
De nouvelles perspectives pour la gestion de la Qualité de Service inter-domaine à grande échelle

Mohamed Boucadair
Pierre Lévis

*France Télécom Division R&D
42, rue des Coutures
BP 6243, F-14066 Caen cedex 4
{mohamed.boucadair, pierre.levis}@francetelecom.com*

RÉSUMÉ. Ce papier ouvre de nouvelles perspectives dans la gestion de la QS (Qualité de Service) IP pour l'Internet, à grande échelle. Nous proposons deux nouveaux aspects dans cette problématique, à partir desquels nous introduisons le concept de Meta-QoS-Class, que nous développons en présentant son application à l'Internet. Nous spécifions ensuite un protocole q-BGP associé à ce concept.

ABSTRACT. This paper opens up new perspectives on QoS (Quality of Service) for the Internet on a large scale. We propose two new issues in this area, from which we introduce the paradigm of Meta-QoS-Class, and then we elaborate on how it can be applied to the Internet. We finally specify a q-BGP protocol associated with this concept.

MOTS-CLÉS : BGP, DiffServ, inter-domaine, Internet, QS

KEYWORDS: BGP, DiffServ, inter-domain, Internet, QoS

1. Introduction

Il est notoire qu'à l'heure actuelle, malgré une dizaine d'années de travaux sur la QS IP, il n'existe toujours pas de déploiement de réseau QS IP à l'échelle de l'inter-domaine (malgré quelques propositions comme celle faite par le LIP6 (Nguyen, 2004), mais qui demeure une proposition partielle, vu qu'elle ne propose pas de framework ni d'architecture de déploiement de QS à l'inter-domaine et se focalise seulement sur la négociation des SLS). Le seul service de transport utilisable à travers l'Internet demeure le best-effort. Les raisons qui ont été invoquées sont multiples : un service premium est actuellement impossible à construire (Teitelbaum, 2001) dû, entre autres, à des contraintes trop fortes et coûteuses sur les routeurs ; la plupart des recherches fonctionnent de manière trop réactive (Crowcroft, 2003) oscillant entre périodes jugées « sans besoin de QS » ou la bande passante est supposée quasi infinie, et périodes jugées « avec besoin de QS » où se profile la menace d'une montée en volume significative des communications ; enfin, la mise en place de la QS dans un réseau, contrairement à ce qui a pu être parfois annoncé, est un processus complexe et coûteux (Bell, 2003).

Les études et les réalisations ont principalement considéré l'utilisateur professionnel, qui demande des garanties fortes et permanentes, pour un nombre de sites parfaitement connus et localisés, de type RPV (Réseau Privé Virtuel). Ce modèle de service n'implique souvent qu'un seul opérateur, ou un nombre très limité d'opérateurs, mais toujours parfaitement identifiés. Par contre, l'utilisateur plus résidentiel n'a été quasiment jamais pris en considération. Cet utilisateur a des demandes de QS beaucoup plus génériques, au moment où il le décide, avec peu d'a priori sur les services qu'il utilisera, quel que soit l'endroit où ces services sont rendus, pour recevoir et émettre de l'information. Ce type de service implique une infrastructure de QS inter-domaine très large, déployée à grande échelle sur tout l'Internet. On pourra se référer à (Huston, 2000) pour des considérations sur la pluralité des approches d'architectures QS IP. Les travaux présentés ici, conduits dans le cadre du projet européen IST MESCAL¹ suite du projet IST TEQUILA² (Trimintzios, 2001), se focalisent plus particulièrement sur ces aspects de QS inter-domaine à grande échelle. Nous nous plaçons dans le cadre générique de DiffServ (Blake, 1998), chaque opérateur ayant déployé des classes de services (I-QC pour : local QoS Class) dans son réseau.

Une solution de transport QS pour l'Internet devrait s'assurer que les mécanismes et techniques qu'elle met en œuvre n'altèrent par l'esprit dans lequel l'Internet s'est développé jusqu'à présent, voir (Clark, 1988). Il ne s'agit bien sûr pas, dans notre idée, d'exclure toute évolution du paradigme Internet (voir (Blumenthal, 2001) et (Kempf, 2004) pour des réflexions sur ces évolutions), mais de conserver un

1. <http://www.mescal.org>

2. <http://www.ist-tequila.org>

ensemble de propriétés qui pourraient être jugées comme fondamentales, et acceptées par tous les acteurs impliqués, afin de construire une Société de l'Information ouverte à tous.

Ce papier se propose d'ouvrir de nouvelles perspectives dans la QS IP à grande échelle, et plus précisément dans le cadre de la problématique inter-domaine. Après un ensemble de définitions, nous présentons deux nouveaux éléments à prendre en compte dans la problématique QS inter-domaine : le dédoublement des l-QC et le piège des chaînes d'AS (Autonomous System), à partir desquels nous introduisons le concept de Meta-QoS-Class, que nous développons en présentant son application à l'Internet. Nous spécifions ensuite un protocole q-BGP associé à ce concept.

2. Définitions

Dans le cadre de ce document les termes suivants signifient:

- *Service Level Specification (SLS)* : ensemble de paramètres techniques négociés pour un service. Il s'agit ici de service de transport à QS garantie, en termes de {D, J, L}.
- *Customer SLS (cSLS)* : SLS établi entre un fournisseur de service et un client.
- *Provider SLS (pSLS)* : SLS établi entre deux opérateurs de réseaux.
- *QoS Class (QC)* : capacité de transport à QS, exprimée en termes de {D, J, L}.
- *Local QC (l-QC)* : QC qui couvre un seul AS (notion similaire au Per Domain Behavior (Nichols, 2001)).
- *Extended QC (e-QC)* : QC qui couvre plusieurs AS. Il s'agit d'une suite ordonnée de l-QC.

3. Deux nouveaux éléments dans la problématique de la QS inter-domaine

3.1. Dédoublement des l-QC

La figure 1 représente une partie de l'Internet avec quatre AS ayant chacun des l-QC en nombre variable. On suppose qu'une entité cliente a négocié un cSLS, avec un fournisseur de service (qui peut être celui qui gère le domaine AS1), donnant deux niveaux de QS différents, pour deux communications, de l'utilisateur U1, connecté sur l'AS1, à l'utilisateur U2, connecté sur l'AS4. Pour chaque communication, une l-QC a été choisie dans chaque AS, afin de sélectionner une e-QC qui satisfasse au niveau de QS exprimé dans le cSLS. Soit au final, l'e-QC1 et l'e-QC2.

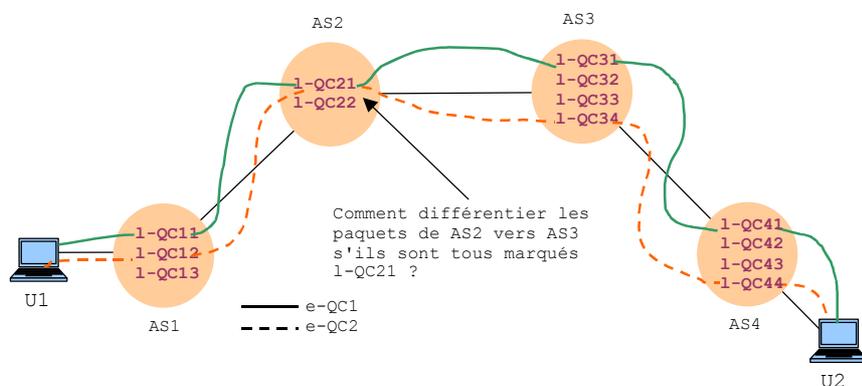


Figure 1. Le problème de dédoublement des I-QC

On remarque qu'en sortie de l'AS2, I-QC21 est aboutée dans AS3 à I-QC31 et à I-QC34. Concrètement, cela signifie que dans l'ensemble des paquets qui subissent I-QC21 dans AS2, certains devront subir I-QC31 dans AS3, alors que d'autres devront subir I-QC34. Ce qui reviendra, si DiffServ est appliqué, à les marquer avec un DSCP (Differentiated Services Code Point) différent. Le problème est ici de savoir comment reconnaître ces deux types de paquets, sachant qu'a priori il n'y a pas de raison pour qu'ils n'aient pas le même DSCP dans AS2 (remarquons que tous ces paquets ont aussi en commun au moins les adresses IP source et destination). On ne peut bien sûr pas affecter un DSCP différent par e-QC dans AS2, ce qui limiterait à 64 le nombre de communications inter-domaines qu'un AS serait capable de traiter ! Il est tout aussi déraisonnable de faire effectuer à l'ASBR (AS Border Router) de sortie de l'AS2, un contrôle systématique de tous les paquets IP (avec ou sans QS) sortants, en examinant les champs protocolaires tels que, au niveau IP : adresse source, adresse destination, protocole, DSCP ; et au niveau UDP ou TCP : port source, port destination. Afin de classer chaque paquet et de rechercher dans une table l'e-QC qui lui correspond.

Une I-QC d'un AS peut être ainsi, a priori, aboutée à plusieurs I-QC dans le domaine suivant. Pour simplifier l'exposé du problème on s'est limité ici au cas d'aboutement sur deux I-QC que l'on appellera *dédoublement*. Ce phénomène de dédoublement de I-QC doit être soit interdit, soit limité et parfaitement contrôlé. Pour d'autres considérations sur la problématique QoS inter-domaine dans le plan des données, voir (Levis, 2004a).

3.2. Menace d'une aire de glaciation

Un opérateur va chercher à étendre ses propres services à QS, implantés par des I-QC, afin de pouvoir offrir à ses clients des services de transport à QS, pour des destinations qui ne sont pas forcément dans son domaine. Pour cela il doit passer des accords, dits de pSLS, avec ses vis-à-vis. Nous allons analyser ici la manière dont ces pSLS peuvent être établis, en mettant en garde contre des risques de blocage.

Supposons que l'opérateur n-1 annonce, à travers un medium quelconque, qu'il offre à ses voisins la possibilité d'utiliser des chemins à QS garantie vers des destinations distantes. Chaque offre peut se décrire sous la forme (Destination, {D, J, L}) où Destination est un préfixe de réseau IP, et {D, J, L} sont les valeurs de QS garanties pour le transport des paquets, du réseau de l'opérateur n-1 jusqu'à Destination. L'opérateur n, qui est un des voisins (voisin au sens BGP) directs de l'opérateur n-1, peut en déduire les performances qu'il pourrait garantir à ses propres clients pour joindre ces Destinations. En effet, il lui suffit d'appliquer le bon opérateur qui combinera les performances de ses propres I-QC, aux performances offertes par l'opérateur n-1, par exemple une simple addition pour ce qui est du paramètre Délai. L'opérateur n achète les (Destination, D, J, L) de l'opérateur n-1 qui lui semblent offrir des opportunités de services intéressantes pour ses clients. Il négocie alors les pSLS correspondants avec l'opérateur n-1.

Les garanties sont fournies de manière récursive. L'opérateur n-1 garantit à l'opérateur n la traversée de toute la chaîne des AS (ASn-1, ASn-2, ASn-3..., AS1), de son propre AS, à l'AS sur lequel se trouve la Destination. En se focalisant sur ASn-3, on voit que l'opérateur n-1 garantit implicitement à l'opérateur n un niveau de QS pour la traversée de ASn-3. ASn-3 est susceptible de faire ainsi partie de nombreuses chaînes d'AS, et de voir sa traversée garantie par plein d'opérateurs dont il ne soupçonne peut être même pas l'existence. D'où deux problèmes connexes que nous mettons en évidence :

– *Le piège des chaînes d'AS* : en tant qu'opérateur, ma liberté de terminer ou renégocier un des mes accords incluant mon propre AS (sans forcément s'y limiter), ainsi que ma liberté de modifier les caractéristiques de mes propres I-QC, est limitée par le nombre d'accords externes qui garantissent la traversée de mon propre AS.

– *La menace d'une aire de glaciation* : on risque ainsi de se retrouver avec une infrastructure Internet à QS, complètement gelée, où personne ne peut renégocier aucun accord.

4. Meta-QoS-Class

4.1. Du pSLS au concept de Meta-QoS-Class

De l'analyse qui précède nous déduisons les deux recommandations suivantes : (1) N'aboutir une I-QC à une I-QC voisine, que sur la seule connaissance de leurs

performances, sans prendre en considération les performances des I-QC distantes, le critère d'aboutement est alors le degré de similitude entre caractéristiques des deux I-QC et (2) Ne garantir par un pSLS que les performances sur la traversée d'un seul et seulement un seul AS.

Il peut apparaître alors un problème de cohérence et de consistance sur une e-QC. L'e-QC : (I-QC_n, I-QC_{n-1}, I-QC_{n-2}..., I-QC₁) est créée par aboutement deux à deux des I-QC voisines ayant des caractéristiques de QS similaires. Il risque de se produire un glissement progressif de ces caractéristiques : entre I-QC_n et I-QC_{n-1}, puis entre I-QC_{n-1} et I-QC_{n-2}, et ainsi de suite jusqu'à I-QC₁, avec in fine un écart significatif entre la première et la dernière I-QC. Pour contraindre cet écart à rester dans des limites acceptables, nous avons besoin d'un outil de classification. Selon cet outil, deux I-QC sont susceptibles d'être aboutées si, et seulement si, elles appartiennent au même type de classe. Nous appelons Meta-QoS-Class (ou MQC en abrégé) chaque classe de cette taxinomie des I-QC. *Deux I-QC ne peuvent être aboutées que si, et seulement si, elles appartiennent à la même MQC.*

4.2. Basé sur les besoins des applications

La philosophie qui sous-tend le concept de MQC repose sur le fait que les besoins des applications sont universels. Quel que soit l'endroit où les utilisateurs sont connectés, ils utilisent à peu près les mêmes types d'applications, dans des contextes professionnels ou résidentiels similaires. Ils sont face aux mêmes difficultés, et sont enclins à exprimer des besoins comparables en termes de QS. De leurs côtés, les opérateurs, confrontés à ces mêmes besoins clients, ont toute chance de déployer chacun des I-QC similaires pour traiter chaque type d'application. Il n'y a pas de raison particulière pour qu'un opérateur du Canada ait déployé une I-QC apte au transport de la VoIP avec des caractéristiques de délai et de gigue faibles, alors qu'un autre opérateur en Suisse aurait déployé une I-QC avec des caractéristiques opposées. Les applications imposent des contraintes sur le réseau, indépendamment d'où le service est offert sur l'Internet. On peut se reporter à (Miras, 2002) pour un inventaire très complet des besoins applicatifs.

4.3. Exemple de Meta-QoS-Class

Il faut bien comprendre qu'une MQC est une notion abstraite, que ce n'est pas une I-QC réelle, déployée dans un réseau réel. Nous donnons ici un exemple de MQC, cet exemple est à considérer uniquement comme une illustration du concept. La définition effective des MQC devrait résulter d'un processus de normalisation.

- Nom : MQC Gold 1
- Utilisation cible : applications sensibles.
- Performance : délai faible, gigue faible, pertes faibles.

- Contrainte sur le trafic : uniquement des flux TCP-friendly. On dissocie deux classes car, parce que l'on autorise la destruction de paquets, un mélange de flux TCP et de flux non-TCP risquerait de pénaliser les flux TCP.

- Autres contraintes : sur chaque interface de sortie de routeur, le trafic de la classe peut être supérieur à la bande passante réservée pour la classe (comportement type AF (Heinanen, 1999)), la différence entre volume de trafic et bande passante ayant un impact sur le taux de pertes.

Une MQC est ainsi définie par un ensemble d'attributs qualifiant la traversée d'un AS. Certains d'entre eux peuvent dépendre du diamètre de l'AS, par exemple un plus long délai peut être accepté pour de grands AS. Les attributs de performances peuvent également être pondérés pour donner une priorité à ceux pour lesquels le service est le plus sensible. La notion de MQC ne doit pas être confondue avec la notion de PDB (Per-Domain Behavior) de DiffServ (Nichols, 2001). Les deux notions ont en commun de spécifier un certain niveau de QS. Elles diffèrent par leurs objectifs. Le but de la définition d'un PDB est d'aider à l'implantation de capacités de transport à QS, à l'intérieur d'un réseau. Le but de la MQC est d'aider à la négociation d'accords entre opérateurs.

4.4. L'Internet comme un ensemble de plans à Meta-QoS-Class

Nous allons voir maintenant l'impact possible de l'application de cette notion sur l'Internet. La figure 2 représente une portion d'Internet réduite, vue sous trois aspects différents : à gauche la vue physique, au centre la vue logique par rapport à la MQC1, à droite la vue logique par rapport à la MQC2.

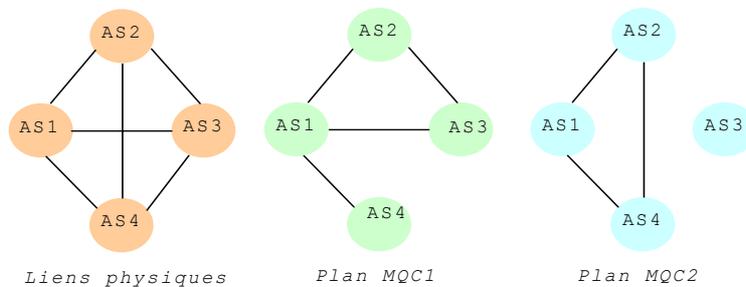


Figure 2. Différentes vues d'une portion d'Internet

Les deux vues logiques représentent les pSLS signés entre ASs, pour la MQC mentionnée. Intéressons nous plus particulièrement au plan MQC2. Supposons que cette MQC2 soit la MQC Gold que nous avons donnée en exemple précédemment. Nous pouvons déduire de ce schéma que les domaines AS1, AS2 et AS4 ont chacun au moins une l-QC qui peut être classifiée dans la catégorie MQC Gold. Par contre,

8 GRES, 28 Février - 02 Mars 2005, Luchon.

nous ne savons rien sur AS3. Nous pouvons déduire également que les couples d'AS : (AS1, AS2), (AS1, AS4), (AS2, AS4), ont négocié chacun un pSLS (que l'on suppose ici bidirectionnel) en aboutant deux I-QC conformes à la MQC Gold. Supposons qu'à son tour, AS3 veuille joindre cette infrastructure Gold. Il faut d'abord qu'il s'assure de la conformité d'une de ses I-QC avec la MQC Gold. Ensuite, il doit négocier un pSLS, avec par exemple AS2. Dès que le pSLS est effectif, AS3 se trouve alors connecté au plan MQC Gold et peut potentiellement bénéficier de toute l'infrastructure de ce plan MQC. Pour un opérateur, contracter avec un autre opérateur un pSLS basé sur le concept de MQC signifie, de manière très simple, rajouter un lien dans le plan de cette MQC. Exactement comme actuellement, pour un opérateur, se raccorder à un autre opérateur signifie rajouter un lien à l'Internet.

Tous les opérateurs doivent avoir la même compréhension de ce qu'est une MQC donnée. Un consensus global doit donc être atteint pour la définition de chaque MQC. Typiquement, il s'agit d'un travail de normalisation à prendre en charge par un organisme international. Le besoin de normalisation pour la QS inter-domaine a déjà été évoqué, par exemple dans (Eder, 2002). Le nombre de MQC à définir, et par conséquent le nombre de plans MQC différents, doit rester très petit, de l'ordre de 3 ou 4, pour éviter une complexification non nécessaire. Les normes définissant les MQC devraient être accompagnés de documents décrivant des tests de conformité.

Pour conserver les propriétés de base que nous avons énoncées en introduction de notre papier, nous choisissons de faire fonctionner chaque plan MQC comme un Internet parallèle. Le routage inter-domaine pourra alors être traité par un protocole basé sur BGP (Border Gateway Protocol) (Rekhter, 1995), que l'on appellera q-BGP. D'un point de vue abstrait, tout se passe comme si chaque plan MQC avait son propre protocole BGP standard. Concrètement, par contre, au niveau implantation, nous avons retenu un seul protocole qui gère le routage de tous les plans MQC. Ce protocole est décrit en détail dans le paragraphe 5. Un plan MQC bénéficie ainsi de toute la souplesse du routage IP, et plus particulièrement de sa propriété de résilience.

Le concept de MQC s'accommode parfaitement des nouvelles applications. Conformément au principe de bout en bout (Saltzer, 1984), une nouvelle application que l'on n'aurait pas su anticiper, doit avoir un impact très faible sur les I-QC existantes, parce que celles-ci doivent avoir été conçues, de façon à permettre à n'importe quelle nouvelle application d'être compatible avec l'infrastructure QS existante. Mais ce problème ne concerne pas les MQC, parce qu'une MQC est un objet abstrait qui n'a aucune existence physique. Il concerne uniquement l'ingénierie des I-QC. Ainsi, une nouvelle application, tout au plus, pourra conduire à la définition d'une nouvelle MQC, et d'un nouveau processus de classification pour les I-QC.

A l'inter-domaine, il sera nécessaire de signaler à quel MQC chaque paquet appartient, ceci peut se faire par exemple par le DSCP. Il n'y a donc plus de problème de *dédoublement de I-QC*, puisque du point de vue abstrait de la MQC, on

aboute toujours une MQC à la même MQC. La correspondance entre la MQC signalée par le DSCP et la l-QC qui traitera effectivement le paquet, restant un problème et une décision complètement locale à chaque AS. A noter que grâce au concept de MQC, nous avons réduit la complexité de signaler au niveau de chaque AS une e-QC parmi toutes, soit potentiellement une communication parmi toutes, ce qui peut faire un nombre énorme dans un AS, à signaler une MQC parmi toutes, soit potentiellement un nombre que l'on peut compter sur les doigts d'une seule main.

La structure de chaque plan MQC est susceptible d'évoluer de manière très dynamique au fur et à mesure de l'apparition de nouveaux accords, ou de la résiliation d'accords existants. Cependant, un pSLS donné est complètement insensible à tout changement portant sur un autre pSLS, soit dans le même plan MQC, soit dans un autre plan MQC. Il n'y a donc pas de risque ici de *piège de chaînes d'AS* ni de *menace d'aire de glaciation*.

Un plan MQC peut avoir des trous, ce qui limite la portée des services offerts. On ne peut disposer de services accessibles à une très large échelle, que si on dispose de grandes étendues connexes, donc uniquement si le concept de MQC est largement utilisé. Cependant, il faut bien noter que ce problème n'est pas spécifique à l'utilisation de MQC, mais qu'il est commun à toute problématique de services QS à grande échelle. Un développement graduel des pSLS, conduit néanmoins à un développement graduel de l'étendue des services.

4.5. Les garanties pour les cSLS

Rappelons que nous avons basé les pSLS sur la base d'un accord à couverture locale. Un pSLS ne garantit que la traversée d'un AS, et ne garantit donc pas le transport plus loin qu'un saut d'AS. La garantie, sur l'ensemble du chemin d'AS, repose donc sur la concaténation de l'ensemble des accords locaux. Ce système fonctionne donc, dans la seule mesure où chaque opérateur joue son rôle et garantit effectivement le niveau de QS qu'il annonce. Cette limitation n'est cependant pas propre à notre approche, cette limitation existe quelle que soit l'approche utilisée, avec ou sans MQC, avec ou sans couverture globale du pSLS.

En ce qui concerne les cSLS, rien n'empêche de fournir des garanties de bout en bout, alors que les pSLS ne fournissent que des garanties limitées chacun à la traversée d'un AS. La figure 3 montre l'étendue des garanties fournies par les deux types de SLS. Il n'y a aucun lien automatique, a priori, entre ce qui est dit dans le pSLS, et ce qui est dit dans le cSLS. Le fournisseur de service s'efforcera simplement qu'il y ait un niveau de cohérence suffisant entre les deux, par rapport à une prise de risque toujours présente. Voir à cet effet (Teitelbaum, 2003) pour une analyse sur l'écart possible entre ce que fournit le réseau et ce qu'offre le fournisseur de service.

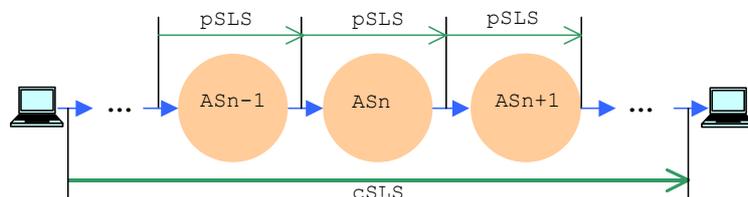


Figure 3. *Étendue des garanties des pSLS et du cSLS*

De manière plus générale, la MQC n'est qu'un concept. Baser les pSLS et les échanges q-BGP sur ce concept n'interdit, en aucune manière, d'utiliser d'autres mécanismes et protocoles, que l'on jugera pertinents pour fournir le niveau de QS désiré, aussi bien dans le plan des Données, le plan de Contrôle, que le plan Service. En particulier : DiffServ, lissage de trafic, ingénierie de trafic, contrôle d'admission, bandwidth broker, facturation, etc., sont complètement légitimes dans notre approche. Le concept de MQC simplement oriente et fédère la façon dont les accords pSLS sont négociés.

5. QoS-Inferred Border Gateway Protocol (q-BGP)

5.1. Motivations

Les opérateurs de réseau ont déployé des protocoles de routage inter-domaine, afin d'échanger des informations de routage, et permettre ainsi d'atteindre des destinations au-delà des frontières de leurs propres domaines. Le protocole BGP est le protocole de facto déployé à cet effet. Les informations véhiculées par BGP ne semblent pas suffisantes pour offrir des services à QS de bout en bout. Si le choix de la route est basé sur l'attribut AS_PATH, le processus de routage minimise le nombre de domaines traversés, ce qui ne garantit pas forcément le chemin avec la meilleure QS. La figure 4 présente un exemple de topologie d'AS. Si BGP est activé entre AS voisins, l'AS1 choisira le chemin {AS2, AS5} pour joindre des destinations annoncées par l'AS6, et non le chemin {AS3, AS4, AS5}. Pourtant, le délai pour traverser les deux AS : AS2 et AS5 est de 250ms, alors que le délai pour traverser les trois AS : AS3, AS4 et AS5 n'est que de 200ms.

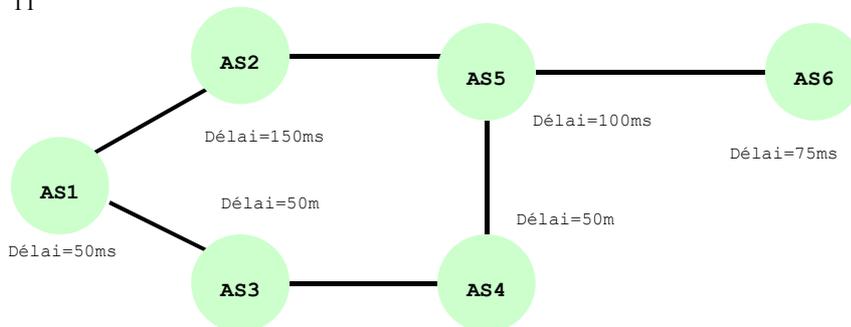


Figure 4. *Choix de chemin d'AS*

Dans cette perspective, les opérateurs réseaux doivent faire évoluer les protocoles de routage inter-domaine qu'ils déploient, afin de répondre aux exigences QS, et ainsi offrir de nouveaux services plus sophistiqués. Dans cette section, nous proposons un enrichissement du protocole BGP.

5.2. Cas d'utilisation : q-BGP avec des MQC

Nous choisissons d'avoir une seule session q-BGP entre deux domaines, capable de gérer le routage pour l'ensemble des plans MQC, y compris la MQC best-effort. Lorsqu'un paquet va circuler à travers plusieurs domaines, il devra suivre un chemin où il n'empruntera que des l-QC correspondant à la même MQC. Il faut donc que q-BGP puisse calculer des routes dans chaque plan MQC. Pour cela, la session q-BGP doit échanger des informations d'accessibilité de réseaux par rapport à chaque MQC. Le protocole q-BGP doit donc être enrichi au minimum d'une information de MQC associée à chaque chemin d'AS annoncé. A cette information de base peuvent se rajouter des informations supplémentaires donnant des valeurs de paramètres QS pour chaque chemin d'AS. Nous distinguons ainsi deux niveaux d'enrichissement de BGP :

- Niveau 1 : échanger seulement l'identifiant de la MQC qui a été négociée durant la phase d'accord d'un pSLS.
- Niveau 2 : échanger en plus de l'identifiant de la MQC négociée, un ensemble de valeurs de paramètres QS tels que le délai minimum, délai maximum, etc.

De plus, q-BGP doit être compatible avec BGP. En particulier, un processus BGP doit savoir lire, dans un message q-BGP, les informations relatives au best-effort (considéré comme une MQC particulière par q-BGP) et de manière plus générale un processus q-BGP et un processus BGP doivent être capables d'établir une session, et d'échanger des informations de routage relatives au best-effort. Enfin,

12 GRES, 28 Février - 02 Mars 2005, Luchon.

les informations rajoutées à BGP le sont expressément dans les limites autorisées par la norme elle-même (Rekhter, 1995).

Il est important de rappeler que q-BGP n'est qu'un élément parmi d'autres dans la fourniture d'un service de transport à QS de bout en bout. Ainsi, aux deux niveaux de complexité de q-BGP proposés, vont correspondre deux groupes de services. Notre proposition q-BGP est donc relativement générique et devrait pouvoir s'appliquer à un grand nombre de solutions de service de transport à QS inter-domaine. Par la suite, nous appelons groupe 1 les solutions reposant sur l'échange des informations q-BGP de niveau 1, et groupe 2 les solutions reposant sur l'échange des informations q-BGP de niveau 2.

5.3. Spécification de q-BGP

Pour offrir des services à QS inter-domaine, q-BGP doit diffuser les informations propres au groupe auquel appartient le service de transport qui l'exploite. Ces informations sont listées ci-dessous :

- Option de service QS (groupes 1 et 2).
- Identifiant de MQC (groupes 1 et 2).
- Performance de QS (groupe 2).

Toutes ces informations sont incluses dans les messages BGP, en utilisant uniquement les nombreuses possibilités d'extension offertes par BGP lui-même.

5.3.1. Option de service QS

Un routeur q-BGP doit connaître les options supportées par son voisin en termes d'extensions q-BGP supportées, et d'attributs qu'il peut décoder. La négociation de ces options est réalisée grâce à l'utilisation du paramètre facultatif «BGP Capability» (Chandra, 2002) défini par le groupe de travail IDR (Inter-Domain Routing) de l'IETF. Ce paramètre est utilisé à l'initialisation d'une session q-BGP, il est inclus dans le message OPEN. Il liste l'ensemble des capacités supportées par l'émetteur. Nous avons défini pour ce paramètre une nouvelle capacité « QoS Service Capability ». Elle indique à quel groupe (groupe 1 ou groupe 2) le service de transport qui utilise q-BGP appartient, et donc, la présence ou non d'informations quantitatives de paramètres de QS. Pour plus d'informations à propos de la structure de cet attribut se reporter à (Howarth, 2004).

5.3.2. Identifiant MQC

Afin de préciser l'identifiant de la MQC dans laquelle se fait l'annonce d'un chemin d'AS, nous avons créé le champ « QoS Class Identifier ». La valeur de ce champ doit être conforme à ce qui a été convenu entre deux pairs lors de la phase de négociation de pSLS. Ce champ est codé sur un octet. Il doit être précisé pour

De nouvelles perspectives pour la gestion de la QS inter-domaine à grande échelle
13

chaque chemin d'AS annoncé. Pour véhiculer cette information, nous avons repris et adapté la proposition d'attribut QoS_NLRI (pour une version antérieure de cet attribut se référer à (Cristallo, 2002). C'est un attribut optionnel transitif.

5.3.3. Performance de QS

Nous avons également intégré les informations de performance de QS dans l'attribut QoS_NLRI. Les types d'informations possibles sont les suivants : variation du délai inter paquet (gigue), débit réservé, débit disponible, taux de perte, délai unidirectionnel minimum, délai unidirectionnel maximum, délai unidirectionnel moyen. Pour la structure précise de l'attribut QoS_NLRI, se reporter à (Howarth, 2004). Une annonce q-BGP peut contenir une ou plusieurs informations de performance de QS.

5.3.4. Processus de choix de route

Les informations de QS étant échangées par des messages UPDATE BGP, le processus de choix de route devrait tenir compte de ces informations. Ce processus est différent selon le groupe auquel appartient le service de transport à QS que sert q-BGP. Nous avons donc modifié le processus de choix de route BGP pour tenir compte des ces informations de QS. Dans le cas du groupe 1, le processus de routage choisit une route par plan MQC de manière très similaire à ce que fait actuellement BGP, c'est-à-dire en sélectionnant le chemin avec le plus petit nombre d'AS. L'optimisation du chemin est basée sur les politiques q-BGP configurées par chaque opérateur en fonction des accords pSLS établis avec ses voisins. Quant au groupe 2, le processus examine les informations de QS véhiculées par q-BGP en fonction de la priorité associée à chaque paramètre de QS. Le processus de routage favorise les routes qui optimisent le paramètre de QS le plus prioritaire. Ci-dessous, le pseudo-code du processus de choix de route.

1. Examiner l'ensemble de routes reçues à destination d'un préfixe donné
2. Sélectionner le paramètre de QS avec la plus haute priorité, et retourner la/les routes qui optimise(nt) le paramètre en question
3. Si une seule route est trouvée, stocker cette route dans la RIB
4. Si plusieurs routes retournées
 - ➔ Exclure le paramètre QS déjà examiné de la liste des paramètres QS
 - ➔ Recommencer l'étape 1.

6. Conclusion

Dans cet article, nous avons présenté de nouvelles perspectives dans le domaine de la QS inter-domaine. Nous avons mis en évidence deux nouveaux aspects à prendre en compte dans cette problématique : le dédoublement de I-QC et le risque de piège des chaînes d'AS. A partir de cette analyse, nous avons introduit la notion

14 GRES, 28 Février - 02 Mars 2005, Luchon.

de Meta-QoS-Class qui devrait faciliter la négociation des pSLS. Nous avons également spécifié un protocole q-BGP qui est un enrichissement de BGP, compatible avec BGP, et qui répond à des besoins de service non rendus par les moyens existants. Dans le cadre du projet MESCAL nous avons réalisé une implantation sous routeurs Linux. Les tests conduits ont principalement permis de valider le protocole sous ses aspects fonctionnels, et ont montré son aptitude à calculer des chemins d'AS répondant aux exigences de QS. Au niveau des performances, le comportement observé montre que les routeurs ne sont pas lourdement impactés par l'introduction des nouveaux plans MQC. Cependant, une étude reste à faire sur des routeurs commerciaux afin de mesurer le niveau exact des performances dans un contexte opérationnel. Ces résultats feront l'objet d'une future publication.

Nous tenons à remercier tous les partenaires du projet MESCAL pour leur participation à ces réflexions.

7. Références

- Almes G., Kalidindi S., Zekauskas M., « A One-way Delay Metric for IPPM », RFC 2679, septembre 1999.
- Almes G., Kalidindi S., Zekauskas M., « A One-way Packet Loss Metric for IPPM », RFC 2680, septembre 1999.
- Bell G., « Failure to Thrive: QoS and the Culture of Operational Networking », *RIPQoS Workshop SIGCOMM2003*, août 2003.
- Blake S., Black D., Carlson M., Davies E., Wang Z., Weiss W., « An Architecture for Differentiated Services », RFC 2475, décembre 1998.
- Blumenthal M., Clark D.D., « Rethinking the design of the Internet: The end to end arguments vs. the brave new world », *ACM Transactions on Internet Technology*, Vol. 1, No. 1, août 2001, pp 70-109.
- Chandra R., Scudder J., Capabilities Advertisement with BGP-4, RFC3392, novembre 2002.
- Clark D.D., « The Design Philosophy of the DARPA Internet Protocols », *ACM SIGCOM 88*.
- Cristallo G., Jacquenet C., « An Approach to Inter-domain Traffic Engineering », *XVIII World Telecommunications Congress (WTC2002)*, septembre 2002.
- Crowcroft J., Hand S., Mortier R., Roscoe T., and Warfield A., « QoS's Downfall: At the bottom, or not at all! », *RIPQoS Workshop SIGCOMM2003*, août 2003.
- Demichelis C., Chimento P., « IP Packet Delay Variation Metric for IP Performance Metrics (IPPM) », RFC 3393, novembre 2002.
- Eder M., Chaskar H., Nag S., « Considerations from the Service Management Research Group (SMRG) on Quality of Service (QoS) in the IP Network », RFC3387, septembre 2002.

De nouvelles perspectives pour la gestion de la QS inter-domaine à grande échelle
15

Heinane J., Baker F., Weiss W., Wroclawski J., « Assured Forwarding PHB Group », RFC2597, juin 1999.

Huston G., « Next Steps for the IP QoS Architecture », RFC2990, novembre 2000.

D1.4 Livrable projet IST MESCAL. Disponible à www.mescal.org

Kempf J., Austein R., « The Rise of the Middle and the Future of End to End: Reflections on the Evolution of the Internet Architecture », RFC3724, mars 2004.

Levis P., Asgari A., Trimintzios P., « Considerations on inter-domain QoS and Traffic Engineering issues through a utopian approach », *Service Assurance with Partial and Intermittent Resources (SAPIR)*, Fortaleza Brazil, août 2004, ©Springer-Verlag.

Levis P., Boucadair M., Morand P., Trimintzios P., « The Meta-QoS-Class concept: a step towards global QoS inter-domain services », *International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2004)*, octobre 2004.

Miras D., « A Survey on Network QoS Needs of Advanced Internet Applications », Internet2 QoS Working Group working document, décembre 2002.

Nguyen M., Boukhatem N., « Négociation du niveau de service avec le protocole COPS-SLS », *Annales des Télécommunications*, Vol. 59, 2004.

Nichols K., Carpenter B., « Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification », RFC3086, avril 2001.

Rekhter Y., Li T., « A Border Gateway Protocol 4 », RFC 1771, mars 1995.

Saltzer J.H., Reed D.P., Clark D.D., « End to End Arguments in System Design », in *ACM Transactions on Computer Systems*, 1984.

Teitelbaum B., « Future Priorities for Internet2 QoS », Internet2 QoS Working Group document, octobre 2001.

Teitelbaum B., Shalunov S., « What QoS Research Hasn't Understood About Risk », *RIPQoS Workshop SIGCOMM2003*, août 2003.

Trimintzios et al., « A Management And Control Architecture for Providing IP Differentiated Services in MPLS-based Networks », *IEEE Communications Magazine*, vol. 39, No 5, pp. 80-88, May 2001.

Wang Z., Crowcroft J., « Quality of Service Routing for Supporting Multimedia Applications », *IEEE Journal of Selected Areas in Communications (JSAC)*, 1996.

Xiao L., Lui K.-S., Wang J., Nahrstedt K., « QoS extension to BGP », *10th IEEE International Conference on Network Protocols (ICNP'02)*, novembre 2002.