

Groupe de travail Réseau  
**Request for Comments : 2121**  
Catégorie : Information

G. Armitage, Bellcore  
mars 1997  
Traduction Claude Brière de L'Isle

## Problèmes affectant la taille de grappe MARS

### Statut de ce mémoire

Le présent mémoire apporte des informations à la communauté de l'Internet. Il ne spécifie aucune sorte de norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

### Résumé

La diffusion groupée IP sur ATM (*Asynchronous Transfer Mode, mode de transfert asynchrone*) utilise actuellement le modèle de MARS (*Multicast Address Resolution Server, serveur de résolution d'adresse de diffusion groupée*) [RFC2022] pour gérer l'utilisation de SVC (*Switched Virtual Circuit, circuit virtuel commuté*) ATM en point à multipoint pour la transmission IP de paquets en diffusion groupée. La portée de tout service MARS est la grappe MARS – normalement la même qu'un sous-réseau logique IPv4 (LIS, *Logical IP Subnet*). Les réseaux IP/ATM actuels sont habituellement construits avec les considérations d'acheminement et transmission en envoi individuel qui dictent la taille des LIS individuels. Cependant, comme la diffusion groupée IP est développée comme un service, la taille d'un LIS ne sera jamais que celle que peut atteindre la grappe MARS. Le présent document porte un regard qualitatif sur les questions qui touchent aux contraintes pesant sur la taille d'une grappe MARS, y compris l'impact des limites de circuit virtuel (VC, *virtual circuit*) dans les commutateurs et les NIC, la distribution géographique des membres d'une grappe, et l'utilisation de maillages de VC ou de modes MCS (*multicast server, serveur de diffusion groupée*) pour prendre en charge les groupes de diffusion groupée.

## 1. Introduction

Une grappe MARS est l'ensemble des interfaces IP/ATM qui sont d'accord pour s'engager dans des circuits virtuels commutés (SVC) point à multipoint ATM directs pour effectuer la transmission de paquets IP en diffusion groupée [RFC2022]. Chaque interface IP/ATM (un client de MARS) doit conserver les informations d'état qui concernent les adresses ATM de chaque nœud d'extrémité (receveur) de chaque SVC en point à multipoint qu'il a ouvert. De plus, chaque client de MARS reçoit des messages MARS\_JOIN et MARS\_LEAVE de la part du MARS chaque fois qu'il est exigé que les clients autour de la grappe mettent à jour leurs SVC point à multipoint pour un groupe de diffusion groupée IP donné.

La définition de la "taille" de grappe peut signifier deux choses - le nombre de clients de MARS qui utilisent un MARS donné, et la distribution géographique des clients du MARS. Le nombre des clients du MARS dans une grappe a un impact sur la quantité d'informations d'état que tout client peut avoir besoin de mémoriser tout en gérant les SVC point à multipoint sortants. Il a aussi un impact sur le taux moyen de trafic JOIN/LEAVE qui est propagé par le MARS sur le circuit virtuel de contrôle de grappe (ClusterControlVC), et le nombre de VC en point à multipoint qui peuvent avoir besoin d'une modification chaque fois qu'un MARS\_JOIN ou MARS\_LEAVE apparaît sur le ClusterControlVC.

La distribution géographique des clients affecte la latence entre un client qui produit un MARS\_JOIN, et elle est finalement ajoutée aux VC en point à multipoint des autres clients de MARS qui transmettent au groupe de diffusion groupée spécifié. (Cette latence est constituée à la fois par le temps de propagation des MARS\_JOIN, et par le délai de la réaction dans le nuage ATM sous-jacent aux messages ADD\_PARTY suivants.)

Lors de la construction d'un réseau IP/ATM, il est important de comprendre le pire cas de limites d'adaptation applicables à vos grappes. Le présent document porte un regard principalement qualitatif aux choix conceptuels qui imposent les contraintes les plus sévères à la taille de grappe. Comme l'accent est mis sur les scénarios de plus mauvais cas, la plus grande partie de l'analyse supposera des groupes de diffusion groupée qui sont fondés sur des maillages de VC et ont tous les membres de la grappe comme sources et receveurs. L'ingénierie qui utilise les pires conditions de frontières, puis applique des optimisations comme les serveurs de diffusion groupée (MCS), donne à la grappe une marge de sécurité. On espère qu'une analyse quantitative plus détaillée des limites de taille de grappe sera accélérée par le présent document.

La Section 2 commente les exigences d'état de VC du modèle de MARS, puis les Sections 3 et 4 identifient la charge du traitement des changements de groupe et les caractéristiques de latence d'une grappe en fonction de sa taille. La Section 5 regarde comment les routeurs de diffusion groupée (aussi bien les architectures conventionnelles que les combinaisons de routeur/commutateur) augmentent l'adaptabilité d'un réseau IP/ATM à capacité de diffusion groupée. Finalement, la Section 6 expose comment l'utilisation de serveurs de diffusion groupée (MCS, *Multicast Server*) peut avoir un impact sur le pire cas de limites de grappe.

## 2. Limitations de l'état du circuit virtuel

Deux caractéristiques des NIC et commutateurs ATM vont limiter le nombre des membres qu'une grappe peut contenir. Ce sont :

Le nombre maximum de VC qui peuvent être générés, ou terminés, sur un accès (VCmax).

Le nombre maximum de nœuds d'extrémité qui peuvent être pris en charge par un nœud racine (LEAFmax).

Nous allons supposer que le nœud de MARS a des valeurs VCmax et LEAFmax similaires à celles des membres de la grappe. VCmax affecte la taille de la grappe à cause de ce qui suit :

Le MARS termine un VC de contrôle en point à point à partir de chaque membre de la grappe, et est à l'origine d'un VC pour ClusterControlVC et pour ServerControlVC.

Lorsque un groupe de diffusion groupée est fondé sur un maillage de VC, un membre du groupe termine un VC à partir de chaque envoyeur au groupe, par groupe.

Lorsque un groupe de diffusion groupée est fondé sur le MCS, le MCS termine un VC à partir de chaque envoyeur à ce groupe.

LEAFmax affecte la taille de la grappe à cause de ce qui suit :

ClusterControlVC provenant du MARS. Il a un nœud d'extrémité par membre de la grappe (client du MARS).

Les SVC de transmission de paquet sortant de chaque client du MARS pour chaque groupe de diffusion groupée IP auquel des envois sont faits. Il y a un nœud d'extrémité pour chaque membre de groupe lorsque un groupe est fondé sur un maillage de VC.

Les SVC de transmission de paquet sortant de chaque MCS pour chaque groupe de diffusion groupée IP auquel des envois sont faits. Il y a un nœud d'extrémité pour chaque membre du groupe quand un groupe est fondé sur un MCS.

Si nous avons N membres de la grappe, et M groupes actifs de diffusion groupé (utilisant le mode de maillage de VC, et de population dense – tous les receveurs sont des envoyeurs) on peut faire les observations suivantes :

Le ClusterControlVC a N nœuds d'extrémité, donc  $N \leq \text{LEAFmax}$ .

Le MARS termine un VC en point à point à partir de chaque membre de la grappe, et est à l'origine de ClusterControlVC et de ServerControlVC, de sorte que  $(N+2) \leq \text{VCmax}$ .

Chaque membre de la grappe est la source d'un VC par groupe, termine (N-1) VC par groupe, génère un VC en point à point avec le MARS, et termine un VC comme extrémité sur ClusterControlVC, de sorte que  $(M*N) + 2 \leq \text{VCmax}$ .

Le VC dont la source est chaque membre de la grappe par groupe va à tous les autres membres de la grappe, de sorte que  $(N-1) \leq \text{LEAFmax}$ .

Comme toutes les conditions ci-dessus doivent être simultanément réunies, on peut voir que l'exigence la plus contraignante est soit  $(M*N) + 2 \leq \text{VCmax}$ , soit  $N \leq \text{LEAFmax}$ .

La limite impliquant VCmax est fondamentalement contrôlée par la consommation de VC des membres du groupe qui utilisent un maillage de VC pour la transmission des données, plutôt que par la terminaison des VC de contrôle en point à point sur le MARS. (Cela va en pratique dépendre beaucoup de la distribution des membres du groupe de diffusion groupée au sein de la grappe.)

La limite LEAFmax vient de ClusterControlVC, et est indépendante de la densité des membres du groupe (ou des ratios d'envoyeurs par rapport aux receveurs) pour les groupes de diffusion groupée actifs au sein de la grappe.

Avec UNI 3.0/3.1 la limite la plus évidente sur LEAFmax est  $2^{15}$  (l'identifiant de nœud d'extrémité fait 15 bits). Cependant, le logiciel de pilote de signalisation pour la plupart des NIC ATM peut imposer une limite beaucoup plus basse que cela - en fonction de la quantité d'informations d'état qu'ils peuvent mémoriser pour chaque nœud d'extrémité (et qu'ils sont capables de mémoriser) pour les SVC en point à multipoint.

VCmax est contraint par le matériel de NIC ATM (par la segmentation disponible ou les instances de réassemblage) ou par la capacité en VC de l'accès de commutateur auquel le NIC est rattaché. VCmax sera le plus petit des deux.

Un client MARS peut imposer ses propres limitations de mémorisation d'état, de telle sorte que la consommation combinée de mémoire d'un client de MARS et du pilote du NIC ATM dans un hôte donné limite à la fois LEAFmax et VCmax aux valeurs inférieures que le NIC ATM seul pourrait avoir été capable de supporter.

Il serait possible de travailler sur les limites de LEAFmax en répartissant les nœuds d'extrémité entre plusieurs SVC point à multipoint fonctionnant en parallèle. Cependant, une telle approche exige d'autres études, et ne résout pas la limitation de VCmax associée à un nœud qui termine trop de VC.

Une observation voisine peut aussi être faite sur le nombre de clients de MARS dans une grappe qui peut être limitée par les contraintes de mémoire du MARS lui-même. Il est obligé de conserver l'état sur tous les groupes auxquels chacun de ses clients de MARS s'est joint. Pour une limite de mémoire donnée, le nombre maximum de clients de MARS doit chuter si le nombre moyen de groupes rejoints par client augmente. Selon le niveau d'adhésion aux groupes, cette limitation peut être plus sévère que LEAFmax.

### 3. Charge de la signalisation

Il va y avoir dans toute grappe un niveau "ambiant" d'activité de MARS\_JOIN/LEAVE. Les caractéristiques dynamiques de cette activité vont dépendre du type des applications de diffusion groupée qui fonctionnent dans la grappe. Pour une distribution relative constante des applications de diffusion groupée, on peut supposer que, lorsque augmente le nombre de clients de MARS d'une grappe donnée, le niveau ambiant d'activité de MARS\_JOIN/LEAVE augmente également. Cela augmente la fréquence moyenne avec laquelle le MARS traite et propage les messages MARS\_JOIN/LEAVE.

L'existence du trafic MARS\_JOIN/LEAVE a aussi par conséquent un impact sur l'activité de signalisation au niveau ATM (à travers les frontières de UNI et de {P}NNI). Pour les groupes qui sont pris en charge par un maillage de VC, chaque MARS\_JOIN ou MARS\_LEAVE propagé sur un ClusterControlVC va résulter en un message ADD\_PARTY ou DROP\_PARTY envoyé à travers les UNI de tous les clients de MARS qui transmettent à un groupe donné. Lorsque augmente le nombre des membres d'une grappe, le nombre moyen des clients de MARS qui déclenchent une activité de signalisation ATM en réponse aux MARS\_JOIN/LEAVE augmente aussi.

La taille d'une grappe doit être choisie de façon à fournir un certain niveau de maîtrise de ce niveau ambiant de signalisation de MARS et de UNI/NNI.

Certains raffinements du comportement du client de MARS peuvent aussi être explorés pour lisser les transitions de signalisation UNI. Il est actuellement demandé aux clients de MARS de n'initier la revalidation des adhésions au groupe que lorsque le client envoie un paquet à un SVC invalidé du groupe. Un client pourrait appliquer un algorithme similaire pour décider quand il devrait produire des ADD\_PARTY. Par exemple, après avoir vu un MARS\_JOIN, attendre qu'il ait effectivement un paquet à envoyer, envoyer le paquet, puis initier le ADD\_PARTY. Il en résulterait que les clients qui transmettent activement mettraient à jour leurs SVC plus tôt que les clients qui transmettent de façon intermittente.

### 4. Latences des changements du groupe

La latence de changement de groupe peut se définir comme le temps nécessaire pour que tous les envoyeurs à un groupe aient correctement mis à jour leurs SVC de transmission après la réception d'un MARS\_JOIN ou MARS\_LEAVE de la part du MARS. Ceci est affecté à la fois par le nombre de membres de la grappe et par la distribution géographique des membres de la grappe. (Les groupes qui sont fondés sur un MCS créent le plus faible impact lorsque de nouveaux membres adhèrent ou quittent, car seul le MCS a besoin de mettre à jour son SVC de transmission.) Dans certaines circonstances, en particulier des environnements de modélisation ou de simulation, les latences de changement de groupe au sein d'une grappe peuvent être une importante caractéristique à contrôler.

Comme noté à la section précédente, la charge de signalisation des ADD\_PARTY/DROP\_PARTY créés par les changements d'adhésion dans les groupes fondés sur le maillage de VC croît en parallèle de la croissance du nombre de membres de la grappe (en supposant le scénario du pire cas où chaque membre de la grappe est un envoyeur au groupe). Avec l'augmentation de la charge d'UNI, le réseau ATM lui-même peut commencer à fournir un traitement de livraison plus lent des événements requis.

Une distribution géographique large des membres de la grappe retarde aussi la propagation des messages MARS\_JOIN/LEAVE et ATM UNI/NNI. Plus les divers membres sont éparpillés, plus il leur prend de temps pour recevoir le trafic de MARS\_JOIN/LEAVE sur le ClusterControlVC, et plus il prend de temps au réseau ATM pour réagir aux demandes ADD\_PARTY et DROP\_PARTY. Si les chemins à longue distance sont peuplés de nombreux commutateurs

ATM, les délais de propagation dus au traitement par commutateur vont ajouter des délais substantiels dus à la vitesse de la lumière.

(Malheureusement, les mécanismes pour lisser la charge de la signalisation ATM transitoire décrits à la section 3 ont pour conséquence d'augmenter la latence du changement de groupe, car le but pour certains envoyeurs est de retarder délibérément la mise à jour de leurs SVC de transmission. C'est un domaine où l'architecte de système doit faire des compromis en fonction de la situation.)

L'effet du traitement interne du MARS lui-même sur la latence des changements du groupe n'apparaît pas de façon évidente, pas plus que la façon dont elle pourrait être impactée par la taille de la grappe. Un composant de la latence du traitement du MARS va dépendre de la mise en œuvre spécifique de la base de données et des algorithmes de recherche tout autant que du nombre de membres du groupe pour le groupe qui est modifié à un instant donné. Comme le nombre maximum de membres d'un groupe pour un groupe donné est égal au nombre de membres de la grappe, il y aura une relation indirecte (même si elle est petite) entre le pire cas de latence du traitement du MARS et la taille de la grappe.

## 5. Grands réseaux IP/ATM utilisant des Mrouteurs

La construction d'un réseau IP sur ATM à capacité de diffusion groupée à grande échelle est un compromis entre les tailles de grappe et le nombre des Mrouteurs. Pour un nombre d'hôtes donné, le nombre de grappes augmente à mesure que les grappes individuelles rétrécissent. Comme les Mrouteurs sont les intersections topologiques entre les grappes, le nombre de Mrouteurs augmente à mesure que la taille des grappes individuelles diminue. (Le nombre réel de Mrouteurs dépend largement de la topologie logique IP qu'on choisit de mettre en œuvre, car un seul Mrouteur physique peut interconnecter plus de deux grappes à la fois.) C'est une question de déploiement local que de déterminer le mélange optimal de grappes et de Mrouteurs.

Actuellement, deux grandes classes de Mrouteurs peuvent être identifiées :

- Ceux qui génèrent des VC uniques dans des grappes cibles, et transmettent/entrelacent des données au niveau du paquet IP (c'est le Mrouteur conventionnel).

- Ceux qui génèrent des VC uniques dans des grappes cibles, mais créent des "raccourcis" internes au niveau de la cellule entre les VC provenant de différentes grappes (par exemple, le routeur de commutation de cellules (*Cell Switch Router*)).

Comment ces Mrouteurs établissent et gèrent les associations de VC et des flux de trafic IP sort du domaine d'application du présent document. Cependant, on va regarder brièvement à leur impact sur la consommation de VC et la charge de signalisation ATM.

### 5.1 Impact du Mrouteur conventionnel

Un Mrouteur conventionnel agit comme un point d'agrégation pour la signalisation et pour les charges du plan des données. Il cache les changements d'adhésion à un groupe spécifiques d'un hôte dans une grappe aux envoyeurs qui sont dans les autres grappes, et protège les membres d'un groupe (receveurs) au sein d'une grappe d'avoir à être les nœuds d'extrémité sur les SVC provenant d'envoyeurs d'autres grappes.

En agissant comme point d'entrée dans une grappe, un Mrouteur conventionnel établit un seul SVC de transmission pour les paquets IP. Ce SVC unique porte les données pour les autres grappes entrelacées au niveau du paquet IP. Ce seul SVC unique doit être modifié en réponse aux changements d'adhésion aux groupes au sein de la grappe cible. Par conséquent, il n'est pas nécessaire que les sources dans les autres grappes soient au courant, ou réagissent au trafic de MARS\_JOIN/LEAVE dans la grappe cible. (La charge de signalisation UNI résultante, identifiée à la section 3 est aussi localisée dans la grappe cible.)

Les clients de MARS au sein de la grappe cible bénéficient aussi de cette agrégation des chemins de données parce qu'ils ne terminent qu'un seul SVC provenant du Mrouteur (par groupe) plutôt que plusieurs SVC générés par les envoyeurs réels dans les autres grappes.

Les Mrouteurs conventionnels aident à contrôler les facteurs limitants décrits aux sections 2, 3, et 4. Une grappe hypothétique de 10 000 nœuds pourrait être coupée en deux grappes de 5000 nœuds, ou quatre grappes de 2500 nœuds, etc., pour réduire la consommation de VC. Ou on peut avoir 200 nœuds sur le total de 10 000 qui sont connus pour se joindre aux groupes et les quitter rapidement, alors que les 9800 autres sont très stables, de sorte qu'on peut déployer des grappes de, respectivement, 200, 2500, 2500, 2500, 2300 hôtes.

## 5.2 Impact du routeur à commutation de cellules (CSR)

Une autre classe de Mrouteur, le routeur de commutation de cellules (CSR, *Cell Switch Router*) tente d'utiliser les informations de flux de niveau IP pour gérer de façon dynamique la commutation des données à travers l'appareil en dessous du niveau IP. Une fois que le CSR a identifié un flux de trafic IP, et l'a associé à un SVC entrant et sortant, il commence à fonctionner comme un appareil de niveau cellule ATM plutôt que comme un appareil de niveau paquet.

Même lorsque il fonctionne dans ce mode, le CSR isole les grappes rattachées des activités MARS\_JOIN/LEAVE de chaque autre, de la même manière qu'un Mrouteur conventionnel. Cela se fait parce que le CSR gère ses SVC de transmission tout comme un client de MARS normal – répondant aux messages MARS\_JOIN/LEAVE au sein de la grappe cible en mettant à jour les arborescences point à multipoint dont la racine est sur ses propres accès ATM.

Cependant, comme les AAL\_SDU AAL5 ne peuvent pas être entrelacés au niveau de la cellule sur un seul SVC, un CSR ne peut pas effectuer simultanément de raccourcis au niveau de la cellule et agréger les flux de paquets IP provenant de plusieurs envoyeurs sur un seul SVC dans une grappe cible. Il en résulte que le CSR doit construire un SVC de transmission distinct dans la grappe cible pour chaque SVC qui est une extrémité dans une grappe source (pour s'assurer que les cellules provenant des sources individuelles ne sont pas entrelacés avant d'atteindre les moteurs de réassemblage des membres de groupe dans la grappe cible).

Il est intéressant de noter que la charge de signalisation UNI offerte au sein de la grappe cible par le CSR est potentiellement supérieure à celle d'un Mrouteur conventionnel. Si il y a N envoyeurs dans la grappe source, le CSR aura construit N SVC point à multipoint identiques vers les membres du groupe au sein de la grappe cible. Si un nouveau MARS\_JOIN est produit au sein de la grappe cible, le CSR doit produire N ADD\_PARTY pour mettre à jour les N SVC dans la grappe cible. (Dans des circonstances similaires, un Mrouteur conventionnel aurait produit seulement un ADD\_PARTY pour son seul SVC dans la grappe cible.)

Donc, sans la capacité à fournir de raccourcis de transmission internes avec les frontières de AAL\_SDU intactes, le CSR ne fournit que l'isolation du trafic de MARS\_JOIN/LEAVE au sein des grappes. Il ne peut pas fournir l'agrégation de chemins des données d'un Mrouteur conventionnel.

## 6. Impact des serveurs de diffusion groupée (MCS)

Comme le présent document se concentre sur les scénarios de plus mauvais cas, la plus grande partie de l'analyse a supposé que les groupes de diffusion groupée sont fondés sur un maillage de VC et que tous les membres de toutes les grappes sont des sources et des receveurs. L'impact de l'utilisation d'un MCS pour prendre en charge un groupe de diffusion groupée peut être dramatique dans le contexte de la consommation des ressources par le groupe, mais l'est beaucoup moins dans le contexte global des limites de taille de grappe.

L'impact intra-grappe, par groupe, d'un MCS est assez analogue à l'impact inter-grappe d'un Mrouteur conventionnel. Le MCS agrège les flux de données (un seul SVC se termine sur chaque membre de groupe, indépendamment du nombre d'envoyeurs) et isole le trafic MARS\_JOIN/LEAVE (qui est passé au ServerControlVC plutôt qu'au ClusterControlVC). Le trafic et la charge de signalisation UNI résultants sont aussi réduits, car seul le SVC qui transmet en sortie du MCS a besoin d'être modifié pour chaque changement d'adhésion dans le groupe pris en charge par le MCS.

Le déploiement d'un mélange de groupes fondés sur un mélange de MCS et d'un maillage de VC va certainement améliorer l'utilisation des ressources. Cependant, la portée réelle de l'amélioration (et par conséquent l'extension possible de la taille de la grappe) va dépendre en grande partie de la dynamique des applications principales et des caractéristiques tirées des sections 2, 3, et 4 qui sont les principales limitations.

Par exemple, si VCmax ou LEAFmax (section 2) sont les principales limitations, on doit se souvenir que chaque MCS souffre lui-même des mêmes limites de NIC que le MARS et les clients de MARS. Même si l'utilisation d'un MCS réduit considérablement le nombre de VC par client de MARS et par groupe, chaque MCS a quand même besoin de terminer un SVC par envoyeur – ce qui donne un potentiel de un SVC provenant de chaque membre de la grappe. (Cela peut devenir un SVC par membre et par groupe si le MCS prend plusieurs groupes en charge simultanément.)

Supposons que nous ayons une grappe où tous les groupes sont fondés sur un MCS, que chaque MCS ne prenne en charge qu'un seul groupe, et que VCmax et LEAFmax s'appliquent tous deux également aux nœuds de MCS comme MARS et aux nœuds de clients de MARS. Si nous avons N membres de grappe, M groupes, et si tous les receveurs sont des envoyeurs pour un groupe pris en charge par un MCS donné, les observations suivantes peuvent être faites :

Chaque SVC de transmission d'un MCS a N nœuds d'extrémité, de sorte que  $N \leq \text{LEAFmax}$ .

Chaque MCS termine un SVC provenant de N envoyeurs, génère un SVC de chemin de transmission, génère un SVC de contrôle en point à point avec le MARS, et termine un SVC comme extrémité sur le ServerControlVC, donc  $N + 3 \leq VC_{max}$ .

Le ClusterControlVC du MARS a N nœuds d'extrémité, de sorte que  $N \leq LEAF_{max}$ .

Le ServerControlVC du MARS a M nœuds d'extrémité, de sorte que  $M \leq LEAF_{max}$ .

Le MARS termine un VC en point à point provenant de chaque membre de la grappe, un VC en point à point provenant de chaque MCS, génère un ClusterControlVC, et génère un ServerControlVC, de sorte que  $N + M + 2 \leq VC_{max}$ .

Chaque membre de grappe est la source d'un VC par groupe, termine un VC par groupe, génère un VC en point à point pour le MARS, et termine un VC comme extrémité sur le ClusterControlVC, de sorte que  $2 * M + 2 \leq VC_{max}$ .

Comme toutes les conditions ci-dessus doivent être réunies simultanément, on peut voir que les exigences les plus contraignantes sont :

$$N + M + 2 \leq VC_{max} \text{ (if } M \leq N \text{)}$$

$$2 * M + 2 \leq VC_{max} \text{ (if } M \geq N \text{)}$$

ou

$$N \leq LEAF_{max}.$$

(En supposant qu'en général  $M+2 > 3$ , la contrainte  $VC_{max}$  à chaque MCS n'est donc pas un facteur limitant.)

On peut avoir une idée des impacts relatifs des groupes à maillage de VC par rapport aux groupes fondés sur le MCS en considérant une grappe où  $M1$  représente le nombre de groupes fondés sur le maillage de VC, et  $M2$  représente le nombre de groupes fondés sur le MCS. Là encore, nous supposons le pire cas de densité de groupe (tous les N membres de la grappe sont receveurs et envoyeurs).

Comme noté à la section 2, la contrainte  $VC_{max}$  en mode maillage de VC vient de chaque client de MARS, et est :

$$N * M1 \leq VC_{max} - 2$$

Pour le cas du MCS, nous avons deux scénarios,  $M2 \leq N$  et  $M2 \geq N$ .

Si  $M2 \leq N$ , on peut voir la consommation de VC par les groupes fondés sur le maillage de VC qui vont devenir la contrainte applicable sur la taille de grappe N lorsque  $N + M2 \leq N * M1$ , c'est-à-dire,  $M1 \leq 1 + (M2/N)$ .

Donc si il y a plus d'un groupe fondé sur un maillage de VC, et moins de groupes fondés sur le MCS que de membres de la grappe ( $M2 < N$ ) la contrainte sur la taille de grappe est dictée par les caractéristiques du maillage de VC :  $N * M1 \leq VC_{max} - 2$ . (Si  $M2 = N$ , il peut alors y avoir deux groupes fondés sur le maillage de 2 avant que les caractéristiques de maillage de VC soient le facteur déterminant.)

Maintenant, si  $M2 > N$  (plus de groupes fondés sur le MCS que de membres de la grappe) le calcul est plus complexe car dans ce cas,  $VC_{max}$  chez le client de MARS est le paramètre limitant pour les deux cas de maillage de VC et de MCS. La limite devient :

$$N * M1 + 2 * M2 \leq VC_{max} - 2$$

Cependant, cette situation est assez étrange car elle implique qu'il y a plus d'entités de MCS que d'interfaces d'hôtes ou de routeur dans la grappe (compte tenu de l'hypothèse d'un groupe par MCS).

L'impact des entités de MCS qui prennent simultanément en charge plusieurs groupes fera l'objet d'une autre étude.

## 7. Questions non résolues

Une large gamme d'analyses qualitatives peut être extraites des scénarios typiques de déploiement de MARS. Le présent document n'essaye pas de développer de modèles numériques de la consommation de VC, des latences de bout en bout, etc.

## 8. Conclusion

Le présent document a présenté une vue générale qualitative des paramètres qui affectent la taille des grappes de MARS. Les limitations du nombre de nœuds d'extrémité qu'un SVC en point à multipoint peut prendre en charge, la taille de la base de données du MARS, les délais de propagation des messages du MARS et de l'UNI, et la fréquence des messages de contrôle du MARS et de l'UNI sont tous identifiés comme des problèmes qui vont exercer des contraintes sur les grappes. Les Mrouteurs conventionnels sont identifiés comme des agrégateurs utiles du trafic de diffusion groupée IP et d'informations de signalisation. Les routeurs de commutation de cellules sont notés comme n'offrant que certains des attributs d'agrégation des Mrouteurs conventionnels. La diffusion groupée IP sur ATM à grande échelle exige une combinaison de Mrouteurs et de grappes de MARS dimensionnées de façon appropriée. Finalement, on a montré que dans une simple grappe où il y a moins de groupes fondés sur MCS que de membres de la grappe, deux groupes fondés sur un maillage de VC, ou plus, sont suffisants pour rendre l'utilisation de serveurs de diffusion groupée inappropriée pour le pire cas de limite de taille de grappe.

## Considérations pour la sécurité

Les questions de sécurité ne sont pas abordées dans le présent mémoire.

## Remerciements

Des remerciements sont adressés à Rajesh Talpade (Georgia Tech) pour son apport spécifique sur les aspects du compromis entre maillage de VC et MCS, et à Joel Halpern (Newbridge) pour ses apports généraux sur l'objet du document.

## Adresse de l'auteur

Grenville Armitage  
Bellcore, 445 South Street  
Morristown, NJ, 07960  
USA  
mél : [gja@thumper.bellcore.com](mailto:gja@thumper.bellcore.com)  
téléphone +1 201 829 2635

## Références

[RFC2022] G. Armitage, "Prise en charge de la [diffusion groupée sur réseaux ATM](#) fondés sur UNI 3.0/3.1", novembre 1996. (P.S.)