

(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
01 octobre 2020 (01.10.2020)



(10) Numéro de publication internationale
WO 2020/193761 A1

(51) Classification internationale des brevets :
G06F 16/27 (2019.01)

VEDY, Philippe ; 44 avenue de la République CS50010,
92326 Chatillon (FR).

(21) Numéro de la demande internationale :
PCT/EP2020/058756

(74) Mandataire : VIDON BREVETS & STRATÉGIE ; P.
O. Box 90333, Technopôle Atalante, 16B rue de Jouanet,
35703 Rennes Cedex 7 (FR).

(22) Date de dépôt international :
27 mars 2020 (27.03.2020)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
1903250 28 mars 2019 (28.03.2019) FR

(71) Déposant : ORANGE [FR/FR] ; 78 rue Olivier de Serres,
75015 Paris (FR).

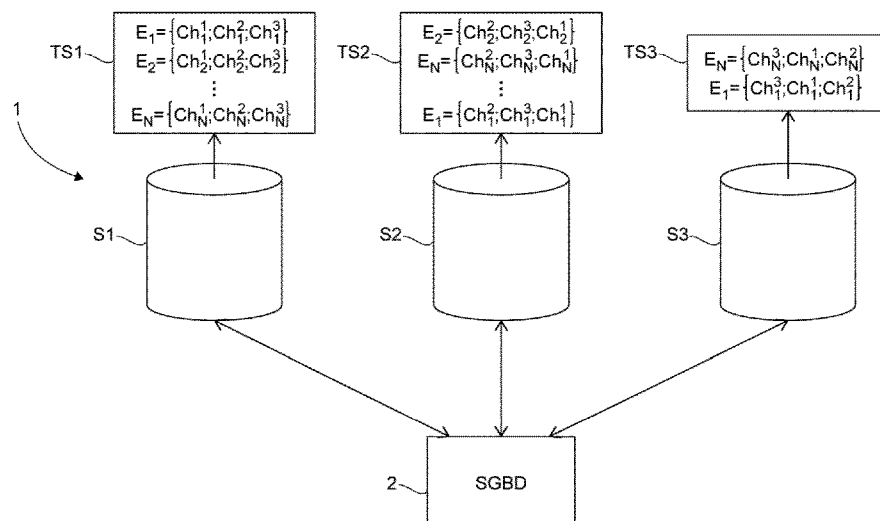
(72) Inventeurs : OBAME MEYE, Pierre ; 44 avenue de la Ré-
publique CS50010, 92326 Chatillon (FR). RAIPIN PAR-

(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, 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: SYSTEM FOR STORING DATA REDUNDANTLY, CORRESPONDING METHOD AND COMPUTER PROGRAM

(54) Titre : SYSTÈME DE STOCKAGE REDONDANT DE DONNÉES, PROCÉDÉ ET PROGRAMME D'ORDINATEUR
CORRESPONDANTS

[Fig 1]



(57) Abstract: The invention relates to a system for storing (1) data redundantly, comprising a plurality of data storage devices (S1, S2, S3), the data being structured in the form of a plurality of records (E₁, E₂,... E_N). According to the invention, a first device (S1) of the plurality of storage devices stores the records according to a first ordered sequence of records (TS1) and at least one second device (S2, S3) of the plurality of storage devices stores at least one subset of the records according to at least one second ordered sequence of records (TS2, TS3) which is separate from the first ordered sequence for the at least one subset.

(57) Abrégé : L'invention concerne un système de stockage (1) redondant de données, comprenant une pluralité de dispositifs de stockage de données (S1, S2, S3), les données étant structurées sous forme d'une pluralité d'enregistrements (E₁, E₂,... E_N). Selon l'invention, un premier dispositif (S1) de la pluralité de dispositifs de stockage

[Suite sur la page suivante]

(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))

stocke les enregistrements selon une première séquence ordonnée d'enregistrements (TS1) et au moins un deuxième dispositif (S2, S3) de la pluralité de dispositifs de stockage stocke au moins un sous-ensemble des enregistrements selon au moins une deuxième séquence ordonnée d'enregistrements (TS2, TS3) distincte de la première séquence ordonnée pour ledit au moins un sous-ensemble.

DESCRIPTION

TITRE : Système de stockage redondant de données, procédé et programme d'ordinateur correspondants.

Domaine technique

5 Le domaine de l'invention est celui du stockage de données ou d'informations dans des dispositifs de stockage, sous forme de bases de données. Plus précisément, l'invention concerne une technique de stockage redondant de données.

Art antérieur

10 Une base de données permet de stocker sous forme structurée, et d'accéder par requête, à un ensemble de données ou d'informations en rapport avec un thème ou une activité. Un tel stockage s'opère le plus souvent sur un ou plusieurs serveurs de stockage, ou disques de stockage, sous forme informatisée.

Afin d'améliorer la rapidité et la fiabilité d'accès aux données contenues dans une base de données, il a été proposé de réaliser un stockage redondant des données : le contenu de la base
15 de données est répliqué intégralement une ou plusieurs fois, et chacune des copies identiques de ce contenu est stockée sur un serveur ou support de stockage distinct. Ces dispositifs de stockage peuvent être ou non géographiquement distribués.

Ce type de réplication présente le double avantage suivant :

- d'une part, il améliore la disponibilité de la base de données. En effet, l'existence de
20 plusieurs copies identiques du contenu de la base de données permet de répondre à un plus grand nombre de requêtes d'accès simultanées aux données de la base ;
- d'autre part, il améliore la durabilité de la base de données, en accroissant la résistance aux pannes qui pourraient affecter l'un des dispositifs de stockage. Même en cas de défaillance de l'un des dispositifs rendant impossible l'accès à la copie qu'il contient, il reste en effet possible de
25 répondre aux requêtes d'accès aux données de la base, en interrogeant un autre dispositif détenant une autre copie identique du contenu de la base de données.

Si cette technique de stockage redondant des données est intéressante, elle pourrait cependant être optimisée selon plusieurs aspects.

Tout d'abord, ce stockage redondant est consommateur en ressources de stockage, tant sur le
30 plan énergétique, qu'en termes d'espace de stockage. Il existe donc un besoin d'une technique de stockage redondant de données qui soit plus économe en ressources.

Ensuite, si la redondance des données permet d'améliorer leur disponibilité, les performances d'accès aux données pourraient encore être améliorées, en réduisant le temps nécessaire à leur récupération depuis les dispositifs de stockage. Il existe donc un besoin d'une technique de
35 stockage redondant des données qui présente des performances accrues d'accès aux données.

En d'autres termes, il existe un besoin d'une technique de stockage redondant de données, qui soit aussi durable que les techniques de l'art antérieur, mais qui soit plus économe, plus performante, tout en offrant une disponibilité accrue par rapport aux solutions existantes.

Exposé de l'invention

5 L'invention répond à ce besoin en proposant un système de stockage redondant de données, comprenant une pluralité de dispositifs de stockage de données, les données étant structurées sous forme d'une pluralité d'enregistrements. Un premier dispositif de la pluralité de dispositifs de stockage stocke les enregistrements selon une première séquence ordonnée d'enregistrements et au moins un deuxième dispositif de la pluralité de dispositifs de stockage stocke au moins un sous-ensemble des enregistrements selon au moins une deuxième séquence ordonnée d'enregistrements distincte de la première séquence ordonnée pour le ou les sous-ensemble(s).

Ainsi, l'invention repose sur une approche tout à fait nouvelle et inventive du stockage redondant de données. En effet, une telle technique de stockage redondant de données repose, 15 classiquement, sur une réplication du contenu de la base de données, mais sur un stockage du contenu répliqué en suivant un ordre d'enregistrements différent de celui du contenu initial. Ainsi, les données sont stockées différemment dans les différents dispositifs de stockage. On améliore ainsi les performances d'accès aux données : en effet, l'ordre de stockage des enregistrements peut influencer sur le temps d'accès aux données qu'ils contiennent, en fonction du type de la requête formulée pour accéder aux données. Offrir plusieurs séquences ordonnées distinctes d'enregistrements, sur différents dispositifs de stockage, permet de maximiser les probabilités de disposer de la séquence ordonnée d'enregistrement la plus adaptée à tout type de requête d'accès reçue par le système.

En outre, la technique de stockage redondant selon un mode de réalisation de l'invention propose, sur au moins certains dispositifs de stockage du système, de ne répliquer qu'une partie 25 des enregistrements (par exemple ceux qui font le plus souvent l'objet de requêtes d'accès), afin d'économiser les ressources du système, en évitant de répliquer inutilement certains enregistrements, par exemple parce qu'ils font l'objet de peu de demandes de consultation. Une telle technique de stockage redondant des données est donc plus économe en ressources que les techniques antérieures.

30 Selon un mode de réalisation particulier, chaque enregistrement comporte au moins deux champs dans chacun desquels est enregistré une donnée : la première séquence ordonnée d'enregistrements est obtenue par tri des enregistrements à partir des données enregistrées dans un premier de ces champs, et la (ou les) deuxième séquence ordonnée d'enregistrements est

obtenue par tri des enregistrements à partir des données enregistrées dans au moins un deuxième de ces deux champs.

On considère par exemple un jeu de données dont chaque enregistrement contient un nom dans un premier champ et une date de naissance dans un deuxième champ. Un premier dispositif de stockage stocke ce jeu de données selon une première séquence ordonnée d'enregistrements obtenue par un tri par ordre alphabétique des noms contenus dans le premier champ ; un deuxième dispositif de stockage stocke ce jeu de données selon une deuxième séquence ordonnée d'enregistrements obtenue par un tri croissant ou décroissant des dates de naissance contenues dans le deuxième champ.

On comprend qu'à réception d'une requête visant à obtenir tous les enregistrements associés aux individus dont le nom commence par la lettre A, l'accès au premier dispositif de stockage offrira une performance accrue par rapport au deuxième dispositif de stockage, en termes de temps d'accès aux données, dans la mesure où l'accès aux données se fera par lecture séquentielle sur le disque de stockage. En revanche, en cas d'accès aux enregistrements stockés par le deuxième dispositif de stockage, dans lequel les enregistrements sont triés par date de naissance, il sera nécessaire, pour obtenir les données requises, de parcourir toute la base de données d'enregistrement si aucun index n'existe, ce qui est beaucoup plus lent ; si un index existe, les données requises seront obtenues par des lectures aléatoires sur le disque, qui sont, par nature, beaucoup plus lentes que les lectures séquentielles.

La séquence ordonnée d'enregistrements proposée par le deuxième dispositif de stockage, à savoir un tri par date de naissance des enregistrements, sera en revanche bien adaptée pour obtenir rapidement les données en réponse à une requête visant à déterminer tous les enregistrements associés aux individus nés entre 1990 et 2000.

Ainsi, selon un aspect particulier, un tel système de stockage comprend un système de gestion de base de données configuré pour déterminer un dispositif de stockage à interroger, parmi le premier et le ou les deuxième(s) dispositifs de stockage, en fonction d'une requête de donnée reçue par le système, pour optimiser un temps de récupération de la donnée requise.

Un tel système de gestion de base de données (ou SGBD) analyse la requête reçue et pointe vers le dispositif de stockage du système dont la séquence ordonnée d'enregistrements est la plus adaptée à une récupération rapide des données requises. On améliore ainsi les performances d'accès aux données stockées.

Selon un autre aspect, un tel système de gestion de base de données est également configuré pour déterminer le ou les sous-ensemble(s) d'enregistrements à stocker dans le ou les deuxième(s) dispositif(s) de stockage, en fonction d'un nombre et/ou d'une fréquence de requêtes portant sur les enregistrements du sous-ensemble.

Par exemple, pour une base de données contenant des enregistrements relatifs à des joueurs de football, le SGBD peut choisir de répliquer, sur un dispositif de stockage, uniquement les enregistrements relatifs aux dix plus grandes stars de la discipline sportive, qui font le plus souvent l'objet de requêtes. L'espace de stockage nécessaire est avantageusement restreint par rapport à celui nécessaire pour une réplique intégrale de la base de données ; la disponibilité des données de la base reste cependant élevée, dans la mesure où la majorité des requêtes reçues porte sur ce sous-ensemble de stars du foot.

Selon encore un autre aspect, un tel système de gestion de base de données est également configuré pour déterminer un nombre de dispositifs de stockage de la pluralité de dispositifs de stockage dans lesquels stocker les enregistrements et des séquences ordonnées d'enregistrements distinctes associées, en fonction d'un type et/ou d'une volumétrie de requêtes susceptibles d'être reçues.

Ainsi, les champs des enregistrements sur lesquels opérer les tris afin de construire des séquences ordonnées d'enregistrements, ainsi que le nombre de répliques des enregistrements à stocker sur différents dispositifs de stockage, sont choisis de manière statique, par rapport aux types de requêtes attendues et à leur volumétrie estimée.

Selon encore un autre aspect, un tel système de gestion de base de données est également configuré pour émettre une commande de stockage d'au moins un sous-ensemble des enregistrements selon une séquence ordonnée d'enregistrements vers un dispositif de stockage de la pluralité de dispositifs de stockage, en fonction d'un type et/ou d'une volumétrie de requêtes reçues.

Ainsi, les choix des champs sur lesquels opérer les tris et du nombre de répliques des enregistrements s'opère de manière dynamique, ce qui permet d'améliorer la disponibilité et la performance du système de stockage sur au moins une partie des données contenues dans le système, et ce, de manière sélective pendant le cycle de vie des données et de la base de données. Ceci est particulièrement avantageux par rapport aux systèmes de stockage de l'art antérieur, selon lesquels le niveau de copie ou réplique est global pour toutes les données enregistrées.

L'invention concerne également un procédé de stockage redondant de données, dans une pluralité de dispositifs de stockage de données, les données étant structurées sous forme d'une pluralité d'enregistrements. Un tel procédé de stockage comprend :

- une étape de stockage des enregistrements dans un premier dispositif de la pluralité de dispositifs de stockage selon une première séquence ordonnée d'enregistrements ;
- une étape de stockage, dans au moins un deuxième dispositif de la pluralité de dispositifs de stockage d'au moins un sous-ensemble des enregistrements selon au moins une deuxième

séquence ordonnée d'enregistrements distincte de la première séquence ordonnée pour ce ou ces sous-ensemble(s).

Selon un aspect particulier, chaque enregistrement comportant au moins deux champs dans chacun desquels est enregistré une donnée, la première séquence ordonnée d'enregistrements est obtenue par tri des enregistrements à partir des données enregistrées dans un premier champ, et la ou les deuxième(s) séquence(s) ordonnée(s) d'enregistrements est obtenue par tri des enregistrements à partir des données enregistrées dans au moins un deuxième champ.

Selon encore un aspect particulier, sur réception d'une requête de donnée, un tel procédé de stockage comprend une étape de détermination d'un dispositif de stockage à interroger, parmi le premier et le ou les deuxième(s) dispositifs de stockage, en fonction de la requête de donnée reçue, pour optimiser un temps de récupération de la donnée requise.

L'invention concerne également un produit programme d'ordinateur comprenant des instructions de code de programme pour la mise en œuvre d'un procédé tel que décrit précédemment ou l'exécution des fonctionnalités d'un système tel que décrit précédemment, lorsqu'il est exécuté par un processeur.

L'invention vise également un support d'enregistrement lisible par un ordinateur sur lequel est enregistré un programme d'ordinateur comprenant des instructions de code de programme pour l'exécution des étapes du procédé de stockage selon l'invention tel que décrit ci-dessus.

Un tel support d'enregistrement peut être n'importe quelle entité ou dispositif capable de stocker le programme. Par exemple, le support peut comporter un moyen de stockage, tel qu'une ROM, par exemple un CD ROM ou une ROM de circuit microélectronique, ou encore un moyen d'enregistrement magnétique, par exemple une clé USB ou un disque dur.

D'autre part, un tel support d'enregistrement peut être un support transmissible tel qu'un signal électrique ou optique, qui peut être acheminé via un câble électrique ou optique, par radio ou par d'autres moyens, de sorte que le programme d'ordinateur qu'il contient est exécutable à distance. Le programme selon l'invention peut être en particulier téléchargé sur un réseau par exemple le réseau Internet.

Alternativement, le support d'enregistrement peut être un circuit intégré dans lequel le programme est incorporé, le circuit étant adapté pour exécuter ou pour être utilisé dans l'exécution du procédé de contrôle d'affichage précité.

Le procédé de stockage et le programme d'ordinateur correspondants précités présentent au moins les mêmes avantages que ceux conférés par le système de stockage selon la présente invention.

Présentation des figures

D'autres buts, caractéristiques et avantages de l'invention apparaîtront plus clairement à la lecture de la description suivante, donnée à titre de simple exemple illustratif, et non limitatif, en relation avec les figures, parmi lesquelles :

[Fig 1] présente un synoptique général d'un système de stockage redondant de données selon un mode de réalisation de l'invention ;

[Fig 2] illustre sous forme d'organigramme les différentes étapes de la phase d'initialisation du système de stockage redondant de Fig 1 ;

[Fig 3] décrit le processus de stockage d'un enregistrement dans un système de stockage redondant de données selon un mode de réalisation de l'invention ;

[Fig 4] présente le processus de lecture d'un enregistrement dans un système de stockage redondant de données selon un mode de réalisation de l'invention ;

[Fig 5] présente la structure matérielle d'un système de gestion de base de données SGBD selon un mode de réalisation de l'invention.

Description détaillée de modes de réalisation de l'invention

Le principe général de l'invention repose sur un stockage redondant de données, i.e. un stockage de plusieurs copies d'une même base de données sur plusieurs dispositifs de stockage distincts, mais dans lequel les différentes copies de la base de données ne sont pas strictement identiques, mais de formes différentes, notamment en termes d'ordre des enregistrements de la base sur le dispositif de stockage.

Selon un aspect secondaire, sur au moins un dispositif du système de stockage, on ne stocke qu'une copie des données les plus fréquemment requises, et non de l'intégralité de la base de données.

On présente désormais, en relation avec Fig 1, un synoptique général d'un système de stockage redondant de données 1 selon un mode de réalisation de l'invention.

Un tel système de stockage 1 comprend une pluralité de dispositifs de stockage référencés S1 à S3, dont seulement trois exemples ont été représentés à titre illustratif sur Fig 1. Le nombre de dispositifs de stockage n'est bien sûr pas limité à trois, et tout autre nombre de dispositifs de stockage supérieur ou égal à deux peut également être envisagé. Notamment, plus le volume de données à stocker est élevé, plus l'accès aux données est fréquemment sollicité, plus le nombre de champs de chaque enregistrement est élevé, et plus il est avantageux de disposer d'un nombre important de supports de stockage distincts, ceci, afin d'améliorer la disponibilité et la durabilité de la base de données.

Les dispositifs de stockage S1 à S3 peuvent être localisés dans un même centre de traitement de données (en anglais « data center »), ou être géographiquement distribués, le cas échéant localisés dans des pays différents.

En outre, de tels dispositifs de stockage peuvent être des serveurs de stockage, des disques durs, des mémoires flash, etc.

Un système de gestion de base de données, ou SGBD, 2 permet de manipuler le contenu des bases de données stockées dans les dispositifs de stockage S1 à S3. Un tel système de gestion de base de données 2 se présente sous forme d'un ensemble de logiciels, et sert à effectuer les opérations ordinaires telles que rechercher, ajouter ou supprimer des enregistrements (*Create, Read, Update, Delete* abrégé *CRUD*), manipuler les index, créer ou copier des bases de données.

On considère par exemple une base de données contenant un ensemble d'enregistrements référencés E_1 à E_N . Chacun de ces enregistrements E_i comprend une pluralité de champs (trois dans l'exemple de Fig 1), notés Ch_i^1 , Ch_i^2 et Ch_i^3 . Par exemple, chaque enregistrement E_i correspond à un individu, dont le nom est enregistré dans un premier champ Ch_i^1 , dont le prénom est contenu dans un deuxième champ Ch_i^2 , et dont la date de naissance est stockée dans un troisième champ Ch_i^3 .

Le premier serveur de stockage S1 est associé à un premier type de stockage TS1, selon lequel les enregistrements E_1 à E_N sont ordonnés par ordre croissant des valeurs des données contenues dans le premier champ Ch_i^1 , soit dans cet exemple par ordre alphabétique du nom. Ainsi, le premier type de stockage TS1 correspond à une séquence ordonnée $(E_1, E_2, \dots, E_N)$.

Le deuxième serveur de stockage S2 est associé à un deuxième type de stockage TS2, selon lequel les enregistrements E_1 à E_N sont ordonnés par ordre croissant des valeurs des données contenues dans le deuxième champ Ch_i^2 , soit dans cet exemple par ordre alphabétique du prénom. Ainsi, le deuxième type de stockage TS2 correspond par exemple à une séquence ordonnée $(E_2, E_N, \dots, E_1)$.

L'intégralité des enregistrements E_1 à E_N est stockée sur chacun des deux serveurs S1 et S2, ce qui permet d'assurer la durabilité du système de stockage : en effet, en cas de panne de l'un des deux serveurs, par exemple S1, les données restent accessibles par adressage d'une requête au serveur S2.

Cependant, on comprend qu'une requête visant à obtenir une liste d'individus dont le nom de famille commence par la lettre A sera plus rapidement traitée si le SGBD 2 l'adresse au premier serveur S1, qui offre un type de stockage TS1 par ordre alphabétique de nom : une simple lecture séquentielle des enregistrements permet en effet d'extraire rapidement du serveur de stockage S1 la liste des individus demandés. Le temps de réponse est donc optimisée.

En revanche, si la requête vise à obtenir une liste d'individus dont le prénom commence par la lettre A, et qu'elle est adressée par le SGBD 2 au premier serveur de stockage S1, son traitement ne sera pas optimisé.

En effet, si aucun index n'existe sur le serveur de stockage S1, il est nécessaire de parcourir toute la liste ordonnée des enregistrements $(E_1, E_2, \dots, E_N)$ pour en extraire ceux dont la valeur du

deuxième champ Ch_i^2 commence par la lettre A. On rappelle qu'un index peut être défini comme un lot de données destiné à accélérer les opérations de recherche de données. La structure de l'index comporte des valeurs associées à des pointeurs où chaque pointeur permet de retrouver la donnée qui a cette valeur.

5 Si donc un index existe sur le serveur de stockage S1, il permet de pointer sur les champs Ch_i^2 dont la valeur commence par la lettre A. Cependant, l'accès aux enregistrements correspondants nécessite d'opérer des lectures aléatoires sur le disque de stockage S1, qui sont par nature beaucoup plus lentes que les lectures séquentielles.

10 A réception d'une requête visant à obtenir une liste d'individus dont le prénom commence par la lettre A, le système de gestion de base de données SGBD 2 pourra avantageusement l'adresser au deuxième serveur de stockage S2, afin d'en optimiser le temps de traitement : la liste demandée pourra être rapidement obtenue par lecture séquentielle du contenu de la copie de la base de données stockée sur le deuxième serveur de stockage S2.

15 Le système de stockage de Fig 1 comprend par ailleurs un troisième support de stockage S3, sur lequel n'est stocké qu'un sous-ensemble des enregistrements de la base de données, par exemple E_N et E_1 . Par exemple, ce stockage suit un type de stockage TS3 qui consiste à classer les enregistrements par ordre croissant de la valeur de la donnée contenue dans le troisième champ Ch_i^3 , soit selon une séquence ordonnée $(E_N; E_1)$ dans cet exemple.

20 Le sous-ensemble (E_1, E_N) d'enregistrements à stocker sur le serveur S3 a été choisi car il s'agissait des enregistrements les plus fréquemment demandés (constat, en cours de fonctionnement du système de stockage 1, d'une volumétrie élevée de requêtes concernant ces enregistrements), ou dont il était prévu qu'ils fassent le plus fréquemment l'objet de requêtes, à l'initialisation du système de stockage 1. De même, le choix du type de stockage TS3 peut résulter d'un choix statique, par rapport aux types de requêtes attendues, à l'initialisation du système de stockage 1.

25 Il peut également résulter d'un choix dynamique, pendant le cycle de vie des données, suite à un constat d'une fréquence élevée de requêtes portant sur le troisième champ des enregistrements. On présente désormais plus en détail, en relation avec les figures suivantes, les étapes de stockage d'un enregistrement dans le système de stockage 1 de Fig 1, et de lecture de cet enregistrement, en réponse à une requête d'accès.

30 Une phase de configuration, ou d'initialisation du système de stockage redondant 1 est nécessaire, avant de pouvoir traiter toute requête. Cette initialisation est illustrée par le procédé de Fig 2.

L'étape référencée 100 correspond à la création de la base de données dans le système de stockage redondant 1. Au cours d'une étape référencée 101, est créé le schéma/modèle de

données de la base, à savoir le type d'enregistrements E_i , les types des champs Ch_i^j qui les composent, etc....

Au cours d'une étape SEL_Ch référencée 102, sont choisis les champs Ch_i du schéma de données sur lesquels les optimisations de stockage seront faites. En effet, tous les champs des enregistrements n'ont pas nécessairement la même importance, et il peut être avantageux de prévoir des copies de la base de données triées selon la valeur de certains des champs des enregistrements, mais pas de certains autres, qui font l'objet de requêtes moins fréquentes.

Au cours d'une étape DEF_TS_i référencée 103, on définit les types de stockage TS_i par rapport aux champs de données Ch_i choisis précédemment lors de l'étape 102. Un type de stockage par défaut pourrait aussi être défini.

Lors de l'étape DEF_REP(TS_i) 104, on définit le facteur de réplication REP selon le type de stockage TS_i. On peut par exemple déterminer qu'il est nécessaire de disposer de trois copies de la base de données selon le type de stockage TS2, de deux copies de la base selon le type de stockage TS3, mais de seulement une copie selon le type de stockage TS1.

On attribue ensuite (étape MATCH(S_i, TS_i) 105) des types de stockage TS_i aux différents supports de stockage S_i (serveurs, disques durs, mémoires flash, etc....) qui stockeront les données de la base de données. Cette attribution peut se faire de manière aléatoire si les supports de stockage S_i sont identiques ou en tenant compte de leurs spécificités et des données à stocker. Par exemple, les supports de stockage S_i les plus performants pourraient se voir attribuer les types de stockage TS_i susceptibles de demander le plus de performance. Cette étape peut être basée sur un profiling des données et des requêtes.

Au cours d'une étape SEL_Req_i référencée 106, on définit le type de requêtes Req_i à optimiser en tenant compte des champs Ch_i à optimiser. L'intégralité ou des sous-ensembles des requêtes Req_i peuvent être déclarés comme étant à optimiser.

Lors de l'étape référencée 107 MATCH (Req_i, TS_i), on procède à la mise en correspondance des types de requêtes Req_i avec les types de stockage TS_i. Puis, lors de l'étape référencée 108 MATCH (E_i, TS_i), on procède à la mise en correspondance des enregistrements E_i avec les types de stockage TS_i.

Toutes les étapes 100 à 108 décrites ci-dessus peuvent être effectuées lorsqu'on a une connaissance plus ou moins approfondie des données à stocker et des requêtes qui seront jouées sur la base de données. Dans le cas contraire, il est possible de ne procéder, à l'initialisation du système de stockage redondant de données, qu'à l'étape 100 de création de la base de données, en se fondant sur la détermination d'un type de stockage par défaut. Lors du cycle de vie de la base de données, par un processus d'analyse ou d'apprentissage, et à partir des connaissances qui seront extraites par ce processus, les étapes référencées 101 à 108 pourraient être mises en

œuvre, une ou plusieurs fois, par exemple par itérations successives, pour adapter le stockage par rapport à l'activité observée sur la base de données.

Le schéma de Fig 3 illustre le processus de stockage d'un enregistrement dans un système de stockage redondant de données selon un mode de réalisation de l'invention.

5 Un tel système de stockage comprend N supports ou serveurs de stockage référencés S1 à SN. Un client CLT envoie au premier serveur de stockage S1 un enregistrement E₁ qui doit être mémorisé dans la base de données. Cet enregistrement E₁ peut être composé de plusieurs champs Ch₁¹, Ch₁², Ch₁³, Ch₁⁴, etc. Cet envoi prend la forme d'une commande ST(E₁), adressée au serveur S1, qui peut être un serveur de stockage de données, ou un serveur de type SGBD 2, qui est en charge de
10 distribuer les requêtes d'enregistrement aux serveurs de stockage appropriés. Dans l'exemple de Fig 3, on considère que le serveur S1 est également un serveur de stockage des données, selon le type de stockage TS1.

A réception de l'enregistrement E₁, le serveur S1 l'analyse, à partir d'informations qu'il a précédemment mémorisées relatives à la mise en correspondance des enregistrements et des
15 types de stockage, et en déduit que cet enregistrement E₁ doit être associé à un type de stockage TS2. Cette analyse peut, dans un mode de réalisation, être effectuée par un serveur de métadonnées dédié.

Le serveur S1 envoie alors (ST(E₁, TS2)) l'enregistrement E₁ aux serveurs S2 et S5, qui sont les serveurs de stockage stockant des répliques de la base de données selon le type de stockage TS2.

20 Comme illustré sur Fig 3, un serveur de stockage S_i peut mémoriser plusieurs copies de la base de données, selon plusieurs types de stockage distincts. Ainsi, le serveur S2 gère des copies de la base de données selon les types de stockage TS2 à TS_n et le serveur de stockage S4 stocke des copies de la base de données selon les types de stockage TS1 à TS_n. En revanche, le serveur S3 ne stocke qu'une copie de la base de données, correspondant à la séquence ordonnée
25 d'enregistrements TS3 ; de même, les serveurs S5 et SN ne conservent qu'un exemplaire de la base de données, respectivement selon le type de stockage TS2 et TS_n.

Dans l'exemple de Fig 3, le serveur S1, qui peut être le système de gestion de base de données SGBD 2, n'envoie donc la requête de stockage de l'enregistrement E₁ qu'aux serveurs S2 et S5 capables de gérer le type de stockage TS2.

30 Le schéma de Fig 4 illustre le principe de lecture d'un enregistrement préalablement stocké dans un système de stockage redondant 1 de données selon un mode de réalisation de l'invention.

Le client CLT adresse au serveur S1 une requête Req(E₁) de lecture de l'enregistrement E₁. Comme précédemment, ce serveur S1 peut être un serveur de stockage de la base de données, ou un simple serveur chargé d'aiguiller les requêtes vers les serveurs de stockage appropriés : il joue
35 alors le rôle de SGBD 2. A réception de la requête Req(E₁), le serveur S1 opère une mise en

correspondance de l'enregistrement E_1 et du type de stockage TS2, et détermine un serveur de stockage duquel il va pouvoir extraire rapidement l'enregistrement E_1 . Cette mise en correspondance peut être effectuée sur la base d'informations à disposition du serveur S1, telles que l'ordre des champs dans E_1 et l'ordre des champs dans les différents types de stockage, ou être effectuée par un serveur de métadonnées dédié.

Dans l'exemple de Fig 4, le serveur S1 détermine que les performances optimales en termes de temps de traitement de la requête seront atteintes en adressant la requête $Req(E_1)$ de lecture de E_1 au serveur S5. Il envoie donc $Req(E_1)$ au serveur S5.

Par simple lecture séquentielle des enregistrements stockés selon la séquence ordonnée correspondant au type de stockage TS2, le serveur S5 peut rapidement extraire les données relatives à l'enregistrement E_1 de son espace de stockage, et l'envoyer en réponse au client CLT, via le serveur S1.

On optimise ainsi la performance du système de stockage redondant 1, en identifiant facilement le serveur associé au type de stockage le plus adapté à la nature de la requête du client.

On notera que, si un seul type de stockage TS était défini dans le système de stockage redondant de Fig 4, ou si la requête $Req(E_1)$ reçue du client CLT n'avait pas été préalablement affectée à un type de stockage TS_i donné, le choix, par le serveur S1, du dispositif de stockage S_i auquel transmettre la requête pourrait se faire de manière aléatoire, ou en tenant compte de la charge des différents supports de stockage S1 à S_N. Ainsi, le choix du support de stockage présentant la charge la plus faible permet également d'optimiser le temps de traitement de la requête, et donc d'améliorer les performances du système 1.

Le système de stockage redondant de données décrit ci-avant dans ses différents modes de réalisation est particulièrement avantageux pour la gestion de bases de données de grande taille, typiquement à partir du milliard d'enregistrements, et présente les avantages suivants :

- la possibilité de stocker de manière différente les différentes répliques de la base de données, en optimisant l'ordre de stockage des enregistrements en tenant compte de l'utilisation qui est faite des données de la base ;
- l'amélioration des performances de manière sélective sur l'ensemble des données, en ne répliquant que les données fortement sollicitées ;
- l'économie des ressources en n'augmentant le niveau de répllication que pour les données qui sont le plus sollicitées par des requêtes ;
- la possibilité d'adapter le niveau de répllication à la demande, de manière dynamique, sur les données stockées, par apprentissage.

On présente désormais, en relation avec Fig 5, la structure matérielle d'un système de gestion de base de données SGBD 2 selon un mode de réalisation de l'invention.

Le terme système de gestion peut correspondre aussi bien à un composant logiciel qu'à un composant matériel ou un ensemble de composants matériels et logiciels, un composant logiciel correspondant lui-même à un ou plusieurs programmes ou sous-programmes d'ordinateur ou de manière plus générale à tout élément d'un programme apte à mettre en œuvre une fonction ou un ensemble de fonctions.

Plus généralement, un tel système de gestion de base de données SGBD comprend une mémoire vive 13 (par exemple une mémoire RAM), une unité de traitement 12 équipée par exemple d'un processeur, et pilotée par un programme d'ordinateur, stocké dans une mémoire morte 11 (par exemple une mémoire ROM ou un disque dur). A l'initialisation, les instructions de code du programme d'ordinateur sont par exemple chargées dans la mémoire vive 13 avant d'être exécutées par le processeur de l'unité de traitement 12. La mémoire vive 13 contient notamment les règles de mise en correspondance des types de requêtes avec les types de stockage, les règles de mise en correspondance des enregistrements avec les types de stockage, un descriptif des types de stockage et de l'ordre de stockage des enregistrements qui leur sont associés, etc. Le processeur de l'unité de traitement 12 pilote l'analyse des requêtes et des enregistrements reçus, pour identifier les types de stockage qui leur correspondent, et les supports de stockage gérant ces types de stockage, à qui adresser les enregistrements à stocker, ou les requêtes d'accès aux enregistrements.

La Fig 5 illustre seulement une manière particulière, parmi plusieurs possibles, de réaliser le système de gestion de base de données SGBD 2, afin qu'il effectue les étapes du procédé détaillé ci-dessus, en relation avec Fig 1 à Fig 4 (dans l'un quelconque des différents modes de réalisation, ou dans une combinaison de ces modes de réalisation). En effet, ces étapes peuvent être réalisées indifféremment sur une machine de calcul reprogrammable (un ordinateur PC, un processeur DSP ou un microcontrôleur) exécutant un programme comprenant une séquence d'instructions, ou sur une machine de calcul dédiée (par exemple un ensemble de portes logiques comme un FPGA ou un ASIC, ou tout autre module matériel).

Dans le cas où le système de gestion de base de données SGBD 2 est réalisé avec une machine de calcul reprogrammable, le programme correspondant (c'est-à-dire la séquence d'instructions) pourra être stocké dans un médium de stockage amovible (tel que par exemple une disquette, un CD-ROM ou un DVD-ROM) ou non, ce médium de stockage étant lisible partiellement ou totalement par un ordinateur ou un processeur.

REVENDICATIONS

1. Système (1) de stockage redondant de données, comprenant une pluralité de dispositifs (S1-SN) de stockage de données, lesdites données étant structurées sous forme d'une pluralité d'enregistrements (E_i),

5 **caractérisé en ce qu'**un premier dispositif (S1) de ladite pluralité de dispositifs de stockage stocke lesdits enregistrements (E_i) selon une première séquence ordonnée d'enregistrements (TS1) **et en ce qu'**au moins un deuxième dispositif (S2-SN) de ladite pluralité de dispositifs de stockage stocke au moins un sous-ensemble desdits enregistrements (E_i) selon au moins une deuxième séquence ordonnée d'enregistrements (TS2-TSn) distincte de ladite première séquence ordonnée pour ledit
10 au moins un sous-ensemble.

2. Système de stockage selon la revendication 1, **caractérisé en ce que**, chaque enregistrement (E_i) comportant au moins deux champs (Ch_i^j) dans chacun desquels est enregistré une donnée, ladite première séquence ordonnée d'enregistrements (TS1) est obtenue par tri desdits enregistrements (E_i) à partir des données enregistrées dans un premier (Ch_i^1) desdits au
15 moins deux champs, et ladite au moins une deuxième séquence ordonnée d'enregistrements (TS2) est obtenue par tri desdits enregistrements à partir des données enregistrées dans au moins un deuxième (Ch_i^2) desdits au moins deux champs.

3. Système de stockage selon l'une quelconque des revendications 1 et 2, **caractérisé en ce que** ledit système de stockage comprend un système de gestion de base de données (2) configuré pour déterminer un dispositif de stockage à interroger, parmi ledit premier et ledit au moins un
20 deuxième dispositifs de stockage, en fonction d'une requête de donnée reçue par ledit système, pour optimiser un temps de récupération de ladite donnée requise.

4. Système de stockage selon la revendication 3, **caractérisé en ce que** ledit système de gestion de base de données est également configuré pour déterminer ledit au moins un sous-ensemble d'enregistrements à stocker dans ledit au moins un deuxième dispositif de stockage, en
25 fonction d'un nombre et/ou d'une fréquence de requêtes portant sur lesdits enregistrements dudit sous-ensemble.

5. Système de stockage selon l'une quelconque des revendications 3 et 4, **caractérisé en ce que** ledit système de gestion de base de données est également configuré pour déterminer un nombre de dispositifs de stockage de ladite pluralité de dispositifs de stockage dans lesquels stocker lesdits enregistrements et des séquences ordonnées d'enregistrements distinctes associées, en fonction d'un type et/ou d'une volumétrie de requêtes susceptibles d'être reçues.

6. Système de stockage selon l'une quelconque des revendications 3 à 5, **caractérisé en ce que** ledit système de gestion de base de données est également configuré pour émettre une
35 commande de stockage d'au moins un sous-ensemble desdits enregistrements selon une

séquence ordonnée d'enregistrements vers un dispositif de stockage de ladite pluralité de dispositifs de stockage, en fonction d'un type et/ou d'une volumétrie de requêtes reçues.

7. Procédé de stockage redondant de données, dans une pluralité de dispositifs de stockage de données, lesdites données étant structurées sous forme d'une pluralité d'enregistrements,

5 **caractérisé en ce qu'il comprend :**

- une étape de stockage desdits enregistrements dans un premier dispositif de ladite pluralité de dispositifs de stockage selon une première séquence ordonnée d'enregistrements ;
 - une étape de stockage, dans au moins un deuxième dispositif de ladite pluralité de dispositifs de stockage d'au moins un sous-ensemble desdits enregistrements selon au moins une
- 10 deuxième séquence ordonnée d'enregistrements distincte de ladite première séquence ordonnée pour ledit au moins un sous-ensemble.

8. Procédé de stockage selon la revendication 7, **caractérisé en ce que**, chaque enregistrement comportant au moins deux champs dans chacun desquels est enregistré une donnée, ladite première séquence ordonnée d'enregistrements est obtenue par tri desdits

15 enregistrements à partir des données enregistrées dans un premier desdits au moins deux champs, et ladite au moins une deuxième séquence ordonnée d'enregistrements est obtenue par tri desdits enregistrements à partir des données enregistrées dans au moins un deuxième desdits au moins deux champs.

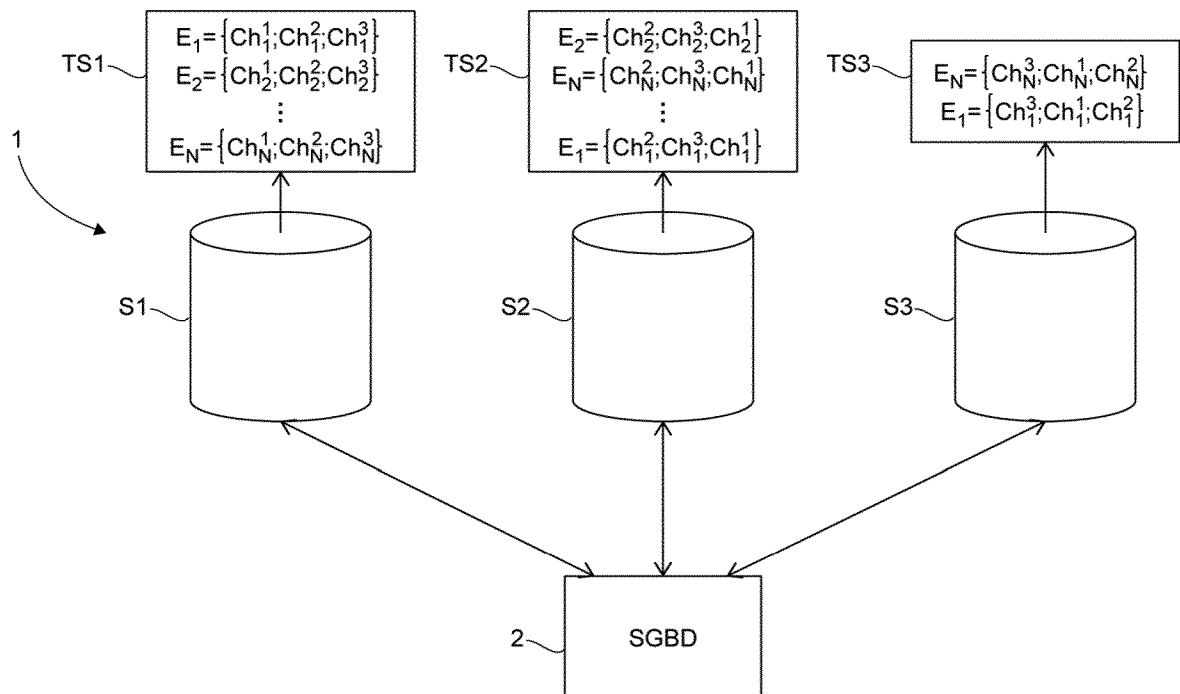
9. Procédé de stockage selon l'une quelconque des revendications 7 et 8, **caractérisé en ce**

20 **que**, sur réception d'une requête de donnée, il comprend une étape de détermination d'un dispositif de stockage à interroger, parmi ledit premier et ledit au moins un deuxième dispositifs de stockage, en fonction de ladite requête de donnée reçue, pour optimiser un temps de récupération de ladite donnée requise.

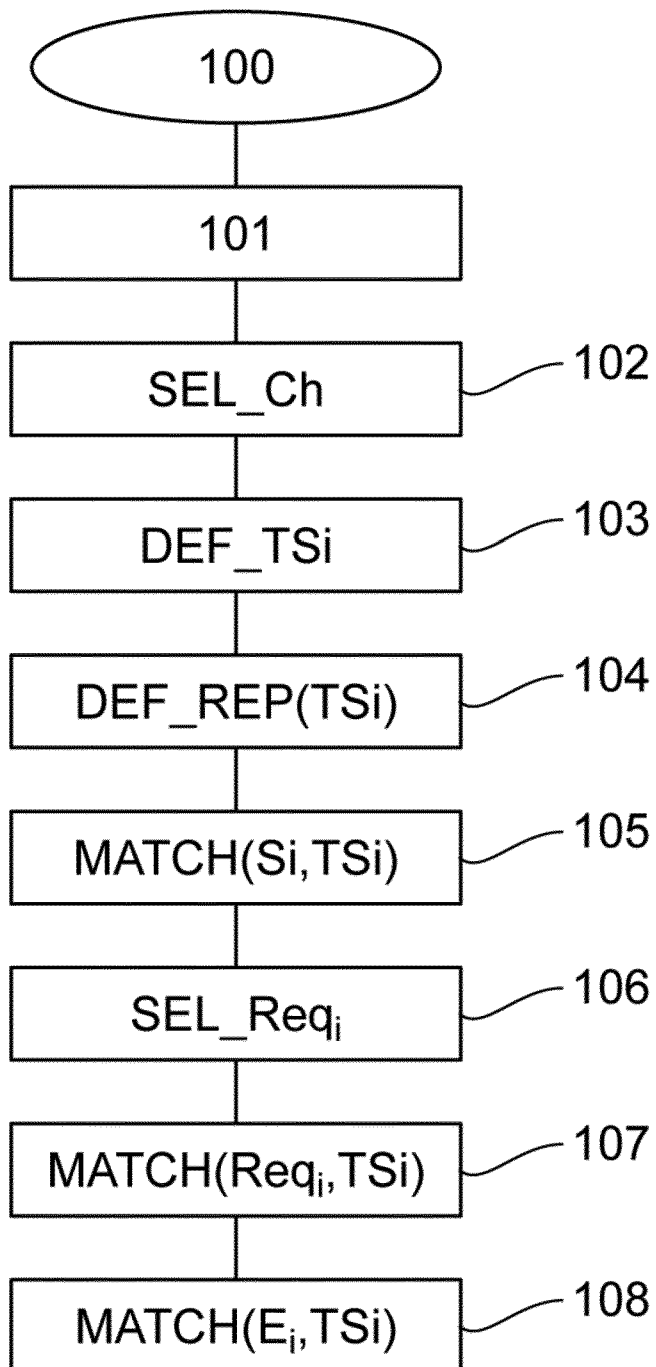
10. Produit programme d'ordinateur comprenant des instructions de code de programme

25 pour la mise en œuvre d'un procédé selon l'une quelconque des revendications 7 à 9, lorsqu'il est exécuté par un processeur.

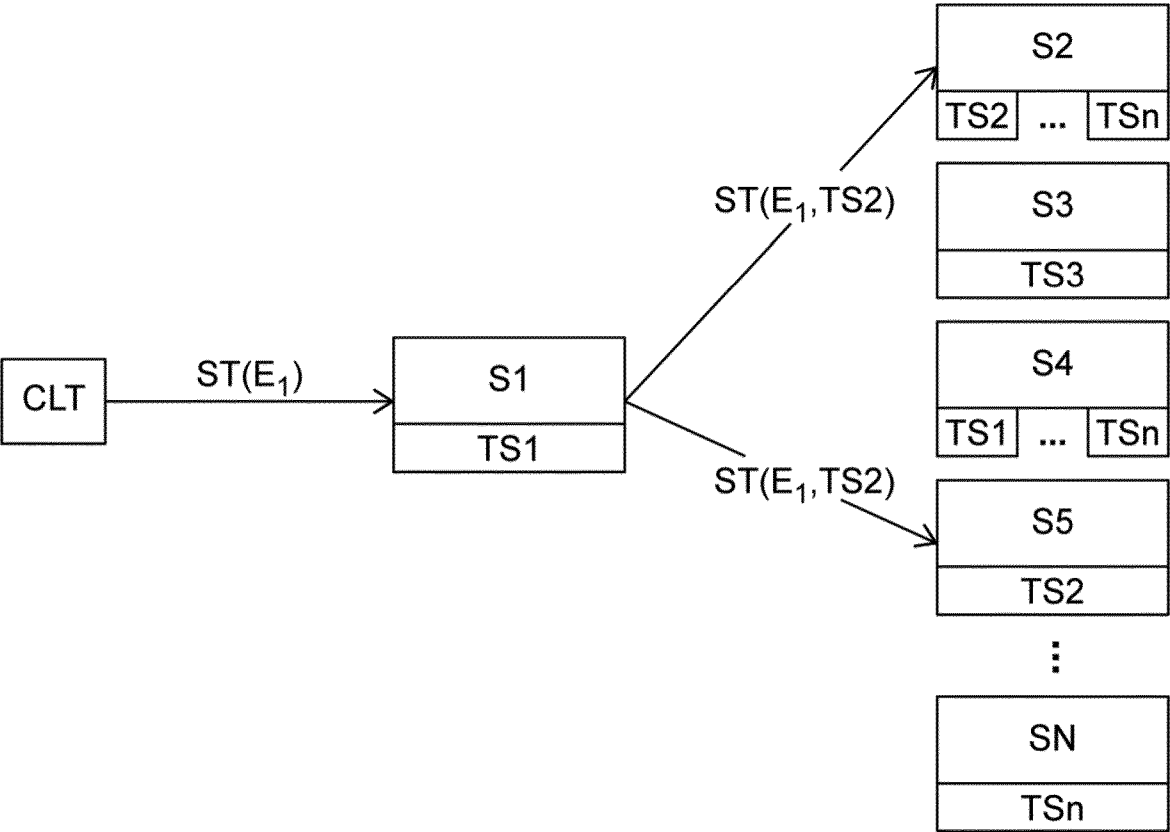
[Fig 1]



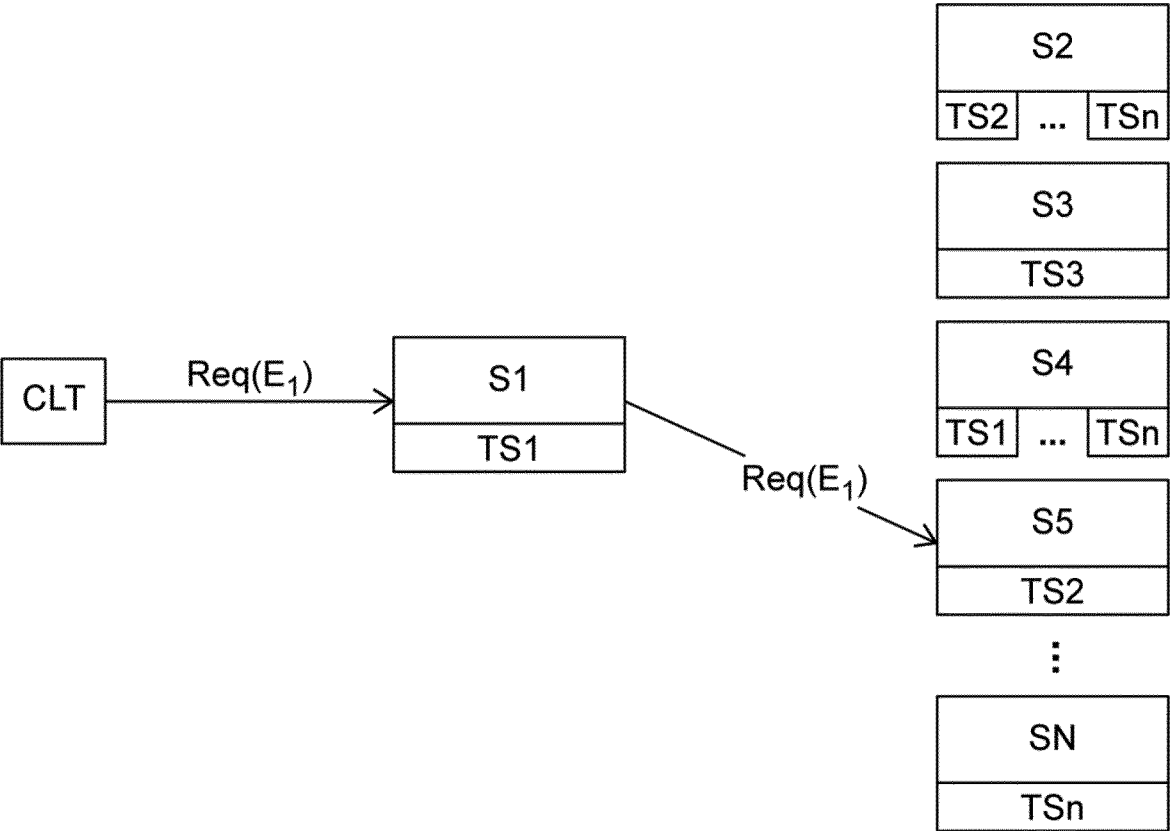
[Fig 2]



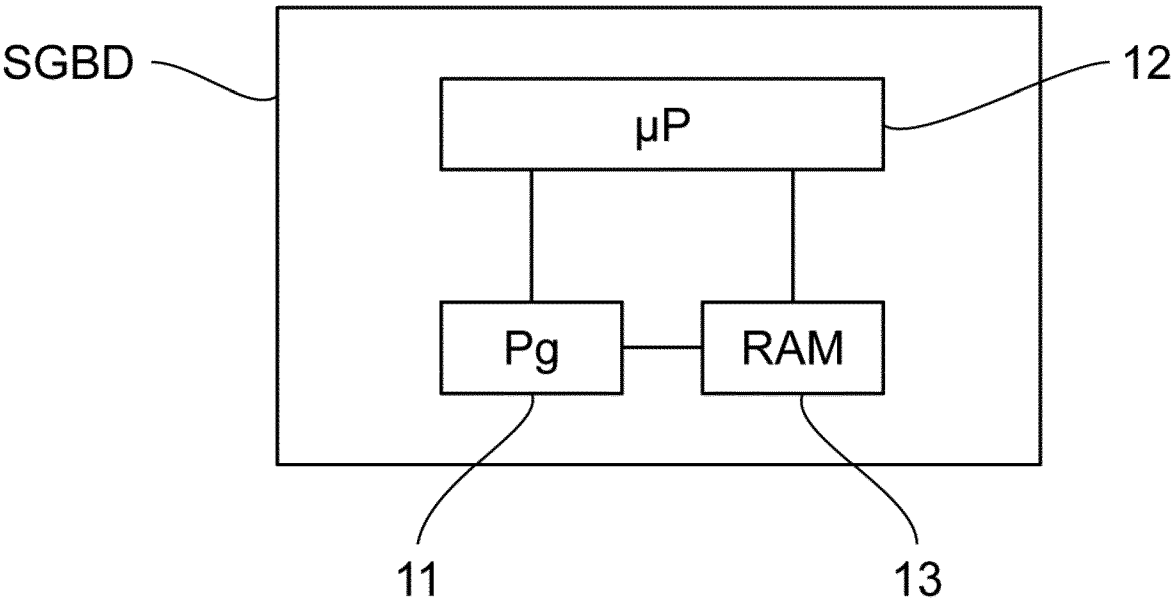
[Fig 3]



[Fig 4]



[Fig 5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2020/058756

A. CLASSIFICATION OF SUBJECT MATTER**G06F 16/27**(2019.01)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, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004098390 A1 (BAYLISS DAVID [US] ET AL) 20 May 2004 (2004-05-20) abstract paragraph [0127] - paragraph [0133]; figures 11A-11C paragraph [0139] - paragraph [0144]; figures 13A-13C	1-10
X	WO 2014000578 A1 (IBM [US]; IBM CHINA CO LTD [CN]) 03 January 2014 (2014-01-03) abstract page 9, line 11 - page 11, line 7; figures 1, 2 page 16, line 5 - page 19, paragraph 15; figure 6	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

25 May 2020

Date of mailing of the international search report

03 June 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

Barieux, Marc

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/EP2020/058756

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2004098390	A1	20 May 2004	NONE			
WO	2014000578	A1	03 January 2014	CN	103514229	A	15 January 2014
				DE	112013003205	T5	02 April 2015
				GB	2517885	A	04 March 2015
				US	2014012810	A1	09 January 2014
				US	2016314178	A1	27 October 2016
				US	2019050470	A1	14 February 2019
				WO	2014000578	A1	03 January 2014

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/EP2020/058756

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
INV. G06F16/27
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, INSPEC

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 2004/098390 A1 (BAYLISS DAVID [US] ET AL) 20 mai 2004 (2004-05-20) abrégé alinéa [0127] - alinéa [0133]; figures 11A-11C alinéa [0139] - alinéa [0144]; figures 13A-13C -----	1-10
X	WO 2014/000578 A1 (IBM [US]; IBM CHINA CO LTD [CN]) 3 janvier 2014 (2014-01-03) abrégé page 9, ligne 11 - page 11, ligne 7; figures 1, 2 page 16, ligne 5 - page 19, alinéa 15; figure 6 -----	1-10



Voir la suite du cadre C pour la fin de la liste des documents



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

25 mai 2020

Date d'expédition du présent rapport de recherche internationale

03/06/2020

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

Barieux, Marc

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/EP2020/058756

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2004098390	A1	20-05-2004	AUCUN	

WO 2014000578	A1	03-01-2014	CN 103514229 A	15-01-2014
			DE 112013003205 T5	02-04-2015
			GB 2517885 A	04-03-2015
			US 2014012810 A1	09-01-2014
			US 2016314178 A1	27-10-2016
			US 2019050470 A1	14-02-2019
			WO 2014000578 A1	03-01-2014
