

(12)

DEMANDE DE BREVET D'INVENTION

A1

(22) Date de dépôt : 11.02.16.

(30) Priorité : 25.02.15 US 14630863.

(43) Date de mise à la disposition du public de la demande : 26.08.16 Bulletin 16/34.

(56) Liste des documents cités dans le rapport de recherche préliminaire : Ce dernier n'a pas été établi à la date de publication de la demande.

(60) Références à d'autres documents nationaux apparentés :

Demande(s) d'extension :

(71) Demandeur(s) : HGST NETHERLANDS B.V. — NL.

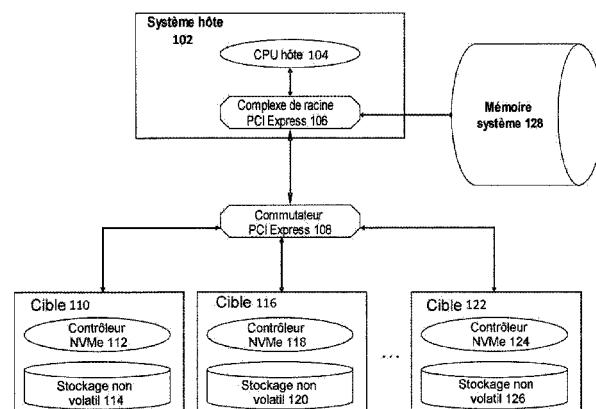
(72) Inventeur(s) : DEWITT DYLAN MARK, ESPESETH ADAM MICHAEL, MCCAMBRIDGE COLIN CHRISTOPHER et DREYER DAVID GEORGE.

(73) Titulaire(s) : HGST NETHERLANDS B.V..

(74) Mandataire(s) : IPSILON FERAY LENNE CONSEIL.

(54) SYSTEME ET PROCEDE DE COPIE A L'ECRITURE SUR UN SSD.

(57) Des techniques d'amélioration de la fonctionnalité de copie à l'écriture dans un SSD sont décrites. Dans certains modes de réalisation, ces techniques peuvent être mises en oeuvre sous la forme d'un Procédé destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD consistant à prévoir, dans la mémoire d'un dispositif, une structure de données d'indirection. La structure de données peut comprendre une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index et une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données. Ces techniques peuvent consister à parcourir, en utilisant un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.



SYSTEME ET PROCEDE DE COPIE A L'ECRITURE SUR UN SSD

Contexte de l'invention

[0001] La spécification de Mémoire Non Volatile express (NVMe) est une spécification concernant l'accès à des dispositifs à semi-conducteurs (SSD, Solid-State Devices) et d'autres dispositifs cibles raccordés par l'intermédiaire d'un bus d'Interconnexion de Composants Périphériques Express (PCIe, Peripheral Component Interconnect Express). L'interface hôte PCIe pour SSD NVMe définit un concept d'Espaces des noms, qui sont analogues à des volumes logiques pris en charge par des adaptateurs SAS RAID (Redundant Array of Independent Disks, Matrice Redondante de Disques Indépendants). La fonctionnalité de copie à l'écriture dans un SSD peut être mise en œuvre en utilisant des espaces des noms. Les espaces des noms sont généralement mis en œuvre en tant qu'abstraction au-dessus de l'espace d'Adresses de Blocs Logiques (LBA, Logical Block Address) global poursuivi dans le système d'indirection d'un SSD.

[0002] Les métadonnées LBA n'indiquent qu'une seule LBA d'hôte et elles ne comprennent pas de compte de référence. L'inclusion ou l'ajout d'un compte de référence dans les métadonnées conduirait à des écritures supplémentaires pour écraser les données avec de nouvelles métadonnées, cela étant une mauvaise solution. Sans de tels comptes de références dans les métadonnées LBA, il n'existe aucun mécanisme permettant de déterminer s'il existe des copies clonées supplémentaires (par exemple, le fait que des LBA supplémentaires pointent sur les mêmes données). La gestion de multiples copies de données clonées sur un SSD se heurte par conséquent à des difficultés particulières en ce qui concerne la récupération d'espace mémoire. A titre d'exemple, lorsqu'un hôte modifie la LBA "source" après une opération de copie, il peut conduire à des difficultés liées à la récupération d'espace mémoire. La copie source peut en fait être la copie "maîtresse" ayant été écrite avant que l'opération de copie ait été effectuée afin de dupliquer les données sur une LBA d'hôte supplémentaire ou plus. Lorsque cette LBA maîtresse est modifiée, un algorithme de récupération d'espace mémoire non sensible à la copie peut libérer les données physiques à la prochaine occasion, étant donné qu'il n'existe aucun procédé permettant de modifier efficacement les métadonnées de ces données pour indiquer qu'un plus grand nombre de LBA d'hôte pointent sur ces données.

Résumé de l'invention

[0003] Des techniques d'amélioration de la fonctionnalité de copie à l'écriture dans un SSD sont décrites. Dans certains modes de réalisation, les techniques peuvent être mises en œuvre sous la forme d'un Procédé destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD, consistant à prévoir, dans la mémoire d'un dispositif PCIe, une structure de données d'indirection. La structure de données peut comprendre une entrée maîtresse pour une copie originale ou source des données clonées, l'entrée maîtresse comportant une référence à un index maître et une référence à un index suivant, une entrée de clone pour les données clonées, l'entrée clonée comportant une référence à l'index maître et une référence à un index suivant. Ces techniques peuvent consister à parcourir, au moyen d'un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

[0004] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, le dispositif hôte peut comprendre au moins l'un des éléments suivants : un serveur d'entreprise, un serveur de base de données, une station de travail, et un ordinateur.

[0005] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, la structure de données d'indirection peut comprendre une pluralité d'adresses physiques.

[0006] Conformément à d'autres aspects de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée circulaire, l'entrée maîtresse pour des données clonées comprenant une référence à un index maître et une référence à un index suivant.

[0007] Conformément à d'autres aspects de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée circulaire, l'entrée de clone pour les données clonées comprenant une référence à l'index maître et une référence à un index suivant.

[0008] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée à une seule extrémité, une entrée contenue dans un index fournissant une indication du fait que l'index est un index maître.

[0009] Conformément à d'autres aspects de cet exemple de mode de réalisation, les références peuvent consister en des entrées dans une table d'indirection absolue pour l'adressage de blocs logiques.

5 [0010] Conformément à d'autres aspects de cet exemple de mode de réalisation, les références peuvent comprendre des entrées dans une structure de données arborescente pour un adressage de blocs logiques.

[0011] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, la fonctionnalité améliorée de copie à l'écriture peut comprendre une fonctionnalité améliorée de copie d'espace des noms.

10 [0012] Conformément à d'autres aspects de cet exemple de mode de réalisation, les techniques peuvent consister à positionner un indicateur pour un ou plusieurs blocs logiques condensés pour indiquer que lesdits un ou plusieurs blocs logiques condensés sont clonés.

[0013] Conformément à d'autres aspects de cet exemple de mode de réalisation, un index maître de l'entrée maîtresse peut pointer sur l'entrée maîtresse.

15 [0014] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, l'index maître de l'entrée clonée peut pointer sur l'entrée maîtresse.

[0015] Conformément à d'autres aspects de cet exemple de mode de réalisation, l'index suivant d'une dernière entrée clonée dans une structure de données peut pointer sur l'entrée maîtresse.

20 [0016] Conformément à d'autres aspects de cet exemple de mode de réalisation, les techniques peuvent consister à déterminer que l'entrée de clone pour les données clonées est une entrée de clone seule, la détermination consistant à déterminer que l'index suivant de l'entrée clonée concorde avec l'index maître de l'entrée clonée, déterminer que l'index suivant de l'entrée maîtresse pointe sur l'entrée de clone, décloner l'entrée de clone des données clonées en réglant l'index suivant de l'entrée de clone à une entrée d'indirection indiquant un bloc logique condensé et régler l'entrée d'index maître à une entrée d'indirection indiquant un bloc logique condensé, et décloner l'entrée maîtresse des données clonées en réglant l'index suivant de l'entrée maîtresse à une première entrée d'indirection indiquant un premier bloc logique condensé d'une entrée

maîtresse originale et régler l'index maître de l'entrée maîtresse à une seconde entrée d'indirection indiquant un second bloc logique condensé de l'entrée maîtresse originale.

[0017] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, les techniques peuvent consister à déterminer que l'entrée de clone pour les données clonées est l'une d'une pluralité d'entrées de clones, la détermination consistant à déterminer au moins l'un des points suivants : que l'index suivant de l'entrée clonée ne concorde pas avec l'index maître de l'entrée clonée, et que l'index suivant de l'entrée maîtresse ne pointe pas sur l'entrée de clone, et décloner l'entrée de clone des données clonées en réglant l'index suivant d'une entrée précédente afin qu'il pointe sur une entrée indiquée par l'index suivant de l'entrée de clone.

[0018] Conformément à d'autres aspects de cet exemple de mode de réalisation, les techniques peuvent consister à examiner une entrée pendant un processus de récupération d'espace mémoire, déterminer que l'entrée contient un indicateur cloné, et déterminer que l'entrée impliquée dans le processus de récupération d'espace mémoire est une entrée valide ne devant pas être supprimée sur la base de la détermination du fait que l'entrée contient l'indicateur cloné.

[0019] Dans d'autres modes de réalisation, les techniques peuvent être mises en œuvre sous la forme d'un produit de programme informatique constitué d'une série d'instructions exécutables sur un ordinateur. Le produit de programme informatique peut mettre en œuvre un processus destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD. Le programme informatique peut mettre en œuvre les étapes consistant à prévoir, dans la mémoire d'un dispositif, une structure de données d'indirection comprenant une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index, une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données, et parcourir, au moyen d'un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

[0020] Dans encore d'autres modes de réalisation, les techniques peuvent être mises en œuvre sous la forme d'un système destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD. Le système peut comprendre un premier dispositif, dans lequel le premier dispositif comprenant des instructions stockées, qui sont stockées en mémoire. Les

instructions peuvent comprendre une instruction destinée à fournir, dans la mémoire du premier dispositif, une structure de données d'indirection comprenant une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index, une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments

5 suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données, et parcourir, au moyen d'un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

10 [0021] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, la structure de données d'indirection peut comprendre une pluralité d'adresses physiques.

[0022] Conformément à d'autres aspects de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée circulaire, l'entrée maîtresse pour des données clonées comprenant une référence à un index maître et une référence à un index suivant.

15 [0023] Conformément à d'autres aspects de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée circulaire, l'entrée de clone pour les données clonées comprenant une référence à l'index maître et une référence à un index suivant.

20 [0024] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, la structure de données d'indirection peut faire partie d'une liste chaînée à une seule extrémité, une entrée contenue dans un index fournissant une indication du fait que l'index est un index maître.

[0025] Conformément à d'autres aspects de cet exemple de mode de réalisation, les références peuvent comprendre des entrées dans une table d'indirection absolue pour un adressage de blocs logiques.

25 [0026] Conformément à d'autres aspects de cet exemple de mode de réalisation, le premier dispositif peut comprendre un dispositif d'Interconnexion de Composants Périphériques Express (PCIe).

[0027] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, les techniques peuvent en outre comprendre une instruction destinée à positionner un indicateur

pour un ou plusieurs blocs logiques condensés pour indiquer que lesdits un ou plusieurs blocs logiques condensés sont clonés.

[0028] Conformément à d'autres aspects de cet exemple de mode de réalisation, l'index maître de l'entrée maîtresse et l'index maître de l'entrée clonée peuvent pointer sur l'entrée 5 maîtresse et l'index suivant d'une dernière entrée clonée dans une structure de données qui pointe sur l'entrée maîtresse.

[0029] Conformément à des aspects supplémentaires de cet exemple de mode de réalisation, le dispositif cible (par exemple, un dispositif PCIe) peut comprendre au moins l'un des éléments suivants : une unité de traitement graphique, une carte d'acquisition audio/vidéo, un 10 disque dur, un adaptateur de bus hôte, et un contrôleur de Mémoire Non Volatile express (NVMe, Non-Volatile Memory express). Selon certains modes de réalisation, le dispositif cible peut être un dispositif respectant la norme NVMe.

[0030] La présente invention va maintenant être décrite de manière plus détaillée en référence à des exemples de modes de réalisation de celle-ci, comme illustré dans les dessins annexés. 15 Bien que la présente invention soit décrite ci-après en référence à des exemples de modes de réalisation, il est à noter que la présente invention n'est pas limitée à ceux-ci. L'homme du métier normalement compétent ayant accès aux directives présentées ici saura envisager d'autres formes de réalisation, variantes, et modes de réalisation, ainsi que d'autres domaines 20 d'utilisation, qui entrent dans le cadre de la présente invention telle qu'elle est décrite ici, et vis-à-vis desquels la présente invention peut être d'une grande utilité.

Brève description des dessins

[0031] Pour favoriser une compréhension plus approfondie de la présente invention, on se référera à présent aux dessins annexés, dans lesquels des éléments identiques sont désignés par des références numériques identiques. Ces dessins ne doivent pas être considérés comme 25 limitant la présente invention, mais sont présentés à seul titre d'exemple.

[0032] La figure 1 représente un exemple de schéma fonctionnel illustrant une pluralité de dispositifs PCIe en communication avec un dispositif hôte, conformément à un mode de réalisation de la présente invention.

[0033] La figure 2 illustre une structure de données représentant des blocs logiques condensés non clonés, conformément à un mode de réalisation de la présente invention.

[0034] La figure 3 illustre des tables représentant des entrées de structure de données d'indirection maîtresse et clonée, conformément à un mode de réalisation de la présente invention.

[0035] La figure 4 représente un exemple de module d'amélioration de la fonctionnalité de copie à l'écriture dans un SSD, conformément à un mode de réalisation de la présente invention.

[0036] La figure 5 représente un organigramme illustrant une fonctionnalité améliorée de copie à l'écriture dans un SSD, conformément à un mode de réalisation de la présente invention.

[0037] La figure 6a représente une structure de données de formats d'entrées de structures de données de Segments C Maîtres, conformément à un mode de réalisation de la présente invention.

[0038] La figure 6B représente une structure de données de formats d'entrées de structures de données de Segments C de Clones, conformément à un mode de réalisation de la présente invention.

Description

[0039] La présente invention concerne une fonctionnalité améliorée de copie à l'écriture. Dans certains modes de réalisation, cette fonctionnalité de copie à l'écriture peut comprendre des copies de l'espace des noms. L'interface hôte PCIe pour SSD NVMe définit un concept d'Espaces des noms, qui sont analogues à des volumes logiques pris en charge par des adaptateurs SAS RAID (Redundant Array of Independent Disks). Un espace des noms peut être dédié à une Machine Virtuelle (VM, Virtual Machine). Dans un SSD, les espaces des noms peuvent être logiquement isolés les uns des autres et peuvent être effacés de manière sécurisée et reconvertis sans affecter d'autres Espaces des noms.

[0040] Un identifiant d'espace des noms peut être inclus dans une commande d'accès au support émise par l'hôte, en association avec la LBA dans cet espace des noms. Le SSD peut utiliser une structure de données (par exemple, une table de consultation, un arbre, une table de hachage, un topogramme binaire, etc.) afin de traduire cette combinaison d'espace des noms et

de LBA en une LBA globale utilisée de manière interne pour accéder au SSD. Selon certains modes de réalisation, les références faites à une LBA peuvent désigner cette LBA globale.

5 [0041] Des modes de réalisation de la présente invention décrivent un système et un procédé de mise en œuvre d'une fonction efficace de 'Copie d'Espace des noms' qui évite de dupliquer les données sur le SSD. Cela réduit l'amplification d'écriture se produisant dans le SSD, ce qui prolongeant ainsi la durée de vie du SSD tout en conduisant à de meilleures performances.

10 [0042] Les copies d'espace des noms sont une forme de fonctionnalité de "copie à l'écriture". Sur la fonction de copie, un pointeur généré pointe sur la copie unique sur le support. Une nouvelle copie sur le support est générée et mise à jour lors d'une écriture. Une fonction de copie de l'espace des noms nécessite une mise en œuvre efficace de la "copie à l'écriture" sur un SSD. Des modes de réalisation de la présente invention peuvent être appliqués à des copies d'espace des noms. Des modes de réalisation de la présente invention peuvent également être appliqués à d'autres formes de réalisation de la "copie à l'écriture" pour un SSD. A titre d'exemple, une copie "instantanée" peut être utilisée pour créer des images en un point dans le temps d'un espace des noms et des formes de réalisation du présent mode de réalisation peuvent être utilisées pour effectuer le suivi de copies instantanées.

20 [0043] Des modes de réalisation de la présente invention fournissent un système d'indirection de SSD (par exemple une table de LBA absolue) ou un procédé permettant d'inclure de multiples entrées qui pointent sur le même emplacement physique. Une telle forme de réalisation peut permettre une récupération d'espace mémoire efficace lorsqu'il existe de multiples références. Le suivi de multiples références ou la manipulation de multiples pointeurs (par exemple, sur des données flash NON-ET, NAND) peut améliorer la récupération d'espace mémoire. La récupération d'espace mémoire peut être effectuée en utilisant des métadonnées sur l'unité de stockage non volatile (par exemple, une mémoire flash NON-ET, une mémoire flash NON-OU, etc.) qui comprend la LBA d'hôte pour les données. L'algorithme de récupération d'espace mémoire peut déterminer que des secteurs de l'hôte sont encore valides en consultant ces adresses dans la structure de données d'indirection (par exemple, une table, un arbre, une table de hachage, un topogramme binaire, etc.) pour vérifier si la structure de données pointe encore sur l'emplacement physique. Si cela n'est pas le cas, l'algorithme libère le bloc.

30 [0044] Un ou plusieurs modes de réalisation décrits ici fournissent une représentation efficace d'une entrée d'indirection dupliquée en utilisant un indicateur unique et un autre format

d'entrée d'indirection qui effectue le suivi d'une ou plusieurs adresses LBA d'hôte dupliquées. Un ou plusieurs modes de réalisation peuvent utiliser une structure de données de consultation d'indirection absolue pour le suivi de multiples adresses de blocs logiques pointant sur une même adresse physique. D'autres modes de réalisation peuvent être mis en œuvre en utilisant 5 une table de hachage, un arbre, ou un système à base de composition destiné au suivi d'adresses LBA dupliquées.

[0045] Des techniques d'amélioration de la fonctionnalité de copie à l'écriture dans un SSD sont présentées plus en détail ci-après.

[0046] Se référant à présent aux dessins, la figure 1 est un exemple de schéma fonctionnel illustrant un dispositif PCIe en communication avec un dispositif hôte, conformément à un mode de réalisation de la présente invention. Des améliorations de la fonctionnalité de copie à l'écriture peuvent être mises en œuvre dans une ou plusieurs technologies informatiques, par exemple sous la forme d'un système hôte 102, d'une CPU hôte 104, et d'un complexe de racines PCI express (PCIe root complex) 106. Un commutateur PCI express 108 peut coupler par voie de communication une pluralité de cibles (par exemple, des dispositifs PCIe tels que des cibles à base de NVMe) tels que les Cibles 110, 116 et 122 au système hôte 102 par l'intermédiaire d'un complexe de racines PCI express 106.

[0047] La cible 110 peut contenir un contrôleur NVMe 112 et une unité de stockage non volatile 114. La cible 116 peut contenir un contrôleur NVMe 118 et une unité de stockage non volatile 120. La cible 122 peut contenir un contrôleur NVMe 124 et une unité de stockage non volatile 126.

[0048] La mémoire système 128 peut contenir des ressources à base de mémoire accessibles au Système Hôte 102 par l'intermédiaire d'une interface mémoire (par exemple, une mémoire vive dynamique synchrone de type trois à débit de données double (DDR3 SDRAM)). La 25 mémoire système 128 peut se présenter sous n'importe quelle forme appropriée, par exemple, sans aucune limitation à celles-ci, une mémoire à semi-conducteur (par exemple, une mémoire flash, ou un dispositif à semi-conducteur (SSD)), une mémoire optique, et une mémoire magnétique. La mémoire système 128 peut être une mémoire volatile ou non volatile. La mémoire système 128 peut contenir une ou plusieurs structures de données.

[0049] Selon certains modes de réalisation, des normes d'interfaces autres que le PCIe peuvent être utilisées pour une ou plusieurs parties, parmi lesquelles, sans aucune limitation à celles-ci, la Technologie Avancée de Connexion Série (SATA, Serial Advanced Technology Attachment), la Technologie Avancée de Connexion (ATA, Advanced Technology Attachment),

5 l'Interface pour Petits Systèmes Informatiques (SCSI, Small Computer System Interface), le PCI étendu (PCI-X), le Fibre Channel, le SCSI Connecté en Série (SAS, Serial Attached SCSI), le Secure Digital (SD), la Carte MultiMédia Intégrée (EMMC, Embedded Multi-Media Card), et l'Unité de Stockage Flash Universelle (UFS, Universal Flash Storage).

[0050] Le système hôte 102 peut se présenter sous n'importe quelle forme appropriée, par 10 exemple, sans aucune limitation à celles-ci, un serveur d'entreprise, un hôte de base de données, une station de travail, un ordinateur personnel, un téléphone mobile, un dispositif de jeu, un assistant numérique personnel (PDA, Personal Digital Assistant), un dispositif de messagerie par courriel/texte, un appareil photo numérique, un lecteur multimédia numérique (par exemple MP3), un dispositif de navigation GPS, et un système de télévision.

[0051] Le système hôte 102 et le dispositif cible peuvent comprendre des composants supplémentaires, qui ne sont pas représentés sur la figure 1 pour simplifier le dessin. Par ailleurs, dans certains modes de réalisation, tous les composants représentés ne sont pas présents. De plus, les divers contrôleurs, blocs, et interfaces peuvent être mis en œuvre de n'importe quelle façon appropriée. A titre d'exemple, un contrôleur peut présenter la forme d'un ou plusieurs 20 d'un microprocesseur ou d'un processeur et d'un support lisible par ordinateur qui stocke un code de programme lisible par ordinateur (par exemple, un logiciel ou un micrologiciel) exécutable par le (micro)processeur, par exemple, des portes logiques, des commutateurs, un circuit intégré spécifique d'applications (ASIC, Application Specific Integrated Circuit), un automate programmable, et un microcontrôleur intégré.

[0052] Se référant à la figure 2, une structure de données représentant des blocs logiques condensés non clonés est illustrée, conformément à un mode de réalisation de la présente invention. Huit entrées de blocs logiques condensés de 4 kilo-octets sont illustrées dans ce mode de réalisation. Ces blocs logiques condensés (PLB, Packed Logical Bloc) non clonés peuvent mapper des Adresses de Blocs Logiques (LBA) sur une adresse physique. Comme illustré, les 30 LBA 0-7 peuvent être mappées sur le Bloc d'Adresses Physiques 0, les LBA 7-15 peuvent être mappées sur le Bloc d'Adresses Physiques 1, les LBA 16-23 peuvent être mappées sur le Bloc

d'Adresses Physiques 2, les LBA 24-31 peuvent être mappées sur le Bloc d'Adresses Physiques 3, les LBA 32-39 peuvent être mappées sur le Bloc d'Adresses Physiques 4, les LBA 40-47 peuvent être mappées sur le Bloc d'Adresses Physiques 5, les LBA 48-55 peuvent être mappées sur le Bloc d'Adresses Physiques 6, et les adresses LBA 56-63 peuvent être mappées sur le Bloc 5 d'Adresses Physiques 7.

[0053] La figure 3 illustre des tables représentant des entrées de structure de données d'indirection maîtresse et clonée, conformément à un mode de réalisation de la présente invention. Comme décrit en référence à la figure 3, dans certains modes de réalisation, des améliorations de la copie à l'écriture peuvent être mises en œuvre dans un système d'indirection 10 absolu. Un PLB peut être un "bloc logique condensé" dans une structure de données de consultation (par exemple, une table de consultation à indirection absolue). Un ensemble de 8 PLB consécutifs alignés sur 8 PLB impliqués dans une opération de clonage peuvent être désignés sous le nom de "Segment C" (C-Chunk). Un Segment C dont les données sur le support sont marquées par les PLB qui conviennent pour ce Segment C (c'est-à-dire la copie 15 originale) peut être appelé "Segment C Maître". Les entrées de la structure de données d'indirection pour tous les PLB dans un Segment C peuvent être regroupées pour former une "Entrée C" unique. Une Entrée C correspondant à un Segment C dont les données se trouvent sur le support peut être appelée Entrée C Maîtresse. Les Entrées C qui décrivent des plages de LBA qui sont des copies de la Plage de LBA maîtresses peuvent être désignées sous le nom 20 d'Entrées C de Clones. Un ou plusieurs SSD peuvent définir un PLB pour effectuer le suivi de 4 ko de données de clients plus les métadonnées (par exemple des secteurs de 8x512 o). Dans les modes de réalisation présentés ci-après, un PLB peut simplement être une entrée dans une table. Une entrée de structure de données d'indirection peut être étendue (par exemple, d'un bit) pour faciliter la copie à l'écriture. Un bit supplémentaire peut être un bit de "suivi de clone" qui peut 25 être positionné à 1 pour indiquer soit qu'il y a d'autres PLB pour lesquels ce PLB joue le rôle de copie maîtresse, soit qu'il s'agit d'un clone qui comporte certains autres PLB en tant que sa copie maîtresse. Les bits restants de l'entrée de la structure de données d'indirection dont le bit de suivi de clone est positionné peuvent ou non contenir une adresse de mémoire NON-ET (par exemple, comme cela peut être le cas d'une entrée dont le bit n'est pas positionné). L'autre structure de 30 données possible pour 'suivi de clone = 1' est suivie avec une granularité plus grossière que l'entrée de PLB type, et comprend des champs permettant de créer une liste chaînée d'entrées clonées et un pointeur pointant sur l'entrée maîtresse. L'espace nécessaire pour ces champs supplémentaires peut être obtenu en utilisant une Entrée C unique pour décrire un plus grand

(par exemple 2x) segment d'adresses LBA qu'une entrée d'indirection non clonée. Ce compromis est raisonnable car les données clonées tendent à mettre en jeu des partitions ou des ensembles de fichiers de grandes tailles et non pas des adresses LBA d'hôte individuelles.

[0054] Les adresses physiques des adresses LBA d'hôte individuelles sont réparties de telle manière qu'une consultation supplémentaire est requise pour certaines LBA. Cela permet de libérer de la place pour inclure les pointeurs maître et clone dans chaque entrée de maître et de clone. Cependant, le nombre d'accès à la DRAM pour rechercher les adresses physiques n'est pas augmenté de manière significative. Comme illustré sur la figure 3, l'Entrée C Maîtresse peut contenir les adresses LBA mappées sur toutes les adresses physiques 8-15, mais ne contient que les mappages effectifs sur les adresses physiques 8-11, 14, et 15 (le mappage sur les adresses physiques 12 et 13 peut être obtenu à partir de l'Entrée C de Clone). Comme illustré dans l'Entrée C de Clone de la figure 3, des mappages peuvent être fournis pour les adresses LBA correspondant aux adresses physiques 8-13. Les deux blocs physiques non mappés dans l'Entrée C Maîtresse (Adresses physiques 12 et 13) fournissent l'espace nécessaire pour un Index Maître et un Index Suivant. Les deux blocs physiques non mappés dans l'Entrée C de Clone (Adresses Physiques 14 et 15) fournissent l'espace nécessaire pour un Index Maître et un Index Suivant. Un Index Maître dans une Entrée C Maîtresse pointe toujours sur lui-même. Un Index Maître dans une Entrée C de Clone pointe toujours sur l'Entrée C Maîtresse. Un Index Suivant de l'Entrée C Maîtresse pointe sur une première Entrée C de Clone dans une chaîne. Un Index Suivant d'une Entrée C de Clone pointe sur une Entrée C de Clone suivante s'il y a plus d'une Entrée C de Clone. S'il n'y a qu'une seule Entrée C de Clone ou s'il s'agit de la dernière Entrée C de Clone, l'Index Suivant peut pointer en retour sur l'index C Maître (ou peut pointer sur une valeur spéciale indiquant une extrémité de la liste comme par exemple un pointeur nul). Cela peut permettre de parcourir une Entrée C Maîtresse et une ou plusieurs entrées C de Clones.

[0055] Dans les modes de réalisation utilisant une structure de données à indirection absolue, un PLB (bloc d'emplacements physiques) désigne une entrée individuelle contenant une adresse de mémoire NON-ET physique dans une table de consultation unique. Une granularité de PLB type réserve 4 ko (par exemple des secteurs de 8*512 o) à une entrée unique de la structure de données. A titre d'exemple, on considérera une taille de segment de clone de 8 blocs PLB--le nombre moyen d'accès à la DRAM requis pour des accès par secteurs uniques purement aléatoires à une plage clonée est de 1,25. Ce nombre peut être plus faible avec une taille de

segment de clone plus élevée, le compromis étant des frontières entre clones moins granulaires et davantage d'accès NON-ET requis pour 'décloner' un segment de blocs PLB.

[0056] Si une consultation de PLB en lecture par un hôte pointe sur une entrée clonée, le SSD doit disposer 1) de l'adresse physique et 2) de la LBA ayant été utilisée lorsque les données ont été écrites. L'adresse physique est répartie entre les entrées maîtresse et clonée, facultativement en remplissant toutes les entrées PLB disponibles de données dupliquées afin de réduire la probabilité que la LBA ciblée nécessite une seconde lecture de DRAM. Pour un nombre égal à 2, la LBA maîtresse peut être calculée sur la base du pointeur maître dans chaque entrée de clone--cela ne nécessite pas d'accès supplémentaire à la DRAM par conception, et ce pointeur maître (en association avec le pointeur de clone suivant) peut être mis en cache à partir de la consultation de PLB originale.

[0057] Dans certains modes de réalisation, pour faire tenir des informations supplémentaires dans une structure de données d'indirection existante, le clonage ne peut être poursuivi qu'avec une granularité égale à un multiple quelconque de la taille du PLB. Dans certains modes de réalisation, la granularité peut être choisie comme étant un multiple d'une puissance de deux de la taille du PLB pour permettre un calcul efficace de l'index du Segment C correspondant à une LBA donnée. A titre d'exemple, le multiplicateur peut être égal à 8, mais des granularités de clonage plus importantes peuvent être utilisées. Le clonage peut mettre en jeu le clonage de grandes plages d'adresses LBA en une seule fois (par exemple, la totalité d'un espace des noms), de manière à ce que la pénalité due à l'utilisation d'une plus grande granularité soit minime.

[0058] Dans certains modes de réalisation, la structure de données d'indirection peut comporter, intégrée à celle-ci, une liste chaînée circulaire d'un ou plusieurs Segments C qui se réfèrent aux mêmes données. Dans d'autres modes de réalisation, d'autres formes de listes chaînées (par exemple, une liste chaînée à une seule extrémité) peuvent être utilisées. Les adresses physiques décrivant les données du Segment C peuvent être réparties parmi des entrées d'indirection pour cette liste de Segments C.

[0059] Un Segment C dont les données se trouvent physiquement sur le support, marquées par les PLB qui conviennent pour ce Segment C (c'est-à-dire la copie originale) peut être appelé "Segment C Maître". D'autres Segments C qui se réfèrent à l'instant courant aux mêmes données sans stocker de copie sur le support peuvent être appelés "Segments C de Clones".

[0060] Les entrées de la structure de données d'indirection pour tous les PLB dans un Segment C sont regroupées pour former une "Entrée C" unique. Pour un Segment C Maître, l'Entrée C peut avoir le format illustré sur la figure 6A. Comme illustré, un segment C Maître peut contenir les adresses physiques 0-3 (les quatre premières adresses physiques d'une plage), suivies des adresses physiques 6 et 7. Les deux adresses physiques manquantes peuvent être mappées dans une Entrée C de Clone et les deux emplacements supplémentaires peuvent être utilisés pour fournir un Index Maître et un Index Suivant. Comme illustré, un indicateur de "suivi de clone" peut être positionné à 1. Cela permet d'indiquer à un processus de récupération d'espace mémoire que l'entrée C doit être ignorée.

5

[0061] Pour un Segment C de Clone, l'Entrée C peut avoir le format illustré sur la figure 6B. Comme illustré, les adresses physiques 0-5 peuvent être mappées et les adresses physiques 6 et 7 peuvent être manquantes. Les deux adresses physiques manquantes peuvent être mappées dans l'Entrée C Maîtresse et les deux emplacements supplémentaires peuvent être utilisés pour fournir un Index Maître et un Index Suivant. Comme illustré, un indicateur de "suivi de clone" peut être positionné à 1. Cela permet d'indiquer à un processus de récupération d'espace mémoire que l'entrée C doit être ignorée.

10

[0062] Dans un ou plusieurs modes de réalisation, l'Adresseⁱ NON-ET peut être l'adresse NON-ET pour un i^{ème} PLB du Segment C Maître (où i = 0 représente le premier PLB). Un Index Maître peut être un index de structure de données d'indirection de l'Entrée C Maîtresse (divisée par 8 puisqu'elle occupe l'espace de 8 entrées PLB). Un Index Suivant peut être un index de structure de données d'indirection pour l'Entrée C de Clone suivante pointant sur les mêmes données (divisée par 8 puisqu'elle occupe l'espace de 8 entrées PLB). S'il n'y a plus d'Entrées C de Clones à représenter, cet Index Suivant peut pointer en retour sur l'Entrée C Maîtresse. Dans certains modes de réalisation, cet Index Suivant peut pointer sur une valeur indiquant la fin de la liste (par exemple, un pointeur nul).

20

[0063] Dans certains modes de réalisation, un ou plusieurs tests peuvent être utilisés pour déterminer si un Segment C est une entrée maîtresse. A titre d'exemple, un Segment C peut être le maître si et seulement si l'Index Maître de son Entrée C pointe sur l'Entrée C elle-même.

[0064] La relation entre une Entrée C Maîtresse et une Entrée C de Clone par l'intermédiaire d'Index Maîtres et d'Index Suivants permet l'exécution efficace d'une ou plusieurs des opérations suivantes.

25

[0065] Pour cloner un ensemble de 8 PLB consécutifs alignés sur 8 PLB, un ou plusieurs procédés peuvent être utilisés. A titre d'exemple, dans un ou plusieurs modes de réalisation, le clonage d'un ensemble de blocs PLB peut consister à :

1. Lire les 8 Adresses NON-ET pour les PLB de la copie originale à laquelle il doit être fait référence lors de la création des Entrées C Maîtresse et de Clone ;
- 5 2. Créer une Entrée C Maîtresse au lieu des entrées d'indirection des 8 blocs PLB originaux. L'Index Suivant peut pointer sur la nouvelle Entrée C de Clone ; et
- 10 3. Créer une Entrée C de Clone au lieu des entrées d'indirection des 8 blocs PLB qui vont alors jouer le rôle de clone. L'Index Suivant peut pointer sur la nouvelle Entrée C Maîtresse.

[0066] Pour cloner un Segment C en un nouveau Segment C, un ou plusieurs procédés peuvent être utilisés. A titre d'exemple, dans un ou plusieurs modes de réalisation, le clonage d'un Segment C en un nouveau Segment C peut consister à :

1. Si le Segment C source est un maître, suivre un Index Suivant d'une Entrée C source pour trouver une Entrée C de Clone et la copier à l'emplacement de la nouvelle Entrée C. Si le Segment C source n'est pas un maître, copier l'Entrée C source à l'emplacement de la nouvelle Entrée C ;
- 15 2. Mettre à jour l'Index Suivant de la nouvelle Entrée C avec Index Suivant courant de l'Entrée C Maîtresse ;
- 20 3. Mettre à jour l'Index Suivant de l'Entrée C Maîtresse afin qu'elle pointe sur la nouvelle Entrée C.

[0067] Pour effectuer une consultation de lecture sur un PLB dont l'entrée d'indirection comporte un bit de suivi de clone positionné, un ou plusieurs procédés peuvent être utilisés. A titre d'exemple, dans un ou plusieurs modes de réalisation, l'exécution d'une consultation de lecture sur un PLB ayant un bit d'indirection positionné peut consister à :

1. Observer l'Entrée C qui contient l'entrée d'indirection de ce PLB. Déterminer si cette Entrée C est ou non une Entrée C Maîtresse. Sur la base de la

détermination, déterminer en outre si l'adresse NON-ET pour le PLB Maître souhaité est stockée dans cette Entrée C.

2. Si l'adresse NON-ET est stockée dans la première Entrée C lue, cette adresse NON-ET et l'Index Maître peuvent être renvoyés. L'Index Maître peut être utilisé pour déterminer avec quel nombre de PLB les données vont être marquées, sans effectuer de consultations d'indirection supplémentaires.
3. Si l'adresse NON-ET n'est pas stockée dans la première Entrée C lue et si la première Entrée C lue est une Entrée C Maîtresse, l'Index Suivant peut être suivi pour trouver une Entrée C de Clone qui contient l'adresse NON-ET nécessaire. Cette adresse NON-ET et l'Index Maître peuvent être renvoyés.
4. Si l'adresse NON-ET n'est pas stockée dans l'Entrée C lue et s'il s'agit d'une Entrée C de Clone, l'Index Maître peut être suivi pour trouver une Entrée C Maîtresse qui contient l'adresse NON-ET nécessaire. Cette adresse NON-ET et l'Index Maître peuvent être renvoyés.

15 [0068] Pour "décloner" un Segment C de Clone, une ou plusieurs techniques peuvent être utilisées. A titre d'exemple, dans certains modes de réalisation, ces techniques peuvent consister à :

1. Déterminer s'il n'y a qu'un seul clone. A titre d'exemple, si l'Index Suivant de l'Entrée C cible concorde avec son Index Maître et si l'Index Suivant de l'Entrée C Maîtresse pointe sur l'Entrée C cible, alors il n'y a qu'un seul clone. Lire les Entrées C à la fois pour le Maître et le Clone et écraser l'Entrée C Maîtresse avec des entrées d'indirection "normales" pour ces blocs PLB.
2. Si l'Index Suivant de l'Entrée C cible ne concorde pas avec son Index Maître ou si l'Index Suivant de l'Entrée C Maîtresse ne pointe pas sur l'Entrée C cible, alors d'autres clones sont présents. Eliminer cette Entrée C de sa liste circulaire.
3. Dans tous les cas, lire les données pour le Segment C Maître et les copier dans le Segment C de Clone, en écrasant les Entrées C de Clones avec des entrées d'indirection normales indiquant les adresses NON-ET de la nouvelle copie.

[0069] Pour "décloner" un segment C Maître une ou plusieurs techniques peuvent être utilisées. A titre d'exemple, dans certains modes de réalisation ces techniques peuvent consister à :

1. Lire l'Entrée C Maîtresse et la première Entrée C de Clone ; et
- 5 2. Lire les données depuis un support de stockage pour le Segment C Maître et les copier dans le premier Segment C de Clone du PLB en les écrivant à un nouvel emplacement du support de stockage et associer les métadonnées LBA de l'hôte au Segment C de Clone correspondant ;
- 10 3. Si ce premier clone est le seul clone (c'est-à-dire que son Index Suivant pointe en retour sur le maître), alors écraser la première Entrée C de Clone avec des entrées d'indirection normales pointant sur les adresses NON-ET pour la copie qui vient d'être créée. Ecraser l'Entrée C Maîtresse avec des entrées d'indirection normales pointant sur des PLB du Segment C Maître original ;
- 15 4. Si le premier clone n'était pas le seul clone, éliminer le maître de la liste chaînée circulaire. Ecraser la première Entrée C de Clone avec une nouvelle Entrée C Maîtresse pour la copie qui vient d'être créée. Ecraser d'autres Entrées C de Clones avec une nouvelle Entrée C de Clone pour la copie qui vient d'être créée. Cela permet de favoriser le fait que la première Entrée C de Clone de l'ancienne Entrée C Maîtresse devienne la nouvelle Entrée C Maîtresse pour elle-même et les Entrées C de clones restantes.
- 20

[0070] Lorsqu'une mise à jour d'indirection est requise pour un PLB cible dont le bit de suivi de clone est positionné, ces techniques peuvent consister, de manière atomique, à :

1. Décloner le Segment C du PLB ; et
2. Effectuer la mise à jour d'indirection normalement.

25 [0071] La récupération d'espace mémoire peut être mise en œuvre de telle manière que l'algorithme de récupération d'espace mémoire ne rejette jamais les données pour des données physiques quelconques dont le PLB est marqué par le bit de suivi de clone dans le système d'indirection. Plus précisément, l'algorithme de récupération d'espace mémoire peut considérer

tous les PLB marqués par un bit de suivi de clone comme étant des données "valides" qui nécessitent un déplacement plutôt qu'un effacement.

[0072] Dans un ou plusieurs modes de réalisation, un petit filtre Bloom de comptage pourrait être stocké en SRAM pour effectuer le suivi de Segments C présents dans le système.

5 Lors d'une écriture, si le filtre Bloom peut indiquer qu'il n'y a aucun risque que le PLB fasse partie d'un Segment C, alors une mise à jour directe du système d'indirection peut être effectuée en toute sécurité sans lire en premier lieu l'entrée de structure de données courante. Comme les clones ont tendance à être de grandes plages séquentielles, la fonction de hachage utilisée pour un filtre Bloom peut être du type : $f(\text{Index Entrée C}) = \text{Index Entrée C} / \text{Taille du}$

10 Filtre plutôt qu'une fonction aléatoire.

[0073] La figure 4 illustre un exemple de module d'amélioration de la fonctionnalité de copie à l'écriture dans un SSD, conformément à un mode de réalisation de la présente invention. Comme illustré sur la figure 4, le module de copie à l'écriture 410 peut contenir un module de création d'indirection 412, un module de gestion d'indirection 414, et un module de manipulation d'erreur 416.

[0074] Le module de création d'indirection 412 peut créer une ou plusieurs structures de données pour le suivi de copies obtenues par copie à l'écriture. Un PLB peut simplement être une entrée dans une table. Une entrée de structure de données d'indirection peut être étendue (par exemple, d'un bit) pour faciliter la copie à l'écriture. Un bit supplémentaire peut être un bit de "suivi de clone" qui peut être positionné à 1 pour indiquer soit qu'il y a d'autres PLB pour lesquels ce PLB joue le rôle de copie maîtresse, soit qu'il s'agit d'un clone qui comporte certains autres PLB en tant que sa copie maîtresse. Les bits restants de l'entrée de la structure de données d'indirection dont le bit de suivi de clone est positionné peuvent ou non contenir une adresse NON-ET (par exemple, comme cela peut être le cas d'une entrée dont le bit n'est pas positionné).

20 L'autre structure de données possible pour "suivi de clone = 1" comprend des champs permettant de créer une liste chaînée d'entrées clonées et un pointeur pointant sur l'entrée maîtresse. L'espace nécessaire pour ces champs supplémentaires est obtenu en utilisant une entrée unique pour décrire un plus grand (par exemple 2x) segment d'adresses LBA qu'une entrée d'indirection non clonée.

25 [0075] Le module de gestion d'indirection 414 peut mettre en œuvre une ou plusieurs opérations utilisant une structure de données d'indirection. Le module de gestion d'indirection

414 peut favoriser le clonage de données, la lecture de données clonées, le déclonage de données, et faciliter une récupération d'espace mémoire sûre et efficace en utilisant un ou plusieurs des procédés présentés ci-dessus en référence à la figure 3.

[0076] Un module de gestion d'erreur 416 peut piéger, consigner, signaler, et/ou traiter une 5 ou plusieurs erreurs associées à la gestion de données clonées.

[0077] La figure 5 représente un organigramme illustrant une fonctionnalité améliorée de copie à l'écriture dans un SSD, conformément à un mode de réalisation de la présente invention. Cependant, le processus 500, n'est présenté qu'à titre d'exemple. Le processus 500 peut être modifié, par exemple, en faisant en sorte que des étapes soient ajoutées, changées, éliminées, 10 ou réorganisées. A l'étape 502, le processus peut commencer.

[0078] A l'étape 504, une Entrée C Maîtresse peut être créée. Un Segment C dont les données se trouvent physiquement sur le support, marquées par les PLB qui conviennent pour ce Segment C (c'est-à-dire la copie originale) peut être appelé "Segment C Maître". D'autres Segments C qui se réfèrent à l'instant courant aux mêmes données sans stocker de copie sur le 15 support peuvent être appelés "Segments C de Clones".

[0079] Les entrées de la structure de données d'indirection pour tous les PLB dans un Segment C sont regroupées pour former une "Entrée C" unique. Pour un segment C Maître, l'Entrée C peut avoir le format illustré sur la figure 6A.

[0080] A l'étape 506, une entrée C de clone peut être créée. Pour un Segment C de Clone, 20 l'Entrée C peut avoir le format illustré sur la figure 6B.

[0081] Dans un ou plusieurs modes de réalisation, l'Adresseⁱ NON-ET peut être l'adresse NON-ET pour un *i*^{ème} PLB du Segment C Maître (où *i* = 0 représente le premier PLB).

[0082] A l'étape 508, un bloc peut être attribué pour indiquer un Index Maître. Un Index Maître peut être un index de structure de données d'indirection de l'Entrée C Maîtresse 25 (divisée par 8 puisqu'elle occupe l'espace de 8 entrées PLB). L'Index Maître tant d'une Entrée C Maîtresse que d'une Entrée C de Clone peut pointer sur une Entrée C Maîtresse.

[0083] A l'étape 510, un bloc peut être attribué pour indiquer un Index Suivant. Un Index Suivant peut être un index de structure de données d'indirection pour l'Entrée C de Clone

suivante pointant sur les mêmes données (divisée par 8 puisqu'elle occupe l'espace de 8 entrées PLB). S'il n'y a plus d'Entrées C de Clones à représenter, cet Index Suivant peut pointer en retour sur l'Entrée C Maîtresse.

5 [0084] Dans certains modes de réalisation, un ou plusieurs tests peuvent être utilisés pour déterminer si un Segment C est une entrée maîtresse. A titre d'exemple, un Segment C peut être le maître si et seulement si l'Index Maître de son Entrée C pointe sur l'Entrée C elle-même.

[0085] A l'étape 512, le procédé 500 peut se terminer.

10 [0086] D'autres modes de réalisation entrent dans le cadre et l'esprit de l'invention. A titre d'exemple, la fonctionnalité décrite ci-dessus peut être mise en œuvre au moyen d'un logiciel, d'un matériel, d'un micrologiciel, d'un câblage fixe, ou de combinaisons de n'importe 15 lesquelles de ces formes. Un ou plusieurs processeurs d'ordinateurs fonctionnant conformément à des instructions peuvent mettre en œuvre les fonctions associées à la fonctionnalité améliorée de copie à l'écriture dans un SSD conformément à la présente invention telle que décrite ci-dessus. Si cela est le cas, il entre dans le cadre de la présente 20 invention que de telles instructions soient stockées sur un ou plusieurs supports de stockage non transitoires lisibles par un processeur (par exemple, un disque magnétique ou un autre support de stockage). En outre, des modules mettant en œuvre des fonctions peuvent également être physiquement placés à diverses positions, cela comprenant le fait qu'ils soient répartis de telle manière que des parties des fonctions soient mises en œuvre à différents emplacements physiques.

25 [0087] L'étendue de la présente invention ne doit pas être considérée comme étant limitée aux modes de réalisation particuliers décrits ici. En effet, divers autres modes de réalisation et diverses autres variantes de la présente invention, en plus de ceux décrits ici, apparaîtront à l'homme du métier normalement compétent au vu de la description présentée ci-dessus et des dessins annexés. Par conséquent, ces autres modes de réalisation et variantes devront être considérés comme entrant dans le cadre de la présente invention. De plus, bien que la présente 30 invention ait été décrite ici dans le contexte d'une forme de réalisation particulière dans un environnement particulier pour une application particulière, l'homme du métier normalement compétent notera que son utilité n'est pas limitée à ceux-ci et que la présente invention peut être avantageusement mise en œuvre dans un nombre quelconque d'environnements pour un nombre quelconque d'applications. En conséquence, les revendications présentées ci-après

devront être appréciées selon leur acception la plus large et dans le respect de l'esprit de la présente invention telle que décrite ici.

REVENDICATIONS

1. Procédé destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD consistant à :

5 prévoir, dans la mémoire d'un dispositif, une structure de données d'indirection comprenant :

une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index ;

10 une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données ; et

parcourir, en utilisant un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

2. Procédé selon la revendication 1, dans lequel la structure de données d'indirection comprend une pluralité d'adresses physiques.

3. Procédé selon la revendication 1, dans lequel la structure de données d'indirection fait partie d'une liste chaînée circulaire, l'entrée maîtresse pour des données clonées comprenant une référence à un index maître et une référence à un index suivant.

20 4. Procédé selon la revendication 1, dans lequel la structure de données d'indirection fait partie d'une liste chaînée circulaire, l'entrée de clone pour les données clonées comprenant une référence à l'index maître et une référence à un index suivant.

25 5. Procédé selon la revendication 1, dans lequel la structure de données d'indirection fait partie d'une liste chaînée à une seule extrémité, une entrée contenue dans un index fournissant une indication du fait que l'index est un index maître.

30 6. Procédé selon la revendication 1, dans lequel les références comprennent des entrées dans une table d'indirection absolue pour un adressage de blocs logiques.

7. Procédé selon la revendication 1, dans lequel les références comprennent des entrées dans une structure de données arborescente pour un adressage de blocs logiques.

8. Procédé selon la revendication 1, dans lequel la fonctionnalité améliorée de copie à l'écriture comprend une fonctionnalité améliorée de copie d'espace des noms.
9. Procédé selon la revendication 1, consistant en outre à positionner un indicateur pour 5 un ou plusieurs blocs logiques condensés pour indiquer que lesdits un ou plusieurs blocs logiques condensés sont clonés.
10. Procédé selon la revendication 1, dans lequel un index maître de l'entrée maîtresse pointe sur l'entrée maîtresse.
11. Procédé selon la revendication 3, dans lequel l'index maître de l'entrée clonée pointe sur l'entrée maîtresse.
12. Procédé selon la revendication 4, dans lequel l'index suivant d'une dernière entrée 15 clonée dans une structure de données pointe sur l'entrée maîtresse.
13. Procédé selon la revendication 4, consistant en outre à :
déterminer que l'entrée de clone pour les données clonées est une entrée de clone seule, la détermination consistant à :
20 déterminer que l'index suivant de l'entrée clonée concorde avec l'index maître de l'entrée clonée ;
déterminer que l'index suivant de l'entrée maîtresse pointe sur l'entrée de clone ;
décloner l'entrée de clone des données clonées en réglant l'index suivant de l'entrée de clone à une entrée d'indirection indiquant un bloc logique condensé et régler 25 l'entrée d'index maître à une entrée d'indirection indiquant un bloc logique condensé ; et
décloner l'entrée maîtresse des données clonées en réglant l'index suivant de l'entrée maîtresse à une première entrée d'indirection indiquant un premier bloc logique condensé d'une entrée maîtresse originale et régler l'index maître de l'entrée maîtresse à une seconde entrée d'indirection indiquant un second bloc logique condensé de l'entrée maîtresse 30 originale.
14. Procédé selon la revendication 4, consistant en outre à :
déterminer que l'entrée de clone pour les données clonées est l'une d'une pluralité

d'entrées de clones, la détermination consistant à déterminer au moins l'un des points suivants :

que l'index suivant de l'entrée clonée ne concorde pas avec l'index maître de

l'entrée clonée ; et

que l'index suivant de l'entrée maîtresse ne pointe pas sur l'entrée de clone ; et

5 décloner l'entrée de clone des données clonées en réglant l'index suivant d'une entrée précédente afin qu'elle pointe sur une entrée indiquée par l'index suivant de l'entrée de clone.

15. Procédé selon la revendication 1, consistant en outre à :

examiner une entrée pendant un processus de récupération d'espace mémoire ;

10 déterminer que l'entrée contient un indicateur cloné ; et

déterminer que l'entrée impliquée dans le processus de récupération d'espace mémoire est une entrée valide ne devant pas être supprimée sur la base de la détermination du fait que l'entrée contient l'indicateur cloné.

15 16. Produit de programme informatique constitué d'une série d'instructions exécutables sur un ordinateur, le produit de programme informatique exécutant un processus destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un SSD ; le programme informatique mettant en œuvre les étapes consistant à :

20 prévoir, dans la mémoire d'un dispositif, une structure de données d'indirection comprenant :

une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index ;

25 une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données ; et

parcourir, en utilisant un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

17. Système destiné à fournir une fonctionnalité améliorée de copie à l'écriture dans un 30 SSD, le système comprenant :

un premier dispositif ;

le premier dispositif comprenant des instructions stockées, qui sont stockées en mémoire, les instructions comprenant :

une instruction destinée à fournir, dans la mémoire du premier dispositif, une structure de données d'indirection comprenant :

une entrée maîtresse pour des données clonées, l'entrée maîtresse comportant une référence à un ou plusieurs index ;

5 une entrée de clone pour les données clonées, l'entrée clonée comportant au moins l'un des éléments suivants : une référence à un index maître, une référence à un index suivant, et une valeur indiquant une extrémité d'une structure de données ; et

parcourir, en utilisant un processeur d'ordinateur, une ou plusieurs copies des données clonées en utilisant une ou plusieurs des références.

10 18. Système selon la revendication 17, dans lequel la structure de données d'indirection comprend une pluralité d'adresses physiques.

19. Système selon la revendication 17, dans lequel la structure de données d'indirection fait partie d'une liste chaînée circulaire, l'entrée maîtresse pour des données clonées comprenant une référence à un index maître et une référence à un index suivant.

15 20. Système selon la revendication 17, dans lequel la structure de données d'indirection fait partie d'une liste chaînée circulaire, l'entrée de clone pour les données clonées comprenant une référence à l'index maître et une référence à un index suivant.

20 21. Système selon la revendication 17, dans lequel la structure de données d'indirection fait partie d'une liste chaînée à une seule extrémité, une entrée contenue dans un index fournissant une indication du fait que l'index est un index maître.

22. Système selon la revendication 17, dans lequel les références comprennent des entrées 25 dans une table d'indirection absolue pour un adressage de blocs logiques.

23. Système selon la revendication 17, dans lequel le premier dispositif comprend un dispositif d'Interconnexion de Composants Périphériques Express (PCIe).

30 24. Système selon la revendication 17, comprenant en outre une instruction destinée à positionner un indicateur pour un ou plusieurs blocs logiques condensés pour indiquer que lesdits un ou plusieurs blocs logiques condensés sont clonés.

25. Système selon la revendication 20, dans lequel l'index maître de l'entrée maîtresse et l'index maître de l'entrée clonée pointent sur l'entrée maîtresse et l'index suivant d'une dernière entrée clonée dans une structure de données pointe sur l'entrée maîtresse.

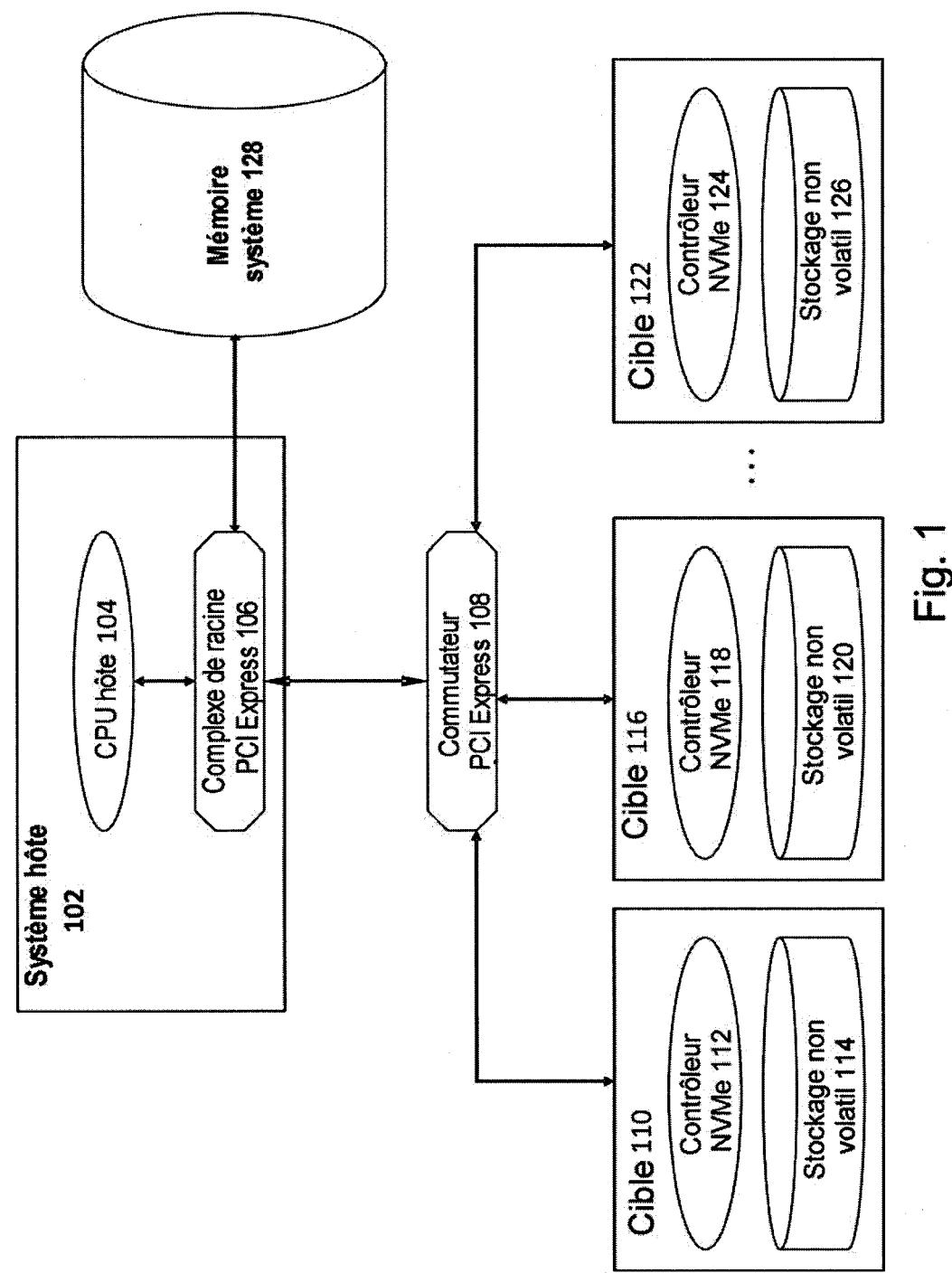
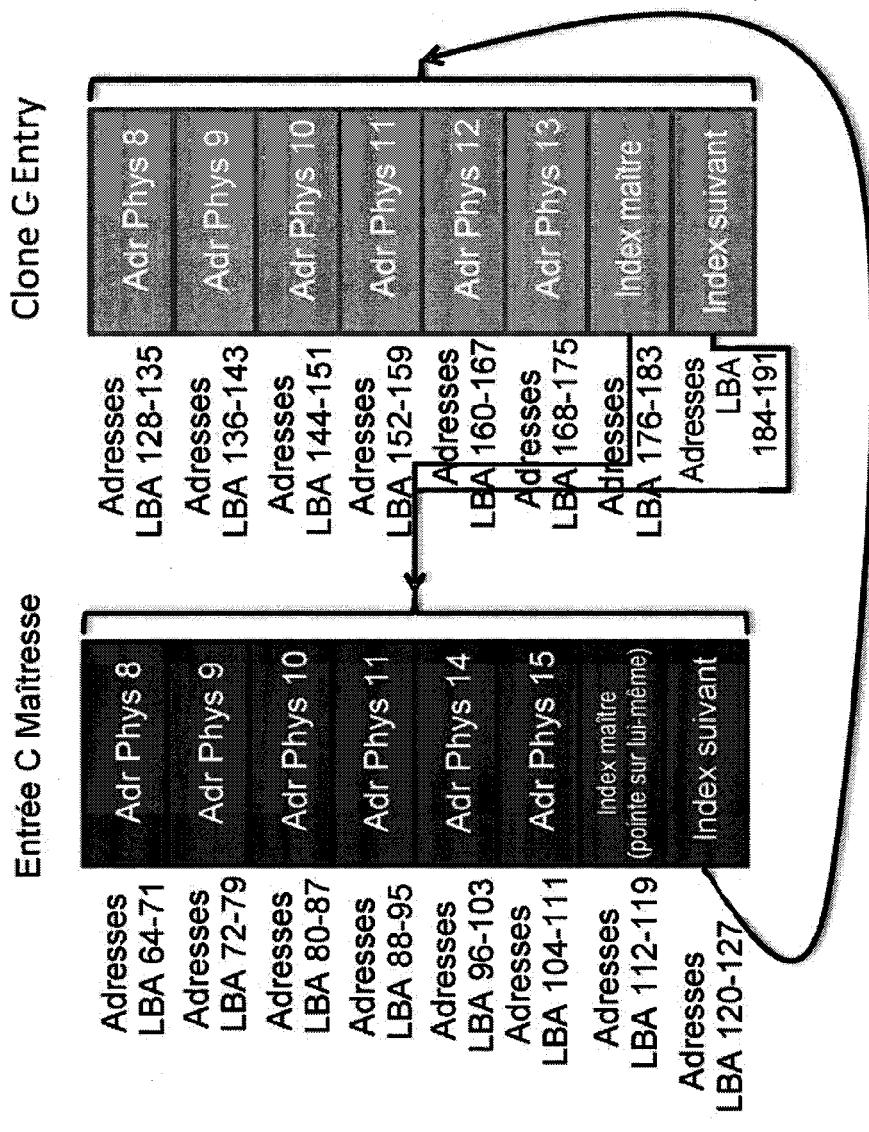


Fig. 1

Blocs PLB non clonés

Adresses	LBA 0-7	Adr Phys 0
Adresses	LBA 8-15	Adr Phys 1
Adresses	LBA 16-23	Adr Phys 2
Adresses	LBA 24-31	Adr Phys 3
Adresses	LBA 32-39	Adr Phys 4
Adresses	LBA 40-47	Adr Phys 5
Adresses	LBA 48-55	Adr Phys 6
Adresses	LBA 56-63	Adr Phys 7

Fig. 2



L'exemple utilise des entrées PLB de 4 ko et une taille de Segment C égale à un multiple de 8. L'"Index Suivant" forme une liste chaînée de manière unique de segments C maîtres et tous clonés.

Fig. 3

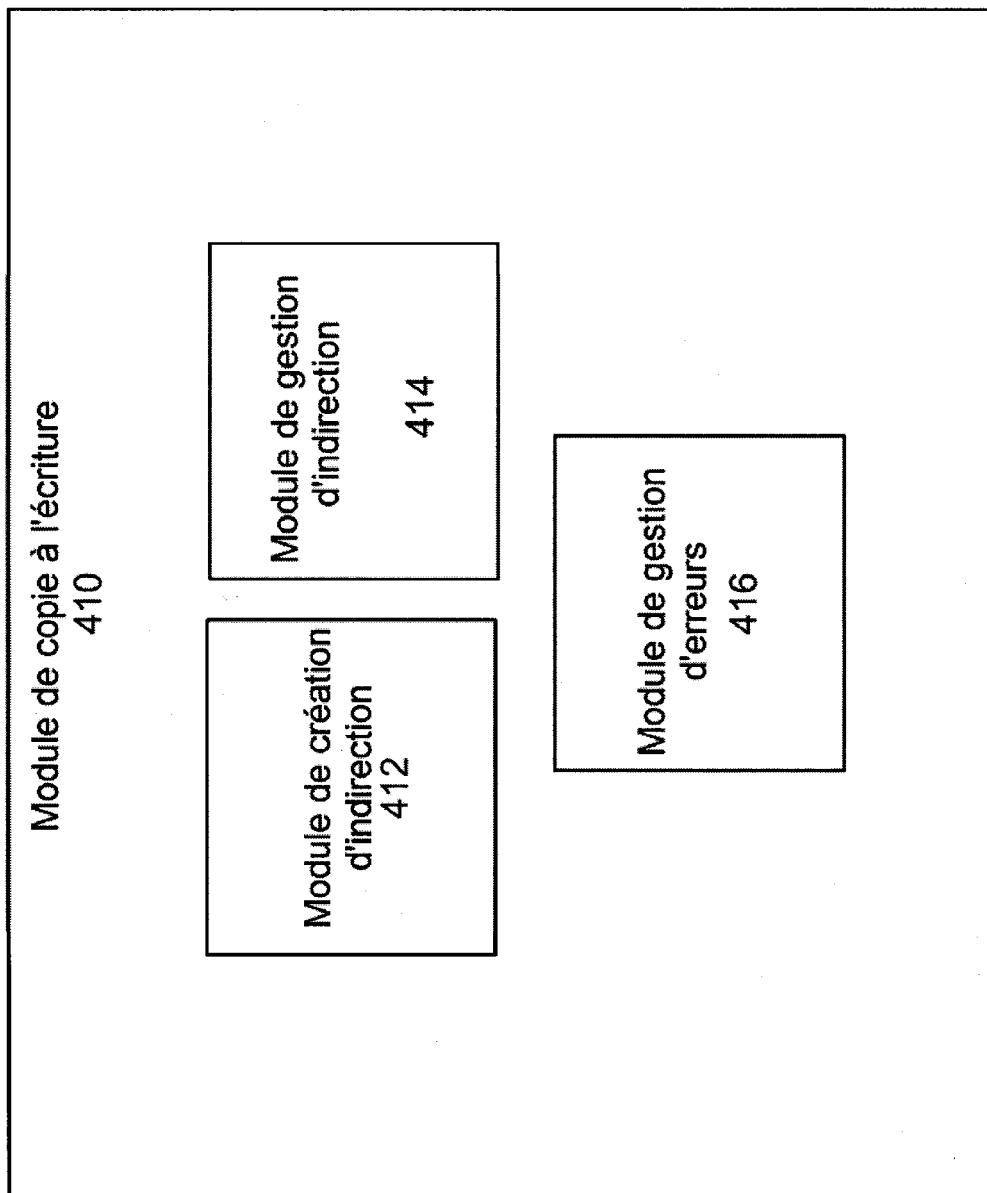


Fig. 4

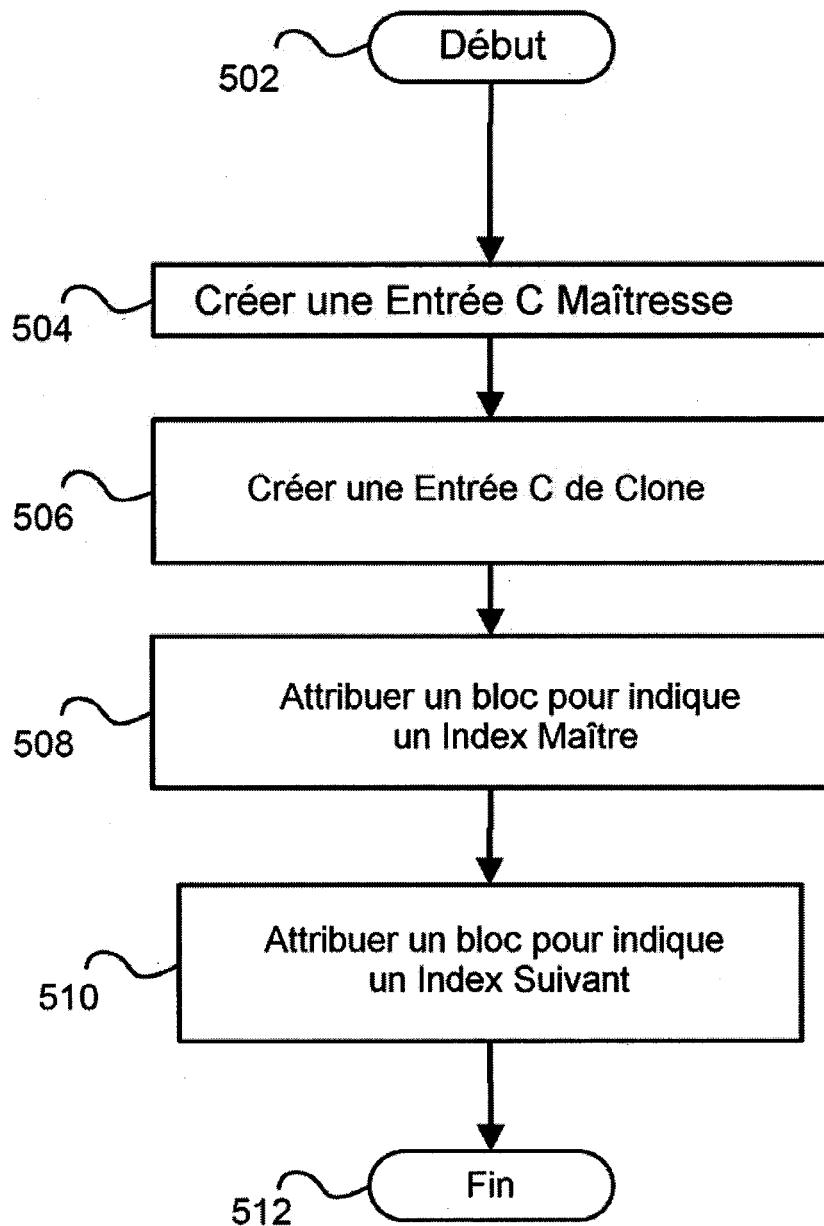


Fig. 5

6/6

Adresse ₀ NON-ET	Suivi Clone = 1
Adresse ₁ NON-ET	Suivi Clone = 1
Adresse ₂ NON-ET	Suivi Clone = 1
Adresse ₃ NON-ET	Suivi Clone = 1
Adresse ₄ NON-ET	Suivi Clone = 1
Adresse ₅ NON-ET	Suivi Clone = 1
Index Maître	Suivi Clone = 1
Index Suivant	Suivi Clone = 1

Fig. 6A

Adresse ₀ NON-ET	Suivi Clone = 1
Adresse ₁ NON-ET	Suivi Clone = 1
Adresse ₂ NON-ET	Suivi Clone = 1
Adresse ₃ NON-ET	Suivi Clone = 1
Adresse ₄ NON-ET	Suivi Clone = 1
Adresse ₅ NON-ET	Suivi Clone = 1
Index Maître	Suivi Clone = 1
Index Suivant	Suivi Clone = 1

Fig. 6B