un candidat prochain bond de remplacement et du SRLG_set(N, D) pour le voisin atteint via ce candidat prochain bond est utilisée pour déterminer la protection de SRLG.

Les paragraphes qui suivent donnent les informations requises pour le calcul des LFA. Les paragraphes 3.1 à 3.4 définissent les différents types de conditions de LFA. Le paragraphe 3.5 décrit les contraintes imposées par la fonction de surcharge IS-IS et de routeur de bout OSPF. Le paragraphe 3.6 définit l'algorithme résumé pour le calcul de LFA en utilisant les définitions des paragraphes précédents.

3.1 Condition sans boucle de base

Les prochains bonds de remplacement utilisés par les mises en œuvre qui respectent la présente spécification DOIVENT se conformer au moins à la condition d'absence de boucle déclarée dans l'inégalité 1. Cette condition garantit que la transmission du trafic à une LFA ne va pas résulter en une boucle après une défaillance de liaison.

D'autres conditions peuvent être appliquées quand on détermine les prochains bonds de remplacement de protection de liaison et/ou de nœud comme décrit aux paragraphes 3.2 et 3.3.

3.2 Prochains bonds de remplacement protégeant le nœud

Pour qu'un prochain bond de remplacement N protège contre la défaillance d'un nœud d'un voisin principal E pour la destination D, N doit être sans boucle par rapport à E et D. En d'autres termes, le chemin de N à D ne doit pas passer par E. C'est le cas si l'inégalité 3 est vraie, lorsque N est le voisin qui fournit un remplaçant sans boucle.

Distance_opt(N, D) < Distance_opt(N, E) + Distance_opt(E, D)

Inégalité 3 : Critère pour un nœud de protection de remplaçant sans boucle

Si Distance_opt(N,D) = Distance_opt(N, E) + Distance_opt(E, D), il est possible que N ait des chemins de coût égal et que l'un d'eux puisse fournir la protection contre la défaillance du nœud E. Cependant, il est également possible d'un des chemins de N passe par E, et que le routeur qui calcule n'ait pas de moyen d'influencer la décision de N de l'utiliser. Donc, il DEVRAIT être supposé qu'un prochain bond de remplacement n'offre pas de protection de nœud si l'inégalité 3 n'est pas satisfaite.

3.3 Liaisons multi-accès de diffusion et non diffusion (NBMA)

La vérification de la propriété de protection de la liaison d'un prochain bond dans le cas d'une liaison de diffusion est plus élaborée que pour une liaison point à point. C'est parce que une liaison de diffusion est représentée comme un pseudo-nœud avec des liaisons de coût zéro qui la connectent aux autres nœuds.

Parce que la défaillance d'une interface rattachée à un segment de diffusion peut signifier la perte de connexité du segment entier, la condition décrite pour la protection d'une liaison de diffusion est pessimiste et exige que le remplaçant soit sans boucle à l'égard du pseudo-nœud. Considérons l'exemple de la Figure 3.

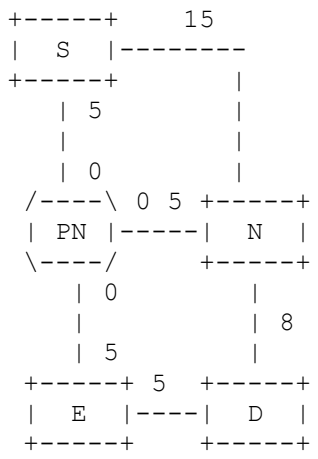


Figure 3 : Remplaçant sans boucle qui protège la liaison

Dans la Figure 3, N offre un remplaçant sans boucle qui protège la liaison. Si le prochain bond principal utilise une liaison de diffusion, alors un remplaçant DEVRAIT être sans boucle à l'égard du pseudo-nœud (PN) de cette liaison pour fournir la protection de la liaison. Cette exigence est décrite dans l'inégalité 4 :

$$D\_opt(N, D) < D\_opt(N, PN) + D\_opt(PN, D)$$

Inégalité 4 : Critère de protection de liaison sans boucle pour les liaisons de diffusion

Parce que le plus court chemin provenant du pseudo-nœud passe par E, si un remplaçant sans boucle provenant d'un voisin N est de protection de nœud, le remplaçant va aussi être de protection de liaison sauf si le routeur S peut seulement atteindre le voisin N de remplacement via le même pseudo-nœud. Comme c'est le seul cas pour lequel un LFA de protection de nœud n'est pas de protection de liaison, cela implique que pour les interfaces point à point, un LFA qui est de protection de nœud est toujours de protection de liaison. Parce que S peut diriger le traffic hors du plus court chemin pour utiliser le remplaçant N, le trafic pourrait passer à travers la même liaison de diffusion que celle qu'il aurait utilisé quand S envoie le trafic au principal E. Donc, un LFA provenant de N qui est de protection de nœud n'est pas automatiquement de protection de liaison pour une liaison en diffusion ou NBMA.

Pour obtenir la protection de liaison, il est nécessaire à la fois que le chemin provenant du prochain bond de remplacement choisi ne traverse pas la liaison intéressante et que la liaison utilisée provenant de S pour atteindre ce prochain bond de remplacement ne soit pas la liaison intéressante. Cette dernière condition peut seulement se produire avec des liaisons non en point à point. Donc, si le prochain bond principal est à travers une interface de diffusion ou NBMA, il est nécessaire de considérer la protection de liaison durant le choix du remplaçant. Pour préciser, considérons la topologie de la Figure 3. Pour que N fournisse la protection de liaison, il est d'abord nécessaire que le plus court chemin de N à D ne traverse pas le pseudo-nœud PN. Ensuite, il est nécessaire que le prochain bond de remplacement choisi par S ne traverse pas PN. Dans cet exemple, le plus court chemin de S à N est via le pseudo-nœud. Donc, pour obtenir la protection de liaison, S doit trouver un prochain bond pour N (la liaison point à point de S à N dans cet exemple) qui évite le pseudo-nœud PN.

Une considération similaire de la liaison de S au prochain bond de remplacement choisi ainsi que du chemin du prochain bond de remplacement choisi est aussi nécessaire pour la protection de SRLG. Le plus court chemin de S au voisin choisi N peut n'être pas acceptable comme prochain bond de remplacement pour fournir la protection de SRLG, même si le chemin de N à D peut fournir la protection de SRLG.

3.4. ECMP et solutions de remplacement

Avec les chemins multiples de coût égal (ECMP, *Equal-Cost Multi-Path*) un préfixe peut avoir plusieurs prochains bonds principaux qui sont utilisés pour transmettre du trafic. Quand un prochain bond principal particulier a une défaillance, les prochains bonds de remplacement devraient être utilisés pour préserver le trafic. Ces prochains bonds de remplacement peuvent eux-mêmes être aussi des prochains bonds principaux, mais ne le sont pas nécessairement. Il n'est pas garanti que d'autres prochains bonds principaux fournissent la protection contre les scénarios de défaillance concernés.

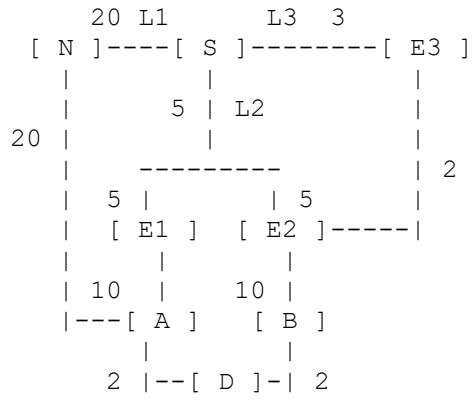


Figure 4 : ECMP où les prochains bonds principaux fournissent une protection limitée

Dans la Figure 4, S a trois prochains bonds principaux pour atteindre D ; ce sont L2 pour E1, L2 pour E2, et L3 pour E3. Le prochain bond principal L2 pour E1 peut obtenir la protection de liaison et de nœud de L3 à E3, qui est un des autres prochains bonds principaux ; L2 à E1 ne peut pas obtenir la protection de liaison de l'autre prochain bond L2 à E2. De même, le prochain bond principal L2 à E2 peut seulement obtenir la protection de nœud de L2 à E1 et peut seulement obtenir la protection de liaison de L3 à E3. Le troisième prochain bond principal de L3 à E3 peut obtenir la protection de liaison et de nœud de L2 à E1 et de L2 à E2. Il est possible aux deux prochains bonds principaux de L2 à E2 et de L2 à E1 d'obtenir un prochain bond de remplacement qui fournisse la protection de liaison et de nœud en utilisant L1.

Les prochains bonds de remplacement sont déterminés pour chaque prochain bond principal séparément. Comme avec le choix de remplaçant dans le cas non ECMP, ces prochains bonds de remplacement devraient maximiser la couverture des cas de défaillance.

3.5 Interactions avec des liaisons IS-IS surchargées, RFC 3137, et de coût élevé

Comme décrit dans la [RFC3137], il y a des cas où il est désirable de ne pas avoir un routeur utilisé comme nœud de transit. Dans ces cas, il est aussi désirable de ne pas avoir le routeur utilisé sur un chemin de remplacement.

Pour calculer un remplaçant, un routeur NE DOIT PAS utiliser un prochain bond de remplacement qui est sur une liaison dont le coût ou le coût inverse est LSInfinity (pour OSPF) ou le coût maximum (pour IS-IS) ou qui a le bit Surcharge établi (pour IS-IS). Pour une liaison de diffusion, le coût inverse associé à un prochain bond de remplacement potentiel est le coût sur le pseudo-nœud annoncé par le routeur de prochain bond. Pour les liaisons en point à point, si une liaison spécifique provenant du routeur de prochain bond ne peut pas être associée à une liaison particulière, alors le coût inverse considéré est celui du coût minimum de liaison du routeur de prochain bond jusqu'au retour à S.

Dans le cas de OSPF, si toutes les liaisons provenant du routeur S à un voisin N_i ont un coût inverse de LSInfinity, alors le routeur S NE DOIT PAS utiliser N_i comme remplaçant.

De même, dans le cas de IS-IS, si N_i a le bit Surcharge établi, alors S NE DOIT PAS considérer d'utiliser N_i comme remplaçant.

Cela préserve le comportement désiré de dérouter le trafic d'un routeur qui suit la [RFC3137], et cela préserve aussi le comportement désiré quand un opérateur règle le coût d'une liaison à LSInfinity pour la maintenance qui ne permet pas le trafic à travers cette liaison sauf si il n'y a pas d'autre chemin.

Si une liaison ou un routeur qui est écarté à cause de son coût était le seul remplaçant possible pour protéger le trafic provenant d'un routeur particulier S pour une destination particulière, il ne devrait alors pas y avoir de remplaçant fourni pour la protection.

3.5.1 Interactions avec les attributs de liaison IS-IS

La [RFC5029] décrit plusieurs fanions dont les interactions avec les LFA doivent être définies. Un routeur NE DEVRAIT PAS spécifier le fanion "protection locale disponible" comme résultat d'avoir des LFA. Un routeur NE DEVRAIT PAS utiliser un prochain bond de remplacement qui est le long d'une liaison pour laquelle la liaison a été annoncée avec l'attribut "liaison exclue du chemin local de protection" ou avec l'attribut "maintenance locale requise".

3.6 Procédure de sélection

Un routeur qui prend en charge la présente spécification DEVRAIT tenter de choisir au moins un prochain bond de remplacement sans boucle pour chaque prochain bond principal utilisé pour un préfixe donné. Un routeur PEUT décider de ne pas utiliser un prochain bond de remplacement sans boucle disponible. Une raison d'une telle décision pourrait être que le prochain bond de remplacement sans boucle ne fournit pas de protection pour le scénario de défaillance intéressant.

Le choix de remplaçant devrait maximiser la couverture des cas de défaillance.

Quand on calcule les prochains bonds de remplacement, le routeur calculant S applique les règles suivantes.

1. S DEVRAIT choisir un prochain bond de remplacement sans boucle de protection de nœud, si il en est de disponible. Si aucun remplaçant sans boucle de protection de nœud n'est disponible, alors S PEUT choisir un remplaçant sans boucle de protection de liaison.

2. Si S a le choix entre un remplaçant sans boucle de protection de liaison et de nœud et un remplaçant sans boucle de protection de nœud qui n'est pas de protection de liaison, S DEVRAIT choisir un remplaçant sans boucle de protection de liaison et de nœud. Cela peut survenir comme expliqué au paragraphe 3.3.

3. Si S a plusieurs prochains bonds principaux, alors S DEVRAIT choisir comme remplaçant sans boucle soit un des autres prochains bonds principaux, soit un remplaçant sans boucle de protection de nœud si il en est de disponible. Si aucun remplaçant sans boucle de protection de nœud n'est disponible et si aucun autre prochain bond principal ne peut fournir de protection de liaison, alors S DEVRAIT choisir un remplaçant sans boucle de protection de liaison.

4. Les mises en œuvres DEVRAIENT prendre en charge un mode où d'autres prochains bonds principaux satisfaisant la condition de base sans boucle et fournissant au moins la protection de liaison ou de nœud sont préférés à tout remplaçant non principal. Ce mode est fourni pour permettre à l'administrateur de préserver les schémas de trafic sur la base du comportement ECMP régulier.

5. Les mises en œuvres considérant les SRLG PEUVENT utiliser la protection de SRLG pour déterminer qu'un remplaçant de protection de nœud ou de liaison n'est pas disponible à l'utilisation.

Suivre les règles ci-dessus maximise le niveau de protection et l'utilisation des prochains bonds principaux (ECMP).

Chaque prochain bond est associé à un ensemble de caractéristiques non mutuellement exclusives en fonction de si il est utilisé comme prochain bond principal pour une destination D particulière, et du type de protection qu'il peut fournir à l'égard d'un prochain bond principal E spécifique :

Chemin principal - le prochain bond est utilisé par S comme principal.

Remplaçant sans boucle de protection de nœud - ce prochain bond satisfait l'inégalité 1 et l'inégalité 3. Le chemin évite S, le voisin principal E de S, et la liaison de S à E.

Remplaçant sans boucle de protection de liaison - ce prochain bond satisfait l'inégalité 1 mais pas l'inégalité 3. Si le prochain bond principal utilise une liaison de diffusion, alors ce prochain bond satisfait l'inégalité 4.

Un chemin de remplacement peut aussi ne fournir aucune protection de SRLG, ou un peu, ou une protection complète aussi bien que la protection de nœud et de liaison ou la protection de liaison. Par exemple, un liaison peut appartenir à deux SRLG, G1 et G2. Le chemin de remplacement pourrait éviter d'autres liaisons dans G1 mais pas G2, et dans ce cas, le remplaçant va seulement fournir une protection partielle du SRLG.

On donne ci-dessous un algorithme qui peut être utilisé pour calculer des prochains bonds de remplacement sans boucle. L'algorithme est donné à titre d'information, et les mises en œuvre sont libres d'utiliser tout autre algorithme pour autant qu'il satisfait aux règles décrites ci-dessus.

La procédure suivante décrit comment choisir un prochain bond de remplacement. La procédure est décrite pour déterminer les prochains bonds de remplacement à utiliser pour atteindre chaque routeur dans la topologie. Les préfixes qui sont annoncés par un seul routeur peuvent utiliser le prochain bond de remplacement calculé pour le routeur auquel ils sont rattachés. La même procédure peut être utilisée pour atteindre un préfixe qui est annoncé par plus d'un routeur quand la

transformation topologique logique décrite au paragraphe 6.1 est utilisée.

S est le routeur qui calcule. S a des voisins N_1 à N_j. Un candidat prochain bond est indiqué par (liaison sortante, voisin) et la liaison sortante doit être connectée en bidirectionnel, comme déterminé par l'IGP. Les candidats prochains bonds de S sont énumérés comme H_1 à H_k. On rappelle que S peut avoir plusieurs prochains bonds sur différentes interfaces à un voisin. H_i.liaison se réfère à la liaison sortante de ce prochain bond et H_i.voisin se réfère au voisin de ce prochain bond.

Pour un routeur de destination D particulier, S a déjà calculé D_opt(S, D) et pour chaque voisin N_i, D_opt(N_i, D), D_opt(N_i, S), et D_opt(N_i, N_j), la distance de N_i à chaque autre voisin N_j, et l'ensemble des SRLG traversés par le chemin D_opt(N_i, D). S devrait suivre la procédure ci-dessous pour chaque prochain bond principal choisi pour atteindre D. Cet ensemble de prochains bonds principaux est représenté par P_1 à P_p. Cette procédure trouve le ou les prochains bonds de remplacement pour P_i.

D'abord, on initialise les informations de remplaçant pour P_i comme suit :
P_i.alt_prochains_bonds = {}
P_i.alt_type = AUCUN
P_i.alt_liaison-protection = FAUX
P_i.alt_nœud-protection = FAUX
P_i.alt_srlg-protection = {}

Pour chaque candidat prochain bond H_h,

1. Initialiser les variables comme suit :
   cand_type = AUCUNE
   cand_liaison-protection = FAUX
   cand_nœud-protection = FAUX
   cand_srlg-protection = {}

2. Si H_h est P_i, le sauter et continuer au prochain candidat prochain bond.

3. Si l'utilisation de H_h.liaison est administrativement permise comme remplaçant, et si le coût de H_h.liaison est moins que le maximum, et si le coût inverse de H_h est moins que le maximum, et si H_h.voisin n'est pas surchargé (pour IS-IS) et si H_h.liaison est bidirectionnelle, alors H_h peut être considéré comme un remplaçant. Autrement, le sauter et continuer au prochain candidat prochain bond.

4. Si D_opt(H_h.voisin, D) >= D_opt(H_h.voisin, S) + D_opt(S, D), alors H_h n'est pas sans boucle. Le sauter et continuer au prochain candidat prochain bond.

5. cand_type = sans boucle.

6. Si H_h est un prochain bond principal, régler cand_type à PRINCIPAL.

7. Si H_h.liaison n'est pas P_i.liaison, régler cand_liaison-protection à VRAI.

8. Si D_opt(H_h.voisin, D) < D_opt(H_h.voisin, P_i.voisin) + D_opt(P_i.voisin, D), régler cand_nœud-protection à VRAI.

9. Si le routeur prend en charge les SRLG, régler alors le cand_srlg-protection à l'ensemble de SRLG traversés sur le chemin provenant de S via P_i.liaison à P_i.voisin. Supprimer de cand_srlg-protection l'ensemble de SRLG auquel H_h appartient. Supprimer de cand_srlg-protection l'ensemble de SRLG traversés sur le chemin de H_h.voisin à D. Maintenant, cand_srlg-protection contient l'ensemble de SRLG auxquels P_i appartient et qui ne sont pas traversés sur le chemin de S via H_h à D.

10. Si cand_type est PRINCIPAL, le routeur préfère d'autres prochains bonds principaux à utiliser comme remplaçants, et le P_i.alt_type n'est pas PRINCIPAL, aller à l'étape 20.

11. Si cand_type n'est pas PRINCIPAL, P_i.alt_type est PRINCIPAL, et le routeur préfère d'autres prochains bonds principaux comme remplaçant, alors continuer au prochain candidat prochain bond.

12. Si cand_nœud-protection est VRAI et si P_i.alt_ nœud-protection est FAUX, aller à l'étape 20.

13. Si cand_liaison-protection est VRAI et P_i.alt_liaison-protection est FAUX, aller à l'étape 20.

14. Si cand_srlg-protection a un meilleur ensemble de SRLG que P_i.alt_srlg-protection, aller à l'étape 20.

15. Si cand_srlg-protection est différent de P_i.alt_srlg-protection, choisir alors entre H_h et P_i.alt_prochains_bonds sur la base de la distance, des adresses IP, ou de tout départage local au routeur. Si H_h est préféré, alors aller à l'étape 20. Si P_i.alt_prochains_bonds est préféré, sauter H_h et continuer au prochain candidat prochain bond.

16 Si D_opt(H_h.voisin, D) < D_opt(P_i.voisin, D) et D_opt(P_i.alt_prochains_bonds, D) ≥ D_opt(P_i.voisin, D), alors H_h est un remplaçant vers l'aval et P_i.alt_prochains_bonds est simplement un LFA. Préférer H_h et aller à l'étape 20.

17. Sur la base des types de remplaçants, des distances de remplaçant, des adresses IP, ou autres départages, décider si H_h est préféré à P_i.alt_prochains_bonds. Si il en est ainsi, aller à l'étape 20.

18. Décider si P_i.alt_prochains_bonds est préféré à H_h. Si il en est ainsi, sauter alors H_h et continuer au prochain candidat prochain bond.

19. Ajouter H_h dans P_i.alt_prochains_bonds. Régler P_i.alt_type au meilleur type de H_h.alt_type et P_i.alt_type. Continuer au prochain candidat prochain bond .

20. Remplacer l'ensemble de prochains bonds de remplacement P_i par H_h comme suit :
    P_i.alt_prochains_bonds = {H_h}
    P_i.alt_type = cand_type
    P_i.alt_liaison-protection = cand_liaison-protection
    P_i.alt_nœud-protection = cand_nœud-protection
    P_i.alt_srlg-protection = cand_srlg-protection

Continuer au prochain candidat prochain bond.

3.7 Types de LFA et compromis

Les LFA peuvent fournir des quantités de protection différentes, et la décision quant au type à préférer dépend de la topologie de réseau et autres techniques en usage dans le réseau. Ce paragraphe décrit les différents niveaux de protection et les compromis associés à chacun.

1. Prochain bond principal : quand il y a des prochains bonds principaux de coût égal, en utiliser un comme remplaçant garantit de ne pas causer de micro boucle impliquant S. Les flux de trafic à travers les chemins vers lesquels le réseau va converger, mais de l'encombrement peut se rencontrer sur les chemins principaux car le trafic est envoyé sur moins de chemins. Tous les prochains bonds principaux sont des chemins vers l'aval.

2. Chemins vers l'aval : un chemin vers l'aval, à la différence d'un LFA, garantit de ne pas causer de micro boucle impliquant S sans considération de la défaillance détectée réelle. Cependant, la couverture attendue de tels remplaçants dans un réseau est supposée être faible. Tous les chemins vers l'aval sont des LFA.

3. LFA : un LFA peut avoir une bonne couverture d'un réseau, selon la topologie. Cependant, il est possible d'avoir des micro boucles impliquant S si une défaillance non protégée se produit (par exemple, une défaillance de nœud quand le LFA était seulement de protection de liaison).

Les différents types de protection sont abrégés en LP (protection de liaison) NP (protection de nœud) et SP (protection de SRLG).

a. LP, NP, et SP : Si un tel remplaçant existe, il donne la protection contre toutes les défaillances.

b. LP et NP seulement : de nombreux réseaux peuvent traiter les défaillances de SRLG via une autre méthode ou peuvent se concentrer sus les défaillances de nœud et de liaison qui sont les plus courantes.

c. LP seulement : un réseau peut traiter les défaillances de nœud via une technique largement disponible et être concerné principalement par la protection des cas les plus courants de défaillance de liaison.

d. NP seulement : cela n'existe que sur les interfaces qui ne sont pas en point à point. Si la protection de liaison est traitée

dans une couche différente, alors un remplaçant NP peut être acceptable.

3.8 Simplification : LFA par prochains bonds

Il est possible de simplifier le calcul et l'utilisation des LFA quand seule la protection de liaison est désirée en considérant et calculant seulement un LFA de protection de liaison pour chaque prochain bond connecté au routeur. Tous les préfixes qui utilisent ce prochain bond comme principal vont utiliser le LFA calculé pour ce prochain bond comme leur LFA.

Même un préfixe avec plusieurs prochains bonds principaux va avoir chaque prochain bond principal protégé individuellement par le LFA associé au prochain bond principal. Ce LFA associé pourrait ou non être un autre des prochains bonds principaux du préfixe.

Cette simplification peut réduire la couverture dans un réseau. En plus de limiter la protection pour les préfixes multi-rattachements (voir le paragraphe 6.1) le calcul par prochain bond peut aussi ne pas trouver de LFA quand un pourrait être trouvé pour un des préfixes qui utilisent ce prochain bond.

Par exemple, considérons la Figure 4 où S a trois prochains bonds ECMP, E1, E2, et E3 pour atteindre D. Pour le préfixe D, E3 peut donner la protection de liaison pour les prochains bonds E1 et E2 ; E1 et E2 peuvent donner la protection de liaison pour les prochains bonds E3. Cependant, si on utilise cette simplification pour calculer les LFA pour E1, E2, et E3 individuellement, il n'y a pas de LFA de protection de liaison pour E1. E3 et E2 peuvent se protéger l'un l'autre.

4. Utilisation d'une solution de remplacement

Si un prochain bond de remplacement est disponible, le routeur redirige le trafic sur le prochain bond de remplacement en cas de défaillance d'un prochain bond principal comme suit.

Quand une défaillance de prochain bond est détectée via une défaillance de l'interface locale ou autres mécanismes de détection de défaillance (voir la [RFC5714]) le routeur DEVRAIT :

1. Supprimer le prochain bond principal associé à la défaillance.
2. Installer le remplaçant sans boucle calculé pour le prochain bond défaillant si il n'est pas déjà installé (par exemple, le remplaçant est aussi un prochain bond principal).

Noter que le routeur PEUT supprimer d'autres prochains bonds si il estime (via une analyse du SRLG) qu'ils peuvent avoir été affectés par la même défaillance, même si ce n'est pas visible au moment de la détection de la défaillance.

Le prochain bond de remplacement DOIT être utilisé seulement pour les types de trafic qui sont acheminés en accord avec le plus court chemin. Le trafic de diffusion groupée est spécifiquement hors du domaine d'application de la présente spécification.

4.1 Terminaison de l'utilisation d'un remplaçant

Un routeur DOIT limiter la durée pendant laquelle un prochain bond de remplacement est utilisé après que le prochain bond principal est devenu indisponible. Cela assure que le routeur va commencer à utiliser les nouveaux prochains bonds principaux. Cela assure que toutes les conditions transitoires possibles sont supprimées et que le réseau converge en accord avec le protocole d'acheminement déployé.

Des techniques sont disponibles pour traiter les micro-boucles de transmission qui peuvent se produire dans un réseautage durant la convergence.

Un routeur qui met en œuvre [MICROLOOP] DEVRAIT suivre les règles qui y sont données pour terminer l'utilisation d'un remplaçant.

Un routeur qui met en œuvre la [RFC6976] DEVRAIT suivre les règles qui y sont données pour terminer l'utilisation d'un remplaçant.

Il est souhaitable d'éviter les micro-boucles de transmission qui impliquent S. Un exemple qui illustre le problème est donné par la Figure 5. Si la liaison allant de S à E est défaillante, S va utiliser N1 comme remplaçant et S va calculer N2 comme nouveau prochain bond principal pour attendre D. Si S commence en utilisant N2 aussitôt que S peut calculer et

installer son nouveau principal, il est probable que N2 n'aura pas encore installé son nouveau prochain bond principal. Cela va causer une boucle de trafic qui va être éliminé jusqu'à ce que N2 ait installé la nouvelle topologie. Cela peut être évité si S retarde son installation et laisse le trafic sur le prochain bond de remplacement.

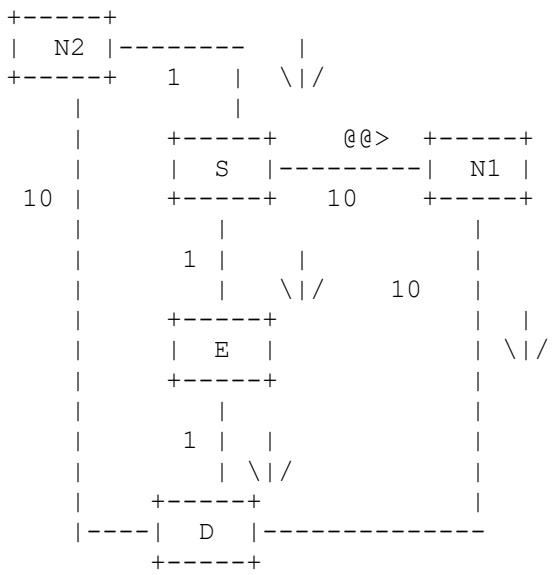


Figure 5 : Exemple où la poursuite de l'utilisation d'un remplacement est désirable

Ceci est un exemple du cas où le nouveau principal n'est pas un remplaçant sans boucle avant la défaillance et donc peut avoir transmis du trafic à travers S. Cela va se produire quand le chemin via un nœud précédemment en amont est plus court que le chemin via un remplaçant sans boucle voisin. Dans ces cas, il est utile de donner un temps suffisant pour assurer que le nouveau voisin principal et les autres nœuds sur le nouveau chemin principal sont passés sur le nouveau chemin.

Si le nouveau principal choisi était sans boucle avant la défaillance, il est alors sûr de passer immédiatement à ce nouveau principal ; le nouveau principal ne dépendait pas de la défaillance et donc son chemin n'aura pas changé.

Étant donné qu'il y a un remplaçant qui fournit une protection appropriée et quand l'hypothèse d'une seule défaillance tient, il est sûr de retarder l'installation des nouveaux principaux ; cela ne va pas créer de boucle de transmission parce que le chemin de remplacement à la destination est connu pour ne pas passer via S ou par l'élément défaillant et ne va donc pas être affecté par la défaillance.

Une mise en œuvre DEVRAIT continuer d'utiliser les prochains bonds de remplacement pour la transmission de paquets même après que les informations sur le nouvel acheminement sont disponibles sur la base de la nouvelle topologie du réseau. L'utilisation des prochains bonds de remplacement pour la transmission de paquets DEVRAIT se terminer :

a. si le nouveau prochain bond principal était sans boucle avant le changement de topologie, ou
b. si un maintien configuré, qui représente une limite de plus mauvais cas sur la longueur de la transition de convergence du réseau, a expiré, ou
c. si la notification d'un changement topologique sans rapport dans le réseau est reçue.

5. Exigences pour le mode LDP

Comme le trafic de LDP [RFC5036] va suivre le chemin spécifié par l'IGP, il est aussi possible que le trafic de LDP suive les remplaçants sans boucle indiqués par l'IGP. Pour ce faire, il est nécessaire que LDP ait les étiquettes appropriées disponibles pour le remplaçant afin que les segments appropriés puissent être installés dans le plan de transmission avant que la défaillance se produise.

Cela signifie qu'un routeur de commutation d'étiquettes (LSR, *Label Switching Router*) fonctionnant avec LDP doit distribuer ses étiquettes pour les classes d'équivalence de transmission (FEC, *Forwarding Equivalence Class*) qu'il peut fournir à tous ses voisins, sans considération de si ils sont ou non en amont. De plus, LDP doit agir de manière libérale dans son mode de rétention afin que les étiquettes qui correspondent aux voisins qui ne sont pas actuellement le voisin principal

soient mémorisées. De même, LDP devrait être en mode non sollicité vers l'aval, afin que les étiquettes pour la FEC soient distribuées autrement que le long du SPT.

Si ces exigences sont satisfaites, alors LDP peut utiliser les remplaçants sans boucle sans exiger de sessions ciblées ou d'extensions de signalisation pour cela.

6. Aspects d'acheminement

6.1 Préfixes multi-rattachements

Un calcul de style SPF est effectué pour chaque topologie, qui correspond à une zone OSPF ou niveau IS-IS particulier. L'IGP doit déterminer les remplaçants sans boucle aux chemins multi-rattachements. Les chemins multi-rattachements se produisent pour les chemins obtenus de l'extérieur du domaine d'acheminement par plusieurs routeurs, pour les sous réseaux sur les liaisons où le sous réseau est annoncé à partir de plusieurs extrémités de la liaison, et pour les chemins annoncés par plusieurs routeurs pour assurer la résilience.

La Figure 6 montre une telle topologie. Dans cet exemple, le plus court chemin pour atteindre le préfixe p est via E. Le préfixe p va avoir la liaison à E comme prochain bond principal. Si le prochain bond de remplacement pour le préfixe p est simplement hérité du routeur qui l'annonce sur le plus court chemin pour p, alors le prochain bond de remplacement du préfixe p va être la liaison pour C. Cela va assurer la protection de liaison, mais pas la protection de nœud qui est possible via A.

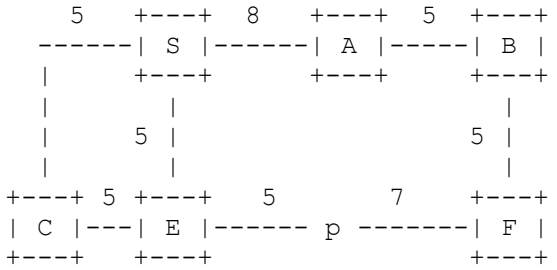


Figure 6 : Préfixe multi-rattachements

Pour déterminer la meilleure protection possible, le préfixe p peut être traité dans les calculs de SPF comme un nœud avec des liaisons unidirectionnelles pour lui à partir des routeurs qui ont annoncé le préfixe. Un tel nœud n'a jamais besoin que ses liaisons soient explorées, car il n'a pas de liaison sortante.

Si il existe plusieurs préfixes multi-rattachements qui partagent la même connexité et différence de métrique à ces routeurs, alors un seul nœud peut être utilisé pour représenter l'ensemble. Par exemple, si dans la Figure 6 il y avait un autre préfixe X connecté à E avec une métrique de 1 et à F avec une métrique de 3, alors ce préfixe X pourrait utiliser le même prochain bond de remplacement que calculé pour le préfixe p.

Un routeur DEVRAIT calculer le prochain bond de remplacement pour un préfixe IGP multi rattachements en examinant les chemins de remplacement via tous les routeurs qui ont annoncé ce préfixe.

Dans tous les cas, un routeur PEUT en toute sécurité simplifier le calcul de préfixe multi rattachements (MHP) en supposant que le MHP est seulement rattaché au routeur qui était son point de rattachement optimal avant la défaillance. Cependant, il peut en résulter qu'un préfixe ne soit pas considéré comme réparable alors que le calcul complet aurait montré qu'une réparation était possible.

6.2 IS-IS

L'applicabilité et les interactions des LFA avec IS-IS multi-topologies [RFC5120] sortent du domaine d'application de la présente spécification.

6.3 OSPF

OSPF introduit certaines complications parce qu'il est possible que le chemin du trafic sorte d'une zone et ensuite rentre

dans cette zone. Cela peut se produire chaque fois qu'un routeur envisage le même chemin à partir de plusieurs zones. Il y a plusieurs cas où des problèmes comme celui là peuvent se produire. Cela arrive quand une autre zone permet pour connecter deux ABR un plus court chemin qui est disponible dans la zone où le LFA a été calculé. Pour préciser, un exemple de topologie est donné à l'Appendice A.

a. Liaisons virtuelles : elles permettent aux chemins de quitter la zone cœur et traverser la zone de transit. Le chemin fourni via la zone de transit peut sortir via tout ABR. Le chemin pris n'est pas le plus court chemin déterminé en faisant un SPF dans la zone cœur.

b. ABR de remplacement [RFC3509] : quand un ABR n'est pas connecté au cœur de réseau, il examine les résumés inter-zones provenant de plusieurs zones. L'ABR A peut déterminer d'utiliser la zone 2 mais ce chemin pourrait traverser un autre ABR de remplacement B qui décide d'utiliser la zone 1. Cela peut conduire à des scénarios similaires à celui illustré à la Figure 7.

c. Résumés d'ASBR : un ASBR peut lui-même être un ABR et peut être annoncé dans plusieurs zones. Cela présente aux autres ABR une décision à prendre quant à la zone à utiliser. C'est l'exemple illustré dans la Figure 7.

d. Préfixes externes d'AS : un préfixe peut être annoncé par plusieurs ASBR dans différentes zones et/ou avec plusieurs adresses de transmission dans différentes zones, qui sont connectés via au moins un ABR commun. Cela impose à ces ABR une décision quant à la zone à utiliser pour atteindre le préfixe.

Les remplaçants sans boucle ne devraient pas être utilisés dans une zone où un des problèmes ci dessus affecte cette zone.

6.3.1 Acheminement OSPF externe

Quand une adresse de transmission est établie dans une annonce d'état de liaison (LSA, *Link State Advertisement*) OSPF d'AS externe, tous les routeurs dans le réseau calculent leurs prochains bonds pour le préfixe externe en faisant une recherche sur l'adresse de transmission dans le tableau d'acheminements, plutôt qu'en utilisant les prochains bonds calculés pour l'ASBR. Dans ce cas, les prochains bonds de remplacement DEVRAIENT être calculés en choisissant parmi les chemins de remplacement à la ou aux liaison de transmission plutôt que parmi les chemins de remplacement à l'ASBR.

6.3.2 OSPF multi-topologies

L'applicabilité et les interactions des LFA avec OSPF multi-topologies [RFC4915] [RFC5838] sortent du domaine d'application de la présente spécification.

6.4 Synchronisation de prochain bond BGP

Normalement, les préfixes BGP sont annoncés avec l'identifiant de routeur du routeur de sortie d'AS comme prochain bond BGP, et les routeurs de sortie d'AS sont atteints au moyen des chemins d'IGP. BGP résout son prochain bond annoncé en le prochain bond immédiat par de potentielles recherches récurrentes dans la base de données d'acheminement. Le réacheminement rapide IP calcule les prochains bonds de remplacement pour toutes les destinations IGP, qui incluent les prochains bonds de remplacement à l'identifiant de routeur du routeur de sortie de l'AS. BGP hérite simplement du prochain bond de remplacement provenant de l'IGP. Le processus de décision BGP n'est pas altéré ; BGP continue d'utiliser la distance IGP optimale pour trouver le plus proche routeur de sortie. Les chemins de diffusion groupée BGP (MBGP, *Multicast BGP*) n'ont pas besoin de copier les prochains bonds de remplacement.

Il est possible de fournir la protection d'ASBR si BGP a choisi un ensemble de prochains bonds BGP et a permis à l'IGP de déterminer le prochain bond principal et les prochains bonds de remplacement comme si le chemin BGP était un préfixe multi rattachements. Ceci fera l'objet d'études futures.

6.5 Considérations de diffusion groupée

Le trafic de diffusion groupée sort du domaine d'application de la présente spécification de réacheminement rapide IP. Les prochains bonds de remplacement NE DEVRAIENT PAS être utilisés pour les vérifications de transmission de diffusion groupée sur le chemin inverse (RPF, *Reverse Path Forwarding*).

7. Considérations sur la sécurité

Le mécanisme décrit dans le présent document ne modifie aucun message de protocole d'acheminement, et donc aucune nouvelle menace relative aux attaques de modifications de paquet ou de répétition n'est introduite. Le trafic pour certaines destinations peut être temporairement acheminé via des routeurs de prochain bond qui ne seraient pas utilisés avec le même changement de topologie si ce mécanisme n'était pas employé. Cependant, ces routeurs de prochain bond peuvent être utilisés de toutes façons quand un changement topologique différent se produit, et donc cela ne peut pas être vu comme une nouvelle menace pour la sécurité.

Dans LDP, la plus large distribution des informations d'étiquette de FEC est toujours avec des voisins avec lesquels une session LDP de confiance a été établie. Cette plus large distribution et la recommandation d'utiliser un mode libéral de rétention d'étiquettes sont estimés n'avoir pas d'impact significatif sur la sécurité.

8. Remerciements

Les auteurs tiennent à remercier Joel Halpern, Mike Shand, Stewart Bryant, et Stefano Previdi de leur assistance et de leur utile relecture.

9. Références

9.1 Références normatives

[RFC2119] S. Bradner, "Mots clés à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. *(MàJ par* RFC8174)

[RFC2328] J. Moy, "OSPF version 2", STD 54, avril 1998. *(MàJ par la* RFC6549, RFC8042, RFC9355)

[RFC2740] R. Coltun, D. Ferguson, J. Moy, "OSPF pour IPv6", décembre 1999. (*Obsolète, voir* RFC5340) *(P.S.)*

[RFC5036] L. Andersson, I. Minei et B. Thomas, éditeurs, "Spécification de LDP", janvier 2001. (*Remplace* RFC3036) (*MàJ par les* RFC6720 , RFC6790, RFC7552,) *(D.S)*

9.2 Références pour information

[MICROLOOP] Zinin, A., "Analysis and Minimization of Microloops in Link-state Routing Protocols", Travail en cours, octobre 2005.

[RFC1195] R. Callon, "Utilisation de l'IS-IS OSI pour l'acheminement dans les environnements TCP/IP et duels", décembre 1990. *(Mise à jour par les* RFC 1349, 5302, 5304*)*

[RFC2966] T. Li, T. Przygienda, H. Smit, "Distribution de préfixes sur un domaine avec IS-IS à deux niveaux", octobre 2000. (*Obsolète, voir* RFC5302) *(Information)*

[RFC3137] A. Retana et autres, "Annonce de routeur OSPF de bout", juin 2001. *(Information) (Remplacée par* RFC6987)

[RFC3509] A. Zinin, A. Lindem, D. Yeung, "Solution de remplacement à la mise en œuvre de routeurs de zone frontière OSPF", avril 2003. *(Information)*

[RFC4203] K. Kompella et autres, "Extensions OSPF pour la prise en charge de la commutation généralisée d'étiquettes multi-protocoles (GMPLS)", octobre 2005. (*MàJ* RFC3630) *(P.S.)*

[RFC4205] K. Kompella et Y. Rekhter, éd., "Extensions de système intermédiaire à système intermédiaire (IS-IS) pour la prise en charge de la commutation généralisée d'étiquettes multiprotocoles (GMPLS)", octobre 2005. (*Obsolète, voir* RFC5307) *(MàJ* RFC3784)

[RFC4915] P. Psenak et autres, "Acheminement multi topologies (MT) dans OSPF", juin 2007. *(P.S.)*

[RFC5029] JP. Vasseur, S. Previdi, "Définition d'un sous TLV d'attribut de liaison IS-IS", septembre 2007. *(P.S.)*

[RFC5120] T. Przygienda, N. Shen, N. Sheth, "M-ISIS : acheminement multi topologies (MT) de système intermédiaire à système intermédiaire (IS-IS)", février 2008. *(P.S.)*

[RFC5340] R. Coltun et autres, "OSPF pour IPv6", juillet 2008. (*P.S. ; Remplace* RFC2740 ; *MàJ par* RFC8362)

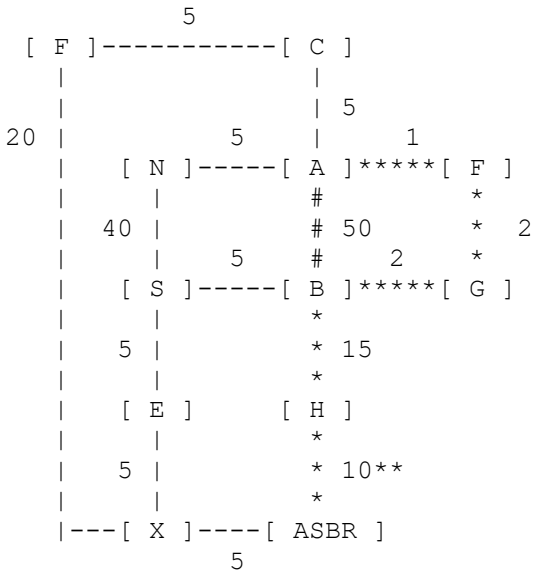
[RFC5714] M. Shand, S. Bryant, "Cadre du réacheminement rapide dans IP", janvier 2010. *(Information)*

[RFC5838] A. Lindem, S. Mirtorabi, A. Roy, M. Barnes, R. Aggarwal, "Prise en charge des familles d'adresses dans OSPFv3", avril 2010. *(P. S. : MàJ par* RFC7949, RFC8362)

[RFC6976] M. Shand et autres, "Cadre d'une convergence sans boucle avec l'approche de la base d'informations de transmission ordonnée (oFIB)", juillet 2013. *(Information)*

Appendice A. Exemple OSPF où le LFA fondé sur la topologie de zone locale est insuffisant

Cet Appendice donne un exemple de scénario où la topologie de zone locale ne suffit pas pour déterminer qu'un LFA est disponible. Comme décrit au paragraphe 6.3, un scénario de problème est celui de résumés d'ASBR où l'ASBR est disponible dans deux zones via des chemins intra-zone et où il y a au moins un ABR ou un ABR de remplacement qui est dans les deux zones. La Figure 7 illustre ce cas :



---- Liaison dans la zone 1
**** Liaison dans la 2
Liaison dans la zone cœur 0

Figure 7 : Topologie avec ASBR multi-zones causant un transit de zone

Dans la Figure 7, l'ASBR est aussi un ABR et est annoncé dans les deux zones 1 et 2. A et B sont tous deux des ABR qui sont aussi connectés à la zone cœur. S détermine que N peut fournir un remplaçant sans boucle pour atteindre l'ASBR. Le chemin de N passe par A. A voit aussi un chemin intra-zone pour l'ASBR via la zone 2 ; le coût du chemin dans la zone 2 est 30, qui est moins que 35, le coût du chemin dans la zone 1. Donc, A utilise le chemin provenant de la zone 2 et dirige le trafic sur F. Le chemin provenant de F dans la zone 2 passe par B. B est aussi un ABR et apprend l'ASBR des deux zones 1 et 2 ; le chemin de B via la zone 1 est plus court (coût 20) que le chemin de B via la zone 2 (coût 25). Donc, B utilise le chemin provenant de la zone 1 qui le connecte à S.

Adresse des auteurs

Alia K. Atlas
BT
mél : alia.atlas@bt.com

Christian Martin
iPath Technologies
mél : chris@ichemin.net

Brent Imhoff
Juniper Networks
mél : bimhoff@planetspork.com

Alex Zinin
Alcatel-Lucent
750D Chai Chee Rd, #06-06
Technopark@ChaiChee
Singapore 469004
mél : alex.zinin@alcatel-lucent.com

Raveendra Torvi
FutureWei Technologies Inc.
1700 Alma Dr. Suite 100
Plano, TX 75075
USA
mél : traveendra@huawei.com

Gagan Choudhury
AT&T
200 Laurel Avenue, Room D5-3C21
Middletown, NJ 07748
USA
mél : gchoudhury@att.com

Don Fedyk
Nortel Networks
600 Technology Park
Billerica, MA 01821
USA
mél : dwfedyk@nortelnetworks.com

Déclaration complète de droits de reproduction

Remerciement
Le financement de la fonction d'édition des RFC est fourni par l'activité de soutien administratif de l'IETF (IASA).