

Le multicast

(Z:\Polys\Multicast\Le_multicast-intro.fm- 12 décembre 2018 17:09)

PLAN

- Généralités
- L'adressage multicast
- Les arbres multicasts
- Construction d'arbres multicasts
- Conclusion

Bibliographie

- S. Paul. Multicasting on the Internet. Kluwer academic publishers, 1998
- C. Huitema. Le routage dans l'Internet. Eyrolles, 1995
- W. Stallings. High Speed Networks. Prentice Hall, 1998
- C. Comer. TCP/IP: architectures, protocoles, applications. InterEditions, 1998
- R. Wittmann, M. Zitterbart. Multicast Communication : Protocols and Applications. Morgan Kaufmann publishers, 2001

1. Généralités

1.1. Présentation

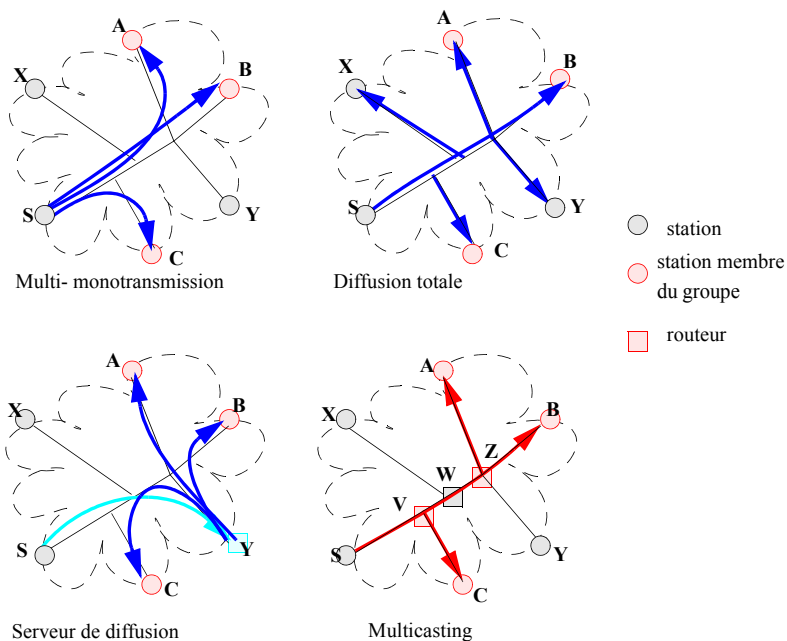
De nombreuses applications envoient les mêmes informations à **plusieurs** destinataires :

- Téléconférence : audioconf., vidéoconf., etc.
. ex : CU-SeeMe, IVS, Netmeeting
- Outils collaboratifs : "whiteboard", éditeur partagé, etc.
- Radio-diffusion, web phone, etc.
- Distribution d'information : messagerie électronique, applications réparties, etc.

Bâtir le réseau de distribution :

- soit au niveau **applicatif** :
 - . nombreuses recopies, cheminement sous-optimal, à base de liaisons bi-points
- ex : "reflectors" de CU-SeeMe, la diffusion en "Peer to Peer". Pb : le stress des liens
- soit au niveau du **réseau** de transmission :
 - . notion de liaison multipoint
 - . le réseau connaît le meilleur chemin !
 - . bonne efficacité

Quelques techniques de diffusion :



1.2. Les différents types de protocoles multicasts

Les protocoles de **routage multicast**

- au niveau de la couche Réseau
- comment acheminer les données multicasts à travers les routeurs du réseau ?
- ex : (IGMP), DVMRP, MOSPF, CBT, PIM, etc.

Les protocoles chargés de la **transmission des données multicasts**

- au niveau de la couche Transport
- comment s'assurer que les données ont été bien transmises ?
 - contrôle de flux et récupération des erreurs
- **EX** : UDP !, ST2 (Stream Transport v2), XTP (eXpress Transport Protocol), MTP (Multicast Transport Protocol), RMP (Reliable Multicast Protocol), SRP (Scalable Reliable Protocol), RMTP (Reliable Multicast Transport Protocol), RAMP (Reliable Adapative Multicast Protocol), etc.

Autres protocoles :

- les protocoles chargés de la gestion des groupes :
 - eX : SAP (Session Announcement Protocol) , SIP (Session Initiation Protocol) utilise HTML, SDP (Session Description Protocol), etc
- les protocoles chargés de la QoS
 - eX : RSVP (ReSerVation Protocol), RTP & RTCP (Real Time & Control Protocol), etc
- les protocoles applicatifs
 - eX : MFTP (Multicast File Transfer Protocol) : MCP + MDP (Control + Data)
- les protocoles chargés de la gestion des adresses multicasts :
 - eX : MSDP (Multicast Source Distribution Protocol)

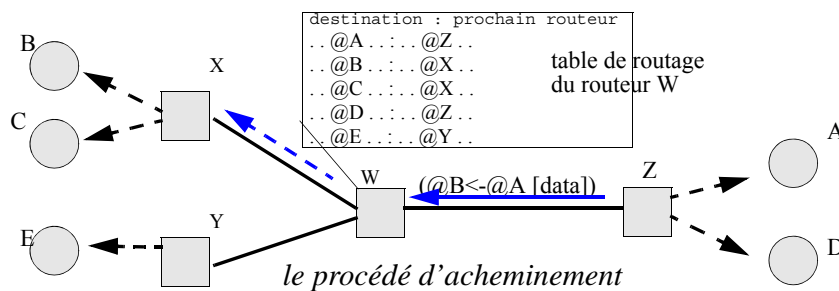
Quelques applications ou utilitaires :

- SDR (Session Directory) utilise les protocoles SAP+SDP,
- Mouted (multicast routing daemon),
- Mrinfo : affiche la configuration d'un routeur multicast,
- Mtrace : donne des statistiques sur la route suivie pour atteindre une destination

=> les protocoles de routage multicast

1.3. Le routage unicast

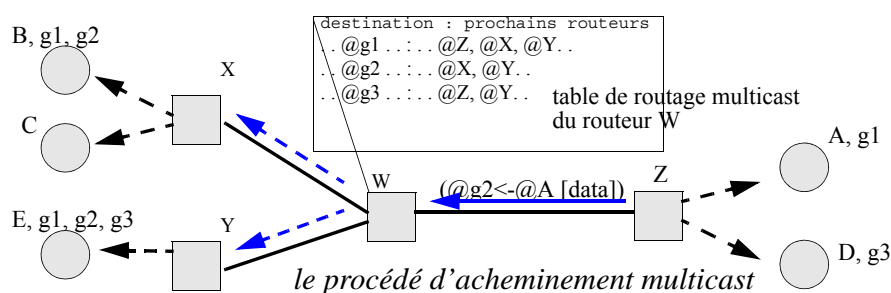
- le procédé d'acheminement ("datagram forwarding")
 - la re-transmission des datagrammes IP.
 - . choix du prochain routeur ("next-hop")



- le procédé d'échange d'information permettant la mise en oeuvre de l'acheminement
 - les protocoles de routage
 - . mise à jour des tables de routage

1.4. Routage multicast

- le procédé d'acheminement multicast (“**multicast forwarding**”)
 - la re-transmission des datagrammes multicast d'IP.
 - . choix des prochains routeurs (“next-hop”)



- le procédé d'échange d'information permettant la mise en oeuvre de l'acheminement
 - les **protocoles de routage multicast**
 - . mise à jour des tables de routage multicast

1.5. Le facteur d'échelle

Les routeurs doivent traiter des milliers de paquets par seconde

- optimisation de l'algorithme d'acheminement
- sa complexité est fonction de la taille de la table de routage
 - algorithme de recherche : $O(n \log(n))$
 - n : nombre de destinations dans le réseau --> potentiellement des milliards !

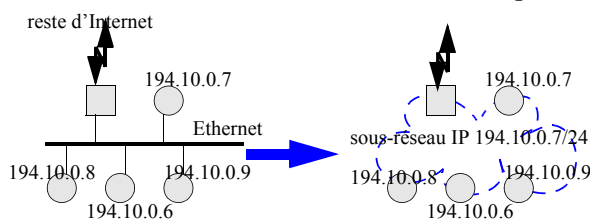
Solutions :

- Agrégation
- Hiérachisation

1.6. Agrégation

- Sous-réseaux IP

- . toutes les stations ayant même préfixe d'adresse forment un sous-réseau IP identifié par ce préfixe ("netid")
- . par exemple : toutes les stations d'un LAN Ethernet peuvent former un sous-réseau IP



- CIDR

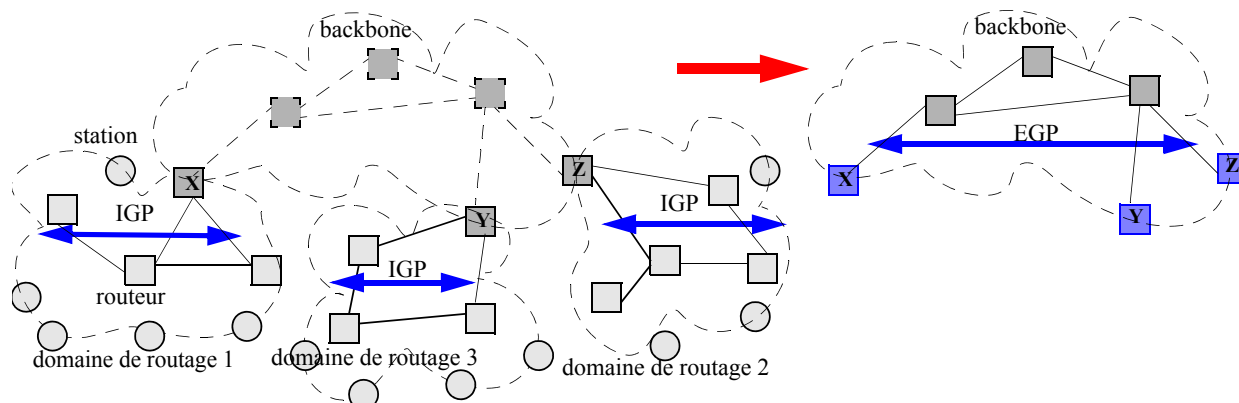
- extension de la notion précédente
- une seule entrée de la table de routage pour toutes les destinations ayant des adresses consécutives et même prochain routeur
 - . par ex : préfixe européen $194.0.0.0/7 = [194.0.0.0, 195.255.255.255]$

⇒ inapplicable pour l'acheminement multicast : pas de localité !

1.7. Les niveaux hiérarchiques de routage

On discerne 2 niveaux hiérarchiques de routage :

- les protocoles de routage IGP régissent les informations de routage échangées **entre routeurs** au sein d'un domaine de routage (**interne**).
 - les routeurs d'un domaine connaissent toutes les destinations du domaine
 - par exemple : RIP, OSPF
- les protocoles de routage EGP régissent les informations de routage échangées **entre routeurs** entre domaines de routage (**externe**, sur le "backbone").
 - les routeurs du backbone ne connaissent que les domaines de routage
 - par exemple : BGP, EGP



⇒ Cette hiérarchie favorise la scalabilité du réseau.

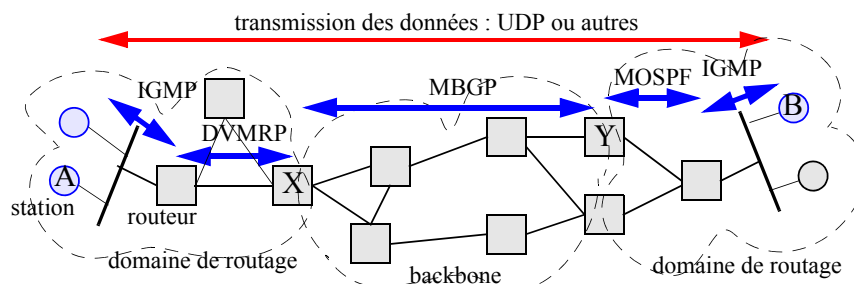
Nota : certains protocoles introduisent un niveau hiérarchique intermédiaire

- “routing area” : sous-domaine de routage
- par ex. OSPF

1.8. Les niveaux hiérarchiques de routage multicast

On discerne 3 niveaux de routage multicast :

- le protocole IGMP régit les informations de localisation des stations appartenant à un groupe, informations échangées **entre les stations et les routeurs** multicasts.
- les protocoles de routage multicast (par exemple : DVMRP ou MOSPF) régissent les informations de routage échangées **entre routeurs** au sein d’un domaine de routage (**interne**).
- les protocoles de routage multicast (par exemple : MBGP) régissent les informations de routage échangées **entre routeurs** multicasts entre domaines de routage (**externe**, sur le “backbone”).



⇒ Cette hiérarchie favorise la scalabilité du réseau.

2. L'adressage multicast

2.1. Introduction

Une adresse multicast d'Internet **identifie un groupe** (de stations)

- Adresse IP de classe D, de préfixe “1110₂”,
- Adresse comprise entre 224.0.0.0 et 239.255.255.255

Réception :

- Toutes les stations possédant cette adresse multicast reçoivent une copie de chaque paquet ayant pour adresse de destination cette adresse multicast.

Emission :

- Pour émettre un datagramme multicast une station n'a pas besoin d'appartenir au groupe.

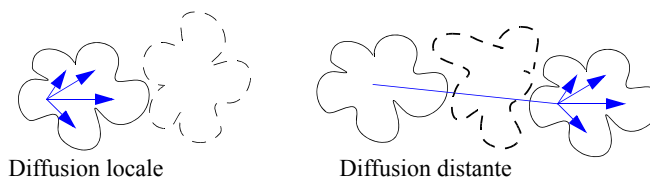
Attention :

- Utilisation interdite des adresses multicasts comme adresse d'émission.
- Contrôle des règles de gestion du groupe : hors du domaine d'IP.
- Ne pas faire l'hypothèse que les datagrammes multicasts émis ne sont uniquement reçus que par les récepteurs envisagés. Le partage de la même adresse multicast par plusieurs applications est possible et doit être résolu au niveau applicatif.

2.2. Adresses multicasts réservées

2.2.1 Diffusion

- Plein “1” !
 - **Diffusion locale** : 255.255.255.255
 - . diffusion à toutes les stations du réseau IP local (à l'émetteur)
 - **Diffusion distante** : netid-A.255.255.255, netid-B.255.255, netid-C.255
 - . diffusion à toutes les stations sur réseau IP distant identifié par “netid”



Note : dû à une implémentation historique, les adresses comportant l'hostid à 0 peuvent être interprétées comme requérant la diffusion.

2.2.2 Les adresses de groupes particuliers

Chaque adresse identifie un ensemble de stations (routeurs) rendant un service particulier

- Optimisation :
 - supprime les interruptions intempestives du processeur lorsque la station reçoit des messages qui ne lui sont pas destinés

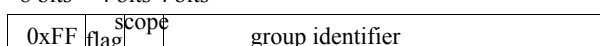
Exemples d'adresses de groupes particuliers :

- 224.0.0.1 : toutes les machines (stations + routeurs) appartenant à un groupe quelconque du réseau IP local
- 224.0.0.2 : tous les routeurs du réseau IP local
- 224.0.0.9 : tous les routeurs RIP-2 appartenant au réseau IP local
- etc.

2.3. Adressage multicast IPv6

L'adressage IPv6 offre un type d'adresse multicast (**préfixe** spécifique : FF_{16}) :

IPv6 address : 16 octets
 8 bits 4 bits 4 bits 112 bits



- Champ **Flag** [4 bits] :
 - adresse affectée de manière permanente ou non
 - . T=0 : attribution permanente de l'adresse multicast
 - . T=1 : attribution non-permanente de l'adresse multicast
- Champ **Scope** [4 bits]
 - 1 : étendue restreinte au noeud
 - 2 : étendue restreinte à la liaison
 - 5 : étendue restreinte au site
 - 8 : étendue restreinte à l'organisme
 - E : étendue mondiale
 - . 0, F : réservé; 3, 4, 6, 7, 9, A, B, C, D: non utilisé
 - . le champ TTL était utilisé à cette fin par IPv4.
- Champ **Group identifier** [14 octets]

- Exemples d'adresses multicasts réservées :

. Aucune station n'appartient à ce groupe réservé :

- FF0s::0

. Tous les noeuds (stations + routeurs) :

- format général : FF0s::1 avec $s=\{1, 2, 5, 8, E\}$

- exemple : tous les noeuds de l'organisme= FF08::1

. Toutes les stations :

- format général : FF0s::2 avec $s=\{1, 2, 5, 8, E\}$

- exemple : toutes les interfaces du noeud (de type station)= FF01::2

. Tous les routeurs :

- format général : FF0s::3 avec $s=\{1, 2, 5, 8, E\}$

- exemple : tous les routeurs du site = FF05::3

. Tous les serveurs NTP :

- format général : FF0s::43 avec $s=\{1, 2, 5, 8, E\}$

But

. Optimisation des algorithmes de découverte :

- minimisation de l'exploration.

3. Arbres multicasts

Le routage multicast s'appuie sur la présence d'**arbres de diffusion** (ou arbre multicast) :

- à chaque groupe multicast on associe un arbre de diffusion
- la station émettrice du paquet multicast forme (souvent) la racine de l'arbre
- les stations membres du groupe multicast forment (souvent) les feuilles de l'arbre et sont les destinataires du paquet multicast
- certains routeurs forment les noeuds d'embranchement de l'arbre

Construction d'un **arbre multicast optimum** :

- arbre multicast **des plus courts chemins** ("shortest path tree")
- arbre multicast **de coût optimum**

3.1. Arbre multicast de coût optimum

Arbre optimum couvrant l'ensemble des membres du groupe : arbre couvrant total

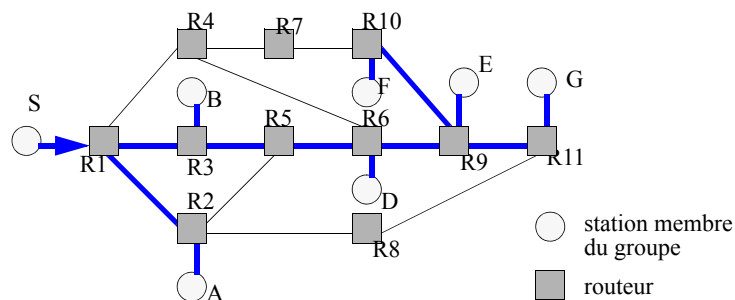
- si tous les noeuds du graphe appartiennent au groupe :
 - ⇒ Arbre recouvrant optimum (“Optimum spanning tree”)
- certains noeuds du graphe n'appartiennent pas au groupe : arbre couvrant partiel
 - ⇒ Arbre de Steiner (“Optimum Steiner tree”)
 - certains noeuds de l'arbre peuvent ne pas appartenir au groupe

Optimisation globale des ressources du réseau utilisées pour la diffusion.

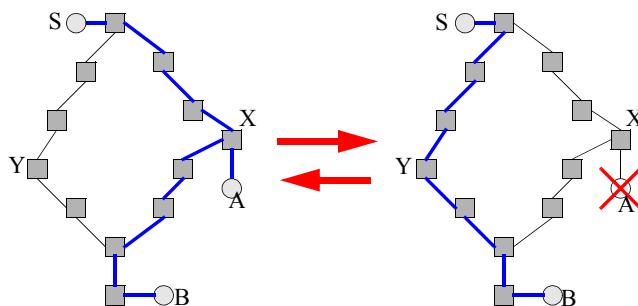
Complexité des algorithmes de construction d'arbres partiels optimaux :

- problème “NP-complet” !

Exemple : fonction de coût unitaire, coût total = 7+7, longueur maximum = 5+2



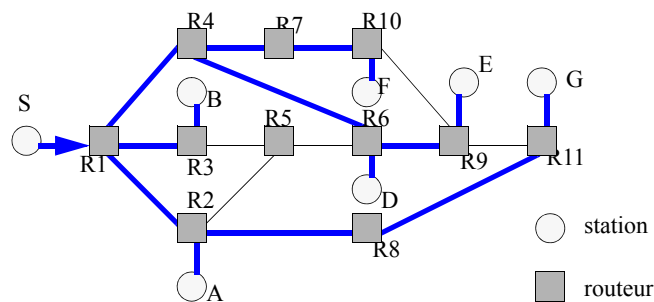
Instabilité de l'arbre de Steiner :



3.2. Arbre multicast des chemins les plus courts

SPT ("shortest path tree") :

- le paquet multicast emprunte le chemin le plus court entre l'émetteur et chaque destinataire
- **minimise le délai** (favorise les besoins des applications)
- on utilise l'algorithme de Dijkstra ("le plus court chemin d'abord")
- Exemple : fonction de coût unitaire, coût = 9+7, longueur max = 3+2



⇒ technique utilisée par les protocoles de routage des réseaux actuels

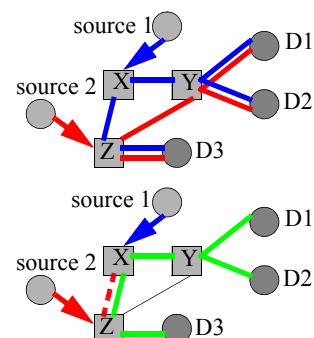
3.3. Partage des arbres multicasts

Partage de l'arbre multicast entre les groupes :

- un même arbre pour tous les groupes multicasts
 - arbre de recouvrement total
 - ex : "Spanning tree" pour l'interconnexion des réseaux locaux
 - inconvénient : cheminement non optimal, accumulation de trafic vers la racine
- un arbre par groupe
 - ex : en général, le multicast d'Internet
 - inconvénient : un grand nombre d'arbres

Partage de l'arbre multicast au sein d'un groupe :

- un groupe peut avoir plusieurs émetteurs dispersés
- "source based multicast tree" : $\langle S, G \rangle$
 - précis mais coûteux
- "shared multicast tree" : $\langle *, G \rangle$
 - rentable mais chemin non-optimal

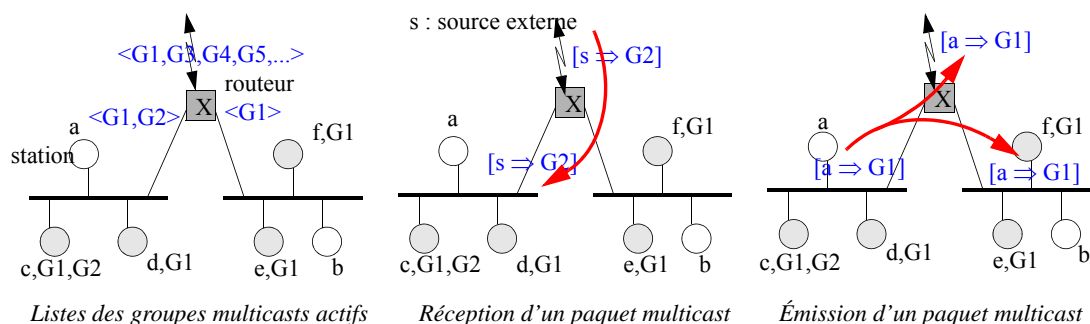


3.4. Routage local des datagrammes multicasts

Un routeur recevant un datagramme muni d'une adresse multicast identifiant un groupe le diffuse sur tous les interfaces où le groupe est actif (sauf celui d'où provient le datagramme).

Un routeur multicast n'a besoin de ne connaître que l'activité d'un groupe. Il ne mémorise que cette seule information.

- chaque routeur pour chacune de ses interfaces maintient la liste des groupes actifs grâce à IGMP.



4. Construction d'arbres multicasts

La construction d'un arbre de diffusion multicast est complexe et coûteuse.

Quelques solutions :

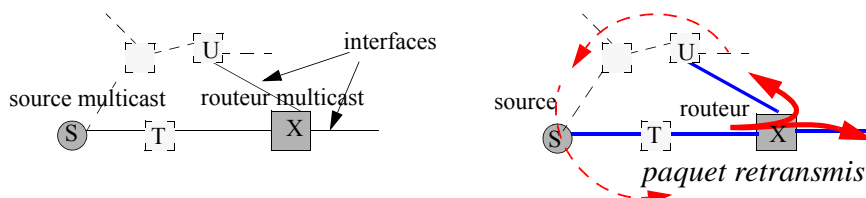
- utilisation de la connaissance topologique du réseau
 - information issue naturellement des protocoles de routage unicast de type "link state algorithm"
 - ex. : OSPF \Rightarrow MOSPF
- (ré-)utilisation des informations de routage unicast :
 - information déjà présente dans les tables de routage
 - 2 grandes techniques :
 - . "Reverse Path Forwarding"
 - . "Core Base Tree"

\Rightarrow RPM (\Rightarrow DVMRP)

4.1. L'inondation

• diffusion **par inondation**

- chaque paquet reçu est retransmis vers toutes les autres interfaces (hormis celle d'origine)
 - . simple, fiable, optimal (!)
 - . indépendant des destinataires
 - . encombrement, boucles et copies multiples !
 - . non-sélectif !



4.2. La méthode "Reverse Path Forwarding"

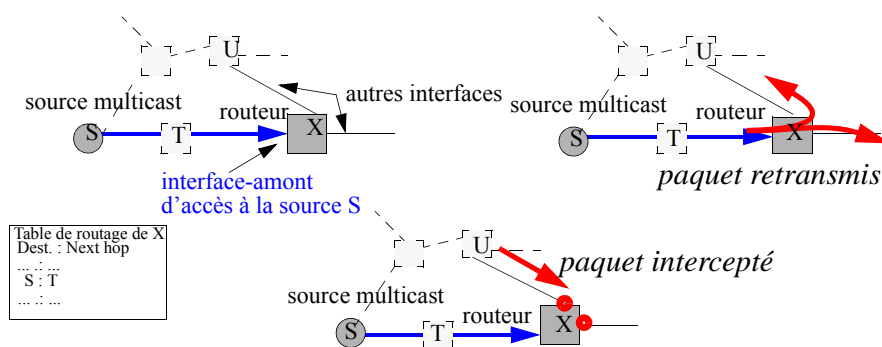
But : suppression des boucles générées par la diffusion par inondation.

Principe :

- chaque **paquet reçu sur l'interface-amont** (qui offre l'unique accès vers la source) **est retransmis vers toutes les autres interfaces**.
- les paquets reçus sur une autre interface sont détruits.

Pour chaque source une interface du routeur est distinguée :

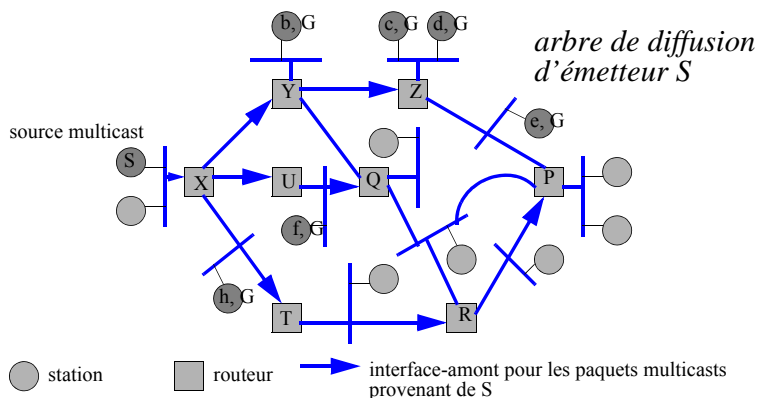
- celle désignée (par le protocole de routage unicast) comme l'interface donnant accès à la source par le meilleur chemin.



4.3. L'arbre de diffusion

Si chacun des routeurs du réseau traite les paquets reçus en appliquant le RPF, on obtient :

- un algorithme de **diffusion** sans rebouclage : le RPB
- sans état supplémentaire, par rapport aux routeurs unicasts
- tous les routeurs, tous les réseaux locaux, toutes les stations reçoivent (au moins) une copie du paquet



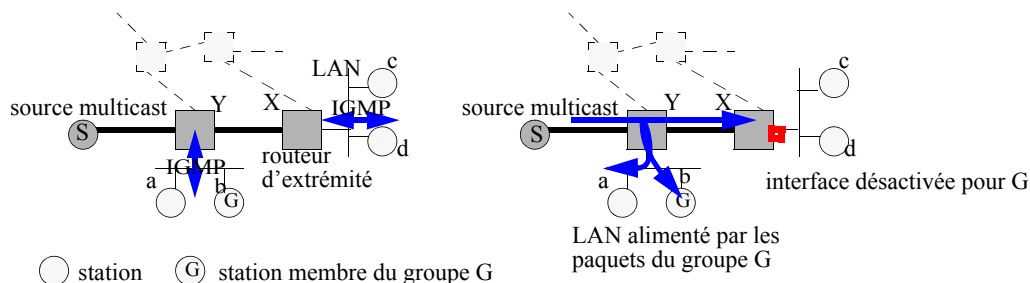
⇒ “Reverse Path Broadcasting”

4.4. Troncature de l'arbre de diffusion

But : supprimer de l'arbre de diffusion les réseaux locaux où aucune station n'est abonné au groupe.

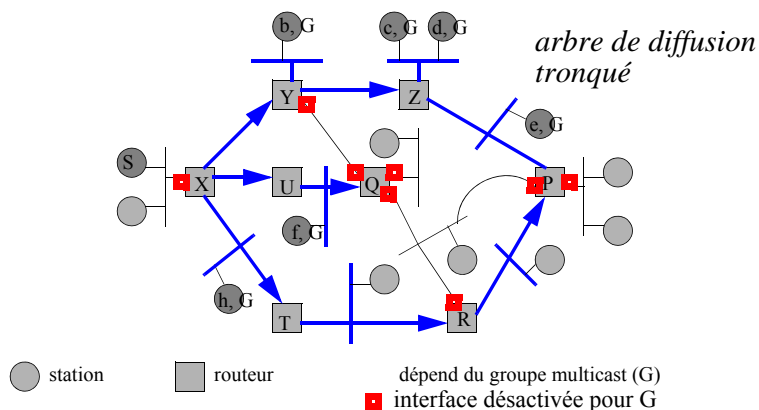
Grâce à la connaissance de l'absence d'activité d'un groupe sur chacune de ces interfaces, un routeur d'extrémité peut en **déduire l'inutilité de transmettre les paquets** de ce groupe, et donc désactiver cette interface pour ce groupe multicast.

- Le protocole IGMP permet à un routeur d'extrémité d'avoir cette connaissance.



Traitement similaire pour les simples liens (n'hébergeant pas de stations) et n'appartenant pas à l'arbre de diffusion.

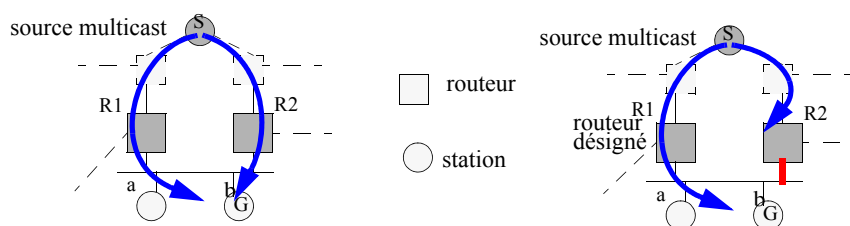
4.5. Arbre de diffusion tronqué



4.6. Les LAN à accès multiples

Certains LAN sont connectés par plusieurs routeurs :

- il faut éviter que les stations situées sur ces LAN reçoivent plusieurs copies du même paquet.



Un **routeur unique est désigné** pour alimenter le LAN partagé, celui ayant :

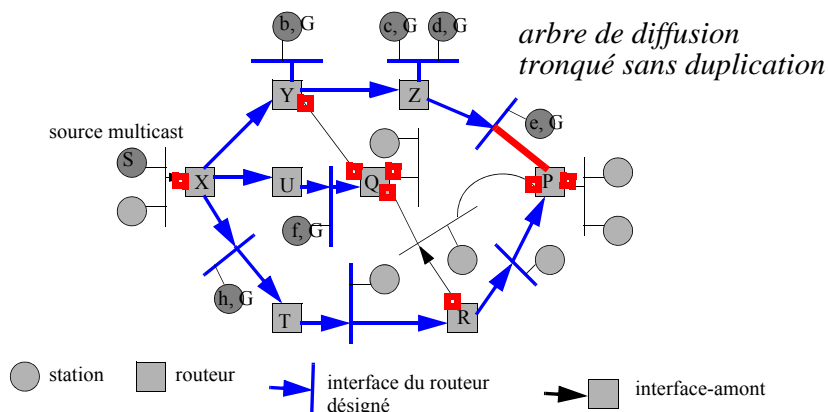
- **le chemin le plus court** vers la source multicast
- **l'adresse la plus petite** (en cas d'égalité)

Les routeurs peuvent détecter la présence d'un autre routeur sur le même LAN par son activité (échange de messages) :

- protocole de routage : problème de cohérence entre protocoles hétérogènes
- protocole IGMP (version 2 et 3)

4.7. L'arbre de diffusion tronqué sans duplication

⇒ "Truncated-Reverse Path Broadcasting"



4.8. Élagage

But : pour un groupe donné, certains branches de l'arbre de diffusion ne donnent accès à aucune station membre du groupe. **Ces branches sont inutiles !**

- En effet, l'arbre construit par T-RPB est un arbre couvrant la totalité des routeurs du réseau.

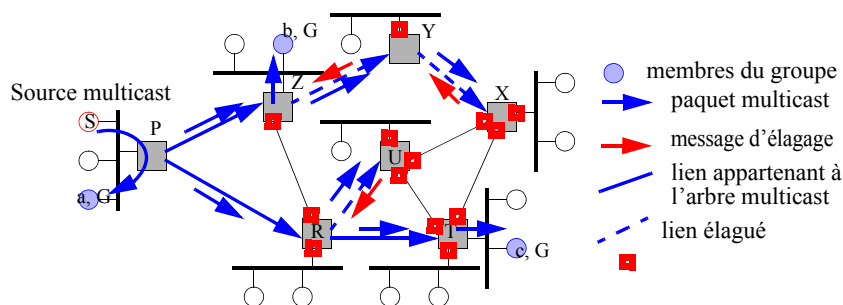
Des messages spécifiques vont permettre l'**élagage** ("pruning").

Si tous les interfaces d'un routeur donne accès à des LAN où aucune station n'est active,

- le routeur génère un message d'élagage sur l'interface-amont.

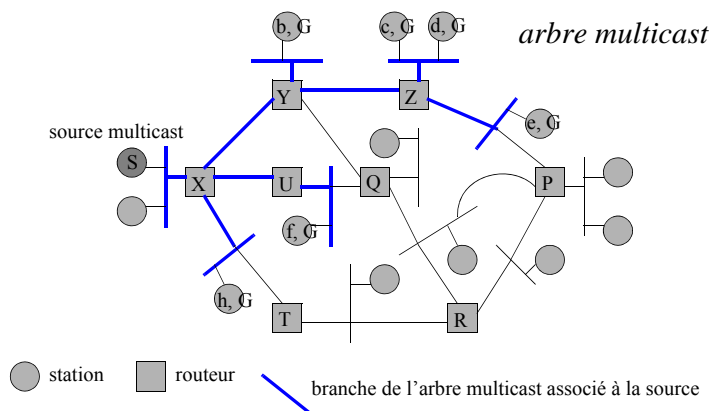
Si un routeur reçoit sur tous ses interfaces-aval actifs un message d'élagage :

- le routeur propage un message d'élagage sur l'interface-amont.



4.9. L'arbre multicast

=> "Reverse Path Multicasting"



4.10. Gestion de l'arbre multicast

Construction de l'arbre pour le groupe G et sa maintenance :

- Sur un routeur, la création d'une entrée multicast pour un groupe G :
 - déclenchée par l'apparition du premier paquet à destination de G.
- L'élagage de l'arbre de G :
 - déclenchée par l'apparition du premier paquet à destination de G alors qu'il n'y a aucun interface actif pour ce routeur et ce groupe.
 - puis par réception d'un message de prune (G).
- Une entrée multicast G est supprimée si l'entrée est inactive pendant trop longtemps :
 - lors de absence de réception de paquet multicast associée à G,
 - un temporisateur est associée à chaque entrée multicast.

5. Conclusion

Le service de multicast est un service utile.

Le procédé d'acheminement et l'adressage IP permet aisément à un paquet d'atteindre ses destinataires, pour peu que les tables de routage soient correctement configurées.

On peut construire plusieurs types d'arbres :

- arbre de Steiner
- arbre des plus courts chemins

On a décrit une technique de construction de l'arbre des plus courts chemins inverses, à partir de diffusions contrôlées par «reverse path».

Actuellement le multicast se développe (Cf. IPTV)

Autres protocoles :

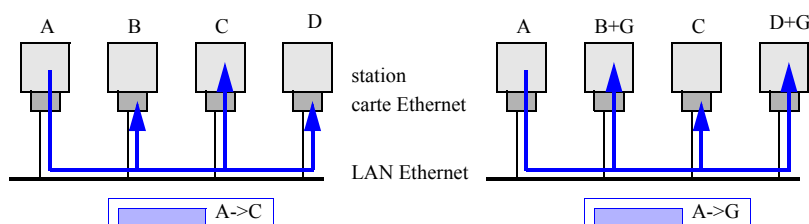
- protocole de routage
- de transport de données fiables (TCP est inadapté au multicast)
- de gestion des groupes, etc.

6. Quelques information supplémentaires

6.1. La diffusion dans les réseaux locaux

Le partage du même support physique entre toutes les stations fait que **le service de diffusion est naturel** dans les réseaux locaux.

- Que les trames soient adressées à une, plusieurs ou toutes les stations, toutes les cartes d'accès au réseau local **reçoivent systématiquement une copie de toutes les trames** circulant sur le RL.
- Même coût entre une simple transmission ou une diffusion partielle ou totale
- Les cartes d'accès recevant une trame qui ne leur est pas adressée, l'ignorent simplement!
- note : les cartes d'accès n'ayant que cela à faire, les traitements intempestifs et superflus ne sont pas pénalisants



On rappelle que les LAN disposent d'un adressage avec la notion d'adresse de groupe (bit G).

6.2. Résolution directe d'adresse

- adresse multicast IP \Rightarrow adresse MAC de groupe

- adresse IP **multicast** :

1110 xxxx **abc defg hijk lmno pqrs tuvw**

23 bits de poids faibles

- adresse IEEE 802 **de groupe** :

0000 0001 0000 0000 0101 1110 0**abc defg hijk lmno pqrs tuvw**

Note :

- . plusieurs (32) adresses IP multicasts sont associées avec la même adresse IEEE 802
par exemple : 224.1.2.16 ou 225.129.2.16 \Rightarrow 01 00 5E 01 02 10