

Meilleures pratiques pour défragmenter un espace de stockage à allocation dynamique



Meilleures pratiques pour défragmenter un espace de stockage à allocation dynamique

Pourquoi défragmenter ?

Avant de considérer et de recommander des configurations dans des environnements de stockage à allocation dynamique, il est important de revoir pourquoi la défragmentation des systèmes d'exploitation Windows s'avère primordiale dans une machine virtuelle et/ou un environnement de stockage virtualisé. Lorsque des données sont fragmentées dans un système de fichiers sur disque local, tel que NTFS, le système d'exploitation génère des requêtes E/S supplémentaires. Pour chaque fragment « logique » du système de fichiers, un paquet de requête E/S distinct doit être créé et transmis aux niveaux de stockage sous-jacents.

Ainsi, par exemple, un fichier comportant 100 fragments générerait 100 entrées/sorties distinctes de plus petite taille, plutôt qu'une seule entrée/sortie globale.

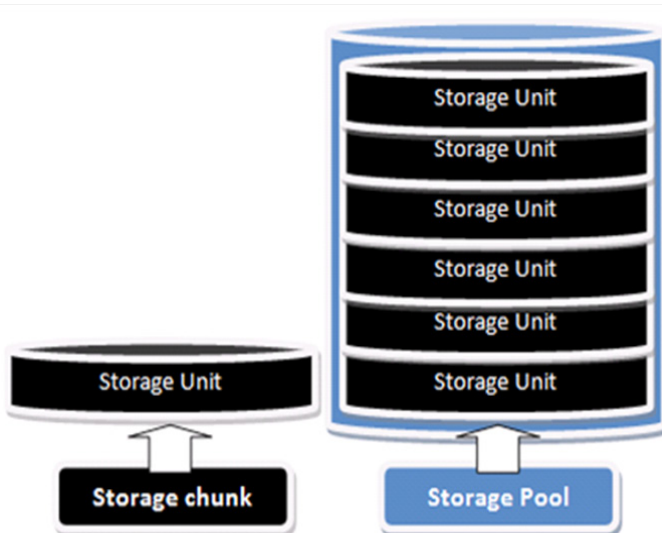
Dans ce cas, un système d'exploitation traite davantage de trafic E/S inutile, ce qui augmente la demande en processeur et mémoire. Dans de nombreux cas, ces E/S supplémentaires sont transmises à un réseau SAN et/ou une plate-forme de virtualisation, ce qui génère des coûts additionnels inutiles.

Dans certains cas, les données qui se trouvent dans des clusters en séquence contiguë d'un système de fichiers sur disque local deviendront physiquement contiguës sur le support de stockage réel tel que le disque ou la baie de stockage.

Cela représente généralement un autre avantage précieux, mais en rien nécessaire à la défragmentation pour améliorer considérablement la performance. Certains systèmes de fichiers (par exemple, un système de fichiers en journaux) utilisés dans des réseaux SAN peuvent délibérément fragmenter des données au niveau du bloc.

Ils peuvent combiner des écritures aléatoires du système d'exploitation en écritures séquentielles au sein du système de stockage. Tandis que cela minimise l'activité E/S dans le réseau SAN, il devient alors plus probable que les données présentes dans ces bandes écrites en séquence soient physiquement fragmentées, car le processus de fusion ne repose pas sur la réorganisation des blocs lors de leur corrélation dans un fichier commun. Les données sont simplement copiées sur le support. Pour ces environnements, vous devez vérifier auprès de votre fournisseur de stockage quelles sont les solutions de défragmentation propriétaires adaptées à son réseau SAN.

« Nous utilisons le logiciel d'optimisation Diskeeper® dans notre vaste système SQL (8 processeurs, des centaines de gigaoctets d'espace sur un SAN, 16 gigaoctets de mémoire vive). La performance des disques a augmenté d'un facteur d'environ 8 à 9. Nous souhaitons ajouter davantage de disques à notre réseau SAN pour résoudre certains problèmes liés aux E/S de disque, mais ce merveilleux logiciel l'a fait pour nous », Dave Underwood, ingénieur principal, CustomScoop.



Indépendamment de la proximité spatiale, l'avantage d'un système de fichiers sur disque local sans fragmentation réside dans le fait que votre système d'exploitation et vos plates-formes de virtualisation ne traitent pas les E/S supplémentaires générées, en raison de la fragmentation. Ces derniers pourront donc héberger un plus grand nombre de systèmes d'exploitation et traiter davantage de données, et cela plus rapidement.

Allocation dynamique 101

Le processus d'allocation dynamique alloue les ressources à partir d'un pool de stockage global, divisé principalement en unités affectables communément appelées « blocs ». L'allocation de stockage dans des environnements « dynamiques » s'effectue au niveau des blocs qui sont prélevés du pool de stockage disponible et encore non alloué.

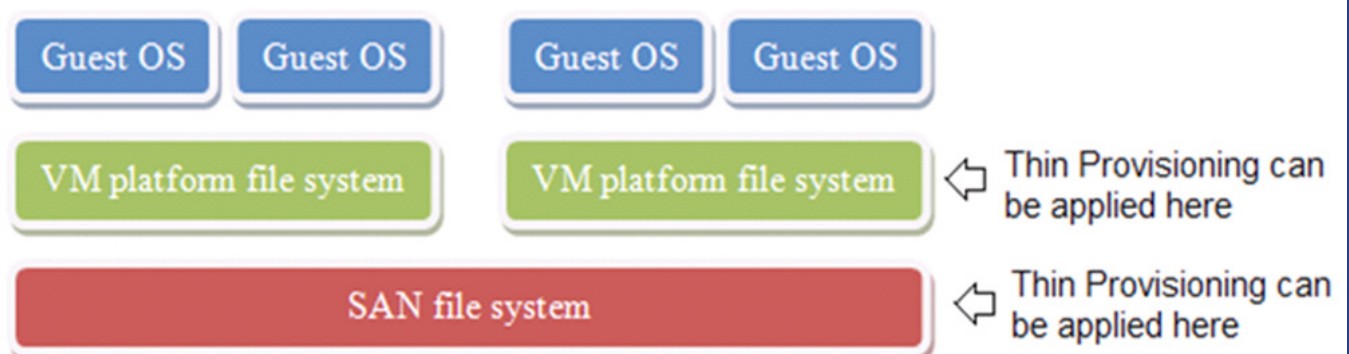
Le terme « dynamique » fait référence à l'utilisation de la technologie d'allocation dynamique au niveau de la couche de la plate-forme virtuelle et de la baie de stockage.

Les pools de stockage se composent de nombreuses et petites unités de stockage (blocs).

Alors que des données sont ajoutées à un conteneur à allocation dynamique, tel qu'un disque virtuel dynamique ou une LUN, ce conteneur croît en taille, généralement en temps réel, d'un bloc ou de plusieurs de ces blocs, selon le nombre de blocs nécessaires pour stocker toutes les écritures entrantes. Un bloc peut représenter en taille quelques kilooctets ou gigaoctets, et varier d'un fournisseur de technologie d'allocation dynamique à un autre. Dans certains cas, la taille est fixe, mais dans d'autres cas, l'utilisateur définit la taille du bloc.

Le mode et la fréquence d'allocation des blocs varient aussi d'un fournisseur à l'autre.

De nombreuses technologies d'allocation dynamique traitent chaque écriture. Elles surveillent les blocs et notamment les modifications au niveau des blocs. Lorsque de nouvelles données sont écrites, l'espace est ainsi alloué en temps réel et ces dernières sont stockées.



Un autre moyen pour allouer de l'espace repose sur le dernier bloc d'un volume Windows. Le dernier bloc, lorsqu'on se réfère à un volume, est le terme qui décrit le dernier cluster/bloc de données écrit (le numéro de cluster logique, ou LCN, le plus élevé du volume). Tout ce qui se trouve au-delà du dernier bloc est considéré comme insignifiant.

Modèle d'écriture et de suppression NTFS

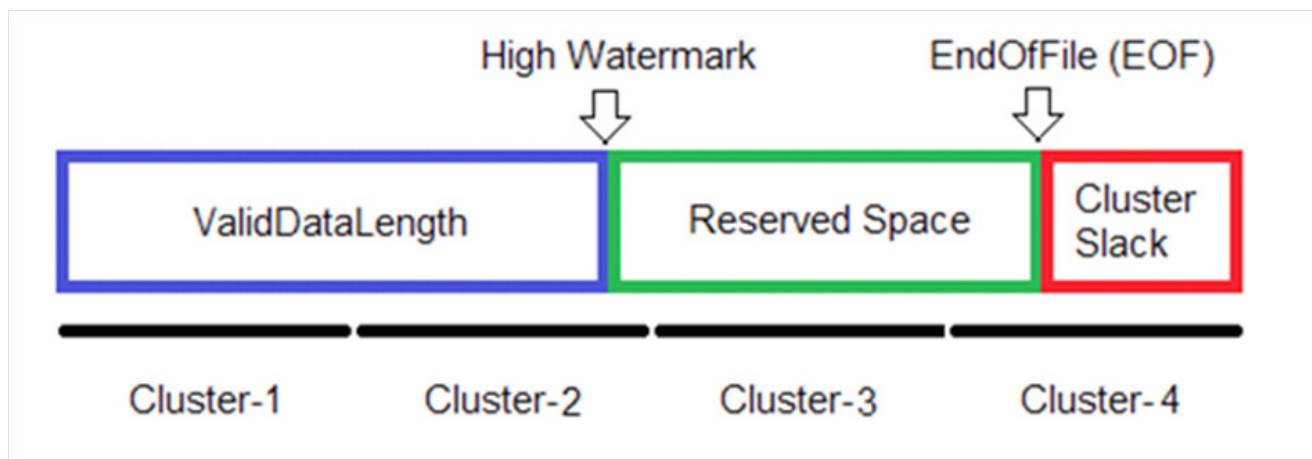
Bien que ne favorisant pas vraiment l'allocation dynamique, NTFS est considéré à tort comme représentant un problème pour les disques/LUN à allocation dynamique. Il a été rapporté que NTFS créait avec négligence en continu de nouveaux numéros de cluster logique plus élevés jusqu'à ce qu'il ait écrit dans chaque cluster du volume, avant de revenir aux clusters libérés suite à des suppressions de fichiers. Cela n'est pas vrai.

Lorsque l'on décrit la conception de NTFS en termes d'allocation de stockage, il faut en premier lieu préciser les différentes tailles de fichier. Il existe trois tailles pour les fichiers NTFS, avec utilisation aussi de derniers blocs.

La longueur de données valide correspond à la distance dans le fichier à laquelle les données ont été effectivement écrites, puisqu'elles résident dans le cache. Cela est reproduit dans la barre bleue du diagramme.

Une longueur de données valide peut inclure des plages clairsemées disséminées entre les données. Le numéro de cluster logique écrit le plus élevé qui constitue la longueur de données valide correspond au dernier bloc de ce fichier. Il n'y a pas de données, du moins liées à ce fichier, qui résident au-delà du dernier bloc. Sans avoir à effectivement écrire des zéros et simplement comme avec les volumes de stockage à dernier bloc, les tentatives de lecture au-delà du dernier bloc renvoient des zéros.

Depuis Windows XP, tous les éléments d'un flux de fichier, dont et jusqu'à la taille d'allocation (l'extrémité du fichier entre la longueur de données valide et la fin du fichier) peuvent être défragmentés



La prochaine étape concerne la taille de fichier. Il s'agit de la longueur de données valide ainsi que d'un espace pré-réservé supplémentaire, qui n'a pas encore fait l'objet d'une écriture (non initialisé), appelé aussi fin ou queue du fichier. Cela constitue la taille logique entière du fichier, indiquée par la combinaison des parties bleue et verte du graphique, et qui prend fin au repère fin de fichier (EndOfFile).

La taille d'allocation vient ensuite. Elle indique la taille physique complète du fichier et est constituée de la longueur de données valide et de l'espace réservé consécutif, jusqu'au dernier cluster que le fichier occupe en partie (peut-être de l'espace perdu). Cela est représenté par la combinaison des parties bleue, verte et rouge du graphique.

Pour faciliter l'écriture de nouvelles données, le pilote du système de fichiers NTFS gère une liste des espaces libres les plus importants sur le volume (à savoir, le numéro de cluster logique de départ et la plage). Lors de la création d'un fichier, l'opération s'effectue dans l'espace libre qui correspond au mieux à la taille des données disponibles à écrire (optimisation).

De plus, on suppose qu'un nouveau fichier crée dépassera la taille actuellement disponible nécessaire pour permettre au système d'exploitation d'écrire. Un espace libre supplémentaire (sur-allocation) est réservé au fichier afin de minimiser la fragmentation (se reporter à l'article ID 228198 de la base de connaissances Microsoft). On suppose aussi que le fichier sera 2, 4, 8, voire 16 fois plus grand que la taille des données actuellement connue. Cela dépend de la quantité de données actuellement disponible pour l'écriture du fichier dans la mémoire cache du système d'exploitation.

Les données du fichier sont écrites sur le volume et le fichier est fermé. Toute sur-allocation est ensuite libérée. L'espace est réintégré au pool d'espace libre et au pilote du système de fichiers NTFS, si considéré comme l'un des

De plus, pour adapter au mieux les tentatives, Windows essaie aussi d'écrire de nouvelles extensions de fichier dans un cluster à proximité d'extensions existantes.

plus grands espaces libres du volume. Dans ce cas, et cela est un point stratégique, NTFS est très favorable à l'allocation dynamique lorsqu'il réserve cette sur-allocation. Il peut le faire sans écrire sur le volume (par une opération d'écriture à blanc).

En définitive, ce processus n'élimine pas la fragmentation dans une quelconque mesure, d'où la nécessité continue de défragmenter le système de fichiers.

Un problème inhérent à NTFS, source de défis universels pour le stockage à allocation dynamique, concerne la capacité à récupérer de l'espace précédemment occupé par les fichiers supprimés.

Ceci est un problème car lorsque les fichiers sont supprimés dans NTFS, le système de fichiers ne fait que mettre à jour ses métadonnées pour indiquer que l'espace occupé peut être réutilisé pour les écritures de nouveaux fichiers. Un fichier supprimé n'est pas véritablement effacé du volume. Ainsi, il se peut que les niveaux de stockage virtuel sous-jacents à NTFS n'aient pas alors connaissance du nouvel espace libre disponible.

Cela soulève un problème quant à l'allocation de stockage dynamique qui, face à des limites de réutilisation d'espace, pourrait éventuellement épuiser tout le stockage dans le pool disponible.

Une solution à ce défi, communément appelée Thin Reclamation (récupération de l'espace alloué), comprend l'identification de l'espace anciennement occupé par les données supprimées et les opérations effectuées ensuite pour récupérer et réapprovisionner cet espace. De nombreuses solutions disponibles permettent de récupérer de l'espace alloué comme l'initialisation des clusters supprimés avec les commandes SCSI UNMAP ou SCSI WRITE_SAME. Ces solutions peuvent varier d'un fournisseur à l'autre.

Défragmentation et allocation dynamique

Comme indiqué précédemment, la défragmentation est un processus vital lorsqu'il s'agit d'atteindre et de maintenir des performances optimales. Lorsque l'allocation dynamique est utilisée pour un système de fichiers partagé de l'hôte virtuel, il s'avère fort probable que les fichiers du disque virtuel dynamique deviennent fragmentés à leur tour, ce qui ajoute des E/S supplémentaires. Dans ces systèmes de stockage, il s'avère primordial de résoudre la fragmentation.

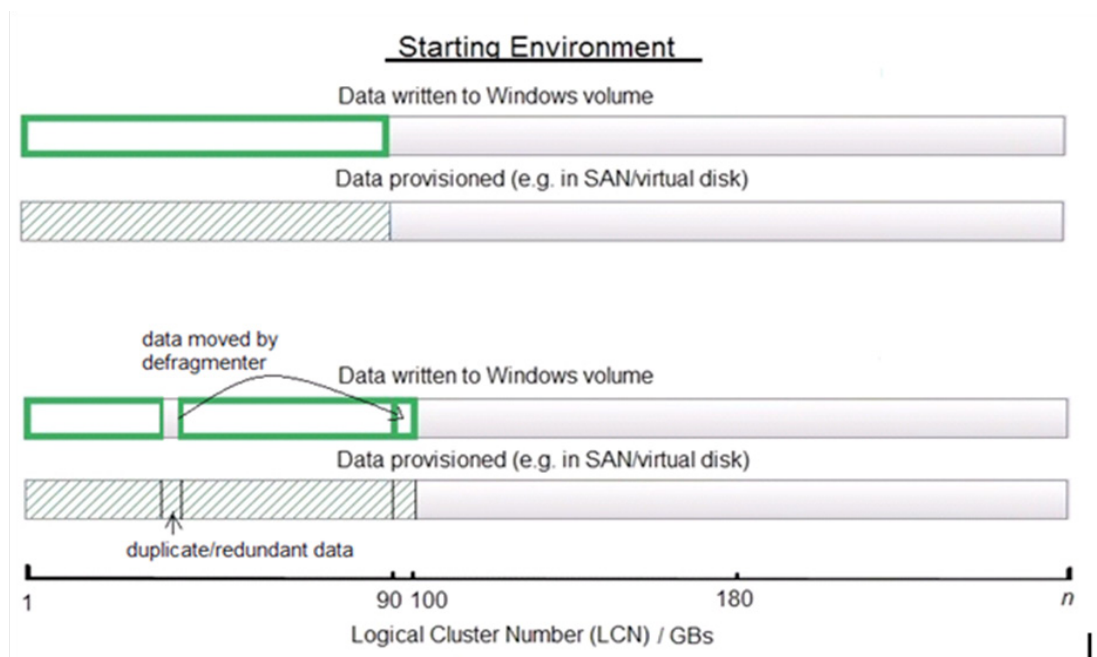
Cependant, malgré tous les avantages de la défragmentation, il est important de considérer aussi les effets indirects potentiels. Ces effets indirects peuvent varier de l'application d'une technologie dynamique à une autre. Ainsi, il est important de savoir comment les deux technologies interagissent.

En utilisant les contrôles d'entrée/sortie dans Windows, la défragmentation consiste essentiellement à transférer des données pour consolider les fragments de fichiers et de regrouper l'espace libre en grandes zones contigües.

Lorsque la technologie alloue de l'espace pour de nouvelles écritures, un processus de défragmentation (qui ne consiste en fait qu'à transférer des données) apparaît sous forme de nouvelles écritures. De plus, les anciens emplacements des données transférées ne sont pas nécessairement identifiés pour une réutilisation. La défragmentation entraîne cependant d'autres exigences en termes de capacité de stockage pour chaque segment de données transférées.

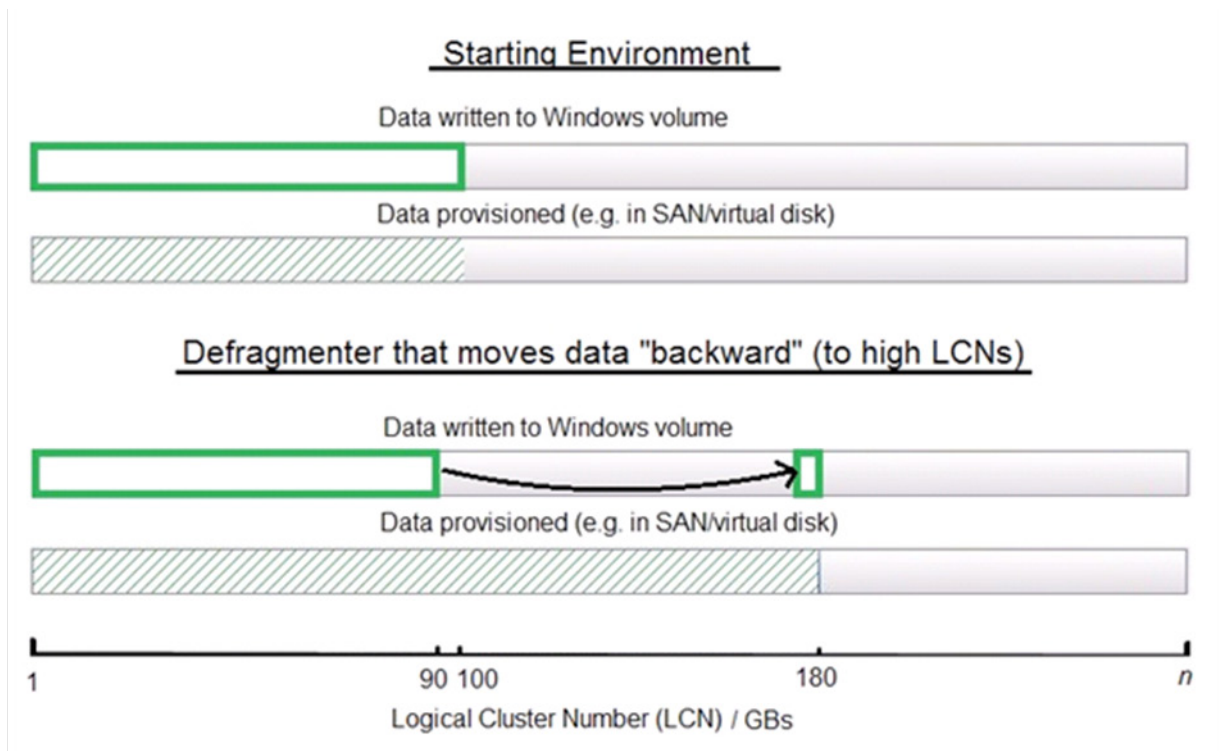
De nombreuses solutions de stockage suivent les modifications au niveau bloc (par exemple, VMFS) pour prendre en charge les fonctionnalités avancées comme les snapshots et la migration à chaud de données. Ainsi, si les blocs du système de fichier utilisés pour la défragmentation sont réduits et que les mêmes sous-ensembles de blocs sont réutilisés par le processus de défragmentation, la croissance du stockage actuel sera probablement infime.

Les nouvelles écritures pourraient faire l'objet d'une allocation redondante, ce qui consommerait de l'espace inutilement



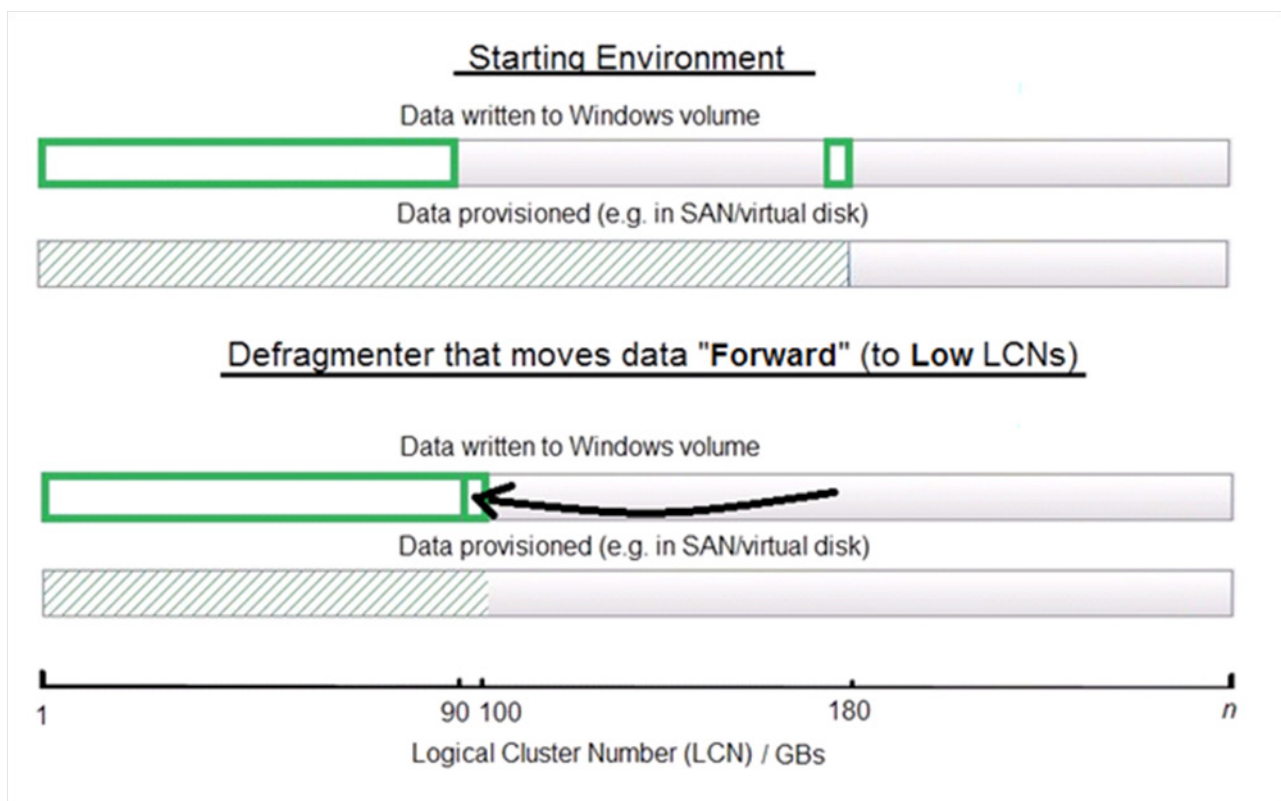
La récupération de l'espace alloué permet de récupérer l'espace gaspillé, comme en cas d'exécution d'un processus de déduplication des données (qui identifierait et supprimerait les données redondantes).

Lorsqu'une allocation de dernier bloc est utilisée, le dernier bloc augmente toujours et ne diminue jamais (sous Windows), ce qui se traduit par moins d'espace disponible et crée donc un problème potentiel. Si un fichier est écrit (ou déplacé par une défragmentation) dans un cluster supérieur, la technologie d'allocation dynamique doit prévoir l'espace destiné à le contenir. Cela s'avère vrai même si le fichier est seulement déplacé temporairement vers un cluster supérieur.



À l'autre extrémité du spectre, le transfert de fichiers en aval peut permettre aux processus de récupération d'espace de mieux récupérer l'espace provisionné (voir le graphique ci-dessous).

Les défragmenteurs peuvent faciliter le processus de compression des fichiers en début de volume.



L'optimiseur de disque de plate-forme virtuelle V-locity®, Diskeeper Pro Premier, Diskeeper Server et les différentes versions de Diskeeper EnterpriseServer utilisent des moteurs spécialisés (les technologies Terabyte Volume Engine™ et Titan Defrag Engine™) conçus généralement pour défragmenter/transférer des fichiers vers un numéro de cluster logique inférieur des volumes Windows.

Prévention proactive de la fragmentation

Il est important d'évaluer les slogans marketing des fournisseurs de solutions de défragmentation concernant l'élimination/la prévention de la majorité de la fragmentation avant qu'elle ne se produise, puisque la technologie à la base de ces slogans peut avoir des conséquences différentes pour le stockage à allocation dynamique.

Les solutions réactives, qui reposent sur une importante consolidation de l'espace libre (compactage de fichiers) pour tirer parti de la fonctionnalité native d'optimisation des systèmes de fichiers NTFS, entraînent une croissance de l'allocation dynamique.

Les technologies proactives, qui ne nécessitent pas d'autre transfert des données pour atteindre leur objectif, ne conduisent pas à l'augmentation de l'espace de stockage à allocation dynamique. Elles offrent l'avantage d'un système de fichiers d'OS en grande partie sans fragmentation qui ne présente aucune conséquence négative pour l'allocation dynamique de stockage.

La technologie en instance de brevet IntelliWrite® de Diskeeper Corporation est une solution véritablement proactive.

IntelliWrite utilise un meilleur processus (par rapport aux sur-allocations en mode natif de NTFS) qui réserve de l'espace après la longueur de données valide en fin de fichier. IntelliWrite est un outil intelligent en ce sens qu'il analyse l'origine des écritures/modifications de fichier et intègre leur mode de fonctionnement au fil du temps. Ce processus heuristique signifie qu'IntelliWrite détermine mieux le volume d'espace de réservation nécessaire à un fichier ouvert pour prévenir la fragmentation. Le fichier peut en effet demander plus d'espace que ce que NTFS allouerait en mode natif, ou moins.

Le résultat des sur-allocations intelligentes d'IntelliWrite se traduit par un degré inégalé de prévention efficace de la fragmentation (un taux de réussite de 85 % ou plus).

Meilleures pratiques

- Utilisez une technologie de prévention proactive de la fragmentation, telle qu'IntelliWrite de Diskeeper Corporation.
- Demandez à votre fournisseur privilégié comment fonctionne l'allocation dynamique et quelles solutions il offre pour la récupération d'espace (dynamique).
 - o Dans des environnements à allocation dynamique étagée, la récupération d'espace à un niveau (par exemple, un disque virtuel dynamique) n'englobe pas nécessairement les autres espaces de stockage alloués à des niveaux inférieurs (par exemple une LUN).
- Défragmentez des volumes à allocation dynamique lorsque la croissance du stockage correspondante peut être gérée (par exemple, avec un processus de récupération dynamique/de déduplication).
- Concernant l'allocation de dernier bloc, utilisez un défragmenteur qui déplace les fichiers vers des numéros de cluster logique inférieurs (processus amont).

- Utilisez un défragmenteur OS/GOS, ou un mode de défragmentation axé sur la performance et non sur l'esthétique.
- Adoptez les outils du fournisseur SAN/VM pour éliminer la fragmentation selon ses pratiques recommandées pour ses systèmes de fichiers en cluster propriétaires.
- Les technologies de classement/organisation de fichiers utilisées par les défragmenteurs des systèmes d'exploitation d'entreprise peuvent s'avérer vraiment précieuses dans de nombreux environnements, notamment pour des solutions axées sur la performance en cas de stockage à attachement direct (Direct Attached Storage). Cependant, elles peuvent conduire les technologies d'allocation dynamique de stockage à une croissance excessive en raison des transferts supplémentaires de données. Ainsi, il est recommandé de les désactiver ou de ne les exécuter que si elles permettent de traiter les effets (à savoir la croissance du stockage).

Diskeeper Corporation

Shaw House, Pegler Way

Crawley, West Sussex RH11 7AF, Royaume-Uni

Téléphone 01 82 88 26 00 Fax +44(0)1293763392 Web: www.diskeeper.fr

Comptant parmi ses clients détenteurs de licences en volume plus de 80 % des entreprises figurant dans le classement Fortune 1000 aux Etats-Unis et fort de plus de deux décennies placées sous le signe de l'innovation en termes de performance des systèmes et de fiabilité (avec l'accent mis sur l'efficacité du stockage), Diskeeper Corporation est un expert reconnu dans le domaine de l'optimisation du stockage.

© 2010 Diskeeper Corporation. Tous droits réservés. Diskeeper Corporation, IntelliWrite, Terabyte Volume Engine, Titan Defrag Engine, V-locity et Diskeeper sont des marques déposées ou commerciales de Diskeeper Corporation. Toutes les autres marques appartiennent à leur propriétaire respectif.