

(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 2017/055708 A1

(43) Date de la publication internationale
6 avril 2017 (06.04.2017)

- (51) Classification internationale des brevets :
G06F 9/50 (2006.01)
- (21) Numéro de la demande internationale :
PCT/FR2016/052386
- (22) Date de dépôt international :
21 septembre 2016 (21.09.2016)
- (25) Langue de dépôt : français
- (26) Langue de publication : français
- (30) Données relatives à la priorité :
1559255 30 septembre 2015 (30.09.2015) FR
- (71) Déposant : ZCOST MANAGEMENT [FR/FR]; 1903
avenue Jacques Cartier La Fregate, 44800 Saint Herblain
(FR).
- (72) Inventeurs : HOFBAUER, Jacky; 15 rue de Trevaly,
44420 La Turballe (FR). KOCH, Bruno; 88 avenue Léon
Blum, 33110 Le Bouscat (FR). CHEVEREAU, Chris-
tophe; 36 avenue du 11 Novembre 1918, 49240 Avrille
(FR).
- (74) Mandataire : BREESE, Pierre; IP Trust, 2 rue de Clichy,
75009 Paris (FR).

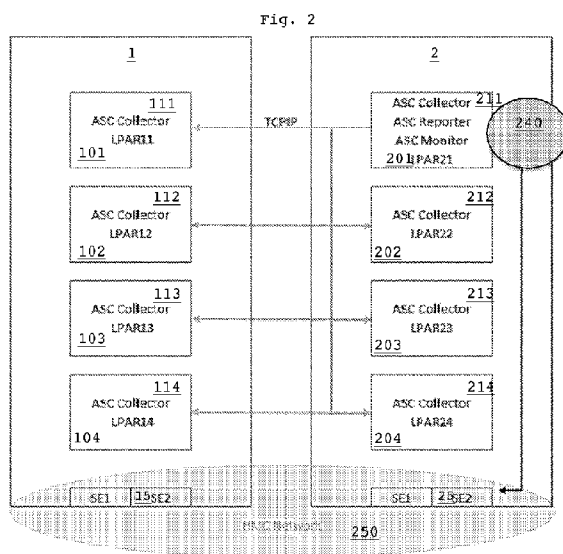
- (81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasien (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))

(54) Title : METHOD FOR MONITORING THE USE CAPACITY OF A PARTITIONED DATA-PROCESSING SYSTEM

(54) Titre : PROCÉDE DE CONTRÔLE DE LA CAPACITE D'UTILISATION D'UN SYSTEME PARTITIONNE DE TRAITEMENT DE DONNEES



(57) Abstract : The invention relates to a method for monitoring the use capacity of a partitioned data-processing system, said system being configured to have a plurality of logical partitions sharing common physical resources, the method consisting of limiting access to the processor resources of said partitions P_i in accordance with the value of the parameters $DC_{P_i}(t)$ for setting an upper limit on the capacity of each of the partitions P_i which are variable over time and are recalculated periodically, said values $DC_{P_i}(t)$ are recalculated periodically in accordance with: the result $NICRP(t)$ of measuring the instantaneous consumption level of the processor resources of all the partitions P_i ; the result $NICT_i(t)$ of measuring the instantaneous workload level of each of the partitions P_i ; the upper limit state of each of the partitions P_i ; and at least one "K_{user}" parameter determined by the user.

(57) Abrégé : L'invention présente un procédé de contrôle de la capacité d'utilisation d'un système partitionné de traitement de données, ledit système étant configuré pour présenter une pluralité de partitions logiques partageant des ressources physiques communes, le procédé consistant

[Suite sur la page suivante]

WO 2017/055708 A1



à limiter l'accès aux ressources processeurs desdites partitions P_i en fonction de la valeur de paramètres $DC_P_i(t)$ de plafonnement de la capacité de chacune des partitions P_i variables dans le temps et recalculée périodiquement, lesdites valeurs $DC_P_i(t)$ sont recalculées périodiquement en fonction : - du résultat $NICRP(t)$ de mesure du niveau instantané de consommation des ressources processeur de l'ensemble des partitions P_i - du résultat $NICT_i(t)$ de mesure du niveau instantané des charges de travail de chacune des partitions P_i , - de l'état de plafonnement de chacune des partitions P_i - d'au moins un paramètre « K_{user} » déterminé par l'utilisateur.

PROCEDE DE CONTROLE DE LA CAPACITE D'UTILISATION D'UN SYSTEME
PARTITIONNE DE TRAITEMENT DE DONNEES

Domaine de l'invention

La présente invention concerne le domaine des serveurs
5 informatique de haute capacité de traitement et de haute
disponibilité, généralement partagés entre plusieurs systèmes
d'informations désignées par le terme de partitions logiques
(LPAR). Ces serveurs comportent des ressources physiques
telles que de la mémoire, des processeurs. Ces serveurs sont
10 paramétrés pour optimiser le partage de ces ressources entre
les différentes partitions logiques et atteindre un niveau de
performance attendu pour chacune de ces partitions.
L'administrateur système affecte des portions de la mémoire à
chaque partition logique (LPAR). Un algorithme d'équilibrage
15 de charge interne assure la répartition de l'affectation des
processeurs aux différentes partitions logiques.

Etat de la technique le plus proche

On connaît dans l'état de la technique le brevet américain
US7752415 décrivant un procédé et un système pour commander
20 l'utilisation de la capacité d'un système de traitement de
données partitionné logiquement, comprenant des ressources
physiques qui sont partagées entre plusieurs partitions
logiques, dans lequel la capacité des partitions logiques est
gérée pour contrôler la consommation des ressources physiques
25 de traitement de données.

On connaît aussi le brevet US7096469 décrivant un procédé pour
faire imposer des plafonnements de capacité telles que celles
imposées par les accords de licence de logiciels dans un
système de traitement de l'information dans lequel une machine
30 physique est divisée en une pluralité de partitions logiques,
dont chacune se voit attribuer une portion définie de

ressources du processeur par un gestionnaire de partition logique.

Un gestionnaire de licence logicielle spécifie une consommation maximale autorisée des ressources du processeur par un programme s'exécutant dans l'une des partitions logiques.

Un gestionnaire de charge de travail s'exécutant également dans la partition mesure la consommation réelle de ressources processeur par la partition logique sur un intervalle de calcul de moyenne spécifié et la compare avec la consommation maximale autorisée.

Si la consommation réelle dépasse la consommation maximale autorisée, le gestionnaire de charge de travail calcule une limite supérieure (« soft capping » en anglais) et interagit avec le gestionnaire de partitions logiques pour plafonner la consommation réelle des ressources du processeur par la partition en conformité avec la limite calculée.

Pour fournir une flexibilité supplémentaire, les partitions se voient assigner un poids fantômes que le gestionnaire de partitions logiques ajoute au poids total de la partition pour déterminer si la partition a dépassé sa part autorisée de ressources processeur dans le contexte de plafonnement des ressources.

Pour les logiciels sous licence, la partition logique devient ainsi un "conteneur" ayant une capacité de traitement inférieure à celle de l'ensemble de la machine.

On connaît encore le brevet US8015335 décrivant un système d'ordinateur hôte partitionné logiquement comprenant processeurs hôte (CPU hôte) divisée en une pluralité de processeurs hôtes (CPU d'hôtes) d'une configuration d'hôtes, un effectuer topologie fonction instruction est exécuté par un

processeur d'hôtes spécifiant un changement de topologie de la configuration d'invité.

Le changement de topologie change préférence la polarisation de processeurs d'hôtes, la polarisation liée à la quantité
5 d'une ressource CPU hôte est fournie à un processeur hôte.

On connaît enfin le brevet US8904405 concernant un serveur dont le système de traitement de données partitionné est divisé en plusieurs partitions logiques.

Dans un système et un procédé comme la gestion d'un système
10 informatique, chaque tâche en cours d'exécution de charge de travail qui peuvent être classés sur la base de la criticité du temps, et les groupes de partitions logiques peuvent être définis librement.

Les limites de capacité de traitement pour les partitions
15 logiques dans un groupe de partitions logiques basées sur des seuils de capacité de traitement définis et sur une détermination itérative de la capacité nécessaire pour les tâches de la charge de travail à durée critique.

La charge de travail peut être équilibrée entre les partitions
20 logiques au sein d'un groupe, pour empêcher la capacité de traitement de surplus étant utilisée pour exécuter la charge de travail critique pas le temps sur une partition logique quand une autre partition logique uniquement les tâches de la charge de travail à durée critique face à un déficit de
25 traitement.

On connaît également dans l'état de la technique –la demande de brevet américain US2014/0373010 qui ne concerne pas le domaine de l'invention revendiquée, mais un système de gestion des ressources pour des machines virtuelles.

Selon cette solution de l'art antérieur, un ordinateur reçoit des informations associées à une ou plusieurs machines virtuelles, comprenant des informations d'utilisation et de charge de travail associé à chaque machine virtuelle des une
5 ou plusieurs machines virtuelles. Dans la solution décrite, l'ordinateur analyse les informations reçues associées à une ou plusieurs machines virtuelles et détermine les machines virtuelles pour la réaffectation des ressources.

Cette solution concerne un contexte particulier, spécifique au
10 système d'exploitation AIX (acronyme de Advanced Interactive eXecutive) orienté vers la virtualisation des processeurs, des disques et du réseau.

L'objet de la solution décrite dans ce document est la maîtrise des performances sur une autre plateforme et non pas
15 le contrôle de la capacité d'utilisation d'un système partitionné pour limiter l'accès aux ressources processeurs desdites partitions P_i en fonction de la valeur de paramètres $DC_{P_i}(t)$ de plafonnement de la capacité de chacune des partitions.

20 Un autre document de l'art antérieur, la demande de brevet américain US20130055277, divulgue un système de traitement de données pour la gestion et l'équilibrage de charge dans plusieurs systèmes gérés dans un système de traitement de données de partitionnement logique. En réponse à une partition
25 logique critique nécessitant des ressources supplémentaires, le mécanisme détermine si un ou plusieurs systèmes gérés ont des ressources disponibles pour satisfaire les besoins en ressources de la partition critique. Le mécanisme effectue au moins une opération de migration de séparation pour déplacer
30 au moins une partition logique entre les systèmes gérés en réponse à la détermination qu'un ou plusieurs systèmes gérés disposent de ressources pour satisfaire les besoins en ressources de la partition critique. Le mécanisme effectue au

moins une opération de partitionnement logique dynamique pour affecter des ressources à au moins une des une ou plusieurs critiques de partitions logiques réagissant à l'exécution de l'opération au moins une migration de partition.

5 Ce document ne concerne pas non plus la gestion dynamique du paramètre de limitation de l'accès aux ressources des partitions du serveur.

On connaît aussi la demande de brevet américain US2008244215 décrivant une solution pour équilibrer l'accès entre plusieurs
10 partitions logiques sur les ressources du système physiques d'un système informatique en utilisant la virtualisation de système. Chacune des partitions logiques est classée, d'abord au cours d'une période de démarrage, en fonction d'un niveau attribué l'utilisation de la fenêtre d'expédition. Les mesures
15 de performance d'une ou plusieurs des ressources du système physique sont déterminés en association avec une ou plusieurs partitions logiques. La détermination des indicateurs de performance est réalisée à un niveau indépendant des interruptions de programmation du matériel.

20 Il s'agit d'un simple arrière plan technologique générale, n'apportant aucun enseignement concernant la gestion dynamique des paramètres de limitation d'accès aux ressources des partitions.

Le brevet américain US7007276 concerne un exemple de gestion
25 de groupes de partitions d'un environnement informatique, comprenant la modification d'une composition d'un groupe d'une pluralité de partitions dudit environnement informatique et de modification dynamique d'une ressource partageable attribué au groupe, utilisable simultanément par plusieurs partitions.

Inconvénients de l'art antérieur

Dans les solutions de l'art antérieur, le partage des ressources ~~de~~ entre chacune des partitions ne permet pas d'assurer des performances satisfaisantes dans toutes les conditions d'utilisation. Par exemple pour satisfaire un besoin exceptionnel très important et transitoire de ressources, le serveur doit être paramétré en fonction du pic de consommation induit, ce qui conduit à sur dimensionner le serveur et donc à augmenter la complexité des moyens matériels, la consommation et l'échauffement produit par le serveur, jusqu'à dégrader gravement le fonctionnement du serveur.

La solution proposée par le brevet US8904405 met en œuvre un procédé de limitation des ressources à l'intérieur de chacune des partitions, fonction de paramètres prenant en compte la priorisation et la classification des charges de travail qui sont déterminées par des traitements automatiques exécutés par le programme de gestion des tâches de travail. Ces données ainsi calculées ne sont pas totalement pertinentes pour assurer une optimisation de la consommation des ressources. Par exemple, si la partition est limitée à une capacité déterminée en fonction d'informations collectées auprès du gestionnaire de tâches, dans une situation où beaucoup de tâches sont classées comme « non prioritaires », l'affectation des ressources sera faussée. En effet, le gestionnaire de tâches se base sur les ressources disponibles et sur l'ensemble des tâches dont il doit assurer l'exécution (quel que soit leur priorité) pour assurer la bonne exécution et la priorisation de ces tâches. Dans les solutions de l'état de la technique, la politique de gestion des ressources est déterminée automatiquement par le seul gestionnaire de tâches (WLM) qui commande l'exécution de tâches non prioritaires en fonction des ressources disponibles, compte tenu du plafonnement global de la consommation des ressources

déterminé pour la totalité des tâches de travail de la partition.

Solution apportée par l'invention

Afin de répondre à ces inconvénients, l'invention concerne
5 selon son acception la plus générale un procédé de contrôle de la capacité d'utilisation d'un système partitionné de traitement de données, ou serveur, ledit système étant configuré pour présenter une pluralité de partitions logiques partageant des ressources physiques communes, le procédé
10 consistant à limiter l'accès aux ressources processeurs desdites partitions P_i en fonction de la valeur de paramètres $DC_{P_i}(t)$ de plafonnement de la capacité de chacune des partitions P_i variables dans le temps, caractérisé en ce que lesdites valeurs $DC_{P_i}(t)$ sont recalculées périodiquement en
15 fonction:

- du résultat $NICRP(t)$ de mesure du niveau instantané de consommation des ressources processeur de l'ensemble des partitions P_i
- du résultat $NICT_i(t)$ de mesure du niveau instantané des
20 charges de travail de chacune des partitions P_i ,
- de l'état de plafonnement de chacune des partitions P_i
- d'au moins un paramètre « K_{user} » déterminé par l'utilisateur.

L'invention permet d'optimiser le partage des ressources
25 informatiques dans un environnement de gros serveurs (« mainframe » en anglais).

Dans la suite du brevet, le « système partitionné » et désigne un ordinateur, généralement un serveur, comprenant un ou plusieurs disques durs organisé en plusieurs partitions dans
30 lesquelles les systèmes d'exploitation présents sur la machine peuvent y gérer leurs informations de manière séparée et privée.

Les ordinateurs exploitent de préférence le système de fichiers de z/OS de la société IBM (noms commerciaux), qui gère les fichiers avec un système de catalogue mais aussi non hiérarchique sans répertoires ni arborescences de fichiers pour les identifiants dsn. Les fichiers sont identifiés par leur identifiant dsn, acronyme signifiant Data Set Name. Un identifiant dsn est composé de plusieurs qualifieurs (d'au plus 8 caractères de long) connectés par un point, dont le premier (HLQ) est généralement défini en tant qu'alias dans un catalogue. Il peut être complété par le second et le troisième, on parle de Multi Level Alias.

Un nom d'un fichier représente une chaîne de caractères (44 maximum) séparés par des points.

Ce nouveau mode de détermination des variables $DC_{P_i}(t)$ permet de fixer un plafonnement aux partitions tout en réduisant les contraintes imposées à la consommation de ressources de ces partitions.

Ce plafonnement est selon une variante encadré par des valeurs minimales et maximales au niveau d'une partition logique et/ou des groupes de partitions logiques, au niveau du système partitionné et au niveau du groupe de systèmes partitionnés.

Avantageusement, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de la valeur antérieure $DC_{P_i}(t-n)$. Selon une variante, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de N valeurs $NICRP(t)$.

Selon une autre variante, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de P valeurs moyennes antérieures de $NICT_i(t)$ calculées chacune sur une fenêtre temporelle d'une durée D prédéterminée.

Selon une autre variante, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de M valeurs antérieures du résultat $NICRP(t)$.

5 Selon une autre variante, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de P valeurs moyennes antérieures de $NICRP(t)$ calculées chacune sur une fenêtre temporelle d'une durée D prédéterminée.

10 Avantageusement, ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction du nombre et de l'état de polarisation verticale $VProc_{P_i}$ représentatif de l'état des processeurs affectés à chacune des partitions P_i .

On entend par « polarisation » au sens du présent brevet le paramètre décrit dans le brevet US8015335 et se référant ~~an~~ au nombre et à l'état des processeurs affectés à une ressources.

15 Selon un premier mode de réalisation, ledit calcul est effectué pour les seules ressources et partitions d'un même serveur.

20 Selon un deuxième mode de réalisation, ledit calcul est effectué pour les ressources et partitions d'une pluralité de serveurs.

Description d'un exemple non limitatif de réalisation de l'invention

25 La présente invention sera mieux comprise à la lecture de la description qui suit, concernant un exemple non limitatif de réalisation illustré par les dessins annexés où :

- la figure 1 représente une vue schématique de l'architecture matérielle d'un serveur selon l'invention

- la figure 2 représente une vue schématique de l'architecture fonctionnelle d'un serveur selon l'invention

- la figure 3 représente l'architecture d'une partition et des procédés de collecte des informations
- la figure 4 représente une vue fonctionnelle schématique d'un moniteur.

5

Architecture matérielle

Dans les systèmes informatiques « gros serveur » (par exemple mainframe d'IBM 390) le nombre de partitions est limité, par exemple 15 partitions logiques. Dans des systèmes plus récents utilisant une architecture de type Z (marque déposée), on prévoit un plafonnement de taille de la mémoire, des interfaces entrée/sortie et de la disponibilité, et la puissance de traitement disponible peut être limitée à un niveau inférieur à ce plafonnement.

La figure 1 représente une vue schématique de l'architecture matérielle d'un système selon l'invention constitué de deux serveurs (1, 2). Bien entendu, le nombre de serveurs peut être supérieur à deux.

Chaque serveur (1, 2) est partitionné en une pluralité de partitions logiques qui se partagent les ressources matérielles, notamment les ressources processeurs. Chaque processeur est caractérisé par un état de polarisation verticale ou horizontale.

Chaque serveur (1, 2) comporte plusieurs processeurs (11, 12, 13 ; 21, 22, 23) ainsi qu'un gestionnaire de processeur (14, 24) de type PR/SM (« Processor Resource/System Manager ») gérant la répartition de la puissance de calcul entre les partitions logiques.

Chaque serveur (1, 2) comporte également une unité de supervision « support element » (15, 25) qui gère la configuration du serveur en particulier le nombre de

processeurs disponibles, la taille de la mémoire et plus généralement l'état physique et l'administration du serveur par l'intermédiaire de consoles (19, 29).

5 Chaque serveur (1, 2) comporte également des cartes de communication (16, 26) avec des mémoires externes (18, 28) et des cartes réseau (17, 27) pour la communication avec des terminaux clients (10, 20).

Architecture fonctionnelle

10 La figure 2 représente un exemple d'architecture fonctionnelle.

Chaque serveur (1, 2) héberge une pluralité de partitions logiques (101 à 104 ; 201 à 204). Chaque partition logique (101 à 104 ; 201 à 204) comprend un collecteur (111 à 114 ; 211 à 214) qui collecte des informations « métriques »
15 comprenant :

- la valeur de paramètres $DC_P_i(t)$ de plafonnement de la capacité de chacune des partitions P_i variables dans le temps et recalculée périodiquement,
- 20 - le résultat $NICRP(t)$ mesurant le niveau instantané de consommation des ressources processeur de l'ensemble des partitions P_i
- le résultat $NICT_i(t)$ mesurant le niveau instantané des charges de travail de chacune des partitions P_i ,
- 25 - l'état de plafonnement de chacune des partitions P_i .

Ces informations sont calculées à partir des données disponibles via le système d'exploitation par des traitements implémentés dans chaque partition.

Ces informations sont, par exemple, enregistrées dans les blocs de contrôle de l'architecture Z d'IBM (noms commerciaux) comme illustré par la figure 3 représentant l'architecture d'une partition (101) et les procédés de collecte des informations. Chaque partition (101) est associée à un collecteur (301).

Le collecteur (301) est constitué par un programme d'ordinateur procédant à la lecture des informations dans les blocs de contrôle (302 à 304) concernant le serveur dans sa globalité, la partition logique (101), la caractérisation de l'affectation des ressources processeurs, en particulier dans les blocs de contrôle (305 à 307) intitulée « z/OS Workload Management (WLM) » dans l'architecture Z d'IBM (noms commerciaux).

Ces données sont enregistrées périodiquement et historisées dans une mémoire (308) du collecteur (301).

Le système comprend soit un seul moniteur (400), ou plusieurs, généralement un par serveur (1, 2).

Le moniteur (400) représenté en figure 4 procède à la lecture périodique des données enregistrées dans la mémoire (308) des collecteurs (301), via un réseau et un protocole TCP/IP. Ces données font l'objet de traitements pour calculer périodiquement, pour chacune des partitions P_i les paramètres $DC_{P_i}(t)$ de plafonnement de la capacité de chacune des partitions P_i .

Ce moniteur (400) comporte un module d'interfaçage (401) éventuellement distinct, générant les instructions numériques pour commander l'application de ces paramètres de plafonnement via l'hyperviseur PR/SM (14, 24) (Processor Resource/System Manager) et les éléments de supervision (15, 25), afin de modifier périodiquement la puissance allouée à chacune des partitions pour lesquelles les données ont été collectées, et,

suivant la configuration du système, le nombre de processeurs actifs.

L'application de ces paramètres peut avoir pour conséquence la modification de la configuration processeur du serveur (1, 2).

5 Cette commande est envoyée à l'hyperviseur (14, 24) selon deux modes possibles :

- le premier mode utilise l'application (240) BCPii (« Base Control Program Internal Interface » (nom commercial)) pour communiquer
10 avec les éléments de supervision (15, 25)
- un deuxième mode utilise le réseau HMC « hardware management consol » (250) pour
communiquer avec les éléments de supervision
(15, 25).

15 Ces données et les identifiants des modifications sont également exploitées pour construire des représentations graphiques ou textuelles représentatives de l'évolution du système. Elles peuvent aussi être exploitées pour générer des messages électroniques, ou des informations de performances,
20 par exemple selon un format universel SMF (« system measurement facility »).

Optionnellement, l'invention peut également prévoir des moyens de changement automatique ou par commande de paramètres de fonctionnement par exemple en fonction d'un gestionnaire
25 d'évènements ou de calendrier.

Elle peut également prévoir des moyens de générations d'alertes, par exemple sous forme de messages électroniques.

Traitement des données

Les données collectées et/ou calculées comprennent :

5 - des données relatives aux serveurs (1, 2) telles que le nom, l'identifiant technique, le type de serveur, la capacité du serveur, le nombre de processeurs physiques et logiques, le nombre de partitions configurées, la somme des poids relatifs des partitions, la somme des plafonnements, la somme des consommations instantanées, la somme des moyennes sur une période temporelle D prédéterminée, le ratio nombre de
10 processeurs physiques / nombre de processeurs logiques,...

15 - des données relatives aux partitions (101 à 104 ; 201 à 204) telles que le nom, le nom du complexe (au sens « syplex » tel qu'exposé dans la notice <http://www.redbooks.ibm.com/redbooks/pdfs/sg244356.pdf>) auquel la partition appartient, son poids relatif, son plafonnement ainsi que l'état de son plafonnement, sa consommation instantanée, la moyenne sur une période temporelle D prédéterminée, le nombre de processeurs logiques, la tendance d'évolution de sa consommation, ...

20 - des données relatives aux composants gestionnaires de charge de travail de chacune des partitions (WLM) destinées à optimiser la charge de la machine et du réseau, telles que les classes de service, la consommation par classe de service et par période,...

25 - des données relatives au nombre et au type de polarisation (moyen, haut, bas) des processeurs, pour chacune des partitions logiques.

Par ailleurs, le traitement met en œuvre des paramètres personnalisables K_{user} tels que :

30 - des paramètres généraux tels que l'intervalle temporel de collecte, les coefficients

modérateurs fixant un seuil minimal de variation de la capacité de chaque partition, les seuils haut et bas de plafonnement autorisé du serveur (CPC_MAX, CPC_MIN,) ...

- 5 - des paramètres d'alerte tels que des seuils de déclenchement de l'envoi d'un message électronique ou d'une action
- des paramètres relatifs aux partitions logiques, tels que des seuils haut et bas de plafonnement autorisé de chaque partition (MSU_MAX, MSU_MIN,), les critères de priorité, la marge de sécurité, une valeur de réactivité, un attribut de protection de chaque partition destinée à contrecarrer le plafonnement en fonction de circonstances définies par l'utilisateur ou calculées automatiquement, l'appartenance à un groupe de partitions logiques,...
- 10
- 15 - des clés de sécurité.

Ces données susvisées font l'objet d'un traitement périodique pour le calcul du plafonnement de chaque partition logique, consistant à :

- réallouer les capacités entre les partitions logiques disponibles
- en cas de surcapacité, réallouer les capacités entre les partitions logiques réelles et une partition logique virtuelle ('fantôme') afin de réduire le plafonnement global des serveurs (1, 2).
- 25

A cet effet, on fixe une limite du plafonnement pour chaque partition logique, calculée en fonction des seuils susvisés haut et bas des partitions (101 à 104 ; 201 à 204) et des serveurs (1, 2), en fonction des paramètres K_{user} susvisés définis par l'utilisateur et des données collectées.

30

Revendications

1 – Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné [serveur] de traitement de données, lesdits systèmes étant configurés pour présenter une pluralité de partitions logiques partageant des ressources physiques communes, le procédé consistant à limiter l'accès aux ressources processeurs desdites partitions P_i en fonction de la valeur de paramètres $DC_{P_i}(t)$ de plafonnement de la capacité de chacune des partitions P_i variables dans le temps et recalculée périodiquement,

caractérisé en ce que lesdites valeurs $DC_{P_i}(t)$ sont recalculées périodiquement en fonction :

- du résultat $NICRP(t)$ de mesure du niveau instantané de consommation des ressources processeur de l'ensemble des partitions P_i
- du résultat $NICT_i(t)$ de mesure du niveau instantané des charges de travail de chacune des partitions P_i ,
- de l'état de plafonnement de chacune des partitions P_i
- d'au moins un paramètre « K_{user} » déterminé par l'utilisateur.

2 – Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 1 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de la valeur antérieure $DC_{P_i}(t-n)$

3 – Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 1 ou 2 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de N valeurs $NICRP(t)$ antérieures.

- 4 - Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 1 ou 2 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de P valeurs moyennes antérieures de $NICT_i(t)$ calculées chacune sur une fenêtre temporelle d'une durée D prédéterminée.
- 5
- 5 - Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 1 ou 2 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction d'un indicateur d'évolution prédictif des valeurs moyennes de $NICT_i(t)$ calculées chacune sur une fenêtre temporelle d'une durée D prédéterminée.
- 10
- 6 - Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 1 ou 2 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de M valeurs antérieures du résultat $NICRP(t)$.
- 15
- 7 - Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon la revendication 3 ou 4 caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction de l'évolution de P valeurs moyennes antérieures de $NICRP(t)$ calculées chacune sur une fenêtre temporelle d'une durée D prédéterminée.
- 20
- 8 - Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données selon l'une au moins des revendications précédentes caractérisé en ce que ledit calcul des valeurs $DC_{P_i}(t)$ est en outre fonction du nombre et de l'état de polarisation verticale $VProc_{P_i}$ représentatif de l'état des processeurs affectés à chacune des partitions P_i .
- 25
- 30

9 – Procédé de contrôle de la capacité d'utilisation d'au moins un système partitionné de traitement de données, selon la revendication 1 caractérisé en ce que ledit calcul est effectué pour les seules ressources et partitions d'un même serveur.

5

10 – Procédé de contrôle de la capacité d'au moins un système partitionné de traitement de données, selon la revendication 1 caractérisé en ce que ledit calcul est effectué pour les ressources et partitions d'une pluralité de serveurs.

10

Fig. 1

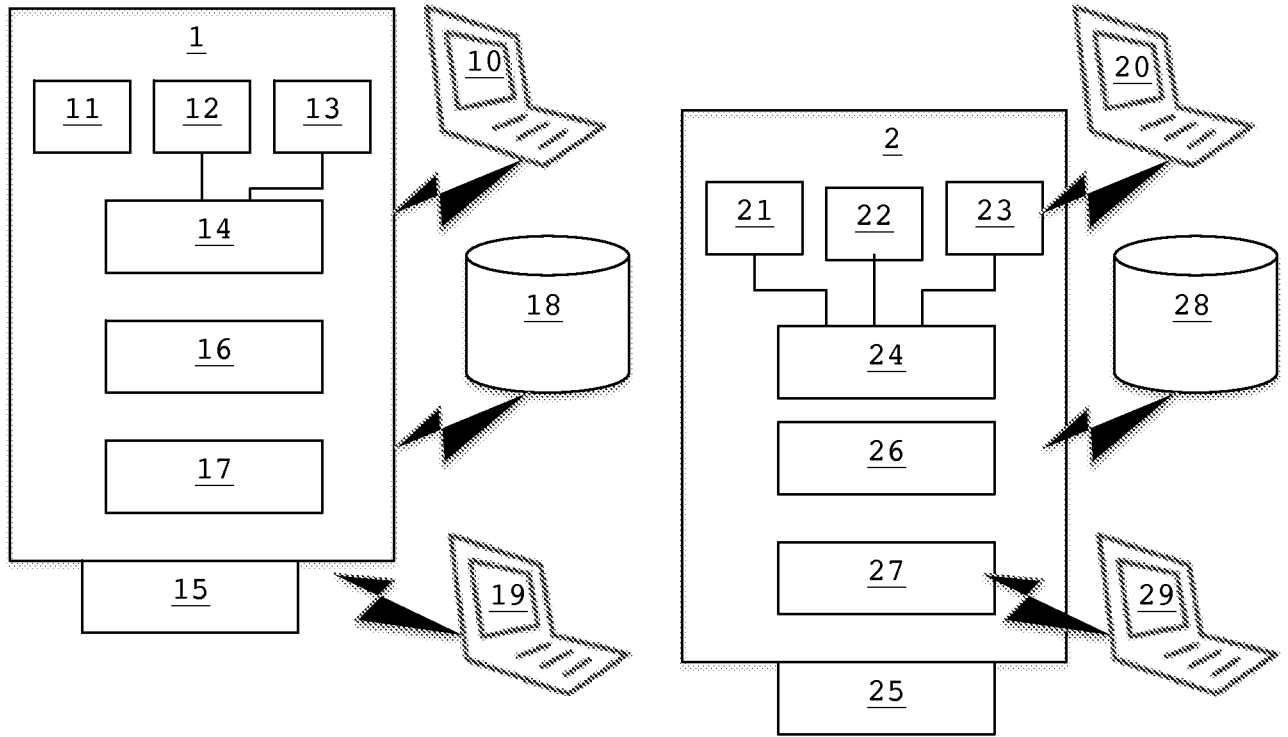


Fig. 2

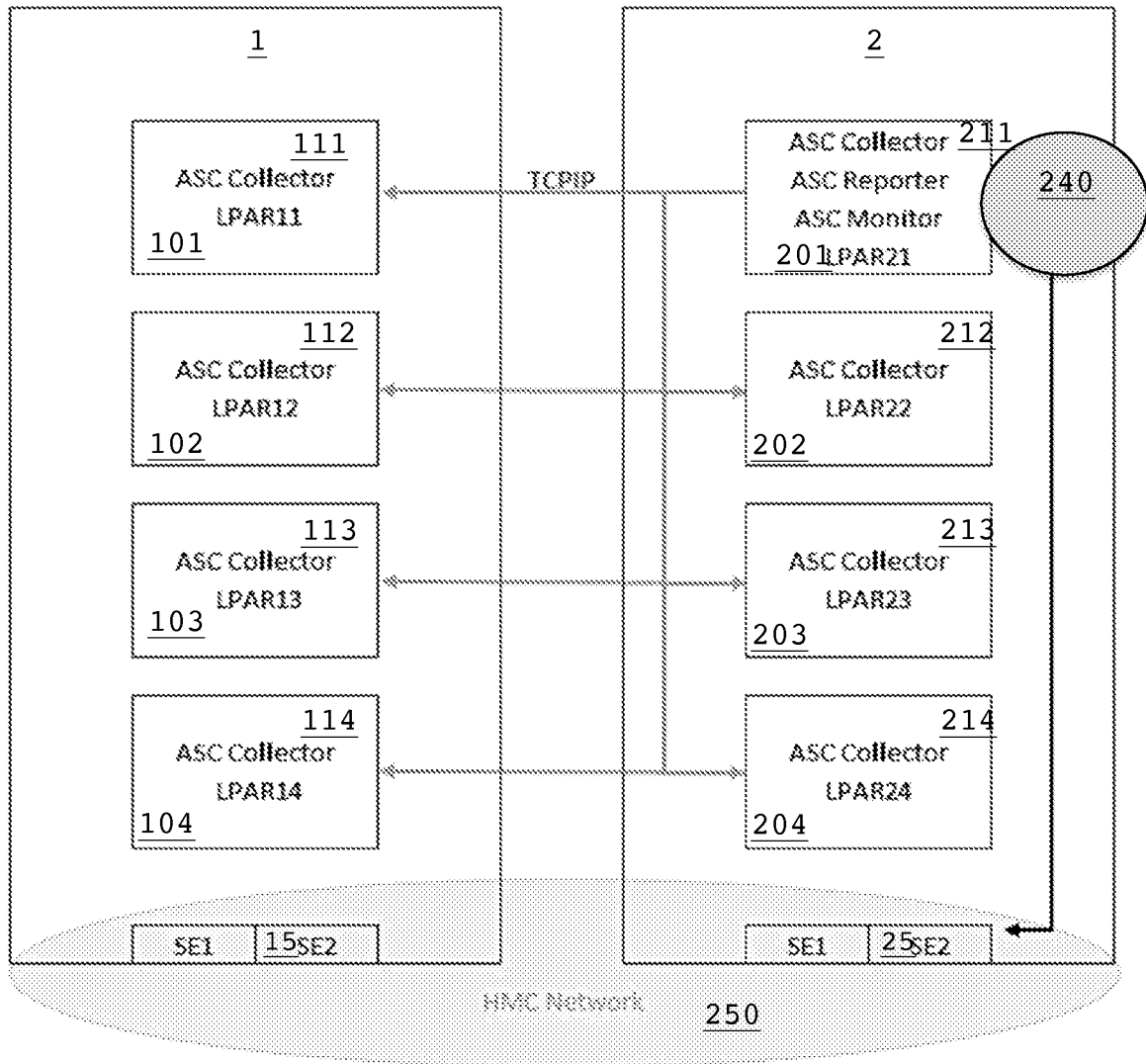


Fig. 3

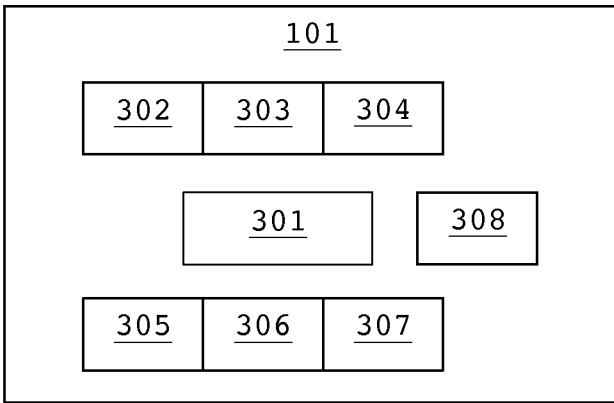
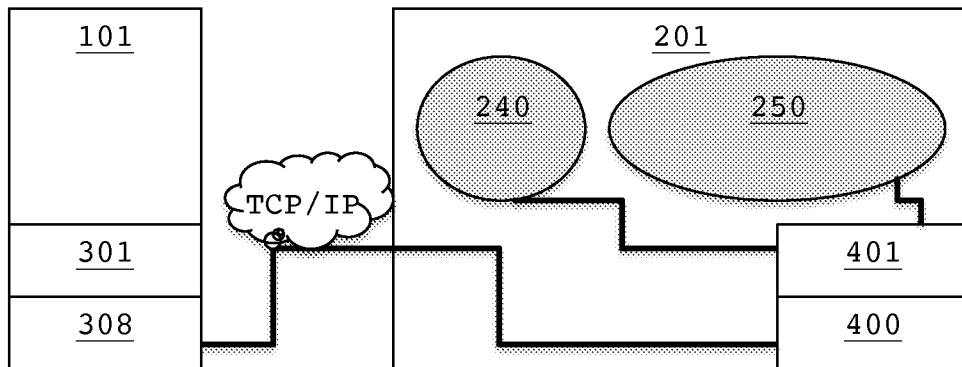


Fig. 4



INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2016/052386

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F9/50
ADD.

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	US 2014/373010 A1 (FOLCO RAFAEL C S [BR] ET AL) 18 December 2014 (2014-12-18) paragraph [0012] paragraph [0027] - paragraph [0031] -----	1-10
X	US 2013/055277 A1 (ASHISH SHIVENDRA [IN] ET AL) 28 February 2013 (2013-02-28) paragraph [0034] - paragraph [0045] ----- -/--	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 26 October 2016	Date of mailing of the international search report 07/11/2016
--	--

Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Dewyn, Torkild
--	--

INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2016/052386

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MACIEJ MLYNSKI: "The Analysis of Influence of IBM pSeries Servers' Virtualization Mechanism on Dynamic Resources Allocation in AIX 5L", PARALLEL AND DISTRIBUTED COMPUTING, 2008. ISPCD '08. INTERNATIONAL SYMPOSIUM ON, IEEE, PISCATAWAY, NJ, USA, 1 July 2008 (2008-07-01), pages 39-46, XP031378726, ISBN: 978-0-7695-3472-5 the whole document	1-10
A	----- US 2008/244215 A1 (FLEMMING DIANE G [US] ET AL) 2 October 2008 (2008-10-02) the whole document	1-10
A	----- US 7 007 276 B1 (KUBALA JEFFREY P [US] ET AL) 28 February 2006 (2006-02-28) the whole document -----	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/FR2016/052386

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014373010	A1	18-12-2014	NONE
US 2013055277	A1	28-02-2013	NONE
US 2008244215	A1	02-10-2008	NONE
US 7007276	B1	28-02-2006	BR 0005052 A 19-06-2001
			CN 1292529 A 25-04-2001
			JP 3807916 B2 09-08-2006
			JP 2001134453 A 18-05-2001
			KR 20010050503 A 15-06-2001
			US 7007276 B1 28-02-2006

<p>A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. G06F9/50 ADD.</p> <p>Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB</p>											
<p>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</p> <p>Documentation minimale consultée (système de classification suivi des symboles de classement) G06F</p> <p>Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche</p> <p>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</p>											
<p>C. DOCUMENTS CONSIDERES COMME PERTINENTS</p> <table border="1"> <thead> <tr> <th>Catégorie*</th> <th>Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents</th> <th>no. des revendications visées</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2014/373010 A1 (FOLCO RAFAEL C S [BR] ET AL) 18 décembre 2014 (2014-12-18) alinéa [0012] alinéa [0027] - alinéa [0031] -----</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>US 2013/055277 A1 (ASHISH SHIVENDRA [IN] ET AL) 28 février 2013 (2013-02-28) alinéa [0034] - alinéa [0045] ----- -/--</td> <td>1-10</td> </tr> </tbody> </table>			Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées	X	US 2014/373010 A1 (FOLCO RAFAEL C S [BR] ET AL) 18 décembre 2014 (2014-12-18) alinéa [0012] alinéa [0027] - alinéa [0031] -----	1-10	X	US 2013/055277 A1 (ASHISH SHIVENDRA [IN] ET AL) 28 février 2013 (2013-02-28) alinéa [0034] - alinéa [0045] ----- -/--	1-10
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées									
X	US 2014/373010 A1 (FOLCO RAFAEL C S [BR] ET AL) 18 décembre 2014 (2014-12-18) alinéa [0012] alinéa [0027] - alinéa [0031] -----	1-10									
X	US 2013/055277 A1 (ASHISH SHIVENDRA [IN] ET AL) 28 février 2013 (2013-02-28) alinéa [0034] - alinéa [0045] ----- -/--	1-10									
<p><input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents</p> <p><input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe</p>											
<p>* Catégories spéciales de documents cités:</p> <p>"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent</p> <p>"E" document antérieur, mais publié à la date de dépôt international ou après cette date</p> <p>"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)</p> <p>"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens</p> <p>"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&" document qui fait partie de la même famille de brevets</p>											
<p>Date à laquelle la recherche internationale a été effectivement achevée</p> <p>26 octobre 2016</p>		<p>Date d'expédition du présent rapport de recherche internationale</p> <p>07/11/2016</p>									
<p>Nom et adresse postale de l'administration chargée de la recherche internationale</p> <p>Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016</p>		<p>Fonctionnaire autorisé</p> <p>Dewyn, Torkild</p>									

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>MACIEJ MLYNSKI: "The Analysis of Influence of IBM pSeries Servers' Virtualization Mechanism on Dynamic Resources Allocation in AIX 5L", PARALLEL AND DISTRIBUTED COMPUTING, 2008. ISPCD '08. INTERNATIONAL SYMPOSIUM ON, IEEE, PISCATAWAY, NJ, USA, 1 juillet 2008 (2008-07-01), pages 39-46, XP031378726, ISBN: 978-0-7695-3472-5 le document en entier -----</p>	1-10
A	<p>US 2008/244215 A1 (FLEMMING DIANE G [US] ET AL) 2 octobre 2008 (2008-10-02) le document en entier -----</p>	1-10
A	<p>US 7 007 276 B1 (KUBALA JEFFREY P [US] ET AL) 28 février 2006 (2006-02-28) le document en entier -----</p>	1-10

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/FR2016/052386

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2014373010	A1	18-12-2014	AUCUN

US 2013055277	A1	28-02-2013	AUCUN

US 2008244215	A1	02-10-2008	AUCUN

US 7007276	B1	28-02-2006	BR 0005052 A 19-06-2001
			CN 1292529 A 25-04-2001
			JP 3807916 B2 09-08-2006
			JP 2001134453 A 18-05-2001
			KR 20010050503 A 15-06-2001
			US 7007276 B1 28-02-2006
