

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
28 janvier 2010 (28.01.2010)

PCT

(10) Numéro de publication internationale
WO 2010/010265 A2

(51) Classification internationale des brevets :
H04L 12/56 (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2009/051300

(22) Date de dépôt international :
3 juillet 2009 (03.07.2009)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0855031 23 juillet 2008 (23.07.2008) FR

(71) Déposant (pour tous les États désignés sauf US) :
FRANCE TELECOM [FR/FR]; 6, place d'Alleray,
F-75015 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **BUOB,**
Marc-Olivier [FR/FR]; 4, place Jacques Madaule,
Résidence Le Rousseau, F-92130 Issy Les Moulineaux
(FR). **DECRAENE, Bruno** [FR/FR]; 3, rue Aristide
Briand, F-92170 Vanves (FR).

(74) Mandataire : **LECOMTE, Isabelle**; France telecom
R&D/PIV/Brevets, 38-40, rue du Général Leclerc,
F-92794 Issy Moulineaux Cedex 9 (FR).

(81) États désignés (sauf indication contraire, pour tout titre
de protection nationale disponible) : AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD,
SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT,
TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

[Suite sur la page suivante]

(54) Title : DISTRIBUTION OF ROUTES IN A NETWORK OF ROUTERS

(54) Titre : DISTRIBUTION DE ROUTES DANS UN RESEAU DE ROUTEURS

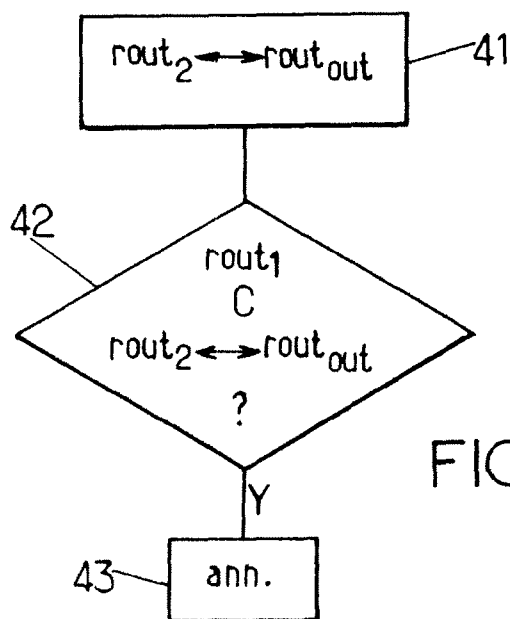


FIG.4.

(57) Abstract : An autonomous system comprises a plurality of routers within a packet transmission network, respective links joining the routers according to a topology of routers. At the level of a first router (rout₁) of the autonomous system which is joined to a second router (rout₂), for a route exiting the autonomous system through an exit router (rout_{out}), a path is determined (41) between the second router and the exit router according to performance criteria. Next, the route is announced (43) to the second router if the first router is situated on the previously determined path (42).

(57) Abrégé : Un système autonome comprend une pluralité de routeurs au sein d'un réseau de transmission de paquets, des liens respectifs reliant les routeurs selon une topologie de routeurs. Au niveau d'un premier routeur (rout₁) du système autonome qui est relié à un second routeur (rout₂), pour une route sortant du système autonome par un routeur de sortie (rout_{out}), on détermine (41) un chemin entre le second routeur et le routeur de sortie selon des critères de performance. Puis, on annonce (43) la route au second routeur si le premier routeur est situé sur le chemin déterminé précédemment (42).



Déclarations en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv)*

Publiée :

— *sans rapport de recherche internationale, sera republiée
dès réception de ce rapport (règle 48.2.g)*

DISTRIBUTION DE ROUTES DANS UN RESEAU DE ROUTEURS

La présente invention est relative au domaine des télécommunications,
5 et plus particulièrement des protocoles de routage et des équipements de
réseau mettant en œuvre de tels protocoles.

Un réseau de télécommunication classique comprend des équipements
de réseau, appelés routeurs, ayant pour fonction d'assurer la transmission des
paquets. La transmission d'un paquet depuis un équipement jusqu'à un autre
10 équipement dans le réseau peut généralement être réalisée suivant plusieurs
chemins. Dans un tel contexte, les routeurs sélectionnent, suivant des critères
déterminés, le meilleur chemin de transmission parmi différents chemins de
transmission possibles. Un routeur sélectionne ainsi un chemin de transmission
sur la base d'informations qui sont généralement reçues selon un protocole de
15 routage et mémorisées dans des tables de routage.

Un protocole de routage fournit des mécanismes aux routeurs d'un
réseau permettant de partager des informations de routage entre eux via des
sessions spécifiques au protocole de routage et, de ce fait, permettant de
construire des tables de routage cohérentes entre elles.

20 Il existe plusieurs protocoles de routage.

Un protocole de routage nommé BGP (« Border Gateway Protocol »)
assure une cohésion du routage Internet. Il est défini par l'IETF, pour 'Internet
Engineering Task Force'.

Un réseau Internet est constitué d'un ensemble de réseaux autonomes
25 ou encore systèmes autonomes, AS (pour 'Autonomous Systems' en anglais)
qui sont connectés entre eux. Un réseau autonome est un ensemble de
routeurs fonctionnant sous une seule administration technique, utilisant un
protocole de routage intérieur de type IGP (pour 'Interior Gateway Protocol'), tel
que par exemple IS-IS (pour 'Intermediate System to Intermediate System' en
30 anglais) ou le protocole OSPF (pour 'Open Shortest Path First' en anglais) et
des informations de routage communes pour router les paquets à l'intérieur du
système autonome. De tels routeurs utilisent par ailleurs un protocole de
routage extérieur, tel que par exemple le protocole BGP-4 dans l'Internet, pour

router les paquets vers les autres systèmes autonomes. Ainsi, deux routeurs d'un même AS peuvent échanger des informations de routage via des sessions selon le protocole BGP4, appelé dans ce cas iBGP, et des routeurs de deux AS respectifs distincts peuvent échanger des informations de routage via des sessions selon le protocole BGP4, appelé dans ce cas eBGP.

Les échanges d'information effectués via des sessions selon le protocole BGP permettent d'obtenir au niveau de chaque routeur d'un AS des informations sur l'ensemble des routeurs de bordure, ou encore routeurs de sortie, par lesquels le trafic peut sortir de l'AS vers une destination externe à l'AS. Le protocole BGP cherche ensuite à associer le meilleur routeur de sortie possible à chaque routeur du réseau. Toutefois, il est à noter qu'en pratique, il n'est pas toujours en mesure de connaître le meilleur routeur de sortie.

Pour mettre en œuvre de tels échanges d'informations selon BGP au sein d'un AS, on peut construire dans un AS une topologie « iBGP full mesh », telle qu'elle est définie dans le RFC 4271 (pour 'Request For Comment' en anglais). Dans ce cas, chaque routeur d'un même AS établit une session iBGP avec chaque autre routeur de l'AS. L'ensemble des routeurs de l'AS est donc entièrement maillé de sessions iBGP. La figure 1 illustre une telle implémentation de BGP au sein d'un AS 11 dans lequel chaque routeur x, b, c, d et e est relié avec tous les autres via des sessions iBGP respectives 13, ces routeurs x-e étant par ailleurs reliés via des sessions eBGP 12 à d'autres routeurs situés au sein d'autres AS.

Pour un AS comprenant un nombre N de routeurs, chaque routeur établit (N-1) sessions iBGP et l'AS comprend un nombre total de $N*(N-1)/2$ sessions iBGP. Un tel mode d'implémentation implique que le nombre de sessions augmente très rapidement avec le nombre de routeurs BGP présents dans l'AS, puisque ce nombre de sessions est fonction du nombre de routeurs au carré. En outre, la configuration d'un nouveau routeur dans l'AS impose la configuration de chacun des routeurs de l'AS, ce qui peut être complexe et lourd à administrer. De plus un routeur ne peut établir qu'un nombre limité de sessions BGP.

Selon une autre implémentation dite 'iBGP Route Reflector' telle que

définie dans le document RFC 4456, il est prévu de centraliser la distribution iBGP des routes sur des routeurs spécifiques référencés en tant que RR (pour 'Route Reflector' en anglais). Ici, chaque routeur de l'AS établit une session iBGP avec un RR, ou de préférence deux RR afin de se protéger contre une panne éventuelle d'un RR. La fonction RR peut également être répartie sur plusieurs routeurs. La figure 2 illustre une telle implémentation. L'AS 11 comprend les routeurs x, b-e. On distingue des routeurs 22 de type "Route Reflector" et des routeurs 23 de type "client". Les sessions selon le protocole iBGP 21 sont alors établies entre chaque routeur RR et ses routeurs clients.

10 Contrairement à un routeur iBGP classique, un RR peut rediffuser des routes iBGP. Il peut ainsi rediffuser:

- les routes iBGP qu'il apprend par ses routeurs clients ; et
- les routes apprises par iBGP à ses routeurs clients.

Enfin, un RR peut être également routeur client d'un autre RR et former ainsi une hiérarchie de RR.

Mais dans le contexte de cette dernière implémentation, le choix des routeurs spécifiques RR dans l'AS n'est pas aisé et un mauvais choix peut être source d'inconvénients divers comme notamment la création de boucles de routage. En outre, un RR ne sélectionne et ne distribue qu'une seule route, une telle restriction pouvant nuire aux performances du routage BGP.

On peut également prévoir d'implémenter de manière partitionnée, soit une implémentation de type iBGP full mesh comme illustré en figure 1, soit encore une implémentation de type Route Reflector comme illustré en figure 2. A cet effet, on prévoit de partitionner l'AS considéré en une pluralité de sous-AS, référencé en tant que 'confédération BGP', et de mettre en œuvre une des deux implémentations illustrées en figure 1 et 2 au sein de chaque sous-AS.

La figure 3 illustre une telle mise en œuvre de l'implémentation de type 'iBGP full mesh' dans un contexte de partitionnement.

Le système autonome 11 est partitionné en une première confédération 31 comprenant les routeurs d et e, et une seconde confédération 32 comprenant les routeurs x, b et c. Les sessions iBGP 13 sont alors établies de

manière 'full mesh' au sein de chaque confédération 31 et 32.

La mise en œuvre de iBGP full mesh dans un tel contexte de partitionnement permet de limiter l'augmentation rapide du nombre de sessions en fonction du nombre de routeurs, mais ne permet pas de résoudre
5 complètement ce problème. Par ailleurs, l'implémentation de iBGP Route Reflector au sein de sous-AS ne permet pas de résoudre les problèmes attachés à cette mise en oeuvre.

En outre, ce partitionnement requiert de déterminer un nombre de sous AS et d'en définir les frontières, ce qui peut s'avérer être complexe.

10 La présente invention vise à améliorer la situation.

Un premier aspect de l'invention propose un procédé de diffusion de routes dans un système autonome comprenant une pluralité de routeurs au sein d'un réseau de transmission de paquets ;
des liens respectifs reliant lesdits routeurs selon une topologie de routeurs ;
15 ledit procédé comprenant les étapes suivantes au niveau d'un premier routeur du système autonome, qui est relié à un second routeur, pour une route sortant dudit système autonome par un routeur de sortie :

/a/déterminer un chemin entre le second routeur et le routeur de sortie selon des critères de performance; et

20 /b/ annoncer la route audit second routeur si le premier routeur est situé sur le chemin déterminé à l'étape /a/.

Grâce à ces dispositions, on est en mesure d'optimiser la diffusion de routes au sein d'un système autonome, ou AS. La distribution ou diffusion d'une route au sein d'un AS est ici avantageusement répartie sur plusieurs
25 routeurs de l'AS, contrairement au cas notamment d'une implémentation 'iBGP Route Reflector' telle que décrite ci-avant.

En outre, dans les conditions d'un mode de réalisation de la présente invention, chaque route n'est avantageusement diffusée qu'une fois, généralement. Tel n'est pas le cas dans une implémentation 'iBGP Route
30 Reflector'. En effet, dans ce type d'implémentation, une même route peut être reçue plusieurs fois par un routeur, engendrant des messages inutiles.

Par ailleurs, le nombre de sessions utilisées pour la diffusion des

informations de routage telles que la diffusion des routes dans l'AS, peut avantageusement être inférieur au nombre des sessions utilisées pour la diffusion des routes dans une implémentation de type 'iBGP full mesh' telle que décrite ci-avant. En effet, il convient de noter ici que l'annonce d'une route est effectuée depuis un premier routeur vers un second routeur qui sont reliés selon une topologie préalablement établie entre les différents routeurs de l'AS considéré. De ce fait, un maillage complet entre les routeurs d'un même AS n'est pas requis. Il suffit de prévoir d'établir une session servant à diffuser les routes entre deux routeurs qui sont déjà reliés selon la topologie considérée.

On entend ici par les termes 'topologie de routeurs' l'ensemble des liens reliant les routeurs deux à deux, deux routeurs étant reliés par un lien étant des routeurs voisins.

Ainsi, l'établissement de sessions iBGP de diffusion de routes peut être avantageusement calé sur les liens déjà définis entre les routeurs selon la topologie de routeurs considérée, chaque routeur connaissant ses routeurs voisins, c'est-à-dire les routeurs avec lesquels il est relié par un lien de la topologie.

Cette caractéristique permet de résoudre des difficultés d'ingénierie, comme énoncées ci-avant, qui sont inhérentes à la mise en œuvre des mécanismes de réflexion de routes lorsque des routeurs RR sont utilisés, telles que la difficulté d'obtenir une distribution performante des informations de routage, la difficulté de choisir un RR qui convienne dans la topologie de routeurs considérée, et qui permette notamment une résistance aux pannes. Elle permet également de garantir que lorsque le réseau considéré est connexe, il existe un chemin de signalisation des routes iBGP. Plus précisément, chaque routeur a toujours connaissance de la meilleure route potentielle.

En outre, le problème d'augmentation rapide du nombre de sessions avec le nombre de routeurs BGP présents dans l'AS dans une topologie « iBGP full mesh », est ici avantageusement écarté.

Dans ces conditions, avantageusement, la diffusion des routes iBGP au sein du réseau suit un chemin optimal entre le routeur de bordure

annonçant la route en iBGP dans l'AS et le routeur utilisant cette route, puisque cette diffusion est basée sur la topologie de routeurs du réseau.

On peut prévoir que la topologie de routeurs considérée ici est issue d'une configuration quelconque, par exemple par configuration d'une liste des
5 liens sur lesquels on souhaite activer les sessions iBGP. La topologie de routeurs peut encore être issue de la mise en œuvre d'un protocole de routage de type IGP de type link state, c'est-à-dire IGP à état de lien qui permet une découverte et diffusion automatique et dynamique de la topologie de routeurs.

Dans ce dernier cas, le chemin optimal est défini au sens d'un
10 protocole de routage de type IGP et est calculé grâce à la topologie du réseau déjà annoncée par l'IGP. Cet IGP de type état de lien est en général déjà utilisé dans les réseaux pour calculer des chemins entre deux routeurs internes à l'AS. Il permet notamment de recalculer efficacement les routages, dans le cas de modification de topologie comme la déconnexion d'un lien physique, ou
15 dans le cas encore de l'arrêt d'un des routeurs de l'AS.

En se calant sur une telle topologie de routeurs pour la diffusion de routes iBGP, on peut bénéficier de ces avantages-là. Dans ces conditions, un changement de topologie tel que ajout ou suppression de routeurs et/ou de liens, ou encore modifications de métriques, est aisé à gérer puisque les
20 sessions de diffusion de routes sont établies sur la base de la topologie de routeurs, qui peut être obtenue par mise en œuvre du protocole de routage IGP considéré. On peut ainsi obtenir un routage qui résiste aux diverses pannes pouvant affecter le système autonome, et quelque soit le scénario de panne, on converge à nouveau vers un routage optimal.

Dans un mode de réalisation de la présente invention, l'étape /a/ est effectuée sur la base d'informations de la topologie de routeurs. En effet, dans certains cas, on peut prévoir que le premier routeur dispose de la topologie du réseau et se base dessus pour déterminer un chemin performant au regard de certains critères de performance, comme par exemple le chemin le plus court.
25

Des sessions de diffusion sont établies entre les routeurs et l'étape /b/ est effectuée via la session de diffusion de routes établie entre les premier et second routeurs.
30

Ainsi, il est donc prévu d'établir des sessions de diffusion de route en se calant sur les liens entre les routeurs définis par la topologie de routeurs, aucune limitation n'étant attachée au type de session de diffusion de routes.

Chaque routeur établit une session avec ses routeurs voisins de l'AS, c'est-à-dire avec les routeurs auxquels il est relié par un lien de la topologie. Un routeur n'a pas besoin de connaître la topologie complète du réseau considéré ; il a uniquement besoin de savoir si ces voisins le traverse pour joindre un point de sortie. Ainsi on peut étendre le domaine d'application de la présente invention à des IGP à vecteur de chemin, comme par exemple RIP.

Les sessions de diffusion peuvent notamment être de type iBGP ou encore de type MP-iBGP (pour 'Multi Protocol Border Gateway Protocol' en anglais). MP-iBGP correspond à une extension du protocole iBGP qui permet aux systèmes autonomes de s'échanger des informations étendues telles que des routes MPLS-VPN (pour 'MultiProtocol Label Switching Virtual Private Network'), ou encore de type Multicast, ou bien IPv6.

Dans un mode de réalisation, on prévoit donc d'établir des sessions de diffusion de routes selon un autre protocole de routage entre les routeurs qui sont reliés selon la topologie considérée, et d'effectuer l'étape /b/ via la session de diffusion établie entre les premier et second routeurs.

En procédant ainsi, les sessions de diffusion de routes sont établies en suivant la topologie du réseau considéré, avec les routeurs voisins de chaque routeur. Ainsi, les sessions de diffusion de routes sont établies uniquement entre deux routeurs qui sont voisins au sens de la topologie de routeurs, qui peut être obtenue par un protocole de routage IGP. De telles conditions permettent à la fois de simplifier, voire d'automatiser, la localisation des sessions iBGP tout en garantissant un routage BGP optimal contrairement à une implémentation de type 'iBGP RR', et de réduire le nombre de sessions de diffusion de routes par rapport à une implémentation de type 'iBGP full mesh', principalement lorsque le nombre de routeurs de l'AS considéré est élevé. De plus, il est à noter que dans la plupart des cas, une route n'est reçue qu'une seule fois, contrairement à une implémentation 'iBGP RR'.

Avantageusement, les étapes du procédé peuvent être appliquées

successivement à chacun des routeurs de la pluralité de routeurs, chacun de ces routeurs pouvant être considéré à son tour en tant que premier routeur. La diffusion de routes peut alors être faite dans tout l'AS.

5 En outre, il apparait avantageux de mettre en œuvre les étapes /a/ et /b/ au niveau du premier routeur pour tous les routeurs qui lui sont voisins, c'est-à-dire pour tous les routeurs qui lui sont reliés directement suivant la topologie considérée. Ces étapes peuvent donc être effectuées successivement pour chacun des seconds routeurs reliés au premier routeur selon la topologie.

10 Dans un mode de réalisation de la présente invention, les critères de performance sont adaptés pour sélectionner le chemin le plus court entre le second routeur et le routeur de sortie. Le protocole de routage IGP permet de déterminer de tels plus courts chemins au sein de la topologie de l'AS obtenue.

15 La topologie de routeurs et des couts respectifs des liens entre les routeurs sont obtenus selon le protocole IGP à état de lien.

Toutefois, aucune limitation du type de critères de performance n'est attachée à la mise en œuvre de l'étape /a/ du procédé de diffusion de routes.

20 Un deuxième aspect de la présente invention propose un routeur dans un système autonome adapté pour mettre en œuvre un procédé de diffusion de routes selon le premier aspect de la présente invention.

Un troisième aspect de la présente invention propose un système autonome comprenant une pluralité de routeurs selon le deuxième aspect de la présente invention.

25 Un quatrième aspect de la présente invention propose un programme d'ordinateur comportant des instructions pour la mise en œuvre du procédé selon le premier aspect de la présente invention lorsque ce programme est exécuté par un processeur.

D'autres aspects, buts et avantages de l'invention apparaîtront à la lecture de la description d'un de ses modes de réalisation.

30 L'invention sera également mieux comprise à l'aide des dessins, sur lesquels :

- les figures 1, 2 et 3 illustrent trois implémentations différentes de

mise en œuvre du protocole BGP dans un système de routeurs selon l'art antérieur ;

- la figure 4 illustre les principales étapes d'un procédé de diffusion de routes selon un mode de réalisation de l'invention ;
- 5 - les figures 5, 6 et 7 illustrent la mise en œuvre d'un procédé de diffusion de routes dans un réseau selon un mode de réalisation de la présente invention pour différents routes ;
- la figure 8 illustre un routeur selon un mode de réalisation de la présente invention.

10 Par la suite, la présente invention est décrite dans le cas où la topologie de routeurs utilisée est obtenue par mise en œuvre d'un protocole de routage de type IGP à état de lien. Toutefois, aucune limitation n'est attachée au mécanisme d'obtention de la topologie de routeurs utilisée.

15 La figure 4 illustre les principales étapes d'un procédé de diffusion de routes selon un mode de réalisation de la présente invention, dans un système autonome comprenant une pluralité de routeurs au sein d'un réseau de transmission de paquets. On peut prévoir, dans un mode de réalisation, que de manière préalable un protocole de routage à état de lien, de type IGP, est mis en œuvre au sein du système autonome considéré. Une topologie du système
20 autonome considéré est alors obtenue.

Cette étape peut ne pas être mise en œuvre, et on peut alors prévoir que la topologie de routeurs, c'est-à-dire l'ensemble des liens reliant les routeurs du réseau entre eux, soit obtenue par configuration.

25 La présente invention couvre toutes les méthodes permettant d'obtenir une topologie de routeurs dans un réseau.

Cette topologie relie certains routeurs de ce système autonome entre eux et de ce fait permet de définir pour un routeur donné un ou plusieurs routeurs voisins, les routeurs voisins étant reliés au routeur donné par des liens directs suivant la topologie considérée.

30 Dans un mode de réalisation de la présente invention il est prévu de prendre en considération successivement les différents routeurs du système autonome considéré et de leur appliquer les étapes suivantes, pour tous les

routeurs voisins au sens de la topologie considérée.

Pour appliquer les étapes du procédé selon un mode de réalisation de la présente invention, on prend en considération un premier routeur $route_1$ parmi les routeurs du système autonome. Puis, suivant la topologie considérée, on détermine un routeur voisin de ce premier routeur, référencé par la suite en tant que second routeur $route_2$. On considère en outre une route de sortie du système autonome par un routeur de sortie $route_{out}$.

Ainsi, au niveau du premier routeur, les étapes suivantes sont appliquées pour le second routeur et pour la route sortant du système autonome.

En premier lieu, à une étape 41, on détermine un chemin performant entre le second routeur et le routeur de sortie selon des critères de performance.

Il convient de noter que de tels critères de performance peuvent correspondre à une détermination du chemin le plus court au sens IGP entre le second routeur et le routeur de sortie. Dans ce cas, le premier routeur dispose d'informations qui lui permettent de calculer lui-même le plus court chemin entre ces deux routeurs.

Puis, à une étape 42, le premier routeur détermine s'il est lui-même situé sur ce chemin performant entre le second routeur, c'est-à-dire son routeur voisin, et le routeur de sortie pour la route considérée.

Si tel est le cas, à une étape 43, le premier routeur annonce cette route de sortie à son routeur voisin $route_2$. Dans le cas contraire, le premier routeur n'annonce pas cette route de sortie, contrairement à une implémentation selon 'iBGP'.

Dans une variante de réalisation, le premier routeur annonce cette route à son routeur voisin $route_2$, si et seulement si le premier routeur est le successeur du routeur $route_2$ sur le chemin performant de $route_2$ à $route_{out}$.

Il est avantageusement prévu d'effectuer ces étapes 41 à 43 au niveau de ce même premier routeur pour tous les routeurs qui sont ses voisins au sens de la topologie considérée, ainsi que pour toutes routes de sortie du système autonome considéré. Puis, avantageusement on peut appliquer ce procédé à

tous les routeurs du système autonome de sorte à diffuser de manière performante les routes dans tout le système autonome.

Il convient donc de noter qu'une session de diffusion de routes selon un mode de réalisation de la présente invention est établie avantageusement entre tous les routeurs voisins au sens de la topologie considérée, telle que par exemple une topologie IGP.

En utilisant la topologie IGP du système autonome, on garantit qu'on évite les boucles de routage et que l'état vers lequel le réseau va converger est optimal. En cas de modification de topologie comme la déconnexion d'un lien physique, ou encore l'arrêt d'un des routeurs de l'AS, on garantit que le nouvel état est de nouveau optimal.

Aucune limitation n'est attachée au type de protocole de routage permettant de fournir une topologie d'un système autonome au regard de la présente invention.

On peut par exemple prévoir d'utiliser un protocole IGP de type 'à état de lien (ou 'link state' en anglais). Ici, chaque routeur annonce aux autres routeurs de l'AS une partie de la topologie de l'AS dont il a eu connaissance de manière directe, c'est-à-dire ses propres préfixes, la liste de ses voisins et des valeurs de coût respectivement associées aux voisins. Sur la base de ces annonces de chemins IGP au sein de l'AS, chaque routeur peut déduire une topologie IGP de cet AS, et calculer les plus courts chemins d'un routeur quelconque de l'AS vers l'ensemble des routeurs de l'AS, notamment en mettant en œuvre un algorithme de type Dijkstra.

Dans un mode de réalisation de la présente invention, chaque routeur établit une session iBGP avec chacun des autres routeurs voisins du même AS, c'est-à-dire avec lequel il est relié directement par un lien IGP.

On peut prévoir que la manière d'annoncer une route par un routeur selon un mode de réalisation de la présente invention à un routeur voisin soit indifférente de la façon dont il a appris lui-même ces routes. Ainsi, par exemple, si ce premier routeur x a appris une route via une session de type eBGP, il peut l'annoncer à une étape 43 selon le procédé de diffusion de routes comme décrit dans la norme BGP, et plus précisément le RFC 4271 (pour 'Request For

Comment' en anglais).

Dans le cas où la route est annoncée pour la première fois selon le protocole iBGP, le premier routeur peut l'annoncer à une étape 43 comme décrit dans la norme BGP Route Reflector, et plus précisément le RFC 4456.

5 La figure 5 illustre une telle mise en œuvre dans un réseau selon un mode de réalisation de la présente invention. Un tel réseau illustré à la figure 5 comprend un AS 11 qui inclut des routeurs x, b-j. Les liens qui sont illustrés en trait continu et non fléché correspondent aux liens IGP qui relient les routeurs de cet AS entre eux, selon la topologie considérée. Ces traits continus
10 correspondent également aux sessions de diffusion de routes selon un mode de réalisation de la présente invention. Les traits fléchés eux représentent la diffusion en elle-même selon le sens de la flèche, à l'étape 43, d'une route depuis un routeur vers un routeur voisin selon le procédé de diffusion de routes d'un mode de réalisation.

15 Dans la topologie considérée, les liens IGP entre deux routeurs sont pondérés de sorte à permettre un calcul de chemin le plus court :

- les liens x-b, x-d, d-e, b-e, b-c, e-i, i-h, i-j, i-g et g-f sont affectés du poids 1 ;
- le lien c-f est affecté du poids 2 ;
- le lien d-h est affecté du poids 4 ; et
- 20 - les liens e-g et g-j sont affectés du poids 9.

Il convient de noter ici que les métriques associées aux liens peuvent être dissymétriques selon la direction du lien. Par exemple, la métrique associée au lien |x,b| peut différer de la métrique associée au lien |b,x|.

25 La diffusion de routes illustrée à la figure 5 est relative à la ou les route(s) sortant de l'AS 11 par le routeur de sortie rout_{out} x, ou encore 'next hop BGP' x.

Dans un mode de réalisation de la présente invention, les routes BGP pour lesquelles le routeur x est le routeur de sortie, sont diffusées comme suit.

30 Chaque routeur de cet AS calcule pour tous ses routeurs voisins le plus court chemin pour joindre le routeur x en tant que 'next hop' de la route BGP, ou encore en tant que routeur de sortie de la route considérée.

Ainsi, par exemple, le routeur x détermine le plus court chemin pour le joindre lui en tant que routeur de sortie, à partir du routeur b et du routeur d.

Depuis le routeur b, le prochain routeur pour atteindre le routeur x est le routeur x lui-même. Par conséquent, le routeur x annonce au routeur b les routes BGP ayant x comme routeur de sortie.

Il en est de même pour le routeur d. En effet, depuis le routeur d, le
5 prochain routeur pour atteindre le routeur x est le routeur x lui-même. Par conséquent, le routeur x annonce au routeur d les routes BGP ayant x comme next hop ou routeur de sortie.

Une fois que le procédé de diffusion de routes a été appliqué au routeur x pour tous ses routeurs voisins, on peut passer au routeur b.

10 Ainsi, de la même manière, le routeur b calcule pour chacun de ses voisins c'est-à-dire ici les routeurs e, c et x leur plus court chemin pour joindre le routeur de sortie x des routes BGP.

A partir du routeur c, le prochain routeur voisin utilisé pour joindre le routeur de sortie x est b, donc le routeur b annonce au routeur c les routes BGP ayant
15 comme next-hop le routeur x.

Le même principe s'applique aisément au routeur voisin e.

Par contre, le routeur x en tant que voisin du routeur b ne répond pas aux mêmes critères puisque le plus court chemin entre le routeur x en tant que voisin du routeur b et le routeur de sortie x, au sens de la topologie présentée, ne passe
20 par le routeur b. Par conséquent, b n'annonce pas, ou ne diffuse pas, les routes BGP au routeur voisin x.

De la même manière, le routeur e calcule pour chacun de ses voisins b, d, g et i les plus courts chemins respectifs pour joindre le next-hop x des routes BGP.

25 A partir du routeur voisin b, le plus court chemin pour joindre le next-hop x passe par x en tant que prochain bond, sans passer par e, donc les routes BGP ayant comme next hop x ne sont pas annoncées au routeur b.

A partir du routeur voisin d, le plus court chemin pour joindre le next-hop x passe par le routeur x en tant que prochain bond et sans passer par le routeur
30 e, donc les routes BGP ne sont pas annoncées au routeur d.

A partir du routeur voisin g, le prochain routeur sur le plus court chemin pour joindre le next-hop x est le routeur i et donc ce n'est pas le routeur e. Par

conséquent, le routeur e n'annonce pas les routes BGP au routeur g.

A partir du routeur i, le prochain routeur pour joindre le next-hop x est le routeur e. Le routeur e annonce donc les routes BGP ayant pour routeur de sortie le routeur x, au routeur g.

5 Ainsi, de proche en proche, les routes BGP ayant comme next-hop x se diffusent au sein de l'AS selon le graphe illustré par les traits continus fléchés en figure 5.

La figure 6 illustre un graphe de diffusion des routes BGP ayant comme routeur de sortie le routeur f dans le même AS 11 que celui illustré à la figure 5.

10 Dans ces conditions, les routes BGP ayant pour routeur de sortie le routeur f sont diffusées :

- par le routeur f aux routeurs c et g ;
- par le routeur c au routeur b ;
- par le routeur b au routeur x ;
- 15 - par le routeur g au routeur i ;
- par le routeur i aux routeurs h et e ; et
- par le routeur e au routeur d.

La figure 7 illustre les graphes de diffusion simultanée de routes BGP ayant respectivement comme next hop le routeur x et le routeur f, selon un mode de
20 réalisation de la présente invention.

Grâce aux dispositions de la présente invention, dans ce contexte là, la diffusion d'une route est effectuée de manière à ce que chaque routeur utilise le routeur de sortie le plus proche. On obtient ainsi une diffusion de route performante et réalisée selon le chemin le plus court pour chaque routeur de
25 l'AS.

Dans un mode de réalisation de la présente invention, lorsqu'un lien de type IGP entre deux routeurs tombe, il est prévu que la session de diffusion de route correspondante ne tombe pas immédiatement après la chute du lien IGP.

En maintenant la session de diffusion de route quelques instants, alors
30 même que le lien IGP correspondant a disparu, on est en mesure d'attendre une convergence de la diffusion de route selon un mode de réalisation de la présente invention selon un nouveau chemin. Ce nouveau chemin est issu de

la convergence du protocole de routage, tel que le protocole IGP, qui permet d'obtenir la topologie sur la base de laquelle la diffusion de routes est réalisée.

Il peut être également avantageux de préciser dans un message de type BGP, et plus précisément au niveau par exemple d'un code d'erreur, que la chute de cette session de diffusion de routes est volontaire, et de ce fait 'normale', dans le sens où elle n'est pas due à un problème BGP.

La figure 8 illustre un routeur selon un mode de réalisation de la présente invention.

Un routeur 80 est relié par un lien, selon la topologie de routeurs considérée, à un autre routeur.

Ce routeur 80 comprend :

- une unité de détermination 81 adaptée pour déterminer, pour un routeur de sortie, si le routeur appartient à un chemin entre cet autre routeur et le routeur de sortie, ledit chemin correspondant à des critères de performance; et
- une unité de diffusion 82 'ann' (pour 'annonce') adaptée pour annoncer la route audit autre routeur si le routeur est situé sur le chemin déterminé par l'unité de détermination.

Il peut en outre comprendre une unité de gestion 83 'mgt' (pour 'management' en anglais) de session de diffusion de routes adaptée pour établir des sessions de diffusion de routes selon un autre protocole de routage avec les routeurs qui lui sont reliés selon la topologie.

L'unité de gestion de session 83 peut en outre être adaptée pour gérer les sessions de diffusion de routes suivant des modifications de la topologie de routeurs.

En cas de panne simple, il est possible qu'une déflexion positive apparaisse durant la convergence des sessions de diffusion de routes selon un mode de réalisation de la présente invention, mais aucune boucle de routage n'apparaît.

Les conséquences de la panne d'un routeur sont similaires à celles de la chute d'un lien IGP et sont exposées ci-avant.

Dans un mode de réalisation de la présente invention, un changement de la topologie IGP du système autonome implique un aménagement correspondant

au niveau des sessions de diffusion de routes.

Par la suite, un tel aménagement est décrit lors de l'apparition d'un nouveau lien IGP entre deux routeurs x et b dans la topologie IGP utilisée pour la mise en œuvre de la présente invention. Un nouveau lien IGP entraîne de
5 préférence l'établissement d'une nouvelle session de diffusion de routes, puis un nouveau calcul de chemin le plus court pour les voisins et une prise de décision de diffusion des routes ou non, selon le procédé décrit ici.

Il convient de prendre en compte le délai d'établissement de cette nouvelle session de diffusion de routes avant de procéder aux annonces de routes
10 relatives à l'étape 43 de ce procédé de diffusion de routes. On peut à cet effet prévoir que les anciens routeurs voisins du routeur x attendent l'expiration d'un temporisateur avant d'effectuer un calcul de plus courts chemins pour leurs routeurs voisins respectifs selon un mode de réalisation de la présente invention.

15 De manière générale, suite à un changement de topologie de type IGP, il n'y a pas d'urgence à supprimer les sessions de diffusion de routes devenues obsolètes.

Il est possible que par rapport à une topologie IGP donnée, un nouveau routeur apparaisse, ou réapparaisse suite à une panne. Dans ce cas, comme
20 l'état de l'art actuel, il est avantageux de diffuser les routes BGP rapidement à ce nouveau routeur. A cet effet, ce nouveau routeur met en œuvre les étapes du procédé de diffusion de routes selon un mode de réalisation de la présente invention. Les routeurs déjà présents dans la topologie considérée, peuvent avantageusement établir une session de diffusion de routes avec ce nouveau
25 routeur, si ce dernier est leur routeur voisin au sens de la topologie de type IGP.

Un avantage supplémentaire de cette invention est que, dans le cas où la session iBGP n'est pas encore établie, comme les sessions iBGP suivent la même topologie que les sessions IGP, dans un mode de réalisation, on peut
30 prévoir d'annoncer dans le contexte protocolaire de type IGP, la topologie des sessions de diffusion de routes pour pouvoir détecter une différence de topologie entre les sessions de diffusion de routes et les liens IGP. Par

exemple, on peut annoncer ce contexte protocolaire de type IGP de la même manière qu'un protocole de type LDP (pour 'Label Distribution Protocol' en anglais) le fait, comme cela est décrit dans le document draft-ietf-mpls-ldp-igp-sync-01.txt "LDP IGP Synchronization".

5 Avantageusement, il convient de noter que la mise en œuvre d'un procédé de diffusion de routes selon un mode de réalisation de la présente invention ne nécessite pas de modification des messages BGP échangés entre les routeurs d'un AS. Simplement certaines réactions du protocole BGP existant sont à adapter.

10 En effet, la ré-émission des messages BGP s'appuie sur le protocole BGP et sur le protocole BGP RR qui sont déjà supportés pour les routeurs BGP.

Un calcul de type IGP SPF est également déjà prévu dans les routeurs existants.

15 Le filtrage des routes BGP lors de l'émission sur une session BGP est déjà défini au sein de la norme BGP et réutilisé dans la norme BGP RT (pour 'Route Target'), tel que cela est défini dans le document RFC 4684. Dans cette variante, les routes BGP sont diffusées conformément aux politiques de routage définies par les voisins.

20 Ici, il peut être prévu qu'un routeur donné calcule, en plus des plus courts chemins partant de lui, les plus courts chemins partant de ses voisins. On tire parti du fait que les plus courts chemins IGP sont rapidement calculés et que de plus ils ne retardent pas la convergence IGP lors d'une panne dans le réseau puisque les calculs effectués au niveau du protocole iBGP sont considérés comme moins prioritaires. Par conséquent, ils ne retardent aucun
25 traitement ou réaction dans un tel AS. Si pour d'autres raisons, de tels calculs sont déjà mis en œuvre, alors aucun calcul supplémentaire n'est requis.

30 La description a été réalisée dans le cas particulier d'un protocole de type 'Link state' ou à état de lien. Il est à noter qu'on peut également utiliser un protocole de type "distance vector", dans lequel pour chaque destination, le second routeur $route_2$ annonce au premier routeur $route_1$ si le premier routeur est sur le chemin de cette destination. Ainsi à l'étape 41, on détermine un chemin performant entre le second routeur et le routeur de sortie selon des critères de

performance en fonction de ces annonces. Il est ici souligné que pour un tel protocole de type "distance vector" déterminer un chemin performant consiste à déterminer un routeur prochain saut ou "next-hop" performant.

REVENDICATIONS

1. Procédé de diffusion de routes dans un système autonome comprenant une pluralité de routeurs au sein d'un réseau de transmission de paquets ;
5 des liens respectifs reliant lesdits routeurs selon une topologie de routeurs ;
ledit procédé comprenant les étapes suivantes au niveau d'un premier routeur (rout_1) du système autonome qui est relié à un second routeur (rout_2), pour une route sortant dudit système autonome par un routeur de sortie (rout_{out}) :
- /a/ déterminer (41) un chemin entre le second routeur et le routeur
10 de sortie selon des critères de performance; et
/b/ annoncer (43) la route audit second routeur si le premier routeur est situé sur le chemin déterminé à l'étape /a/ (42).
2. Procédé de diffusion de routes selon la revendication 1, dans lequel l'étape
15 /a/ est effectuée sur la base d'informations de la topologie de routeurs.
3. Procédé de diffusion de routes selon la revendication 1, dans lequel des sessions de diffusion de routes sont établies entre les routeurs ; et dans lequel l'étape /b/ est effectuée via la session de diffusion de routes établie entre les
20 premier et second routeurs.
4. Procédé de diffusion de routes selon la revendication 1, dans lequel, les étapes /a/ et /b/ sont effectuées lorsque la topologie de routeurs est modifiée.
- 25 5. Procédé de diffusion de routes selon la revendication 1, dans lequel les étapes /a/ et /b/ sont effectuées successivement pour chacun des seconds routeurs reliés au premier routeur selon la topologie de routeurs.
- 30 6. Procédé de diffusion de routes selon la revendication 1, dans lequel les étapes /a/ et /b/ sont appliquées successivement à chacun des routeurs de la pluralité de routeurs, considéré en tant que premier routeur.

7.Procédé de diffusion de routes selon la revendication 1, dans lequel les critères de performance sont adaptés pour sélectionner le chemin le plus court entre le second routeur et le routeur de sortie.

5 8.Procédé de diffusion de routes selon la revendication 1, dans lequel la topologie de routeurs et des coûts respectifs des liens entre les routeurs sont obtenus selon le protocole IGP (pour 'Interior Gateway Protocol') à état de lien.

9.Routeur (80) dans un système autonome comprenant une pluralité de routeurs au sein d'un réseau de transmission de paquets ;
10 ledit routeur étant relié avec un autre routeur par un lien selon une topologie de routeurs ;

ledit routeur comprenant :

- une unité de détermination (81) adaptée pour déterminer, pour un routeur de sortie, si ledit routeur appartient à un chemin entre ledit autre routeur et le routeur de sortie, ledit chemin correspondant à des critères de performance; et
 - une unité de diffusion (82) adaptée pour annoncer la route audit autre routeur si le routeur est situé sur le chemin déterminé par l'unité de détermination.
- 15
20

10.Routeur selon la revendication 9 comprenant en outre :

- une unité de gestion de session de diffusion de routes (83) adaptée pour établir des sessions de diffusion de routes selon un protocole de routage avec les routeurs qui lui sont reliés selon la topologie de routeurs.
- 25

11.Routeur selon la revendication 10, dans lequel l'unité de gestion de session de diffusion de routes (83) est en outre adaptée pour gérer les sessions de diffusion de routes suivant des modifications de la topologie de routeurs.

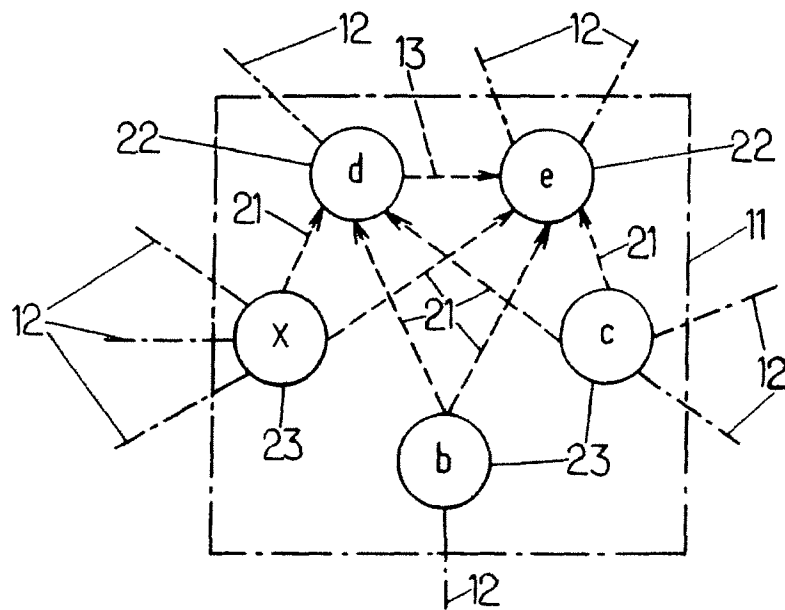
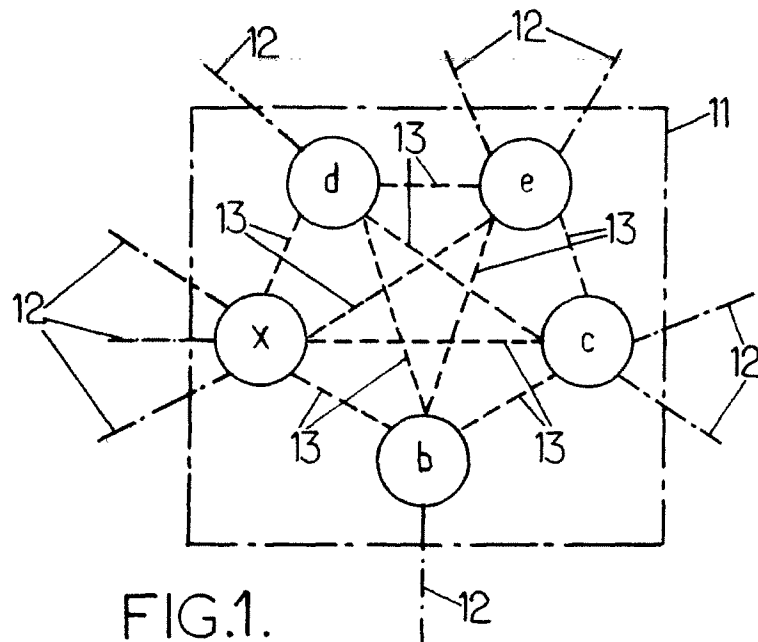
30

12.Système autonome comprenant une pluralité de routeurs au sein d'un réseau de transmission de paquets ; lesdits routeurs étant des routeurs selon

la revendication 9.

13. Programme d'ordinateur comportant des instructions pour la mise en œuvre
du procédé selon la revendication 1 lorsque ce programme est exécuté par un
5 processeur.

1/4



2/4

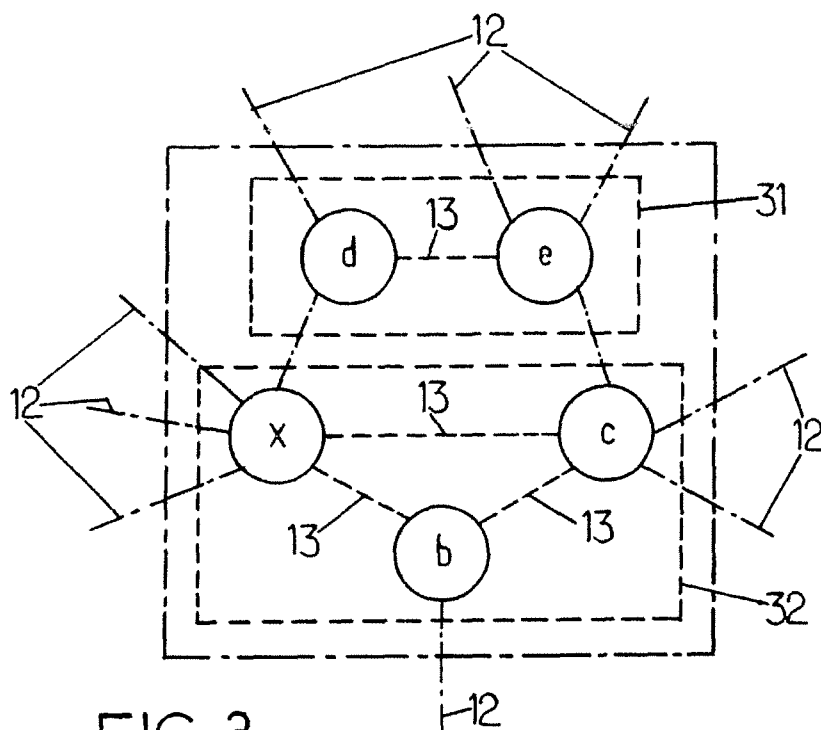


FIG. 3.

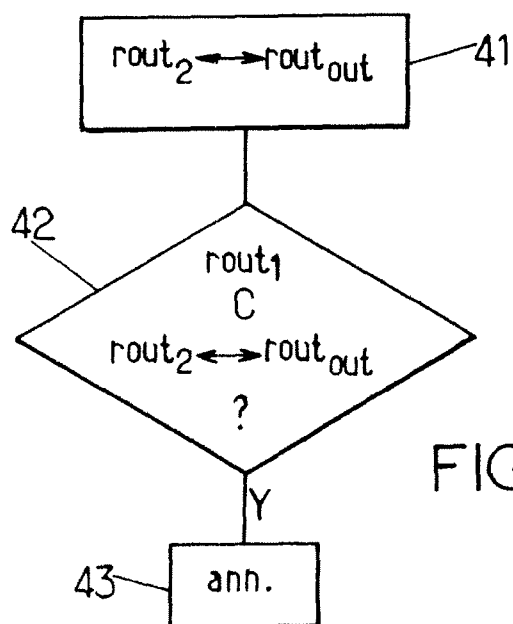


FIG. 4.

3/4

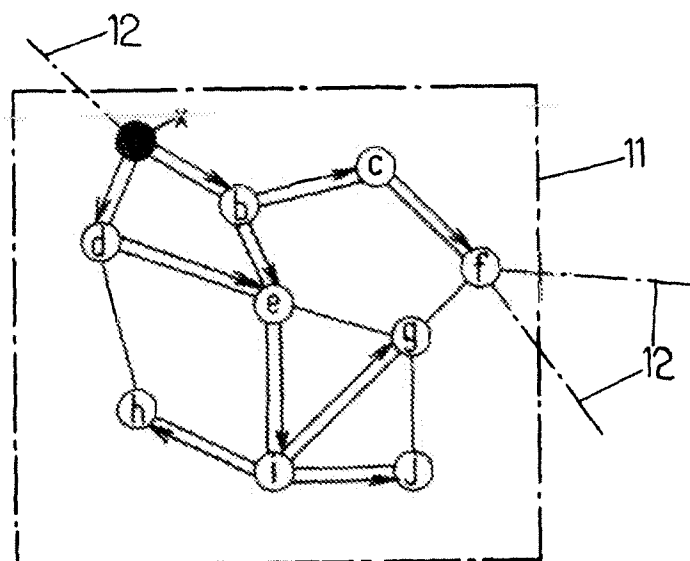


FIG. 5.

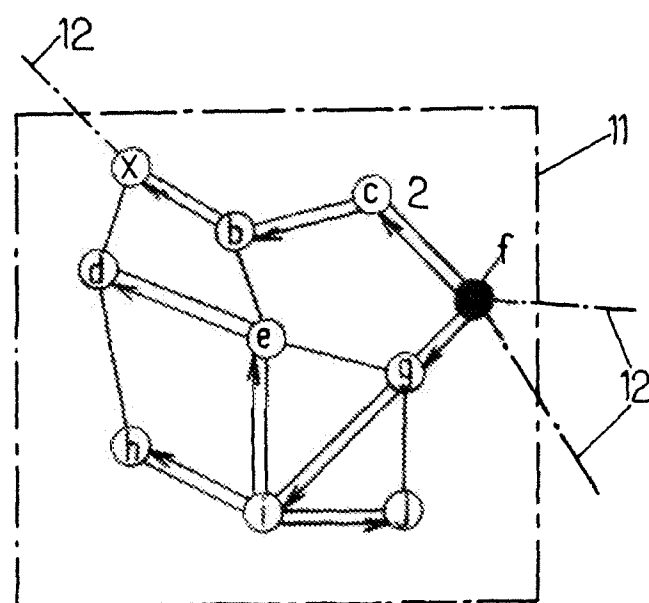


FIG. 6.

4/4

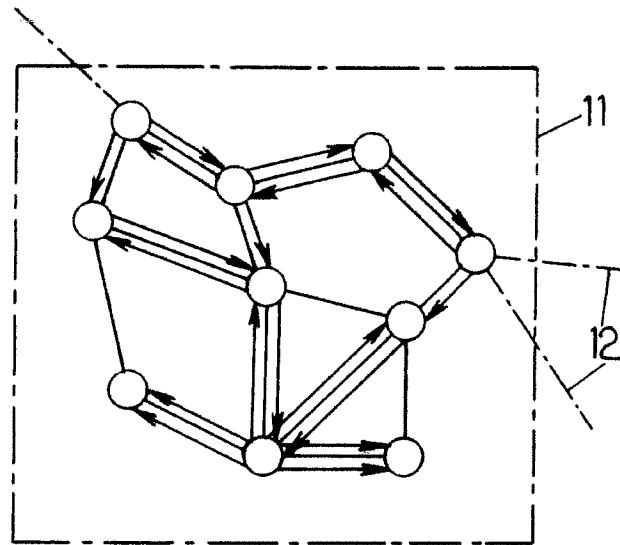


FIG. 7.

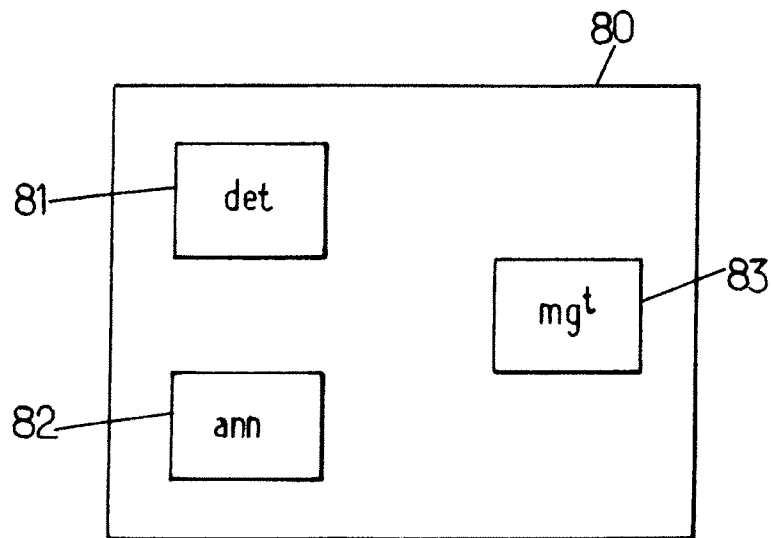


FIG. 8.