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

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

PCT

(43) Date de la publication internationale
16 mars 2006 (16.03.2006)

(10) Numéro de publication internationale
WO 2006/027488 A1

(51) Classification internationale des brevets :
G06F 9/455 (2006.01) *G06F 12/14* (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2005/002196

(22) Date de dépôt international :
2 septembre 2005 (02.09.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0409367 3 septembre 2004 (03.09.2004) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **TRANGO SYSTEMS** [FR/FR]; 5, place Robert Schuman, F-38000 Grenoble (FR).

(72) Inventeur; et

(75) Inventeur/Déposant *(pour US seulement)* : **DEVAUX, Fabrice** [FR/FR]; Chemin de Seloge, F-73800 Les Marches (FR).

(74) Mandataire : **BREESE DERAMBURE MAJEROWICZ**; 38, avenue de l'Opéra, F-75002 Paris (FR).

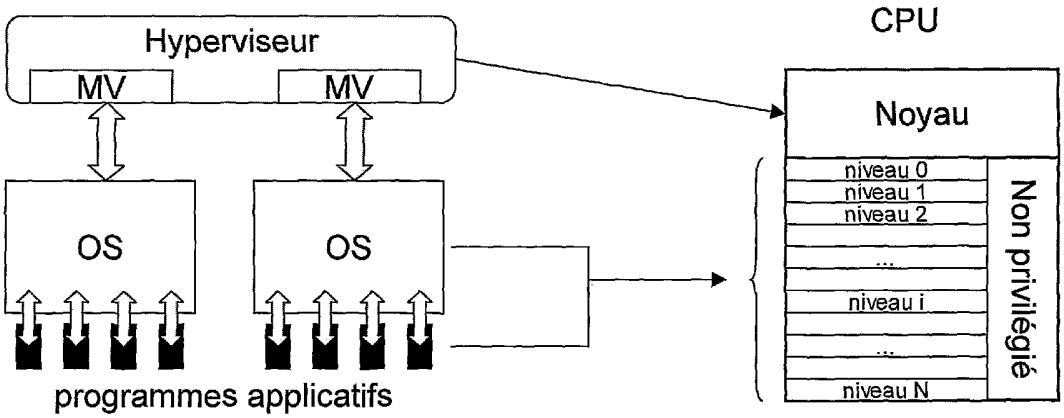
(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK,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

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,

[Suite sur la page suivante]

(54) Title: MECHANISMS FOR CPU VIRTUALIZATION

(54) Titre : MÉCANISMES POUR LA VIRTUALISATION DE CPU



(57) **Abstract:** The invention relates to a method for running, on a processor in a non-privileged mode, different computer programs P while, in a nominal mode, making use of privileged instructions, consisting of running a hypervisor program in a privileged mode of the processor. The hypervisor program furnishes said computer programs P with services equivalent to those available for running in a privileged mode. The source codes of these computer programs P are modified beforehand for replacing the privileged instructions with calls for services furnished by the hypervisor program. The invention is characterized in that the hypervisor program creates at least two privileged submodes organized into a hierarchy within the non-privileged mode and in that the processor comprises only two operating modes.

(57) **Abrégé :** La présente invention se rapporte à un procédé pour l'exécution sur un processeur en mode non privilégié de différents programmes informatiques P faisant en mode nominal usage d'instructions privilégiées, consistant à exécuter un programme hyperviseur en mode privilégié du processeur, ce programme hyperviseur fournissant auxdits programmes informatiques P des services équivalents à ceux disponibles en exécution en mode privilégié, les codes source desdits programmes informatiques P étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par ledit programme hyperviseur, caractérisé en ce que le programme hyperviseur crée au moins deux sous-modes de privilèges hiérarchisés au sein du mode non privilégié et en ce que le processeur dispose de deux modes d'exploitation seulement.

WO 2006/027488 A1

RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *avec rapport de recherche internationale*

— *avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

MÉCANISMES POUR LA VIRTUALISATION DE CPU

La présente invention se rapporte au domaine de l'interfaçage matériel/logiciel dans le monde informatique.

La présente invention se rapporte plus particulièrement à la virtualisation de processeur à deux modes d'exploitation : un mode privilégié et un mode non privilégié. L'invention a pour but d'étendre de façon virtuelle le nombre de niveaux de privilège possibles du processeur.

L'art antérieur connaît déjà, par la publication XEN 2002, University of CAMBRIDGE, Technical Report UCAM-CL-TR-553, un procédé de virtualisation appelé para-virtualisation. Ce procédé met en œuvre des modifications des codes sources des OS à virtualiser de telle façon que les instructions privilégiées soient remplacées par des appels à des services équivalents fournis par un hyperviseur. L'hyperviseur s'exécute dans le mode le plus privilégié du processeur, les OS virtualisés dans un mode moins privilégié que celui de l'hyperviseur, et les processus gérés par lesdits OS virtualisés dans un mode moins privilégié que celui desdits OS. C'est pourquoi ce procédé n'est mis en œuvre uniquement sur des processeurs disposant d'au moins trois niveaux de privilège différents.

La présente invention entend remédier aux inconvénients de l'art antérieur et permettre ainsi de réaliser des para-virtualisations sur des processeurs ne disposant initialement que de deux niveaux de privilèges, en proposant un procédé consistant à sub-diviser le niveau non privilégié du processeur en plusieurs niveaux de privilèges virtuels.

A cet effet, l'invention concerne dans son acception la plus générale un procédé pour l'exécution sur un processeur en mode non privilégié de différents programmes informatiques P faisant en mode nominal usage d'instructions privilégiées, consistant à exécuter un programme hyperviseur en mode privilégié du processeur, ce programme hyperviseur fournissant auxdits programmes informatiques P des services équivalents à ceux disponibles en exécution en mode privilégié, les codes source desdits programmes informatiques P étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par ledit programme hyperviseur,

caractérisé en ce que :

le programme hyperviseur crée au moins deux sous-modes de privilèges hiérarchisés au sein du mode non privilégié et en ce que le processeur dispose de deux modes d'exploitation seulement.

Selon un mode de réalisation, l'un au moins desdits programmes informatiques P est un système d'exploitation (OS) O_i destiné à l'exécution d'au moins un programme applicatif A_{ij} tournant sous lesdits O_i les codes source des OS O_i étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

Selon une variante, l'un au moins desdits programmes informatiques P est un programme applicatif B_k , les codes source dudit programme applicatif B_k étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

Selon une autre variante, les programmes informatiques P comprennent au moins un système d'exploitation (OS) O_i et

au moins un programme applicatif A_{ij} tournant sous lesdits O_i ainsi qu'au moins un programme applicatif B_k, consistant à exécuter un programme hyperviseur en mode privilégié du processeur, ce programme hyperviseur fournissant auxdits OS O_i et programmes applicatifs B_k des services équivalents à ceux disponibles en exécution en mode privilégié, les codes source des OS O_i et desdits programmes applicatifs B_k étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

Selon un mode de mise en œuvre particulier, les programme applicatifs A_{ij} tournent dans au moins un desdits sous-modes moins privilégié que celui de l'OS O_i qui les exécute.

Selon un autre mode de mise en œuvre particulier, les OS O_i tournent dans au moins un desdits sous-modes plus privilégié que celui des programmes applicatifs A_{ij} qu'ils exécutent.

Selon une variante, au moins un des programmes applicatifs A_{ij} est un OS.

Selon une autre variante, un desdits OS est un programme hyperviseur et fournit aux OS qu'il gère des sous-modes de privilèges inférieurs à son propre sous-mode.

Dans un mode de mise en œuvre particulier, ledit hyperviseur met en œuvre des étapes d'adaptation desdits niveaux (sous-modes) de privilège virtuels desdits programmes informatiques P lors d'un appel système dans un niveau de privilège trop faible.

De préférence, lesdites étapes d'adaptation des niveaux de privilège virtuels sont une étape de passage de l'ancien niveau n_{anc} de privilège virtuel d'un programme informatique P_1 à un nouveau niveau n_{nouv} de privilège virtuel et une étape de transfert de l'exécution du service

correspondant audit appel système au gestionnaire d'appel système correspondant audit nouveau niveau de privilège virtuel.

Avantageusement, ladite étape de passage de l'ancien niveau n_{anc} au nouveau niveau n_{nouv} met en œuvre des moyens d'adaptation d'accessibilité aux pages mémoires pour ledit nouveau niveau n_{nouv}.

Dans un mode de réalisation particulier, ledit processeur est à cache physique, lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent une liste d'invalidation créée pour au moins chacun desdits niveaux de privilège virtuels N_i de la machine virtuelle dudit programme informatique P_1 sauf le niveau le moins privilégié, lesdites listes d'invalidation référençant les entrées de translation utilisées pour réaliser les translations d'adresses de niveau de privilège virtuel N_i correspondant, et les entrées de translation desdites listes d'invalidation de niveau de privilège à la fois inférieur ou égal audit ancien niveau n_{anc} et à la fois strictement supérieur audit nouveau niveau n_{nouv} sont invalidées et les listes d'invalidation de niveau de privilège à la fois inférieur ou égal audit niveau n_{anc} et à la fois strictement supérieur audit niveau n_{nouv} sont vidées.

Selon une variante, lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent au moins un sous-ensemble associé à chaque niveau de privilège virtuel, lesdits sous-ensembles regroupant les entrées de translation disponibles pour réaliser les translations audit niveau de privilège virtuel ; et ladite liste d'invalidation de niveau de privilège N_i est restreinte au dit sous-ensemble de même niveau de privilège.

Selon une autre variante, ledit processeur est à cache physique, lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent au moins un sous-ensemble associé à chaque niveau de privilège virtuel, lesdits sous-ensembles regroupant les entrées de translation disponibles pour réaliser les translations audit niveau de privilège virtuel, et les entrées des translations desdits sous-ensembles de niveau de privilège à la fois inférieur ou égal audit ancien niveau n_{anc} et à la fois strictement supérieur audit nouveau niveau n_{nouv} sont invalidées.

Selon une autre variante ledit processeur présente un système matériel d'identification d'espace logique, lesdits ancien et nouveau niveaux (n_{anc}, n_{nouv}) de privilège sont deux niveaux contigus et lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent :

- un numéro unique NL attribué à chaque espace logique virtuel d'une machine virtuelle, ledit numéro NL dépendant de ladite machine virtuelle et dudit espace logique virtuel,
- un ensemble de règles appliquées :

i. pour une translation T de niveau de privilège supérieur ou égal au plus grand niveau de privilège MAX parmi lesdits nouveau et ancien niveaux (n_{anc}, n_{nouv}), ladite translation T est paramétrée comme locale à la valeur de registre d'espace logique NL de ladite machine virtuel et dudit espace logique virtuel courant,

ii. pour une translation T de niveau de privilège inférieur ou égal au plus faible niveau de privilège MIN parmi lesdits nouveau et ancien niveaux (n_{anc}, n_{nouv}), ladite translation T est paramétrée comme globale,

iii. pour une machine virtuelle de niveau de privilège virtuel inférieur audit niveau de privilège MIN, la valeur courante d'espace logique virtuel

contenue dans ledit système matériel d'identification d'espace logique est le numéro global, le numéro global étant commun à toutes les machines virtuelles et différent à tous les numéraux locaux,

iv. pour une machine virtuelle de niveau de privilège virtuel supérieur audit niveau de privilège MAX, la valeur courante d'espace logique virtuel contenue dans ledit système matériel d'identification d'espace logique est ledit numéro NL de ladite machine virtuel et dudit espace logique virtuel courant,

v. lors d'un changement dudit espace virtuel courant, toutes les entrées de TLB réalisant des translations globales sont invalidées,

vi. lors d'un changement de machine virtuelle dans ledit hyperviseur, toutes les entrées de TLB réalisant des translations globales sont invalidées.

Dans un mode de mise en oeuvre, ladite étape (iv) comprend en outre une étape d'invalidation des entrées de translation locales à NL lorsque ledit numéro NL a au moins un synonyme.

Selon un mode de réalisation, le processeur utilise au moins un cache logique et ladite étape (iv) comprend en outre une étape d'invalidation des entrées de cache logiques correspondant aux entrées de translation invalidées.

Selon un autre mode de réalisation, le processeur utilise au moins un cache logique et lesdites étapes (v) et (vi) comprennent en outre une étape d'invalidation des entrées de cache logique correspondant aux translations globales et contenues dans lesdits caches logiques.

Selon une autre variante, ledit processeur est à cache physique et lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent :

- un numéro unique NL attribué à chaque espace logique virtuel d'une machine virtuelle dans un niveau de privilège virtuel donné, ledit numéro NL dépendant de ladite machine virtuel, dudit espace logique virtuel et dudit niveau de privilège virtuel,
- un ensemble de règles appliquées :
  i. pour tout changement soit de ladite machine virtuelle, soit dudit espace logique virtuel ou soit dudit niveau de privilège virtuel, la nouvelle valeur NL est affectée comme valeur courante d'espace virtuel,
  ii. toute translation est effectuée localement à ladite valeur NL courante.
  iii. si la nouvelle valeur NL affectée possède au moins un synonyme, alors toutes les entrées de translation locales à NL sont invalidées.

On comprendra mieux l'invention à l'aide de la description, faite ci-après à titre purement explicatif, d'un mode de réalisation de l'invention, en référence aux figures annexées où :

- la figure 1 représente un système de l'art antérieur reposant sur un hyperviseur ;
- la figure 2 représente un système de l'art antérieur utilisant un processeur disposant de trois niveaux de privilèges différents ;
- la figure 3 représente un schéma structurel l'utilisation du niveau de privilège inférieur du processeur selon la présente invention ; et
- les figures 4 et 5 illustrent la gestion de l'accès aux pages mémoires selon la présente invention.

Les figures 1 et 2 représentent un système selon la publication XEN 2002 qui décrit l'utilisation d'un hyperviseur dans le mode le plus privilégié du processeur (Noyau). Les systèmes d'exploitations virtualisés sur l'hyperviseur effectuent des appels à l'hyperviseur dans un mode de privilège inférieur à celui de l'hyperviseur (dans le cas de la figure 2, le mode intermédiaire) et les processus (1) gérés par les systèmes d'exploitation sont exécutés dans un mode encore moins privilégié que celui des systèmes d'exploitation (mode non privilégié).

La présente invention met en œuvre un procédé de para-virtualisation sur un processeur disposant initialement de deux modes de privilèges : un niveau privilégié du processeur et un niveau non privilégié du processeur.

Un programme, dit programme hyperviseur, est exécuté en mode privilégié du processeur. Ledit hyperviseur gère des machines virtuelles MV_i, chaque machine virtuelle étant associée à un programme informatique P_i exécuté en mode non privilégié du processeur. Une machine virtuelle MV_i est un contexte regroupant les accès aux ressources systèmes ainsi que l'état des registres hardware du processeur pour un programme informatique Pi unique. Lorsque l'hyperviseur passe de l'exécution d'un programme informatique à un autre, les informations (comme le niveau de privilège, les registres du processeur) de la machine virtuelle associée au programme quitté sont sauvegardées tandis que le contexte du nouveau programme est restauré comme paramètres du système.

Dans un mode de réalisation, les programmes Pi sont des systèmes d'exploitation dans lesquels est exécuté au moins un programme applicatif.

Dans une variante, un programme Pi est un hyperviseur fournissant, aux « sous-programmes » qu'il gère, différents modes de privilège hiérarchisés et inférieurs à son propre

sous-mode, et mettant en œuvre le procédé de la présente invention.

Pour chacune des machines virtuelles MV_i, l'hyperviseur fournit N_i sous-modes ou niveaux de privilège virtuel et M_i espaces virtuels d'adressage ainsi qu'un certain nombre de services, accessibles au moyen d'appels systèmes. Comme illustré par la figure 3, ces N_i niveaux de privilège virtuel peuvent, par exemple, être caractérisés par des nombres entiers.

Les programmes informatiques P_i déjà existants et tournant sur les machines virtuelles, sont adaptés pour pouvoir faire appel auxdits services fournis par l'hyperviseur. Cette adaptation est réalisée par une modification des codes source des programmes P_i afin de remplacer les instructions privilégiées par des appels aux services fournis pas l'hyperviseur. Ainsi les programmes Pi ont accès, de façon indirecte et sécurisée, aux instructions privilégiées du système alors même qu'ils sont en mode non privilégié du processeur.

Les N_i niveaux de privilèges virtuels d'une machine virtuelle MV_i sont hiérarchisés entre eux, du niveau le plus privilégié au niveau le moins privilégié.

Lorsqu'un programme Pi associé à une machine virtuelle MV_i, ledit programme Pi étant exécuté dans un niveau de privilège virtuel n_i donné, réalise un appel système, le gestionnaire d'appel système de l'hyperviseur est appelé dans le mode privilégié du processeur. Ce peut être un appel système généré par le programme Pi lui-même ou issu d'un programme applicatif exécuté dans le programme Pi lorsque celui-ci est un système d'exploitation.

Le gestionnaire d'appel système vérifie alors le niveau de privilège virtuel n_i dans lequel l'appel système a

été fait. Soit le niveau de privilège virtuel est suffisamment élevé pour réaliser l'appel système et l'hyperviseur exécute le service correspondant à l'appel système, puis reprend l'exécution de Pi. Soit le niveau de privilège virtuel n'est pas suffisant pour l'exécution du service correspondant à l'appel système et l'hyperviseur passe le niveau de privilège virtuel n_i dans un niveau de privilège virtuel supérieur, puis transfère l'exécution dudit service, en mode non-privilégié du processeur, au gestionnaire d'appel système, correspondant au nouveau niveau de privilège virtuel, du programme Pi. Concrètement, l'hyperviseur dispose d'une variable dans laquelle est stocké le niveau de privilège du programme informatique Pi actif. Pour changer le niveau de privilège, l'hyperviseur modifie cette valeur par une valeur issue de l'ensemble des Ni niveaux de privilèges attribués à la machine virtuelle.

Pour chaque machine virtuelle, l'hyperviseur fournit un service, accessible seulement à partir d'un certain niveau de privilège virtuel, appelé service de commutation d'espace virtuel. Celui-ci indique quel est l'espace virtuel d'adressage courant, la lecture des instructions, la lecture des opérandes desdites instructions, ainsi que l'écriture de leurs résultats, s'effectuant toujours à partir de l'espace virtuel d'adressage courant.

L'accessibilité d'un espace virtuel dépend du niveau de privilège virtuel. L'espace virtuel est entièrement accessible depuis le niveau de privilège virtuel le plus élevé, alors que son accessibilité peut être restreinte depuis des niveaux de privilège virtuels inférieurs. De plus, un espace virtuel accessible depuis un niveau de privilège n_i donné, l'est aussi depuis tous les niveaux de privilège supérieurs à n_i.

La gestion des espaces virtuels se fait par une unité de gestion mémoire MMU qui comprend des entrées de translations. Ces translations correspondent à une association entre une page d'un espace virtuel et une page de mémoire physique de taille identique.

L'hyperviseur contrôle la MMU et, de ce fait, fournit aux machines virtuelles MV_i des services, appelés services de translation, permettant d'effectuer une translation.

Lors de l'appel d'un service de translation, le programme appelant spécifie à partir de quel niveau de privilège virtuel n_i, la page est accessible, la translation résultante étant dite du niveau de privilège virtuel n_i. Le niveau de privilège spécifié ne peut excéder le niveau de privilège virtuel dans lequel le service de translation est appelé.

Avant de programmer de façon adéquate une entrée de translation dans la MMU, les services de translation vérifient, entre autres, que la translation demandée ne risque pas de compromettre l'intégrité de l'hyperviseur ni des autres machines virtuelles.

Habituellement, mais pas exclusivement, un programme P_i fera appel aux services de translation de l'hyperviseur, lors d'un évènement de « translation absente ». Un tel évènement se produit lorsqu'aucune entrée de translation TLB n'indique quelle adresse physique est associée à l'adresse virtuelle générée lors d'un accès.

Cet évènement est traité initialement par l'hyperviseur, en fonction du niveau de privilège virtuel dans lequel l'accès a été tenté et en fonction de l'adresse dudit accès ; soit l'hyperviseur traite directement l'évènement, soit il redirige l'exécution de l'évènement, en

mode non privilégié du processeur, mais avec un privilège virtuel de niveau supérieur, vers une routine de Pi, dédiée au traitement de cet évènement. La routine pouvant être choisie parmi plusieurs routines, en fonction par exemple du niveau de privilège virtuel dans lequel l'accès a été tenté. Ladite routine, par le moyen des services de translation, est à même de corriger l'événement puis de re-exécuter l'instruction qui en était la cause. Ainsi le niveau de privilège virtuel du programme Pi est augmenté jusqu'à être suffisamment élevé pour que le traitement soit réalisé.

Lorsque la gestion des translations se fait au fur et à mesure des évènements de « translation absente », la gestion des translations est dite « à la demande ».

Le passage d'un niveau de privilège à un autre d'un programme Pi au moyen des services fournis par l'hyperviseur nécessite que l'accessibilité des pages mémoires soit adaptée en fonction du nouveau niveau de privilège. On distingue trois modes de réalisation particulièrement avantageux.

Dans un mode de réalisation possible, le processeur du système possède des caches physiques, c'est-à-dire que le cache travaille avec des adresses physiques. Lors d'un accès à la mémoire, l'adresse est translatée par l'unité de gestion mémoire MMU avant d'être présenté au cache.

Lorsqu'un service de translation établit une translation de privilège virtuel n_i, à l'aide d'une entrée de translation TLB_α, il enregistre l'indice α de l'entrée de translation dans une liste correspondant à ce niveau de privilège virtuel n_i, appelée liste d'invalidation n_i. Chaque niveau de privilège, excepté le moins privilégié, possède une liste d'invalidation.

Si cet indice était déjà enregistré dans une quelconque liste d'invalidation n_k, alors il en est retiré puis il est ajouté dans la liste d'invalidation n_i, l'indice d'entrée de translation n'étant alors référencé qu'une seule fois parmi toutes les listes d'invalidation.

Lorsqu'un programme P_i avec un niveau de privilège virtuel n_i abaisse, par l'intermédiaire d'un appel système à l'hyperviseur, son niveau de privilège virtuel au niveau de privilège virtuel n_j, l'hyperviseur invalide toutes les entrées de translation référencées dans les listes d'invalidations correspondant aux niveaux de privilège (n_i, n_j+1), c'est-à-dire du niveau de privilège n_i à celui immédiatement supérieur à n_j, puis re-initialise ces listes d'invalidation de sorte qu'elles deviennent vides. L'invalidation d'une entrée de translation signifie que la translation qu'elle réalisait n'est plus connue, et qu'un accès mémoire nécessitant ladite translation générera un événement de translation absente.

Il n'y a aucune action réalisée sur la liste d'invalidations n_j ce qui explique pourquoi il n'est pas nécessaire d'avoir une liste d'invalidations pour le niveau de privilège le moins privilégié.

Selon une variante, les entrées de translation TLB disponibles sont réparties en plusieurs sous-ensembles, un sous-ensemble étant associé à un niveau de privilège virtuel n_i donné. Une translation de niveau de privilège virtuel n_i est réalisée au moyen d'une entrée de translation choisie dans le sous-ensemble associé à ce niveau de privilège virtuel n_i.

Lorsqu'un programme P_i avec un niveau de privilège virtuel n_i abaisse son niveau de privilège virtuel au niveau du privilège virtuel n_j, l'hyperviseur invalide toutes les entrées de translation des sous-ensembles correspondant aux

niveaux de privilège virtuel (n_i, n_j+1), c'est-à-dire du niveau de privilège n_i à celui immédiatement supérieur à n_j.

Selon une autre variante, la liste d'invalidation associée au niveau de privilège virtuel n_i, telle que décrite ci-dessus, ne peut contenir uniquement que des indices d'entrées de translation du sous-ensemble associé à ce même niveau de privilège virtuel n_i, tel que décrit ci-dessus. On dit que la liste d'invalidation de niveau de privilège n_i est restreinte au sous-ensemble de même niveau de privilège.

Dans un mode de réalisation alternatif, le processeur possède un système matériel d'identification d'espace logique. Avec un système matériel d'identification d'espace logique, on limite les invalidations de MMU sur les processeurs à cache physique ou logique et les invalidations de cache sur les processeurs à cache logique.

Le numéro d'identification d'espace logique, numéro affecté par l'hyperviseur à chaque espace logique utilisé par un processus, est contenu dans un registre matériel associé au processeur, appelé registre d'espace logique courant, et accessible seulement en mode privilégié du processeur, donc uniquement par l'hyperviseur.

Comme illustré par la figure 4, le contenu du registre d'espace logique courant préfixe les adresses logiques générées par un programme. Les adresses étendues résultantes sont utilisées par la MMU et par les caches logiques, en lieu et place des adresses logiques initiales.

Dans la MMU, chaque entrée de translation TLB_i contient un champ supplémentaire permettant d'indiquer pour

quelle valeur du registre d'espace logique courant, la translation est valide.

De plus, pour chaque entrée de translation TLB_i, un mécanisme, par exemple par la mise en place d'un bit d'activation ou de désactivation, permet d'indiquer :

- si la valeur de ce champ doit être ignorée et dans ce cas la validité de la translation est indépendante de la valeur du registre d'espace logique courant, et la translation est dite globale.
- si la valeur de ce champ doit être utilisée et dans ce cas la validité de la translation est dépendante de la valeur du registre d'espace logique courant, et la translation est dite locale.

A titre indicatif, dans la majorité des systèmes actuels où il n'est pas nécessaire de créer des niveaux de privilèges virtuels, le système matériel d'identification d'espace logique permet l'implémentation directe des espaces logiques virtuels en attribuant un numéro d'identification à chacun de ces espaces.

Dans ce mode de réalisation, une valeur du registre d'espace logique courant est attribuée parmi les valeurs possibles, par l'hyperviseur. Cette valeur, appelé le numéro global, est commune à toutes les machines virtuelles MV_i.

De l'ensemble des valeurs possibles du registre d'espace logique courant, retranché du numéro global, l'hyperviseur attribue, de façon exclusive, un numéro à chaque espace logique virtuel de chaque machine virtuelle. NL(i, j) représente le numéro attribué à l'espace logique virtuel j de la machine virtuelle MV_i. Le numéro global est ainsi unique et distinct de tous les numéros NL(i, j) attribués par l'hyperviseur aux différents espaces logiques.

Lorsque toutes les valeurs possibles ont été utilisées, une attribution exclusive d'un numéro ne peut

être effectuée. Un numéro NL(u, v) déjà utilisé est alors attribué à un espace logique virtuel v d'une machine virtuelle MV_u. Dans ce cas, NL(i, j) = NL(u, v) et NL(i, j) et NL(u, v) sont dits synonymes.

Le mécanisme de gestion de la mémoire lors d'une transition, entre deux niveaux de privilèges n_1 et n_2, n_1 étant le niveau immédiatement supérieur à n_2, requise suite à un appel système à l'hyperviseur, est régi par les étapes suivantes qui s'appliquent aux numéros NL(i,j) d'espaces logiques virtuels des machines virtuelles :

i. Lorsqu'un service de translation est appelé, soit le niveau de privilège de la translation demandée est supérieur ou égal a n_1, alors la translation est paramétrée comme locale à NL(i, j), soit le niveau de privilège virtuel de la translation demandée est inférieur ou égal a n_2, alors la translation est paramétrée comme globale.

ii. Lorsque la machine virtuelle MV_i a un niveau de privilège virtuel n inférieur ou égal au niveau de privilège n_2, alors le registre de numéro d'espace logique courant prend la valeur du numéro global (d'espace commun).
Ainsi, les pages mémoires d'une liste d'invalidations accessibles depuis le niveau n_2 et translatées par la MMU, sont visibles indépendamment de la valeur du numéro d'espace logique courant. Le numéro global n'étant jamais utilisé pour réaliser une translation locale, en positionnant le registre d'espace logique courant à la valeur du numéro global, on est sûr de ne jamais pouvoir accéder, par inadvertance, à une page translatée localement (et réservées aux translations

visibles à partir de n_1),ou à des translations de niveau privilégié d'autres machines virtuelles.

iii. Lorsque la machine virtuelle MV_i, ayant pour espace virtuel courant l'espace virtuel j, a un niveau de privilège virtuel n supérieur ou égal au niveau de privilège n_1, le registre d'espace logique courant prend la valeur NL(i, j). Si cette valeur a au moins un synonyme, alors l'ensemble des entrées de translation locales à cette valeur est invalidé, et pour les processeurs ayant des caches logiques, les invalidations correspondantes sont réalisées.
En effet, il n'est pas nécessaire d'invalider les entrées de translation qui ne sont pas concernées par ce synonyme. A chaque fois qu'une valeur NL ayant des synonymes est chargée dans le registre d'espace courant, alors il faut invalider les entrées locales à la valeur de NL, car autrement, des pages seraient visibles depuis d'autres machines virtuelles, d'autres espaces logiques ou d'autres privilèges virtuels.
En l'absence de synonyme, cette règle rend visibles toutes les pages des listes d'invalidations, translatées par la MMU qui ne sont visibles qu'a partir de n_1. La présence de synonymes est gérée assez brutalement, par une invalidation massive des entrées utilisant le synonyme correspondant à la valeur nouvellement chargée dans le registre d'espace courant, entraînant une perte de performance, mais n'entraînant pas de brèche dans la sécurité ni d'incohérence. Le bénéfice de l'invention se réalise lors des transitions de niveau de privilège virtuel ou d'espace virtuel ou de machine virtuelle entre non-synonyme.

iv. Lorsque l'espace virtuel courant change, toutes les entrées de TLB réalisant des translations globales sont invalidées. Pour les processeurs utilisant des caches logiques, les entrées dans ces caches correspondant à des translations globales doivent être invalidées de manière similaire.
Par cette règle, l'hyperviseur interdit que des pages mémoires globales de l'ancien espace virtuel courant soient accessibles depuis le nouvel espace virtuel.

v. Lorsque l'hyperviseur passe de l'exécution d'une machine virtuelle à une autre, toutes les entrées de TLB réalisant des translations globales doivent être invalidées. Pour les processeurs utilisant des caches logiques, les entrées correspondant à des translations globales et contenues dans ces caches doivent être invalidées de manière similaire.

Ceci permet de ne pas rendre les pages mémoires d'une machine virtuelle accessibles aux processus d'une nouvelle machine virtuelle.

Le mécanisme décrit par ces règles permet de gérer efficacement les transitions entre deux niveaux de privilège adjacents.

Ainsi, une transition d'un niveau de privilège virtuel n_0 vers un niveau de privilège virtuel n_3, où n_0 est de privilège supérieur à n_1, n_3 est de privilège inférieur à n_2 et, n_1 et n_2 sont comme décrits précédemment, peut être décomposée en :

- une transition de n_0 vers n_1
- une transition de n_1 vers n_2
- une transition de n_2 vers n_3

Si le mécanisme décrit dans ce mode de réalisation est utilisé pour gérer les transitions entre n_1 et n_2, alors d'autres mécanismes devront être utilisés pour gérer les transitions de n_0 vers n_1 et les transitions de n_2 vers n_3.

Dans la pratique, qui concerne les cas de para-virtualisation les plus courants où il n'y a besoin de créer que deux niveaux de privilège virtuels, ce mécanisme suffit.

Selon un mode de réalisation avantageux et illustré par la figure 5, le processeur possède un système matériel d'identification d'espace logique ainsi que des caches physiques.

A chaque machine virtuelle MV_i possédant N_i niveaux de privilèges et M_i espaces virtuels, N_i x M_i valeurs distinctes d'identifiants d'espace logique sont choisies. NL(i, j, k) représente l'identifiant d'espace logique qui correspond a la machine virtuelle MV_i, ayant pour espace virtuel courant j, et au niveau de privilège virtuel k.

Lors d'un changement concernant soit la machine virtuelle, soit le niveau de privilège de la machine virtuelle, soit l'espace virtuel courant de la machine virtuelle, la nouvelle valeur NL(i, j, k) est écrite dans le registre d'identifiant d'espace courant.

Lorsqu'un service de translation est appelé (lors d'un appel système d'accès à la mémoire), la translation est toujours effectuée localement à la valeur de NL(i, j, k) courante.

Dans ce mode de réalisation, une page mémoire qui est accessible depuis N_i niveaux de privilège virtuels et M_i espaces virtuels d'adressage, pourrait éventuellement avoir

jusqu'à N_i x M_i entrées de translations positionnées de telle façon que la page soit accessible depuis N_i niveaux de privilège virtuels et depuis les M_i espaces virtuels d'adressage. Dans la pratique, la gestion des translations se fait à la demande. Le nombre d'entrées de translations est en moyenne beaucoup plus faible que les N_i x M_i théoriques.

Des synonymes peuvent exister aussi avec ce mécanisme et une gestion des synonymes, telle que décrite précédemment, s'applique. A savoir qu'en cas de synonymie, l'ensemble des entrées de translation locales à la valeur contenue dans le registre d'espace logique courant, est invalidé.

Ce mode de réalisation met, obligatoirement, en œuvre des processeurs à caches physiques. Cette obligation vient du fait que plusieurs entrées de translation peuvent désigner la même page physique que l'on nomme sous le terme de phénomène d'alias. Or les caches logiques deviennent incohérents, c'est-à-dire qu'ils ne fonctionnent pas correctement, en présence d'alias.

L'invention est décrite dans ce qui précède à titre d'exemple. Il est entendu que l'homme du métier est à même de réaliser différentes variantes de l'invention sans pour autant sortir du cadre du brevet.

REVENDICATIONS

1. Procédé pour l'exécution sur un processeur en mode non privilégié de différents programmes informatiques P faisant en mode nominal usage d'instructions privilégiées, consistant à exécuter un programme hyperviseur en mode privilégié du processeur, ce programme hyperviseur fournissant auxdits programmes informatiques P des services équivalents à ceux disponibles en exécution en mode privilégié, les codes source desdits programmes informatiques P étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par ledit programme hyperviseur,

caractérisé en ce que :

le programme hyperviseur crée au moins deux sous-modes de privilèges hiérarchisés au sein du mode non privilégié et en ce que le processeur dispose de deux modes d'exploitation seulement.

2. Procédé selon la revendication 1, caractérisé en ce que l'un au moins desdits programmes informatiques P est un système d'exploitation (OS) O_i destiné à l'exécution d'au moins un programme applicatif A_{ij} tournant sous lesdits O_i les codes source des OS O_i étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

3. Procédé selon la revendication 1, caractérisé en ce que l'un au moins desdits programmes informatiques P est un programme applicatif B_k, les codes source dudit programme applicatif B_k étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

4. Procédé selon la revendication 1, caractérisé en ce que les programmes informatiques P comprennent au moins un système d'exploitation (OS) O_i et au moins un programme applicatif A_{ij} tournant sous lesdits OS O_i ainsi qu'au moins un programme applicatif B_k, consistant à exécuter un programme hyperviseur en mode privilégié du processeur, ce programme hyperviseur fournissant auxdits OS O_i et programmes applicatifs B_k des services équivalents à ceux disponibles en exécution en mode privilégié, les codes source des OS O_i et desdits programmes applicatifs B_k étant préalablement modifiés pour remplacer les instructions privilégiées par des appels aux services fournis par le programme hyperviseur.

5. Procédé selon la revendication 2 ou 4, caractérisé en ce que les programme applicatifs A_{ij} tournent dans au moins un desdits sous-modes moins privilégié que celui de l'OS O_i qui les exécute.

6. Procédé selon la revendication 2 ou 4, caractérisé en ce que les OS O_i tournent dans au moins un desdits sous-modes plus privilégié que celui des programmes applicatifs A_{ij} qu'ils exécutent.

7. Procédé selon la revendication 5 ou 6, caractérisé en ce qu'au moins un des programmes applicatifs A_{ij} est un OS.

8. Procédé selon la revendication 2 ou 4, caractérisé en ce qu'un desdits OS est un programme hyperviseur et fournit aux OS qu'il gère des sous-modes de privilèges inférieurs à son propre sous-mode.

9. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ledit hyperviseur met en œuvre des étapes d'adaptation desdits niveaux (sous-modes) de privilège virtuels desdits programmes informatiques P lors d'un appel système dans un niveau de privilège trop faible.

10. Procédé selon la revendication précédente, caractérisé en ce que lesdites étapes d'adaptation des niveaux de privilège virtuels sont une étape de passage de l'ancien niveau n_{anc} de privilège virtuel d'un programme informatique P_1 à un nouveau niveau n_{nouv} de privilège virtuel et une étape de transfert de l'exécution du service correspondant audit appel système au gestionnaire d'appel système correspondant audit nouveau niveau de privilège virtuel.

11. Procédé selon la revendication précédente, caractérisé en ce que ladite étape de passage de l'ancien niveau n_{anc} au nouveau niveau n_{nouv} met en œuvre des moyens d'adaptation d'accessibilité aux pages mémoires pour ledit nouveau niveau n_{nouv}.

12. Procédé selon la revendication précédente, caractérisé en ce que ledit processeur est à cache physique, lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent une liste d'invalidation créée pour au moins chacun desdits niveaux de privilège virtuels N_i de la machine virtuelle dudit programme informatique P_1 sauf le niveau le moins privilégié, lesdites listes d'invalidation référençant les entrées de translation utilisées pour réaliser les translations d'adresses de niveau de privilège virtuel N_i correspondant, et les entrées de translation desdites listes d'invalidation de niveau de privilège à la fois inférieur ou égal audit ancien niveau n_{anc} et à la fois

strictement supérieur audit nouveau niveau n_{nouv} sont invalidées et les listes d'invalidation de niveau de privilège à la fois inférieur ou égal audit niveau n_{anc} et à la fois strictement supérieur audit niveau n_{nouv} sont vidées.

13. Procédé selon la revendication 12, caractérisé en ce que :

- lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent au moins un sous-ensemble associé à chaque niveau de privilège virtuel, lesdits sous-ensembles regroupant les entrées de translation disponibles pour réaliser les translations audit niveau de privilège virtuel ;

- ladite liste d'invalidation de niveau de privilège N_i est restreinte audit sous-ensemble de même niveau de privilège.

14. Procédé selon la revendication 11, caractérisé en ce que ledit processeur est à cache physique, lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent au moins un sous-ensemble associé à chaque niveau de privilège virtuel, lesdits sous-ensembles regroupant les entrées de translation disponibles pour réaliser les translations audit niveau de privilège virtuel, et les entrées des translations desdits sous-ensembles de niveau de privilège à la fois inférieur ou égal audit ancien niveau n_{anc} et à la fois strictement supérieur audit nouveau niveau n_{nouv} sont invalidées.

15. Procédé selon la revendication 11, caractérisé en ce que ledit processeur présente un système matériel d'identification d'espace logique, lesdits ancien et nouveau niveaux (n_{anc}, n_{nouv}) de privilège sont deux niveaux contigus et lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent :

- un numéro unique NL attribué à chaque espace logique virtuel d'une machine virtuelle, ledit numéro NL dépendant de ladite machine virtuelle et dudit espace logique virtuel,

- un ensemble de règles appliquées :

i. pour une translation T de niveau de privilège supérieur ou égal au plus grand niveau de privilège MAX parmi lesdits nouveau et ancien niveaux (n_{anc}, n_{nouv}), ladite translation T est paramétrée comme locale à la valeur de registre d'espace logique NL de ladite machine virtuel et dudit espace logique virtuel courant,

ii. pour une translation T de niveau de privilège inférieur ou égal au plus faible niveau de privilège MIN parmi lesdits nouveau et ancien niveaux (n_{anc}, n_{nouv}), ladite translation T est paramétrée comme globale,

iii. pour une machine virtuelle de niveau de privilège virtuel inférieur audit niveau de privilège MIN, la valeur courante d'espace logique virtuel contenue dans ledit système matériel d'identification d'espace logique est le numéro global, le numéro global étant commun à toutes les machines virtuelles et différent à tous les numéraux locaux,

iv. pour une machine virtuelle de niveau de privilège virtuel supérieur audit niveau de privilège MAX, la valeur courante d'espace logique virtuel contenue dans ledit système matériel d'identification d'espace logique est ledit numéro NL de ladite machine virtuel et dudit espace logique virtuel courant,

v. lors d'un changement dudit espace virtuel courant, toutes les entrées de TLB réalisant des translations globales sont invalidées,

vi. lors d'un changement de machine virtuelle dans ledit hyperviseur, toutes les entrées de TLB réalisant des translations globales sont invalidées.

16. Procédé selon la revendication précédente, caractérisé en ce que ladite étape (iv) comprend en outre une étape d'invalidation des entrées de translation locales à NL lorsque ledit numéro NL a au moins un synonyme.

17. Procédé selon la revendication précédente, caractérisé en ce que le processeur utilise au moins un cache logique et ladite étape (iv) comprend en outre une étape d'invalidation des entrées de cache logiques correspondantes aux entrées de translation invalidées.

18. Procédé selon la revendication 15, caractérisé en ce que le processeur utilise au moins un cache logique et lesdites étapes (v) et (vi) comprennent en outre une étape d'invalidation des entrées de cache logique correspondant aux translations globales et contenues dans lesdits caches logiques.

19. Procédé selon la revendication 11, caractérisé en ce que ledit processeur est à cache physique et lesdits moyens d'adaptation d'accessibilité aux pages mémoires comportent :

- un numéro unique NL attribué à chaque espace logique virtuel d'une machine virtuelle dans un niveau de privilège virtuel donné, ledit numéro NL dépendant de ladite machine virtuel, dudit espace logique virtuel et dudit niveau de privilège virtuel,

- un ensemble de règles appliquées :

i. pour tout changement soit de ladite machine virtuelle, soit dudit espace logique virtuel ou soit dudit niveau de privilège virtuel, la nouvelle valeur NL est affectée comme valeur courante d'espace virtuel,

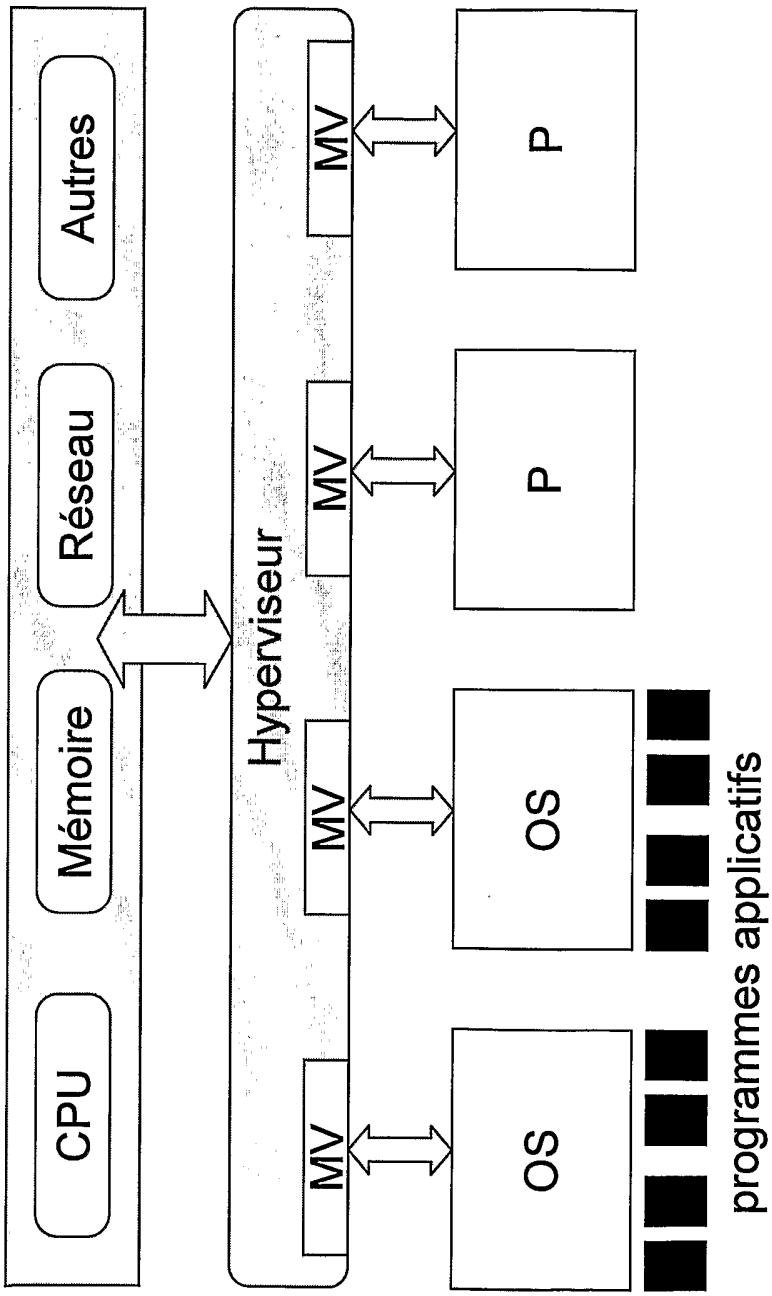
ii. toute translation est effectuée localement à ladite valeur NL courante.

20. Procédé selon la revendication 19, caractérisé en ce que ledit ensemble de règles comprend en outre la règle suivante :

iii. si la nouvelle valeur NL affectée possède au moins un synonyme, alors toutes les entrées de translation locales à NL sont invalidées.

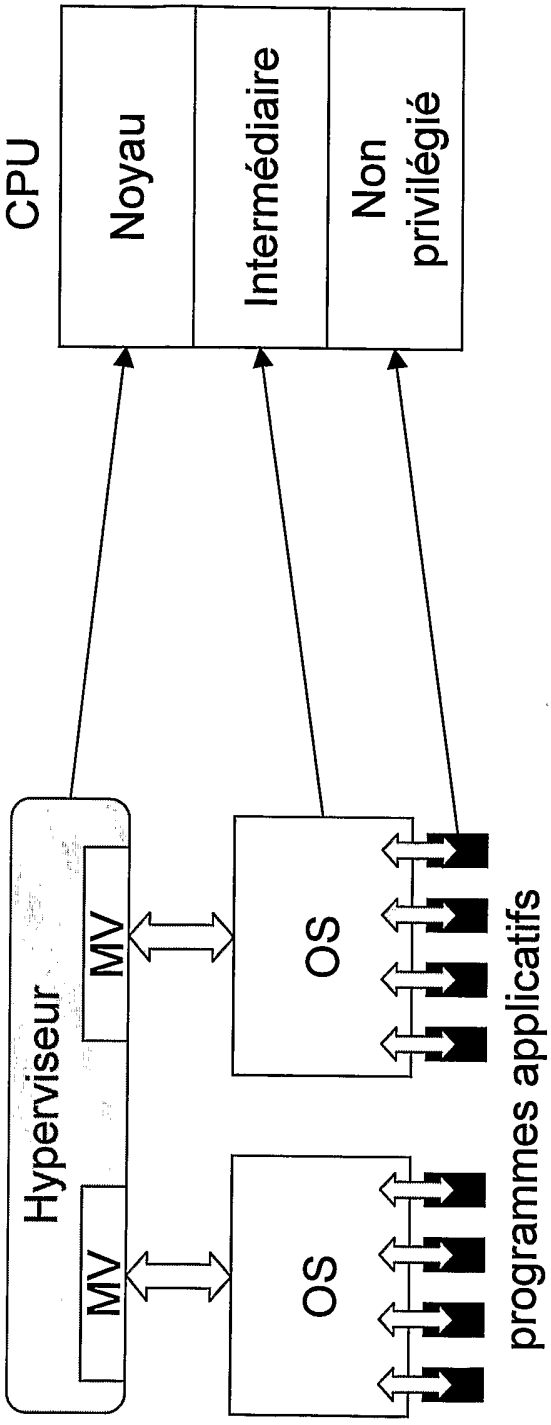
21. Système pour la mise en œuvre du procédé selon l'une quelconque des revendications précédentes, caractérise en ce que le processeur dispose de deux modes d'exploitation.

Figure 1



ART ANTÉRIEUR

Figure 2



ART ANTÉRIEUR

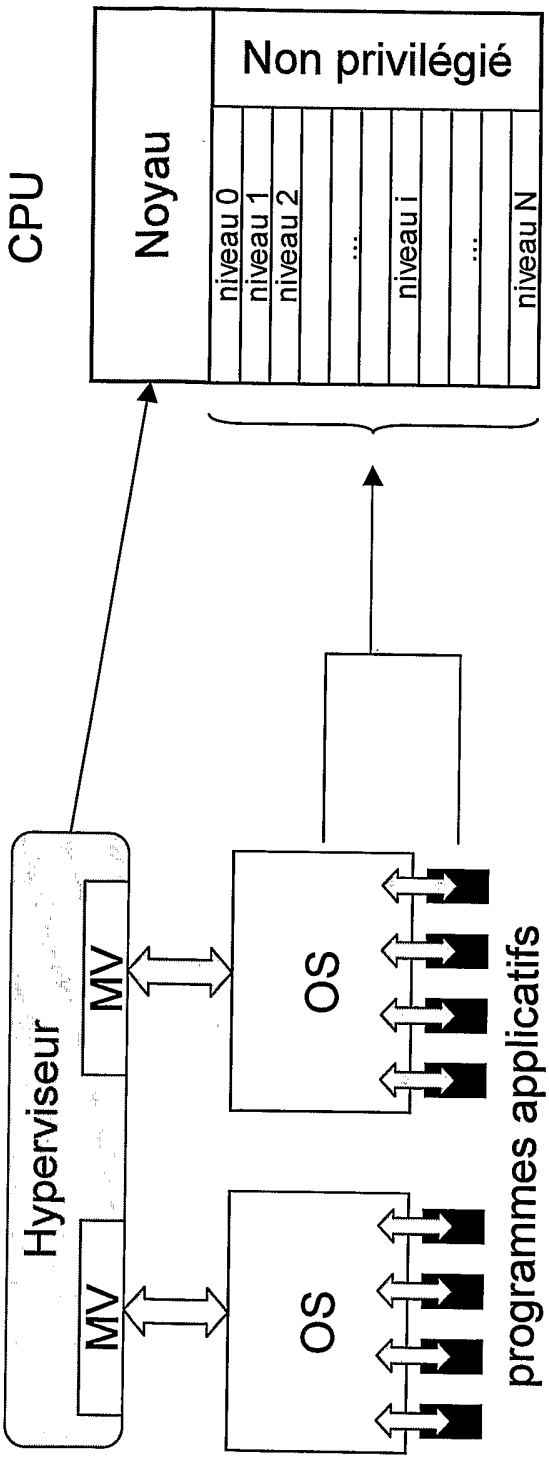

CPU
Noyau
Non privilégié
niveau 0
niveau 1
niveau 2
...
niveau i
...
niveau N
Hyperviseur
MV
MV
OS
OS
programmes applicatifs

Figure 3

Figure 4

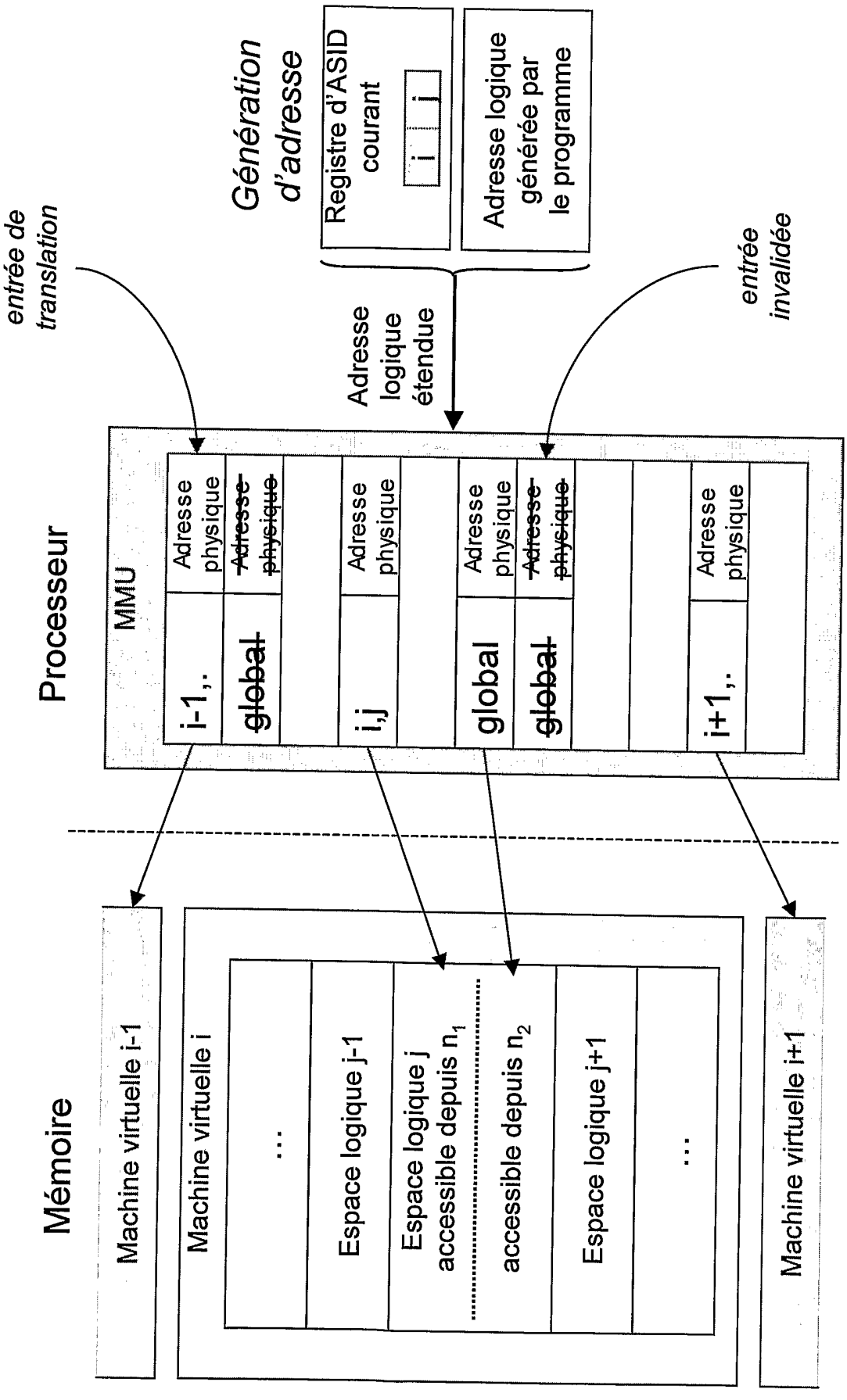
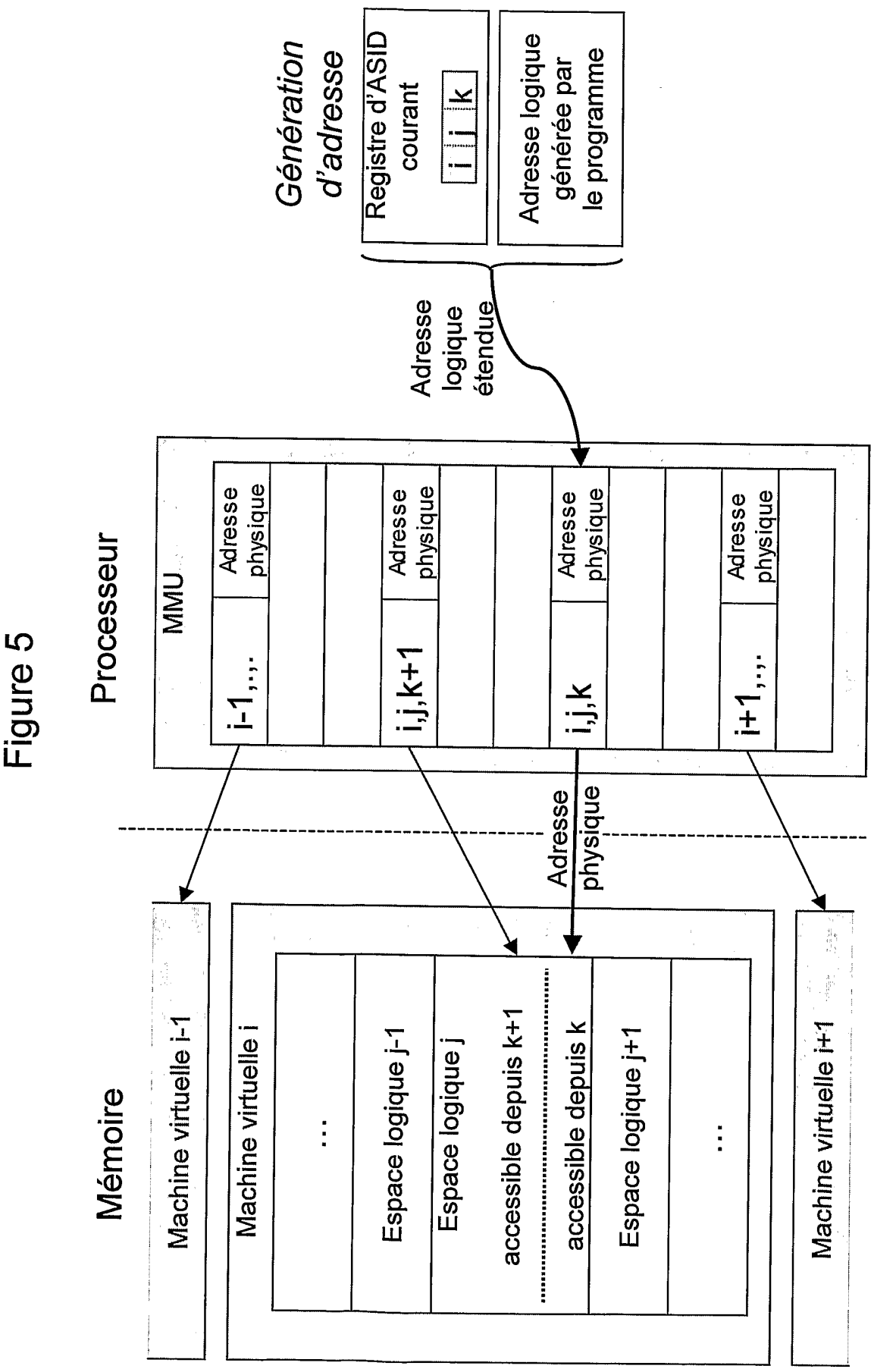

Mémoire
Machine virtuelle i-1
Machine virtuelle i
...
Espace logique j-1
Espace logique j accessible depuis n_1
accessible depuis n_2
Espace logique j+1
...
Machine virtuelle i+1
Processeur
MMU
i-1,,.
Adresse physique
~~Adresse physique~~
~~global~~
i,j
Adresse physique
global
Adresse physique
~~Adresse physique~~
~~global~~
i+1,,.
Adresse physique
entrée de translation
entrée invalidée
Adresse logique étendue
Génération d'adresse
Registre d'ASID courant
i j
Adresse logique générée par le programme

Figure 5


Mémoire
Processeur
Génération d'adresse
Registre d'ASID courant
i j k
Adresse logique générée par le programme
Adresse logique étendue
MMU
Adresse physique
i-1,,,,
i,j,k+1
i,j,k
i+1,,,,
Adresse physique
Machine virtuelle i-1
Machine virtuelle i
...
Espace logique j-1
Espace logique j
accessible depuis k+1
accessible depuis k
Espace logique j+1
...
Machine virtuelle i+1

INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2005/002196

A. CLASSIFICATION OF SUBJECT MATTER
G06F9/455 G06F12/14

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 practical, search terms used)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHANG RONG TSAI ET AL: "ON THE ARCHITECTURAL SUPPORT FOR LOGICAL MACHINE SYSTEMS" MICROPROCESSING AND MICROPROGRAMMING, ELSEVIER SCIENCE PUBLISHERS, BV., AMSTERDAM, NL, vol. 22, no. 2, 1 February 1988 (1988-02-01), pages 81-96, XP000284881 ISSN: 0165-6074 page 84, right-hand column, line 6 - page 86, right-hand column, line 9 page 87, left-hand column, line 14 - last line ----- -/--	1-7, 9-11,21

[X] Further documents are listed in the continuation of Box C.

[X] 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 document 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	Date of mailing of the international search report
30 December 2005	06/02/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Fournier, N

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2005/002196

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	S T KING, G W DUNLAP, P M CHEN: "Operating System Support for Virtual Machines" PROCEEDINGS OF THE 2003 USENIX ANNUAL TECHICAL CONFERENCE, 4 June 2003 (2003-06-04), pages 71-84, XP002332821 page 73, right-hand column, line 14 - page 77, left-hand column, line 8	1-6,21
X	HALL J S ET AL: "VIRTUALIZING THE VAX ARCHITECTURE" COMPUTER ARCHITECTURE NEWS, ASSOCIATION FOR COMPUTING MACHINERY, NEW YORK, US, vol. 19, no. 3, 1 May 1991 (1991-05-01), pages 380-389, XP000229203 ISSN: 0163-5964 * page 381, colonne de gauche: "2 Theory of Virtual Machines" * * page 382, colonne de droite - page 385, colonne de droite: "4 Changes to Support a VAX VMM" * * page 386, colonne de droite - page 388, colonne de gauche: "7.1 Virtualizing Rings" * * page 388, colonne de droite - page 389, colonne de gauche: "8 Conclusion" *	1-11,21
A	US 6 651 132 B1 (TRAUT ERIC P) 18 November 2003 (2003-11-18) column 1, line 58 - column 6, line 14 column 10, lines 29-65	1-21
A	US 6 748 592 B1 (PORTER SWAIN W) 8 June 2004 (2004-06-08) abstract column 3, lines 13-64 column 5, line 50 - column 7, line 13	1-21

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/FR2005/002196

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6651132 B1	18-11-2003	AU 7347901 A	05-02-2002
		EP 1303816 A2	23-04-2003
		WO 0208905 A2	31-01-2002
US 6748592 B1	08-06-2004	AU 3696301 A	27-08-2001
		WO 0161504 A1	23-08-2001

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2005/002196

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
G06F9/455 G06F12/14

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 estréalisable, termes de recherche utilisés)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	SHANG RONG TSAI ET AL: "ON THE ARCHITECTURAL SUPPORT FOR LOGICAL MACHINE SYSTEMS" MICROPROCESSING AND MICROPROGRAMMING, ELSEVIER SCIENCE PUBLISHERS, BV., AMSTERDAM, NL, vol. 22, no. 2, 1 février 1988 (1988-02-01), pages 81-96, XP000284881 ISSN: 0165-6074 page 84, colonne de droite, ligne 6 - page 86, colonne de droite, ligne 9 page 87, colonne de gauche, ligne 14 - dernière ligne ----- -/--	1-7, 9-11,21

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

[X] 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'appartenenant 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'inven tion revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'inven tion 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	Date d'expédition du présent rapport de recherche internationale
30 décembre 2005	06/02/2006

Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Fournier, N

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2005/002196

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	S T KING, G W DUNLAP, P M CHEN: "Operating System Support for Virtual Machines" PROCEEDINGS OF THE 2003 USENIX ANNUAL TECHICAL CONFERENCE, 4 juin 2003 (2003-06-04), pages 71-84, XP002332821 page 73, colonne de droite, ligne 14 - page 77, colonne de gauche, ligne 8	1-6,21
X	HALL J S ET AL: "VIRTUALIZING THE VAX ARCHITECTURE" COMPUTER ARCHITECTURE NEWS, ASSOCIATION FOR COMPUTING MACHINERY, NEW YORK, US, vol. 19, no. 3, 1 mai 1991 (1991-05-01), pages 380-389, XP000229203 ISSN: 0163-5964 * page 381, colonne de gauche: "2 Theory of Virtual Machines" * * page 382, colonne de droite - page 385, colonne de droite: "4 Changes to Support a VAX VMM" * * page 386, colonne de droite - page 388, colonne de gauche: "7.1 Virtualizing Rings" * * page 388, colonne de droite - page 389, colonne de gauche: "8 Conclusion" *	1-11,21
A	US 6 651 132 B1 (TRAUT ERIC P) 18 novembre 2003 (2003-11-18) colonne 1, ligne 58 - colonne 6, ligne 14 colonne 10, ligne 29-65	1-21
A	US 6 748 592 B1 (PORTER SWAIN W) 8 juin 2004 (2004-06-08) abrégé colonne 3, ligne 13-64 colonne 5, ligne 50 - colonne 7, ligne 13	1-21

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°
PCT/FR2005/002196

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 6651132 B1	18-11-2003	AU 7347901 A	05-02-2002
		EP 1303816 A2	23-04-2003
		WO 0208905 A2	31-01-2002
US 6748592 B1	08-06-2004	AU 3696301 A	27-08-2001
		WO 0161504 A1	23-08-2001