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

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

PCT

(43) Date de la publication internationale
29 janvier 2009 (29.01.2009)

(10) Numéro de publication internationale
WO 2009/013428 A2

(51) Classification internationale des brevets :
H04L 29/02 (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2008/051324

(22) Date de dépôt international : 11 juillet 2008 (11.07.2008)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0756500 13 juillet 2007 (13.07.2007) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **INFOVISTA SA** [FR/FR]; 6 rue de la Terre de Feu, F-91940 Les Ulis (FR).

(72) Inventeur; et

(75) Inventeur/Déposant *(pour US seulement)* : **DONIN DE ROSIÈRE, Emmanuel** [FR/FR]; 2 parvis de la bièvre, Apt 61, F-92160 Antony (FR).

(74) Mandataire : **PONTET-ALLANO & ASSOCIES**; 25 rue Jean Rostand, Parc Orsay Université, F-91893 Orsay (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, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ,

[Suite sur la page suivante]

(54) **Title:** METHOD AND SYSTEM FOR DISCOVERING THE LAYOUT OF COMMUNICATIONS BETWEEN APPLICATIONS IN AN INFORMATION NETWORK

(54) **Titre :** PROCEDE ET SYSTEME POUR LA DECOUVERTE DE LA TOPOLOGIE DES COMMUNICATIONS ENTRE APPLICATIONS D'UN RESEAU INFORMATIQUE

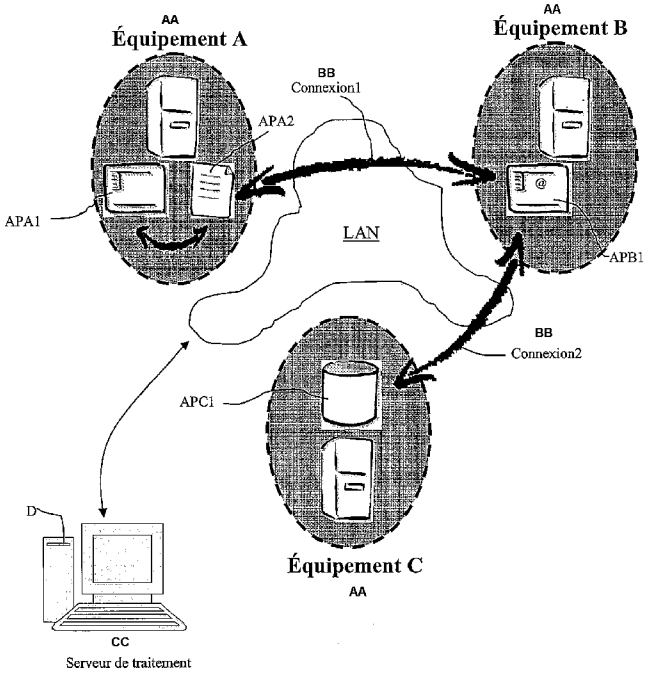

FIGURE 1

AA Piece of equipment A, B, C
BB Connection 1, 2
CC Processing server

(57) **Abstract:** The invention relates to a method for discovering the layout of communications between applications in an information network comprising several pieces of equipment, said method comprising the following steps: a) connection to each piece of equipment in the network, b) acquisition of raw data for each piece of equipment relating to the applications stored in said piece of equipment, c) acquisition of connection data for each piece of equipment relating to each live connection established by an application, d) determining communication paths from the raw and connection data between the respective pairs of applications in said network and e) generating a level 7 layout of said network from said communication paths.

(57) **Abrégé :** L'invention concerne un procédé pour découvrir la topologie des communications entre applications d'un réseau informatique comprenant plusieurs équipements, ce procédé comprenant les étapes suivantes : a) connexion à chacun des équipements du réseau, b) pour chaque équipement, acquisition de données brutes relatives à des applications hébergées dans cet équipement, c) pour chaque équipement, acquisition de données de connexion pour chaque connexion en cours établie par une application, d) à partir des données brutes et des données de connexion ainsi obtenues, on détermine des liens de communication entre respectivement des couples d'applications dudit réseau, et e) on génère une topologie niveau 7 dudit réseau à partir desdits liens de communication.

WO 2009/013428 A2

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), eurasien (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, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclaration 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*

" Procédé et système pour la découverte de la topologie des communications entre applications d'un réseau informatique."

La présente invention se rapporte à un procédé et un système de découverte de la topologie des communications entre applications d'un réseau informatique.

Les mécanismes classiques de découverte de topologie d'un réseau permettent habituellement d'obtenir la topologie d'un point de vue niveau 1 d'après la classification OSI, c'est-à-dire la topologie physique du réseau : les différents éléments constituant le réseau ainsi que la façon dont ils sont reliés entre eux par des câbles par exemple. D'autres mécanismes permettent également d'obtenir la topologie du point de vue de la couche 2 du modèle OSI, c'est-à-dire de définir avec quels équipements un élément donné est capable de communiquer directement. Il existe également d'autres mécanismes permettant de découvrir la topologie de la couche 3 du modèle OSI d'un réseau, c'est-à-dire les différents sous-réseaux existants ainsi que la manière dont ils sont interconnectés.

La présente invention s'intéresse, elle, à la découverte de la topologie de niveau 7 du modèle OSI d'un réseau, c'est-à-dire de la topologie des communications entre les applications exécutées sur les équipements du réseau.

Les procédés classiques de découverte de la topologie des communications entre applications utilisent une sonde passive à l'intérieur du réseau, c'est-à-dire que celle-ci n'effectue aucune interaction avec le réseau. Elle ne fait que capturer les différentes communications effectuées à l'intérieur du réseau et essaye en analysant ces communications, de découvrir la liste des applications sur le réseau ainsi que des liens entre elles. Néanmoins, ce procédé possède de nombreux inconvénients. En effet, la sonde doit être disposée à un point du réseau par lequel toutes les communications transitent afin de pouvoir faire la topologie de tout le réseau. Si un point de ce genre n'existe pas, il devient donc nécessaire de disposer de plusieurs sondes devant collaborer afin de pouvoir entièrement caractériser le réseau. De plus, la ou les sondes n'ont connaissance que des communications entre les applications, elles ne sont donc pas capables de retrouver certaines informations qui ne transitent pas sur le réseau comme

par exemple le nom de l'application, le temps d'exécution, la mémoire utilisée... De la même façon, à cause du manque d'informations, la ou les sondes ne sont pas capables de distinguer deux applications identiques situés sur le même équipement du réseau.

La présente invention a pour but de remédier aux inconvénients précités en proposant un nouveau procédé pour découvrir la topologie des communications des applications sur un réseau.

La présente invention a pour but un procédé permettant de déterminer de façon exhaustive la topologie des communications d'un réseau.

Un autre but de l'invention est de déterminer cette topologie de façon rapide.

On atteint au moins l'un des objectifs précités avec un procédé pour découvrir la topologie des communications entre applications d'un réseau informatique comprenant plusieurs équipements, ce procédé comprenant les étapes suivantes :

a) connexion à chacun des équipements du réseau, en utilisant un et/ou plusieurs protocoles d'administration tels que SNMP, SSH...,

b) pour chaque équipement, acquisition de données brutes relatives à des applications hébergées dans cet équipement,

c) pour chaque équipement, acquisition de données de connexion pour chaque connexion en cours établie par une application, ceci correspondant à des données de niveau 3 ou supérieur dans le modèle OSI,

d) à partir des données brutes et des données de connexion ainsi obtenues, on détermine des liens de communication entre respectivement des couples d'applications dudit réseau ; en d'autres termes on détermine le lien qui existe éventuellement entre les applications prises deux par deux ; et

e) on génère une topologie niveau 7 dudit réseau à partir desdits liens de communication.

Avec le procédé selon l'invention, on scrute chaque équipement du réseau. On répertorie l'ensemble des applications et on établit le lien qu'il peut y avoir entre deux applications données. On réalise une topologie complète du réseau. Cette topologie peut être réalisée de façon rapide

puisqu'on rapatrie l'ensemble des informations récoltées au sein d'un serveur de traitement. Cette topologie peut être sauvegardée sous forme d'un fichier XML que l'on stocke dans une base de données. On peut aussi utiliser toute autre méthode de stockage permanente ou non.

Selon un mode de mise en œuvre avantageux de l'invention, pour déterminer lesdits liens de communication, on réalise une première phase de calcul au cours de laquelle, pour chaque équipement, à chaque connexion on associe l'application correspondante qui a initiée cette connexion ; et on réalise une seconde phase de calcul au cours de laquelle, pour chaque équipement, à chaque connexion on associe l'application correspondante qui a été le destinataire de cette connexion.

On réalise ainsi des correspondances à chaque extrémité d'une connexion.

Avantageusement, les données brutes peuvent comprendre :

- le nom de chaque application,
- l'adresse IP utilisé en temps normal par cette application, et
- le port utilisé en temps normal par cette application,

Selon l'invention, les données de connexion peuvent comprendre :

- le protocole utilisé pour chaque connexion établie,
- l'adresse IP source,
- le port source,
- l'adresse IP destinataire, et
- le port destinataire.

De préférence, les données de connexion comprennent en outre le numéro de processus PID (POUR « Process Identifier » en langue anglaise) utilisé pour chaque communication lors d'une connexion. Lorsqu'on est en mesure de récupérer un PID, à chaque communication d'une connexion, on peut associer l'application correspondante qui a initiée cette communication. De la même manière, à chaque communication d'une connexion, on peut associer l'application correspondante qui a été le destinataire de cette communication.

Selon une caractéristique avantageuse de l'invention, l'association entre une connexion et l'application correspondante dans la première phase

de calcul comprend une étape de comparaison entre d'une part l'adresse IP source et le port source, et d'autre part l'adresse IP utilisé et le port utilisé.

De la même manière, l'association entre une connexion et l'application correspondante dans la seconde phase de calcul comprend une étape de comparaison entre d'une part l'adresse IP destinataire et le port destinataire, et d'autre part l'adresse IP utilisé et le port utilisé.

Avantageusement, les données brutes peuvent comprendre en outre des données de performances de chaque application, telles que des données relatives à l'utilisation mémoire par l'application et la consommation CPU de cette application.

On obtient donc des informations supplémentaires sur les applications, ces informations étant différentes de celles transitant sur le réseau. Cela permet donc d'obtenir une topologie plus riche qu'une topologie obtenue avec un procédé classique selon l'art antérieur.

On peut faire des recoupements parmi l'ensemble des données acquises de façon à obtenir de nombreuses informations sur une communication donnée ; ces informations étant obtenues des deux cotés de cette communication. Cette redondance d'information permet ici de s'affranchir de problème dans l'art antérieur dû à la topographie du réseau au niveau 3 du modèle OSI comme le NAT (pour « Network Address Translation » en langue anglaise). Avec la redondance d'informations obtenue en scrutant la communication des deux cotés, il est possible dans certains cas de reconstruire la communication exacte ayant eue lieu en s'affranchissant des intermédiaires utilisés à cause du NAT.

Selon un mode de mise en œuvre de l'invention, on effectue les étapes a) à e) de façon régulières.

Cette découverte de topologie (scrutation des applications des équipements par les étapes a) à e)) peut également être effectuée de façon planifiée (journalière, hebdomadaire, etc.) afin de suivre l'évolution de la topologie ou la déterminer une fois pour toute à un instant donné.

A titre d'exemple, on peut effectuer les étapes a) à e) en réponse à une consigne prédéterminée. Cela peut être une consigne déclenchée par

un utilisateur ou une alarme générée de façon automatique lors par exemple d'un problème sur le réseau.

Suivant un autre aspect de l'invention, il est proposé un serveur pour découvrir la topologie des communications entre applications d'un réseau informatique comprenant plusieurs équipements, ce serveur comprenant des moyens pour :

a) se connecter à chacun des équipements du réseau,

b) pour chaque équipement, acquérir des données brutes relatives à des applications hébergées dans cet équipement,

c) pour chaque équipement, acquérir des données de connexion pour chaque connexion en cours établie par des applications,

d) à partir des données brutes et des données de connexion ainsi obtenues, déterminer des liens de communication entre respectivement des couples d'applications dudit réseau, et

e) générer une topologie niveau 7 dudit réseau à partir desdits liens de communication.

D'autres avantages et caractéristiques de l'invention apparaîtront à l'examen de la description détaillée d'un mode de mise en œuvre nullement limitatif, et des dessins annexés, sur lesquels :

La figure 1 est une vue générale d'un réseau mettant en œuvre le procédé selon la présent invention,

La figure 2 est un tableau illustrant des données brutes obtenues après scrutation des applications des équipements du réseau, et

La figure 3 est un tableau illustrant des données de connexion obtenues après scrutation des connexions relatives à l'équipement B.

Sur la figure 1 on voit un réseau local LAN comprenant trois équipements A, B et C, ainsi qu'un serveur de traitement D selon l'invention. Ce serveur de traitement est doté de moyens matériels et logiciels conventionnels permettant son bon fonctionnement dans un réseau de communication. Il intègre également une application selon l'invention lui permettant de se connecter à chaque équipement du réseau afin d'effectuer des opérations d'analyse et d'acquisition de données. Chaque

équipement comporte une application permettant au serveur de traitement d'acquérir des données au travers d'un protocole d'administration.

L'équipement est une machine comportant notamment deux application APA1 et APA2. Ces applications peuvent communiquer entre elles et aussi communiquer avec d'autres applications contenues dans d'autres équipements du réseau. L'équipement B comporte l'application APB1. L'équipement C comporte l'application APC1.

On se propose d'établir la topologie du réseau LAN, c'est-à-dire les liens de communication de niveau 7 dans le modèle OSI. Pour ce faire, le serveur de traitement scrute chaque équipement pour récupérer des données brutes. Ces données sont répertoriées dans le tableau de la figure 2.

Ainsi, pour chaque équipement on acquière le nom de chaque application hébergée par cet équipement, l'adresse IP utilisé, le port utilisé ainsi que des données sur la consommation mémoire, la consommation CPU de l'application considérée et le PID de l'application considérée.

Par exemple, sur l'équipement A, on distingue l'application nommée APA1 qui utilise l'adresse IP 212.1.1.1 sur le port 2, ainsi que l'application nommée APA2 qui utilise l'adresse IP 212.1.1.2 sur le port 1.

Sur l'équipement B, on distingue l'application nommée APB1 qui utilise l'adresse IP 213.1.1.4 sur le port 1.

Sur l'équipement C, on distingue l'application nommée APC1 qui utilise l'adresse IP 214.1.1.5 sur le port 1.

Ensuite, à un instant donné ou de façon répétitive, on récupère des données de connexion pour chaque équipement. Sur la figure 3 on distingue les données de connexion obtenues sur l'équipement B à un instant où les connexions connexion1 et connexion2 étaient en cours c'est-à-dire actives.

Connexion1 est caractérisée par une adresse IP source de 213.1.1.4, un port source de 1, une adresse IP destinataire de 212.1.1.2, un port destinataire égal à 1, un protocole TCP et un numéro de processus PID valant 44.

De la même manière, connexion2 est affectée d'une adresse IP source égale à 213.1.1.4, un port source égal à 1, une adresse IP destinataire égale à 214.1.1.5, un port destinataire égale à 1, un protocole UDP et un numéro de processus PID égal à 44.

Avec l'ensemble de données récupérées, le serveur de traitement D effectue une première phase de calcul consistant, pour chaque équipement, à faire le lien entre les applications et les connexions découvertes sur ce même équipement. Pour cela, on compare les couples IP/Port source utilisés par l'application. Le serveur de traitement est alors capable d'attribuer à une application toutes les communications ayant comme source un couple IP/Port utilisé par cette application. Avantageusement, lorsqu'on possède le numéro du processus PID de chaque communication, le serveur de traitement fait le lien entre l'application et les communications associées. Sur le tableau de la figure 3, chaque communication est associé au PID 44 et donc à l'application APB1.

Lors de la seconde phase de calcul, le serveur de traitement fait le lien entre les IP/Port destination des communications et les applications découvertes sur tous les équipements. De la même façon que précédemment, on utilise les données sur les IP/Port découverts lors de l'acquisition des données brutes des applications.

Grâce à ces deux étapes, on fait le lien entre deux applications données d'un même réseau même si ces applications ne se situent pas sur le même équipement.

Bien sûr, l'invention n'est pas limitée aux exemples qui viennent d'être décrits et de nombreux aménagements peuvent être apportés à ces exemples sans sortir du cadre de l'invention.

REVENDICATIONS

1. Procédé pour découvrir la topologie des communications entre applications d'un réseau informatique comprenant plusieurs équipements, ce procédé comprenant les étapes suivantes :

a) connexion à chacun des équipements du réseau,,

b) pour chaque équipement, acquisition de données brutes relatives à des applications hébergées dans cet équipement,

c) pour chaque équipement, acquisition de données de connexion pour chaque connexion en cours établie par une application,

d) à partir des données brutes et des données de connexion ainsi obtenues, on détermine des liens de communication entre respectivement des couples d'applications dudit réseau, et

e) on génère une topologie niveau 7 dudit réseau à partir desdits liens de communication.

2. Procédé selon la revendication 1, caractérisé en ce que pour déterminer lesdits liens de communication, on réalise une première phase de calcul au cours de laquelle, pour chaque équipement, à chaque connexion on associe l'application correspondante qui a initiée cette connexion ; et on réalise une seconde phase de calcul au cours de laquelle, pour chaque équipement, à chaque connexion on associe l'application correspondante qui a été le destinataire de cette connexion.

3. Procédé selon la revendication 1 ou 2, caractérisé en ce que les données brutes comprennent :

- le nom de chaque application,
- l'adresse IP utilisé par cette application, et
- le port utilisé par cette application,

4. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que les données de connexion comprennent :

- le protocole utilisé pour chaque connexion établie,
- l'adresse IP source,
- le port source,
- l'adresse IP destinataire, et

- le port destinataire.

5. Procédé selon les revendications 2 à 4, caractérisé en ce que l'association entre une connexion et l'application correspondante dans la première phase de calcul comprend une étape de comparaison entre d'une part l'adresse IP source et le port source, et d'autre part l'adresse IP utilisé et le port utilisé.

6. Procédé selon les revendications 2 à 4, caractérisé en ce que l'association entre une connexion et l'application correspondante dans la seconde phase de calcul comprend une étape de comparaison entre d'une part l'adresse IP destinataire et le port destinataire, et d'autre part l'adresse IP utilisé et le port utilisé.

7. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que les données brutes comprennent en outre des données de performances de chaque application.

8. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que les données de connexion comprennent en outre le numéro de processus PID utilisé pour chaque communication lors d'une connexion.

9 Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on effectue les étapes a) à e) de façon régulières.

10. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on effectue les étapes a) à e) de façon planifiées.

11. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on effectue les étapes a) à e) en réponse à une consigne prédéterminée.

12. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on stocke la topologie dans une base de données.

13. Serveur pour découvrir la topologie des communications entre applications d'un réseau informatique comprenant plusieurs équipements, ce serveur comprenant des moyens pour :

a) se connecter à chacun des équipements du réseau,

b) pour chaque équipement, acquérir des données brutes relatives à des applications hébergées dans cet équipement,

c) pour chaque équipement, acquérir des données de connexion pour chaque connexion en cours établie par des applications,

d) à partir des données brutes et des données de connexion ainsi obtenues, déterminer des liens de communication entre respectivement des couples d'applications dudit réseau, et

e) générer une topologie niveau 7 dudit réseau à partir desdits liens de communication.

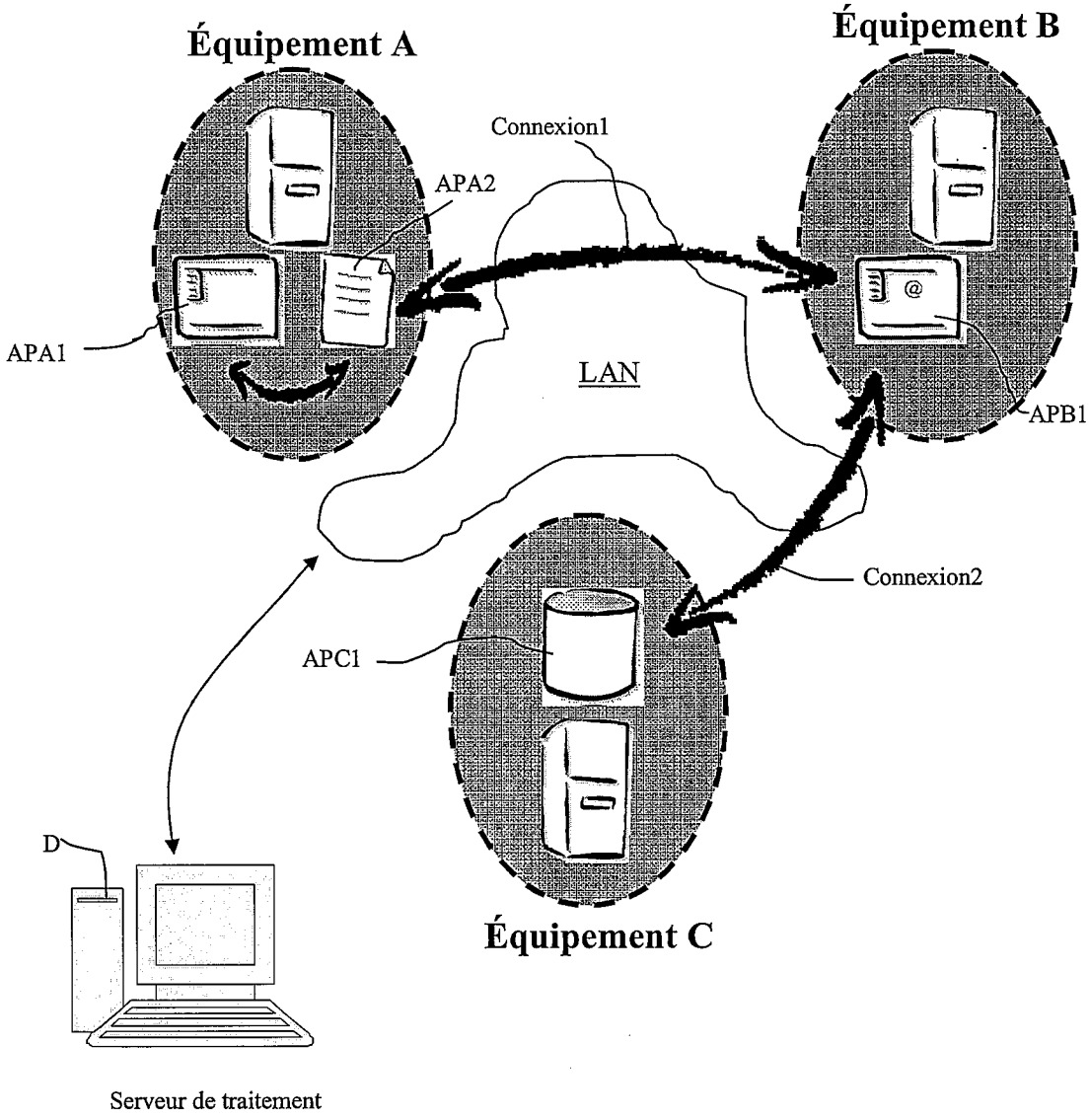
Équipement A
Équipement B
Connexion1
APA2
APA1
LAN
APB1
Connexion2
APC1
D
Équipement C
Serveur de traitement

FIGURE 1

	nom	IP utilisé	Port utilisé	CPU	Mémoire	PID
Équipement A	APA1	212.1.1.1	2	...	...	100
	APA2	212.1.1.2	1	...	...	88
Équipement B	APB1	213.1.1.4	1	...	...	44
Équipement C	APC1	214.1.1.5	1	...	...	22

FIGURE 2

Équipement B						
connexion	IP source	Port source	IP destination	Port destination	Protocole	PID
Connexion1	213.1.1.4	1	212.1.1.2	1	TCP	44
Connexion2	213.1.1.4	1	214.1.1.5	1	UDP	44

FIGURE 3