

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

(19) Organisation Mondiale de la
Propriété Intellectuelle
Bureau international



(10) Numéro de publication internationale
WO 2021/044094 A1

(43) Date de la publication internationale
11 mars 2021 (11.03.2021)

(51) Classification internationale des brevets :
G06F 3/06 (2006.01) G06F 11/14 (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2020/051513

(22) Date de dépôt international :
31 août 2020 (31.08.2020)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
FR1909787 05 septembre 2019 (05.09.2019) FR

(71) Déposant : **BLADE** [FR/FR] ; 151 rue Saint Denis, 75002
PARIS (FR).

(72) Inventeur : **LAMY, Arnaud** ; 2 rue de la concorde, 92600
ASNIERES SUR SEINE (FR).

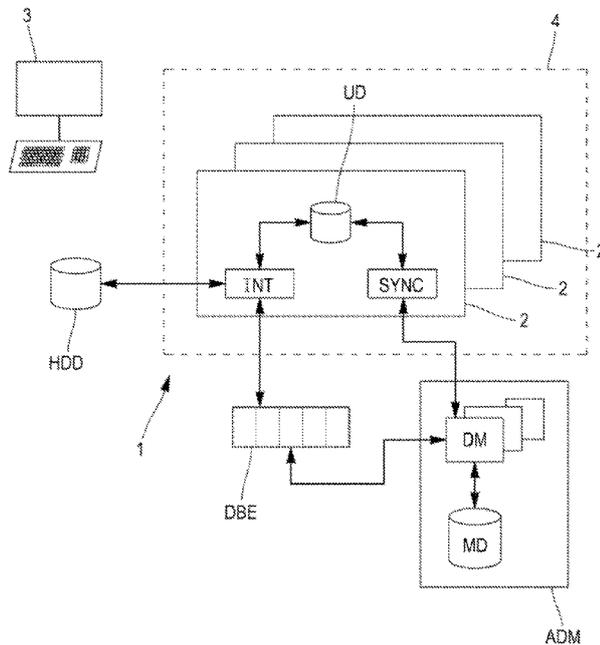
(74) Mandataire : **BREESE, Pierre** ; IP TRUST, 2 rue de Cli-
chy, 75009 Paris (FR).

(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, BN, BR, BW, BY, BZ, CA,
CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ,
EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP,
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: METHOD FOR MANAGING A DATA RECORD IN A COMPUTER SYSTEM

(54) Titre : PROCEDE DE GESTION D'UN GROUPE DE DONNEES DANS UN SYSTEME INFORMATIQUE

Fig. 1



(57) Abstract: The invention relates to a method for managing a data record in a computer system comprising: - at least one computing server (4) for hosting a computer session (2) running with the aid of an operating system having a deduplication index (UD) and managing access to a session storage space (HDD); - a shared storage space (DBE); - an administration server (ADM) for administering the shared storage space (DBE), executing a data management program (DM); the computer session (2) executing an interception program (INT) implementing the following steps: - intercepting a read call to read at least one data record transmitted in the session (2); - accessing the deduplication index (UD) and determining whether the data record is recorded in the shared storage space (DBE); - if so, reading, from the deduplication index (UD), the address of the data record in the shared storage space (DBE) and redirecting the



WO 2021/044094 A1

(84) **États désignés** (*sauf indication contraire, pour tout titre de protection régionale disponible*) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasién (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publiée:

— avec rapport de recherche internationale (Art. 21(3))

read call to this address; - if not, overlooking the read call so that it is processed by the operating system.

(57) **Abrégé** : L'invention porte sur un procédé de gestion d'un groupe de données dans un système informatique comprenant : - au moins un serveur de calcul (4) pour accueillir une session informatique (2) s'exécutant à l'aide d'un système d'exploitation disposant d'un index de déduplication (UD) et gérant l'accès à un espace de stockage de session (HDD); - un espace de stockage partagé (DBE); - un serveur d'administration (ADM) de l'espace de stockage partagé (DBE) exécutant un programme de gestion de données (DM); la session informatique (2) exécutant un programme d'interception (INT) mettant en œuvre les étapes suivantes : - intercepter un appel en lecture d'au moins un groupe de données émis dans la session (2); - accéder à l'index de déduplication (UD) et déterminer si le groupe de données est enregistré dans l'espace de stockage partagé (DBE); - dans un tel cas, lire dans l'index de déduplication (UD) l'adresse du groupe de données dans l'espace de stockage partagé (DBE) et rediriger l'appel en lecture vers cette adresse; - dans le cas contraire, laisser passer l'appel en lecture pour qu'il soit traité par le système d'exploitation.

PROCEDE DE GESTION D'UN GROUPE DE DONNEES DANS UN SYSTEME
INFORMATIQUE

DOMAINE DE L'INVENTION

La présente invention concerne un procédé de gestion d'un groupe de données déduplicqué dans un réseau informatique.

ARRIERE PLAN TECHNOLOGIQUE DE L'INVENTION

Dans un réseau informatique, les utilisateurs peuvent être amenés à manipuler dans leur session respective un même fichier informatique, par exemple une application ou un fichier de données.

Généralement, cette situation conduit à enregistrer le fichier informatique dans l'espace de stockage associé à chaque session, et donc à le dupliquer, ce qui tend à occuper un espace de stockage important. Pour rationaliser cela, on peut prévoir de rendre le fichier informatique disponible dans un espace de stockage partagé, auquel chaque session peut accéder. Toutefois cette solution n'est pas toujours pratique, car elle impose à l'utilisateur d'accéder par lui-même à l'espace de stockage partagé, ce qui lui impose des contraintes. De plus, un accès à un espace de stockage partagé peut introduire une latence d'accès ce qui pénalise l'utilisateur.

La gestion de l'espace de stockage de données et de la vitesse d'accès à ces données se pose tout particulièrement dans une architecture informatique du type « cloud » (traduit parfois en français par le terme « infonuagique »).

Dans une telle architecture, un administrateur met à disposition et mutualise un grand nombre de ressources informatiques, telles que des serveurs de calculs et des dispositifs de stockage, qu'il maintient en service dans un centre de données centralisé. Les utilisateurs accèdent à ces ressources à partir d'un client distant.

Il est important dans un tel contexte de pouvoir éviter la duplication des données afin de limiter l'espace de stockage nécessaire et la consommation énergétique du centre de données. De préférence, la gestion de la déduplication des données doit pouvoir s'effectuer sans affecter notablement la performance du système et sans que cela ne modifie les usages des utilisateurs.

Le document US2012221817A1 divulgue un mécanisme permettant d'éviter la duplication de données dans une architecture informatique. Chaque système de stockage de l'architecture consulte une table locale de hachage et une table globale de hachage pour déterminer si la donnée doit être écrite localement ou si un simple pointeur vers cette donnée présente dans un autre système de stockage de l'architecture est suffisant.

Ce mécanisme conduit à ne stocker qu'une seule fois un bloc de donnée déterminé dans l'architecture informatique et tous les appels à ce bloc sont redirigés vers son unique lieu de stockage, sans jamais créer de données redondantes. Formellement les données ne sont donc pas dédupliquées. Cela présuppose bien entendu que les systèmes de stockages soient tous et en permanence accessibles, pour ne pas perdre l'accès à cette donnée, ce qui n'est pas le cas de toutes les architectures informatiques.

OBJET DE L'INVENTION

Un but de l'invention est de proposer une solution au moins partielle à ce problème. Plus particulièrement, la présente invention vise à dédupliquer des données produites par des sessions informatiques disposant chacune d'un espace de stockage propre et dont l'exécution est intermittente.

BREVE DESCRIPTION DE L'INVENTION

En vue de la réalisation de ce but, l'objet de l'invention propose un procédé de gestion d'un groupe de données dans un système informatique comprenant :

- au moins un serveur de calcul pour accueillir une session informatique s'exécutant à l'aide d'un système d'exploitation disposant d'un index de déduplication et gérant l'accès à un espace de stockage de session ;
- un espace de stockage partagé ;
- un serveur d'administration de l'espace de stockage partagé exécutant un programme de gestion de données ;

la session informatique exécutant un programme de synchronisation mettant en œuvre les étapes suivantes :

- identifier une écriture d'au moins un groupe de données émis dans la session ;
- adresser une requête de synchronisation du groupe de données au programme de gestion de données du serveur d'administration ;
- le cas échéant, recevoir du programme de gestion de données une adresse de stockage du groupe de données dans l'espace de stockage partagé et accéder à l'index de déduplication pour y enregistrer l'adresse du groupe de données dans l'espace de stockage partagé ;

le programme de synchronisation mettant en œuvre également les étapes suivantes :

- recevoir du programme de gestion de données une acceptation de

synchronisation du groupe de données;

- fournir au programme de gestion de données le groupe de données.

Selon d'autres caractéristiques avantageuses et non limitatives de l'invention, prises seules ou selon toute combinaison techniquement réalisable :

- la session informatique exécute également un programme d'interception mettant en œuvre les étapes suivantes :
 - intercepter un appel en lecture d'au moins un groupe de données émis dans la session ;
 - accéder à l'index de déduplication et déterminer si le groupe de données est enregistré dans l'espace de stockage partagé ;
 - dans un tel cas, lire dans l'index de déduplication l'adresse du groupe de données dans l'espace de stockage partagé et rediriger l'appel en lecture vers cette adresse ;
 - dans le cas contraire, laisser passer l'appel en lecture pour qu'il soit traité par le système d'exploitation ;

- le programme d'interception met également en œuvre les étapes suivantes :
 - identifier un échec de lecture du groupe de données dans l'espace de stockage partagé ;
 - requérir au serveur d'administration une nouvelle adresse du groupe de données dans l'espace de stockage partagé;
 - recevoir la nouvelle adresse du groupe de données et accéder à l'index de déduplication pour enregistrer la nouvelle adresse du groupe de données ;
 - rediriger l'appel en lecture vers la nouvelle adresse du groupe de données dans l'espace de stockage partagé ;

Selon un autre aspect, l'invention propose un procédé de gestion d'un groupe de données dans un système informatique comprenant :

- au moins un serveur de calcul pour accueillir au moins une session informatique s'exécutant à l'aide d'un système d'exploitation disposant d'un index de déduplication et gérant l'accès à un espace de stockage de session;

- un espace de stockage partagé ;

- un serveur d'administration de l'espace de stockage partagé hébergeant un index de métadonnées et exécutant un programme de gestion de données ;

le programme de gestion de données mettant en œuvre les étapes suivantes :

- recevoir d'une session informatique une requête de synchronisation d'un groupe de données ;

- consulter l'index de métadonnées pour déterminer si le groupe de données est stocké dans l'espace de stockage partagé ;

- dans l'affirmative, retourner à la session informatique une adresse de stockage du groupe de données dans l'espace de stockage partagé;

- dans le cas contraire, incrémenter un compteur de requête de synchronisation du groupe de données et si le compteur excède un seuil déterminé ;

- adresser à au moins une session informatique une acceptation de synchronisation du groupe de données;

- recevoir de la session informatique le groupe de données et enregistrer le groupe de données dans une adresse de l'espace de stockage partagé ;

- accéder à l'index de métadonnées pour enregistrer l'adresse du groupe de données dans l'espace de stockage partagé.

Avantageusement, l'acceptation de synchronisation du groupe de données est adressée à une pluralité de sessions informatiques.

BREVE DESCRIPTION DES FIGURES

D'autres caractéristiques et avantages de l'invention ressortiront de la description détaillée de l'invention qui va suivre en référence aux figures annexées sur lesquels :

- La figure 1 représente une architecture d'un système informatique compatible avec un procédé conforme à l'invention ;
- La figure 2 représente les étapes mises en œuvre par un programme d'interception d'un procédé conforme à l'invention ;
- La figure 3 représente les étapes mises en œuvre par un programme de synchronisation d'un procédé conforme à l'invention.

DESCRIPTION DETAILLEE DE L'INVENTION

La figure 1 représente une architecture d'un système informatique 1 compatible avec un procédé de gestion d'un groupe de données conforme à l'invention.

Une pluralité de sessions informatiques 2 s'exécute sur au moins un serveur de calcul 4 du système informatique 1. Chaque session 2 est reliée à un client distant 3 avec lequel elle échange des informations de session par l'intermédiaire d'un réseau informatique, tel que le réseau Internet. Les serveurs de calcul 4 peuvent être configurés pour héberger une pluralité de machines virtuelles formant des environnements isolés les uns des autres et accueillant respectivement les sessions informatiques 2 d'utilisateur. Chaque session informatique 2 a donc accès, à l'aide d'un système d'exploitation, à toutes les ressources de la machine virtuelle qui émule parfaitement une machine réelle. La machine virtuelle met notamment à disposition de la session

informatique 2 un espace de stockage de session HDD (apparaissant typiquement dans la session 2 comme un disque dur conventionnel) dans lequel l'utilisateur de la session peut stocker l'ensemble de ses données et de ses applications. A cet effet, le système informatique 1 peut comprendre une pluralité de serveurs de stockage, reliée au serveur de calcul 4, pour accueillir les espaces de stockage de session HDD des sessions 2. Un système informatique 1 peut héberger plusieurs centaines voire plusieurs milliers de sessions informatiques 2.

Comme cela est bien connu en soi, le système d'exploitation de la machine virtuelle qui supporte la session informatique 2 administre les données enregistrées dans l'espace de stockage de session HDD à l'aide d'un système de gestion de fichiers. Il peut par exemple s'agir du système de gestion de fichiers NTFS.

Lorsqu'une application de la session 2 amorce une opération de lecture ou d'écriture d'un fichier ou d'une portion d'un fichier informatique (plus généralement désigné par « groupe de données » dans la présente description) dans l'espace de stockage de session HDD, le système d'exploitation fait appel au système de gestion de fichiers pour effectivement réaliser les opérations de lecture ou d'écriture des données sur le support physique ou logique formant l'espace de stockage de session HDD.

L'architecture informatique de la figure 1, comprend également un espace de stockage partagé DBE comprenant une pluralité d'autres serveurs de stockage en réseau, typiquement du type SAN ou NAS. Des données sont enregistrées dans l'espace de stockage partagé DBE sous le contrôle d'un serveur d'administration ADM du système informatique 1.

Dans le cadre de la présente invention, le système informatique 1 est configuré pour dédupliquer des groupes de données dupliqués

dans les espaces de stockage de session HDD. Il est en effet fréquent qu'une pluralité de sessions informatiques 2 exécute la même application ou manipule le même fichier de données, ou des applications ou des fichiers très similaires entre eux, c'est-à-dire partageant en commun une majorité de groupes de données. De manière conventionnelle, ces groupes de données communs sont enregistrés dans les serveurs de stockage de session HDD autant de fois que le nombre de sessions 2 exploitant ces groupes de données.

Le procédé de gestion des groupes de données dédupliqués qui fait l'objet de la présente description vise à éviter ou à limiter ce phénomène qui est consommateur d'espace de stockage. Plus précisément, le procédé identifie au moins un groupe de données dupliqué dans les espaces de stockage de session HDD et les remplace par un groupe de données unique de l'espace de stockage partagé DBE. Les sessions informatiques 2 accèdent (c'est-à-dire réalisent une opération de lecture ou d'écriture) de manière entièrement transparente aux groupes de données qu'elles exploitent, c'est-à-dire que les groupes de données dédupliqués apparaissent pour ces sessions informatiques 2 et pour les applications qui s'y exécutent comme tout autre groupe de données, enregistré dans l'espace de stockage de session HDD.

Pour atteindre cet objectif, le système d'exploitation de la machine virtuelle sur laquelle s'exécute une session informatique 2 qui gère l'accès à l'espace de stockage de session HDD dispose d'un index de déduplication UD. D'une manière générale, cet index UD permet d'identifier les groupes de données enregistrés dans l'espace de stockage partagé DBE, et de les distinguer de ceux enregistrés dans l'espace de stockage de session HDD. Dans ce but, l'index de déduplication UD associe une clé d'identification ID d'un groupe de données à une adresse

IP de ce groupe de données dans l'espace de stockage partagé DBE.

Les accès en lecture/écriture d'une session aux groupes de données enregistrés dans l'espace de stockage de session HDD peuvent être réalisés sans intervention particulière de l'application, et mis en œuvre de manière conventionnelle par le système d'exploitation de la machine virtuelle. Les accès en lecture/écriture d'une session aux groupes de données enregistrés dans l'espace de stockage partagé DBE sont quant à eux identifiés ou interceptés afin d'engager les traitements adaptés. A cet effet, la session informatique 2 (et plus précisément chaque session informatique 2 du système informatique 1) exécute en tâche de fond sur la machine virtuelle un programme d'interception INT et un programme de synchronisation SYNC qui mettent en œuvre des étapes qui seront décrites dans une section ultérieure de cette description.

L'index de déduplication UD est avantageusement mis en œuvre en exploitant les fonctionnalités du système de gestion de fichiers du système d'exploitation de la machine virtuelle sur laquelle s'exécute la session 2. Un tel système de gestion de fichiers comprend généralement une table de référence, composée d'une pluralité d'enregistrements, chaque enregistrement permettant d'associer un nom de fichier à un ou une pluralité de groupes de données et à une pluralité d'attributs visant à caractériser ces groupes de données.

Dans un exemple de mise en œuvre avantageux de l'invention, la clé d'identification ID d'un groupe de données et l'adresse IP de ce groupe de données dans l'espace de stockage partagé DBE forment des attributs de ce groupe de données dans l'enregistrement correspondant du système de gestion de fichiers. Dans cet exemple de mise en œuvre donc, l'index de

déduplication est parfaitement intégré au système de gestion de fichiers du système d'exploitation de la machine virtuelle sur laquelle s'exécute la session. La clé d'identification ID d'un groupe de données peut correspondre à une empreinte numérique de ce groupe, par exemple calculée à l'aide d'une fonction de hachage telle que la fonction MD5. Une telle fonction permet de représenter un groupe de données sous la forme d'une empreinte de 128 bits (32 caractères hexadécimaux) avec une très forte probabilité que deux groupes différents soient identifiés par deux empreintes différentes. Bien entendu, la présente invention n'est pas limitée à la création d'une clé d'identification de groupe par une telle fonction de hachage, et toute autre méthode permettant d'établir une clé unique par groupe de données peut convenir.

Pour permettre la mise en œuvre d'un procédé de gestion d'un groupe de données conforme à l'invention, le serveur d'administration ADM maintient un index de métadonnées MD. Cet index associe également, le cas échéant, la clé d'identification ID d'un groupe de données à son adresse IP de stockage dans l'espace de stockage partagé DBE. Lorsqu'un groupe de données n'est pas enregistré dans l'espace de stockage partagé DBE, l'index de métadonnées MD ne dispose pas d'entrée correspondant à la clé d'identification ID de ce groupe de données, ou cette entrée est associée à une adresse IP non renseignée.

L'index de métadonnées MD a pour fonction de recenser tous les groupes de données détenus dans l'espace de stockage partagé DBE. Cet index est administré par le programme de gestion de données DM qui s'exécute sur le serveur d'administration ADM. Les index de déduplication UD des sessions informatiques 2 visent quant à eux à répliquer de manière fidèle, mais partielle, l'index de métadonnées MD, uniquement pour les groupes de données auxquelles les sessions ont respectivement besoin d'accéder

(c'est-à-dire pour les groupes de données qui avaient vocation à être enregistrés dans les espaces respectifs de stockage de session HDD). Selon la présente invention, cette réplification de l'index de métadonnées MD dans les index de déduplication UD des sessions 2 est réalisé de manière asynchrone par les programmes de synchronisation SYNC s'exécutant sur les machines virtuelles des sessions 2 et par le programme de gestion des données DM s'exécutant sur le serveur d'administration ADM.

En référence aux figures 2 et 3, on détaille maintenant un mode de mise en œuvre des programmes d'interception INT, de synchronisation SYNC et de gestion de données DM d'un procédé conforme à l'invention.

Ainsi, la figure 2 représente sous forme d'un ordinogramme les traitements mis en œuvre par le programme d'interception INT, lorsqu'une application s'exécutant dans la session informatique 2 cherche à lire au moins un groupe de données enregistré dans l'espace de stockage de session HDD.

Les traitements du programme d'interception INT s'engagent par une première étape 10 interceptant un appel en lecture d'un groupe de données DB par l'application. Dans une étape suivante 20, le programme d'interception détermine la clé d'identification ID du groupe de données DB, par exemple en appliquant à ce groupe la fonction de hachage MD5.

Dans une étape 30, le programme d'interception INT consulte l'index de déduplications UD associé à la session 2 pour déterminer si ce groupe de données DB est marqué comme déduplicé, c'est-à-dire s'il a été enregistré dans l'espace de stockage partagé DBE.

Pour cela il suffit de vérifier que l'index UD présente bien une entrée correspondant à la clé d'identification ID du groupe de données DB en question, et que cette entrée est bien associée à une adresse IP de stockage dans l'espace de stockage partagé DBE.

Si c'est effectivement le cas, on récupère cette adresse IP au cours d'une étape suivante 40 et on redirige l'appel en lecture du groupe vers l'adresse IP dans l'espace de stockage partagé à l'aide de l'interface réseau de la machine virtuelle (étape 50 sur la figure 2). Cet appel est reçu par un contrôleur réseau de l'espace de stockage partagé DBE et le groupe de données correspondant est retourné au programme d'interception INT qui peut, dans une étape finale 60, remettre le groupe de données DB demandé à l'application l'ayant requis.

Dans le cas où la consultation de l'index de déduplication UD conduit à déterminer que le groupe de données DB n'est pas dédupliqué dans l'espace de stockage partagé DBE, le programme d'interception INT interrompt son intervention et laisse l'appel en lecture être traité de manière conventionnelle par le système d'exploitation à l'aide du système de gestion de fichiers. Le groupe de données est alors lu dans l'espace de stockage de session HDD et remis à l'application appelante (étape 41 et 60 de la figure 2).

On note que le traitement d'un appel en lecture d'un groupe de données par une application est traité de manière tout à fait transparente vis-à-vis de l'application appelante, de la session 2 ou de son utilisateur.

La figure 3 représente les traitements mis en œuvre par le programme de synchronisation SYNC et par le programme de gestion de données DM. Ces traitements peuvent être déclenchés de manière

asynchrone après qu'une application s'exécutant dans la session informatique 2 ait écrit au moins un groupe de données dans l'espace de stockage de session HDD. Ils peuvent également être déclenchés lorsqu'un groupe de données est nouvellement dédoublé, et donc enregistré dans une adresse de l'espace de stockage partagé DBE.

En référence à la figure 3, le programme de synchronisation SYNC s'engage au cours d'une première étape 100 d'identification d'une écriture d'au moins un groupe de données DB par une application, dans l'espace de stockage de session HDD. A cet effet, le système d'exploitation de la session 2 peut maintenir une liste ou un journal des opérations d'écriture auquel le programme de synchronisation SYNC peut accéder, de manière asynchrone avec ces opérations d'écriture.

Dans une étape suivante 110, le programme de synchronisation SYNC détermine la clé d'identification ID de ce groupe de données DB, par exemple en appliquant à ce groupe la fonction de hachage MD5.

Dans une troisième étape 120, le programme de synchronisation SYNC adresse une requête de synchronisation au programme de gestion de données DM du serveur d'administration ADM. Cette requête peut comprendre la transmission au programme de gestion de données DM de la clé d'identification ID du groupe de données afin que celui-ci puisse être identifié. La requête peut également comprendre d'autres attributs, tels qu'un identifiant de session U_ID, le nom du fichier correspondant au groupe de données DB, etc.

On peut prévoir, de manière optionnelle, une étape préalable 111 de filtrage au cours de laquelle on cherche à identifier si le groupe de données DB fait partie d'une liste d'exclusion

maintenue par le système d'exploitation de la machine virtuelle sur laquelle s'exécute la session 2. Si c'est bien le cas, le programme de synchronisation SYNC est interrompu.

A la suite de l'envoi de cette requête, au cours de la troisième étape 120, le programme de synchronisation SYNC interrompt son fil d'exécution. Il peut alors traiter une autre instance d'écriture d'un groupe de données, ou attendre la survenance d'un tel évènement.

Le programme de synchronisation SYNC s'exécutant dans une session 2 comprend d'autres fils d'exécution qui peuvent s'engager de manière asynchrone à celui qui vient d'être décrit. Ces fils d'exécution peuvent s'exécuter suite à l'envoi de la requête de synchronisation, dont le traitement par le serveur d'administration ADM sera détaillé dans une partie suivant de cette description.

Ainsi, et selon une première possibilité (SYNC_OK), le programme de gestion de données DM du serveur d'administration ADM peut retourner au programme de synchronisation SYNC une adresse IP de l'espace de stockage partagé DBE dans lequel le groupe de données DB doit être enregistré. Pratiquement, le message reçu par le programme de synchronisation SYNC (étape 130) peut comprendre la clé d'identification ID du groupe de données DB et son adresse IP afin que l'information puisse être décodée par ce programme. Dans une étape suivante 140, le programme de synchronisation SYNC accède à l'index de déduplication UD pour le mettre à jour, c'est-à-dire enregistrer dans cet index l'adresse IP de stockage associé au groupe de données DB dans l'espace de stockage partagé DBE et à la clé d'identification ID de ce groupe.

En conséquence, les futurs appels en lecture de ce groupe de données seront effectivement retournés vers l'adresse IP de

l'espace de stockage partagé DBE, comme cela a été illustré en relation avec la description de la figure 2.

Enfin, le groupe de données DB peut être effacé de l'espace de stockage de session HDD au cours d'une étape 150 d'effacement, tout en conservant bien entendu intacte l'entrée de l'index de déduplication UD associé à ce groupe de données.

Selon une autre possibilité, le serveur d'administration ADM peut retourner au programme de synchronisation SYNC une information NO_SYNC selon laquelle le groupe de données DB ne doit pas être dédupliqué. Comment on le verra dans la suite de cette description, cette possibilité se produit lorsque le serveur d'administration ADM détermine que le groupe de données DB n'est pas suffisamment dupliqué dans les espaces de stockage de session HDD pour être dédupliqué dans l'espace de stockage partagé DBE. La réception d'une telle information de non-déduplication NO_SYNC ne déclenche aucune action nécessaire et la communication de cette information par le programme de gestion de données DM est parfaitement optionnelle.

Selon encore une autre possibilité de réponse à la requête en synchronisation, qui se produit lorsque le groupe de données DB n'est pas encore dédupliqué, mais pour lequel le serveur d'administration ADM détermine qu'il serait souhaitable qu'il le soit, le serveur d'administration ADM adresse au programme de synchronisation SYNC une requête lui indiquant de fournir le groupe de données DB en question. Dans un tel cas, le programme de synchronisation SYNC fournit, au cours d'une étape de transmission 160, au serveur d'administration ADM ce groupe de données DB pour que celui-ci organise son enregistrement dans l'espace de stockage partagé DBE.

Ce transfert de données peut être mis en œuvre par l'ouverture d'une session de transfert partitionné d'un groupe de données entre le serveur d'administration ADM et la session utilisateur. Une telle session de transfert, bien connue en soi, permet de gérer de manière autonome le transfert d'un groupe de données qui peut être volumineux, en partitionnant ce groupe de données, et en traitant les anomalies qui peuvent se produire au cours du transfert (par exemple lorsque la session d'utilisateur 2 est interrompue).

Aussi, après l'envoi du groupe de données DB au serveur d'administration ADM, le fil d'exécution du programme de synchronisation SYNC s'interrompt. À ce stade, le groupe de données DB n'est donc pas encore marqué dans l'index de déduplication UD de la session informatique 2 comme étant déduplicé.

À tout moment, et de manière asynchrone avec les appels en écriture que les applications de la session informatique 2 peuvent initier, le programme de synchronisation SYNC peut recevoir des requêtes SYNC_OK du serveur d'administration ADM (et plus précisément des programmes de gestion de données DM s'exécutant sur ce serveur) indiquant qu'un groupe de données déterminé, identifié par sa clé d'identification ID, a bien été enregistré dans l'espace de stockage partagé DBE à une adresse IP de stockage déterminée.

À la réception de cette requête SYNC_OK au cours de l'étape de réception 130, le programme de synchronisation SYNC exécute, dans une étape suivante, la mise à jour de l'index de déduplication UD (étape 140) en enregistrant et en associant dans cet index la clé d'identification ID et l'adresse IP de stockage du groupe DB et l'efface de l'espace de stockage de session HDD (étape 150). En conséquence, les futurs appels en

lecture de ce groupe de données, interceptés par le programme d'interception INT tel que cela a été décrit en référence à la figure 2, sont effectivement redirigés vers l'adresse IP dans l'espace de stockage partagé DBE et ne sont donc pas lus directement dans l'espace de stockage de session HDD.

On complète la description de la figure 3 en détaillant les traitements mis en œuvre par le programme de gestion de données DM du serveur d'administration ADM.

Celui-ci reçoit les requêtes de synchronisation (étape 200 de réception du programme de gestion de données) qui lui sont adressées par les programmes de synchronisation SYNC des sessions informatiques 2, à l'occasion d'une écriture dans leurs espaces de stockage de sessions HDD respectifs. Ces requêtes sont respectivement accompagnées, comme on l'a vu, des clés d'identification ID des groupes de données pour permettre au programme de gestion de données DM d'identifier le groupe en question et éventuellement d'un identifiant de session U_ID.

Dans une étape suivante (étape 210), le programme de gestion de données DM du serveur administration consulte l'index de métadonnées MD pour déterminer si le groupe de données DB, identifié par sa clé d'identification ID, est bien stocké dans l'espace de stockage partagé DBE. Dans l'affirmative (SYNC_OK), il récupère (étape 220) et retourne (étape 230) à la session informatique 2 l'adresse IP de stockage du groupe en question DB accompagné par la clé d'identification ID pour permettre au programme de synchronisation SYNC de la session requérante U_ID de réaliser ses traitements.

Le programme de gestion des données DM du serveur d'administration ADM peut enregistrer dans l'index de métadonnées

MD l'information selon laquelle la session informatique 2 requérante U_ID a demandé un accès à l'adresse IP.

Si le programme de gestion des données DM détermine que le groupe de données n'est pas détenu dans l'espace de stockage partagé DBE, il incrémente un compteur de déduplication associé à la clé d'identification ID de ce groupe et enregistre l'identifiant de la session informatique 2 requérante U_ID. La valeur du compteur et l'identifiant de la session requérante U_ID peuvent être enregistrés dans des champs de l'index de métadonnées MD associé à la clé d'identification ID du groupe.

Dans une étape suivante 250, le programme de gestion des données DM détermine si le compteur de déduplication excède une valeur seuil prédéterminée. Si ce n'est pas le cas, il peut informer (NO_SYNC), de manière optionnelle, la session informatique 2 requérante que le groupe de données identifié par sa clé d'identification ne sera pas des doublés. Si le compteur de déduplication excède effectivement la valeur seuil prédéterminée, le programme de gestion des données DM déclenche les opérations visant à dédupliquer le groupe de données en question.

Ainsi dans une première étape 260 de ces opérations, il adresse une requête à au moins une et, avantageusement, à une pluralité de sessions informatiques 2 requérantes pour que celles-ci lui fournissent le groupe de données DB en question. Cette requête peut être accompagnée de la clé d'identification ID du groupe de données DB et/ou du nom de fichier associé à ce groupe de données, pour que le programme d'identification puisse retrouver le groupe DB en question. A cet effet, le programme de gestion de données DM peut consulter l'index de métadonnées MD pour déterminer les identifiants U_ID des sessions ayant préalablement requis la synchronisation du groupe de données DB.

Comme on l'a déjà décrit, cette étape 260 de transmission peut comprendre l'ouverture d'une ou d'une pluralité de sessions de transfert partitionné.

À la réception du groupe de données DB requis, le programme de gestion de données DM enregistre temporairement le groupe de données reçu. Il marque ce groupe de données comme nécessitant d'être dédoublé dans l'espace de stockage partagé DBE et, de manière asynchrone, des tâches de fond s'exécutant sur le serveur d'administration ADM procèdent à ce travail (étape 270). Lorsque cette opération est réalisée par une au moins des tâches de fond, le programme de gestion de données DM reçoit de ces tâches l'adresse IP de stockage dans l'espace de stockage partagé DBE du groupe de données. Il met alors à jour, dans une étape suivante 280, l'index de métadonnées MD en y enregistrant l'adresse de stockage IP dans l'entrée associée à la clé d'identification ID du groupe.

On note que le programme de gestion de données DM et plus généralement le serveur d'administration ADM n'a aucun contrôle sur l'exécution des sessions utilisateurs 2 qui s'exécutent sur les serveurs de calcul 4. Celles-ci peuvent donc s'interrompre à tout moment ou n'être pas disponibles pour contribuer à la mise en œuvre du procédé de gestion des groupes de données dédoublés. En adressant une pluralité de requêtes (par exemple plus de 10 requêtes, voire même 100 requêtes) aux sessions 2 qui disposent du groupe de données DB dans leur espace de stockage de session HDD, on s'assure avec un plus grand niveau de confiance d'obtenir au moins une copie de groupe de données au niveau du serveur d'administration ADM. D'une manière plus générale, les programmes de synchronisation SYNC, d'interception INT et de gestion de données DM qui mettent en œuvre un procédé de gestion conforme à l'invention sont sans état (« state less » selon l'expression anglo-saxonne consacrée) et fonctionnent dans

un mode maître (le serveur d'administration ADM)-esclaves (les sessions informatiques 2) qui sont tout à fait tolérantes au fonctionnement et/ou à la disponibilité intermittente des sessions informatiques 2.

Une fois le groupe de données DB déduplicé dans le système informatique 1, et au cours d'une étape suivante, le programme de gestion de données DM adresse à chacune des sessions 2 (plus particulièrement à leur programme de synchronisation) l'information (SYNC_OK) selon laquelle le groupe DB en question est maintenant déduplicé. Cette information peut consister à communiquer à ces programmes de synchronisation SYNC la clé d'identification ID du groupe et son adresse IP de stockage. On rappelle que ces sessions 2 sont recensées et identifiées dans l'index de métadonnées MD.

Le serveur d'administration ADM peuvent être amené à réorganiser la manière dont les groupes de données sont enregistrés dans l'espace de stockage partagé DM. Il peut ainsi s'agir de retirer un des dispositifs de stockage constituant l'espace de stockage partagé DBE pour une raison de maintenance ou de redistribuer les groupes de données sur ces dispositifs pour égaliser leur charge. Dans ce cas, les adresses IP des groupes de données déduplicés peuvent changer au cours du temps et ne plus correspondre à celles enregistrées dans les index de déduplication UD des sessions 2 s'exécutant dans le système informatique 1.

En conséquence, le procédé mis en œuvre par le programme d'interception INT d'une session informatique 2 peut identifier un échec en lecture d'un groupe de données auquel il souhaite accéder dans l'espace de stockage partagé DBE. Dans une telle situation, le programme d'interception INT engage une nouvelle étape dans laquelle il adresse une requête au programme de

gestion de données DM du serveur d'administration ADM pour obtenir une nouvelle adresse IP de stockage du groupe de données déduplicé dans l'espace de stockage partagé DBE. Cette requête comprend la clé d'identification du groupe et un identifiant U_ID de la session informatique 2 requérante. À réception de la nouvelle adresse requise, le programme d'interception INT de la session 2 requérante met à jour l'index de déduplication UD et renouvelle l'appel en lecture du groupe de données à la nouvelle adresse.

Pour augmenter la vitesse d'accès moyenne à ces données dans l'espace de stockage partagé DBE, l'architecture de la figure 1 peut prévoir de constituer cet espace à partir de dispositifs de stockage présentant des temps d'accès différents. L'espace de stockage partagé DBE peut par exemple être constitué de dispositifs à mémoire semi-conducteur émulant le fonctionnement de disques, mais dont les vitesses d'accès sont notoirement très rapides, et de serveurs de stockage plus conventionnels, présentant une grande capacité, mais un temps d'accès plus lent.

Le programme de gestion de données DM qui s'exécute sur le serveur d'administration ADM peut également avoir pour fonction de gérer la duplication ou le transfert des données dans l'espace de stockage partagé DBE de manière à rendre très rapidement accessible aux sessions 2 les données les plus exploitées, dans l'objectif général de minimiser un temps d'accès moyen aux données partagées.

Bien entendu l'invention n'est pas limitée au mode de mise en œuvre décrit et on peut y apporter des variantes de réalisation sans sortir du cadre de l'invention tel que défini par les revendications.

Ainsi, un procédé conforme à l'invention peut être mis en œuvre dans une architecture informatique bien différente de celle présentée sur la figure 1. Il peut s'agir d'une architecture plus conventionnelle, dans laquelle des dispositifs de calcul tels que des ordinateurs personnels sont reliés en réseau. On peut également dans ce cas, et de manière tout à fait transparente aux utilisateurs et aux applications s'exécutant sur ces dispositifs de calcul, dédupliquer les groupes de données présents sur une pluralité de ces dispositifs de calculs pour les enregistrer dans un espace de stockage partagé administré par un serveur d'administration de cet espace de stockage.

REVENDICATIONS

1. Procédé de gestion d'un groupe de données (DB) dans un système informatique comprenant :

- au moins un serveur de calcul (4) pour accueillir une session informatique (2) s'exécutant à l'aide d'un système d'exploitation disposant d'un index de déduplication (UD) et gérant l'accès à un espace de stockage de session (HDD) ;
- un espace de stockage partagé (DBE) ;
- un serveur d'administration (ADM) de l'espace de stockage partagé (DBE) exécutant un programme de gestion de données (DM) ;

la session informatique (2) exécutant un programme de synchronisation (SYNC) mettant en œuvre les étapes suivantes :

- identifier une écriture d'au moins un groupe de données (DB) émis dans la session (2) ;
- adresser une requête de synchronisation du groupe de données (BD) au programme de gestion de données (DM) du serveur d'administration (ADM) ;
- le cas échéant, recevoir du programme de gestion de données (DM) une adresse de stockage du groupe de données (IP) dans l'espace de stockage partagé (DBE) et accéder à l'index de déduplication (UD) pour y enregistrer l'adresse du groupe de données (IP) dans l'espace de stockage partagé (DBE) ;

le programme de synchronisation (SYNC) mettant en œuvre également les étapes suivantes :

- recevoir du programme de gestion de données (DM) une acceptation de synchronisation du groupe de données (DB) ;
- fournir au programme de gestion de données (DM) le groupe de données (DB).

2. Procédé de gestion selon la revendication précédente dans lequel la session informatique (2) exécute également un programme d'interception (INT) mettant en œuvre les étapes suivantes :

- intercepter un appel en lecture d'au moins un groupe de

données (DB) émis dans la session (2) ;

- accéder à l'index de déduplication (UD) et déterminer si le groupe de données (DB) est enregistré dans l'espace de stockage partagé (DBE) ;

- dans un tel cas, lire dans l'index de déduplication (UD) l'adresse du groupe de données (IP) dans l'espace de stockage partagé (DBE) et rediriger l'appel en lecture vers cette adresse (IP) ;

- dans le cas contraire, laisser passer l'appel en lecture pour qu'il soit traité par le système d'exploitation.

3. Procédé de gestion selon la revendication précédente dans lequel le programme d'interception (INT) met également en œuvre les étapes suivantes :

- identifier un échec de lecture du groupe de données (DB) dans l'espace de stockage partagé (DBE) ;

- requérir au serveur d'administration (ADM) une nouvelle adresse du groupe de données dans l'espace de stockage partagé (DBE) ;

- recevoir la nouvelle adresse du groupe de données et accéder à l'index de déduplication (UD) pour enregistrer la nouvelle adresse du groupe de données ;

- rediriger l'appel en lecture vers la nouvelle adresse du groupe de données (DB) dans l'espace de stockage partagé (DBE).

4. Procédé de gestion d'un groupe de données (DB) dans un système informatique comprenant :

- au moins un serveur de calcul (4) pour accueillir au moins une session informatique (2) s'exécutant à l'aide d'un système d'exploitation disposant d'un index de déduplication (UD) et gérant l'accès à un espace de stockage de session (HDD) ;

- un espace de stockage partagé (DBE) ;

- un serveur d'administration (ADM) de l'espace de stockage partagé (DBE) hébergeant un index de métadonnées (MD) et exécutant un programme de gestion de données (DM) ;

le programme de gestion de données (DM) mettant en œuvre

les étapes suivantes :

- recevoir d'une session informatique (2) une requête de synchronisation d'un groupe de données (DB) ;
- consulter l'index de métadatas (MD) pour déterminer si le groupe de données (DB) est stocké dans l'espace de stockage partagé (DBE) ;
- dans l'affirmative, retourner à la session informatique (2) une adresse de stockage du groupe de données (IP) dans l'espace de stockage partagé (DBE) ;
- dans le cas contraire, incrémenter un compteur de requête de synchronisation du groupe de données (DB) et si le compteur excède un seuil déterminé ;
- adresser à au moins une session informatique (2) une acceptation de synchronisation du groupe de données (DB) ;
- recevoir de la session informatique le groupe de données (DB) et enregistrer le groupe de données (DB) dans une adresse (IP) de l'espace de stockage partagé (DBE) ;
- accéder à l'index de métadatas (MD) pour enregistrer l'adresse du groupe de données (IP) dans l'espace de stockage partagé.

5. Procédé de gestion d'un groupe de données (DB) selon la revendication précédente dans lequel l'acceptation de synchronisation du groupe de données (DB) est adressée à une pluralité de sessions informatiques (2).

Fig. 1

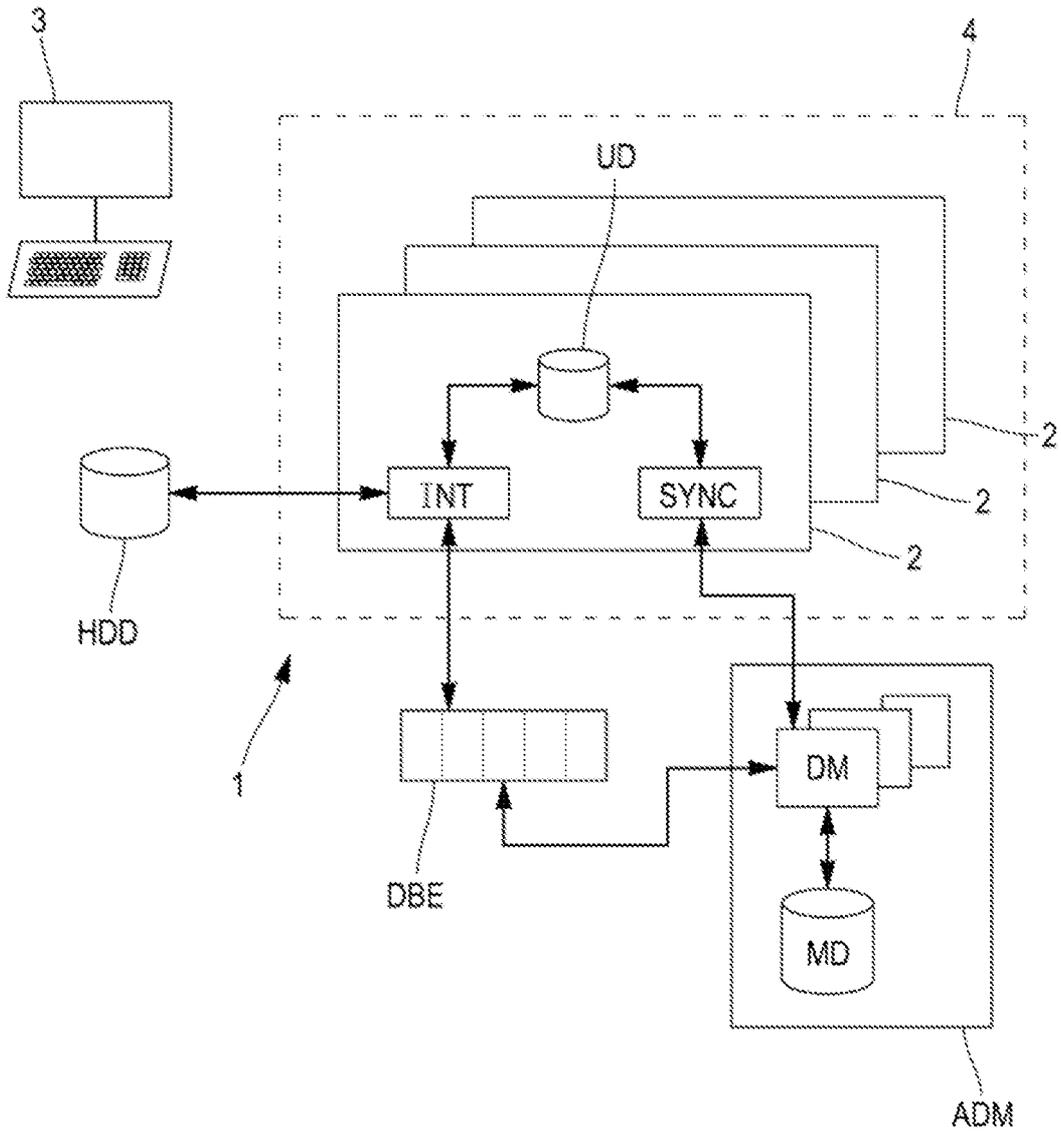


Fig. 2

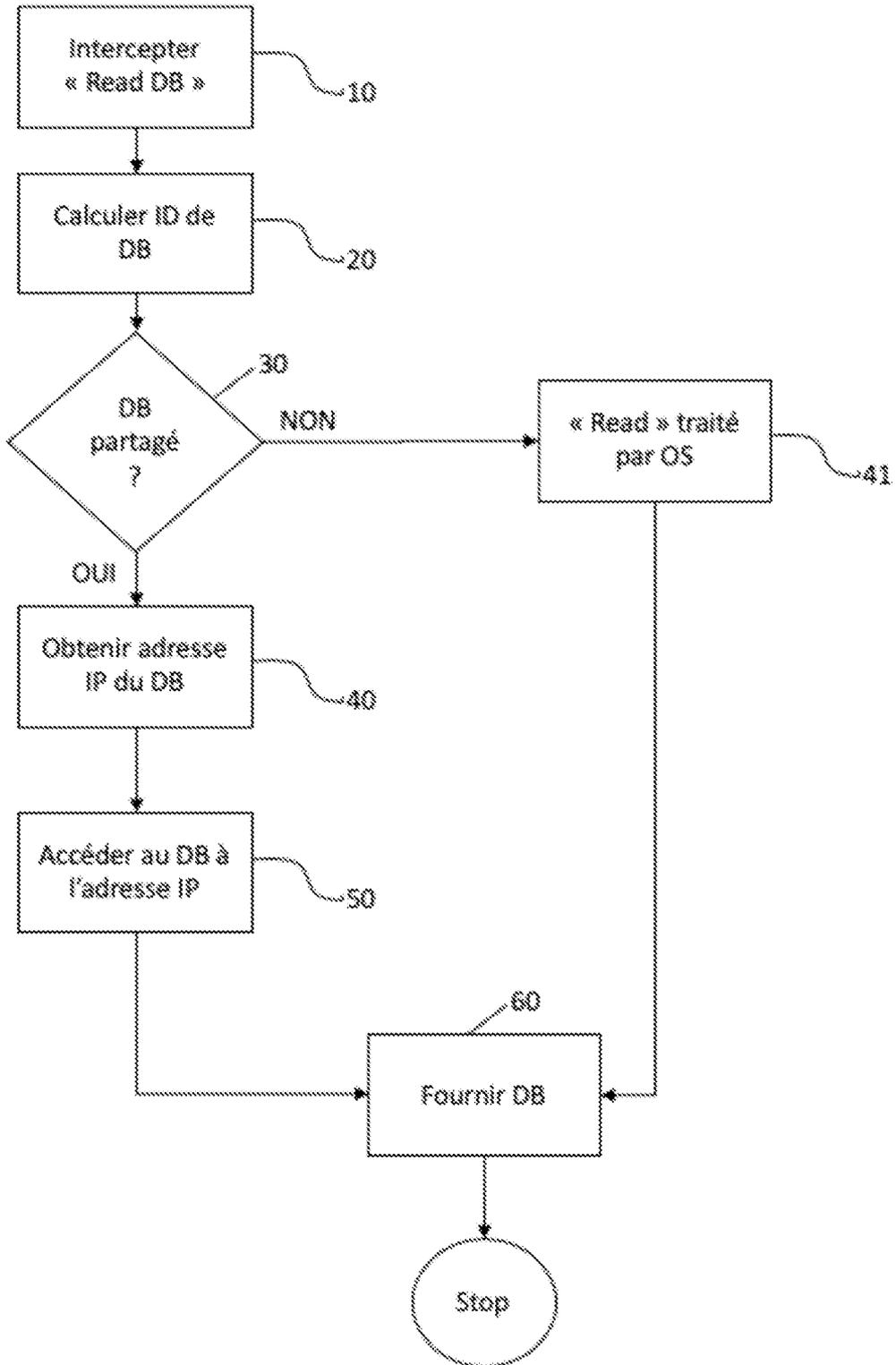
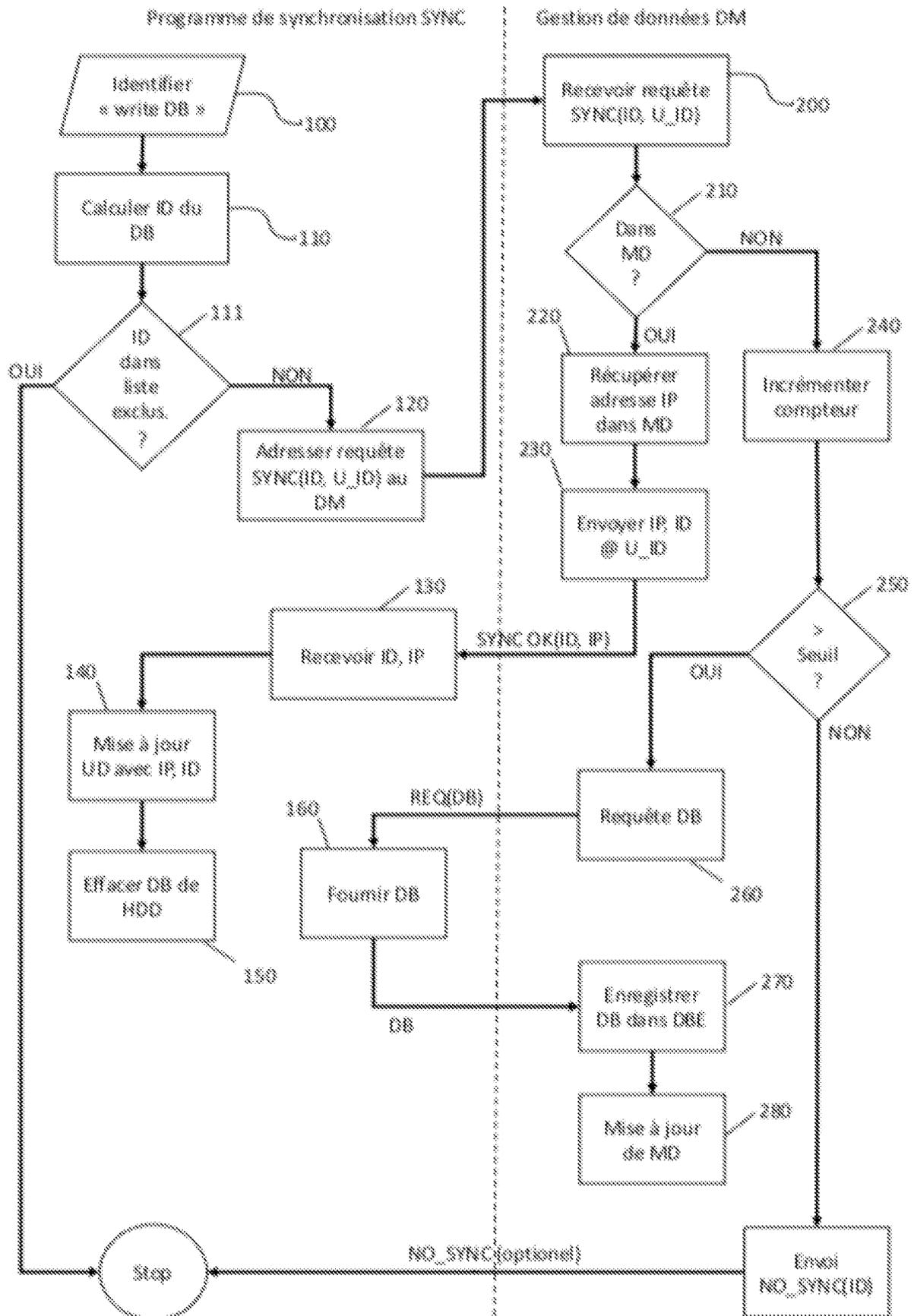


Fig. 3



INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR2020/051513

A. CLASSIFICATION OF SUBJECT MATTER <i>G06F 3/06</i> (2006.01)i; <i>G06F 11/14</i> (2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 2012221817 A1 (YUEH JEDIDIAH [US]) 30 August 2012 (2012-08-30) figures 1,4,7-9 paragraph [0025] - paragraph [0028] paragraph [0043] - paragraph [0067]	1-3 4,5
X A	US 10372683 B1 (WANG XUAN [US] ET AL) 06 August 2019 (2019-08-06) figures 1,5 column 9, line 11 - line 40	4,5 1-3
A	US 2015074065 A1 (CHRIST ACHIM [DE] ET AL) 12 March 2015 (2015-03-12) figure 3 paragraph [0022] - paragraph [0034] paragraph [0076] - paragraph [0083]	1-5
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 04 December 2020		Date of mailing of the international search report 14 December 2020
Name and mailing address of the ISA/EP European Patent Office p.b. 5818, Patentlaan 2, 2280 HV Rijswijk Netherlands Telephone No. (+31-70)340-2040 Facsimile No. (+31-70)340-3016		Authorized officer Alliot, Sylvain Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/FR2020/051513

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2012221817	A1	30 August 2012	US	8190835	B1	29 May 2012
				US	8880824	B1	04 November 2014
				US	2012221817	A1	30 August 2012

US	10372683	B1	06 August 2019	NONE			

US	2015074065	A1	12 March 2015	GB	2518158	A	18 March 2015
				US	2015074065	A1	12 March 2015

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2020/051513

A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. G06F3/06 G06F11/14 ADD.		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE		
Documentation minimale consultée (système de classification suivi des symboles de classement) G06F		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	US 2012/221817 A1 (YUEH JEDIDIAH [US]) 30 août 2012 (2012-08-30)	1-3
A	figures 1,4,7-9 alinéa [0025] - alinéa [0028] alinéa [0043] - alinéa [0067] -----	4,5
X	US 10 372 683 B1 (WANG XUAN [US] ET AL) 6 août 2019 (2019-08-06)	4,5
A	figures 1,5 colonne 9, ligne 11 - ligne 40 -----	1-3
A	US 2015/074065 A1 (CHRIST ACHIM [DE] ET AL) 12 mars 2015 (2015-03-12) figure 3 alinéa [0022] - alinéa [0034] alinéa [0076] - alinéa [0083] -----	1-5
<input type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités:		
"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée		"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets
Date à laquelle la recherche internationale a été effectivement achevée 4 décembre 2020		Date d'expédition du présent rapport de recherche internationale 14/12/2020
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Alliot, Sylvain

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/FR2020/051513

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2012221817 A1	30-08-2012	US 8190835 B1	29-05-2012
		US 8880824 B1	04-11-2014
		US 2012221817 A1	30-08-2012

US 10372683 B1	06-08-2019	AUCUN	

US 2015074065 A1	12-03-2015	GB 2518158 A	18-03-2015
		US 2015074065 A1	12-03-2015
