

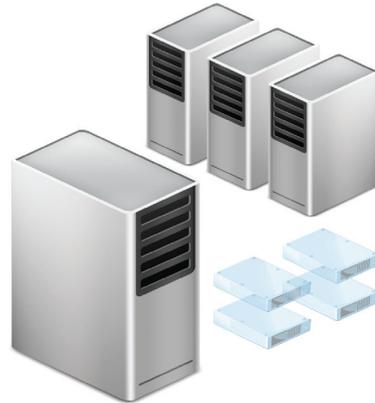
livre blanc

Comment optimiser les performances des plates-formes virtuelles ?

L'intensification de la demande en termes d'efficacité du réseau informatique et de baisse des coûts d'exploitation a suscité un développement phénoménal de la virtualisation ces dix dernières années, sans montrer le moindre signe de ralentissement. Aujourd'hui, beaucoup d'entreprises exécutent davantage de serveurs virtualisés que de serveurs physiques.

Alors que la virtualisation offre des possibilités de consolidation et de meilleure utilisation du matériel, il est extrêmement important de connaître la capacité du matériel et de ne jamais la dépasser.

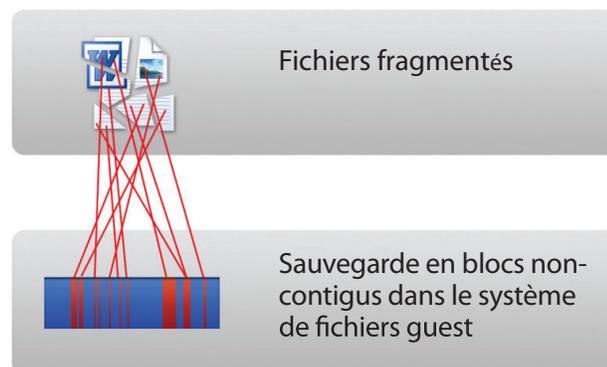
Le besoin de disposer de suffisamment de puissance de calcul et de capacité de stockage est jugée à sa juste valeur, comme en témoignent les nombreux processus et outils de gestion disponibles pour aider à la planification et au juste provisionnement des machines virtuelles pour ces ressources essentielles.



Le trafic E/S, le réseau et les disques sont plus difficiles à comptabiliser dans les environnements virtuels, car ils tendent à être plus imprévisibles

Afin de mieux prendre en charge les E/S disque, la plupart des plates-formes de virtualisation vont mettre en œuvre un réseau de stockage SAN pouvant offrir un débit de données plus important, ainsi qu'un environnement dynamique pour atténuer les fluctuations des demandes E/S.

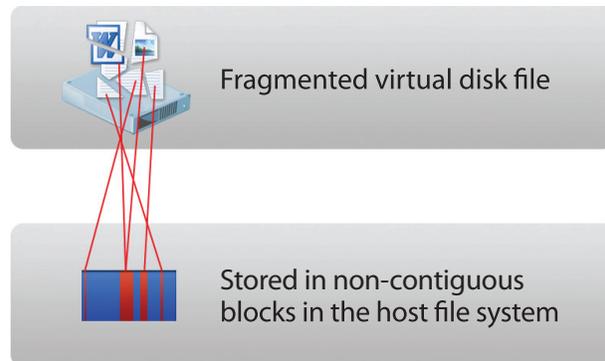
Bien qu'il soit possible de mettre en place une infrastructure de stockage pour répondre aux besoins prévus, des comportements incontrôlables ralentiront toujours les performances.



Fragmentation des fichiers

Parce que les fichiers sont enregistrés sur des systèmes de fichiers de disque local généralistes tels que Windows NTFS, la fragmentation des fichiers s'impose naturellement. Le terme de fragmentation caractérise un état où le flux de données d'un fichier est stocké dans des clusters non contigus au sein du système de fichiers.

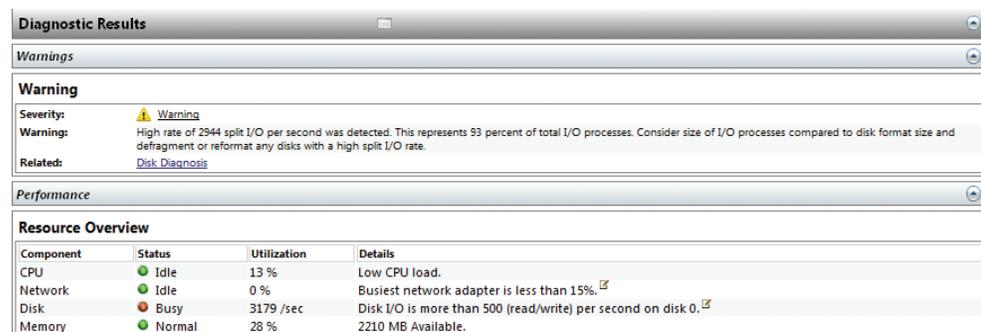
La fragmentation s'applique aux volumes logiques, qu'elle convertit au moyen de pilotes de périphériques en blocs logiques, puis en secteurs physiques résidant sur un périphérique de stockage. On peut se la représenter comme des morceaux de fichiers organisés de manière non contiguë.



La fragmentation des fichiers entraîne une augmentation de la charge d'E/S, ce qui ralentit les performances système pour le système d'exploitation.

Dans le cas des plates-formes virtuelles, un système d'exploitation invité est stocké sous forme de fichier (ensemble de fichiers) sur le système de fichiers des plates-formes virtuelles, en tant que « disque virtuel ». Un disque virtuel est essentiellement un fichier conteneur qui héberge tous les fichiers constituant le système d'exploitation et les données utilisateur d'une machine virtuelle.

Un fichier disque virtuel peut se fragmenter comme tout autre fichier, et ainsi former un disque dur



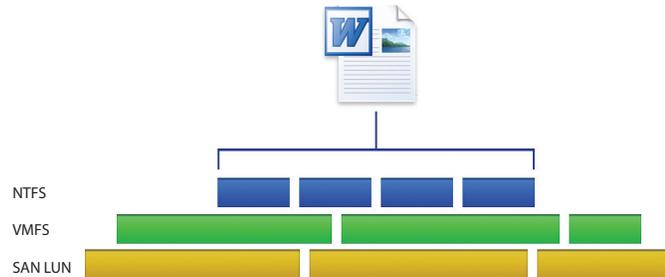
Component	Status	Utilization	Details
CPU	Idle	13 %	Low CPU load.
Network	Idle	0 %	Busiest network adapter is less than 15%. [?]
Disk	Busy	3179 /sec	Disk I/O is more than 500 (read/write) per second on disk 0. [?]
Memory	Normal	28 %	2210 MB Available.

virtuel « logiquement » fragmenté, qui contient toujours une fragmentation type des fichiers. La légende de l'image représentée à droite pourrait être « VirtualServer1.vmdk, 30 Go, en 4 fragments ».

Cette situation équivaut à une fragmentation hiérarchique ou, plus simplement, à une fragmentation dans la fragmentation. Étant donné la nature relativement statique et le grand volume des disques virtuels ainsi que la grande taille d'unité d'allocation du VMFS (généralement 1 Mo), la fragmentation de ces fichiers ne devrait pas provoquer de baisses de performances dans la plupart des cas.

C'est sur le système d'exploitation invité qu'il convient de porter ses efforts pour régler les problèmes de fragmentation.

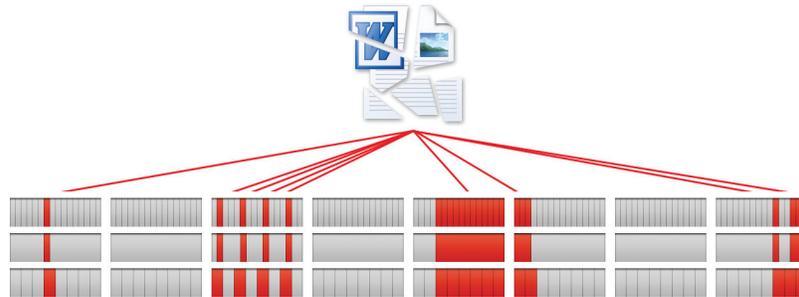
Fragmentation within a Windows VM will cause Windows to generate additional unnecessary I/O. This added I/O traffic can be discovered using Windows Performance Monitor, where it is one of the principal causes for split I/O.



La fragmentation au sein d'une machine virtuelle Windows conduit

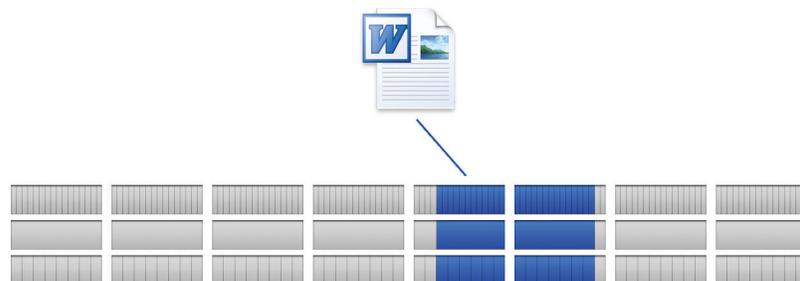
Windows à générer des E/S supplémentaires inutiles. Il est possible de détecter ce surplus de trafic E/S avec l'Analyseur de performances Windows, lorsqu'il constitue l'une des principales causes de fractionnement des E/S.

NTFS (64KB Cluster)
 VMFS (1MB Block)
 SAN LUN (128KB Stripe)



Il existe des technologies de prévention de la fragmentation et de défragmentation permettant d'éliminer la charge d'E/S inutile et d'améliorer les performances du système. La prévention de la fragmentation résout la fragmentation à la source, en provoquant dynamiquement l'écriture contiguë

NTFS (64KB Cluster)
 VMFS (1MB Block)
 SAN LUN (128KB Stripe)



1 VMware guide to proper partition alignment: http://www.vmware.com/pdf/esx3_partition_align.pdf

2 It should be noted that VMFS, in the example above need only read the actual amount of data requested in multiples of 512 byte sectors, and does not need to read an entire 1MB block.

des fichiers par le biais de pilotes de système de fichiers avancés. Il y a défragmentation lorsque les fragments de fichiers sont réalignés dans le système de fichiers en un seul continuum, de sorte qu'une quantité minimale d'E/S disque est nécessaire pour accéder au fichier, d'où l'augmentation de la vitesse d'accès

Alignement des partitions

Selon votre protocole de stockage et votre type de disque virtuel, un mauvais alignement des partitions peut générer des E/S supplémentaires inutiles.

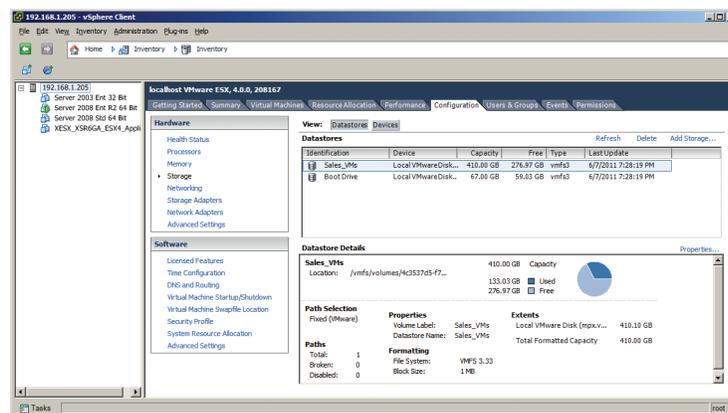
Dans l'exemple de droite, où les volumes ESX et SAN ne sont pas correctement alignés, un fichier Word englobant quatre clusters NTFS génère des E/S supplémentaires inutiles tant dans le système de fichiers VMFS que dans la LUN du SAN.

General	
Manufacturer:	HP
Model:	ProLiant DL385 G2
CPU Cores:	2 CPUs x 2.6 GHz
Processor Type:	Dual-Core AMD Opteron(tm) Processor 2218
License:	vSphere 4 Desktop Host Licensed for up to 100 VMs
Processor Sockets:	1
Cores per Socket:	2
Logical Processors:	2
Hyperthreading:	Inactive
Number of NICs:	4
State:	Connected
Virtual Machines and Templates:	4

Similarités entre alignement de partitions et fragmentation

La fragmentation, à l'instar des partitions incorrectement alignés, peut entraîner un surcroît d'E/S sur plusieurs couches. Alors qu'il est possible d'aligner les partitions une bonne fois pour toute, la fragmentation est permanente et doit être régulièrement contrôlée.

Dans l'exemple ci-dessous, qui suppose un alignement correct des partitions, un fichier en huit fragments dans le système d'exploitation invité génère des E/S supplémentaires sur la couche de la plate-forme de virtualisation2 et sur la LUN.



La défragmentation dans le système d'exploitation invité (de ce fichier) élimine les E/S excédentaires lors de l'accès au fichier du fait que Windows ne génère qu'une E/S. Cette réduction du trafic E/S s'étend au système de fichiers hôte et à la LUN du SAN, garantissant ainsi des gains d'efficacité sur chaque couche.

Bonnes pratiques

La défragmentation du système de fichiers Windows est une solution recommandée par VMware pour améliorer les performances.

L'article 10040043 de la Base de connaissances indique que : « La défragmentation d'un disque est nécessaire pour régler les problèmes rencontrés sur un système d'exploitation à la suite de la fragmentation du système de fichiers. Les problèmes liés à la fragmentation entraînent un ralentissement des performances du système d'exploitation ».

Des tests ont été effectués pour valider les déclarations de VMware.

Environnement de test

Configuration

OS hôte : ESX Server 4.1 avec VMFS (blocs d'1 Mo)

OS client : Windows Server 2008r2 x64 (3 Go de RAM, 1 vCPU)

Logiciel de tests de performance : Iometer (<http://www.iometer.org/>)

Programme de fragmentation : FragmentFile.exe (utilisé pour fragmenter un fichier particulier)

Logiciel de défragmentation : V-locity® (<http://www.diskeeper.com/business/v-locity/>)

Stockage :

Volume de test de 10 Go sur un disque virtuel de 40 Go.

Banque de données VMFS de 410 Go

Contrôleur HP Smart Array P400

RAID 5 (4 disques SCSI de 136 Go à 10 000 tr/min)

Taille d'agrégat par bande de 64 Ko avec décalage de 64 Ko (aligné correctement)

Génération de charge

L'outil de tests de performance au standard de l'industrie a servi à générer une charge d'E/S pour réaliser ces tests.

- Options de configuration Iometer utilisées comme variables pour ces tests :
- Tailles des demandes de transfert : 1 Ko, 4 Ko, 8 Ko, 16 Ko, 32 Ko, 64 Ko, 72 Ko et 128 Ko
- Pourcentage de distribution aléatoire/séquentielle : pour chaque taille de demande de transfert, 0 % et 100 % d'accès aléatoires ont été sélectionnés

- Pourcentage de distribution en lecture/écriture : pour chaque taille de demande de transfert, 0 % et 100 % d'accès en lecture ont été sélectionnés

Paramètres lometer demeurés constants pendant l'ensemble des tests :

- Taille de volume : 10 Go
- Taille du fichier de test lometer (iobw.tst) : 8 131 204 Ko (~7,75 Go)
- Nombre d'opérations E/S en attente : 16
- Temps d'exécution : 4 minutes

Volume Files	
Volume size	10,240 MB
Cluster size	4 KB
Used space	8,023 MB
Free space	2,216 MB
Percent free space	21 %

Low-Performing files percentage	
% of entire volume	77 %
% of used space	98 %

Most Fragmented Files	
Fragments	568,572
File size	7,941 MB
Most fragmented files	\iobw.tst

File fragmentation	
Total files	11
Average file size	724 MB
Total fragmented files	1
Total excess fragments	568,572
Average fragments per file	51,689.36
Files with performance loss	1

Free Space Fragmentation	
Percent low-performing free space:	0%
Total free space extents:	3
Largest free space extent:	911 MB
Average free space extent size:	739 MB
Percent free space	21 %

- Délai d'accélération : 60 secondes
- Nombre d'employés pour la génération automatique : 1

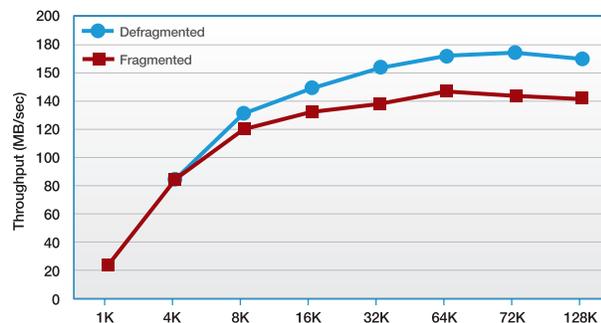
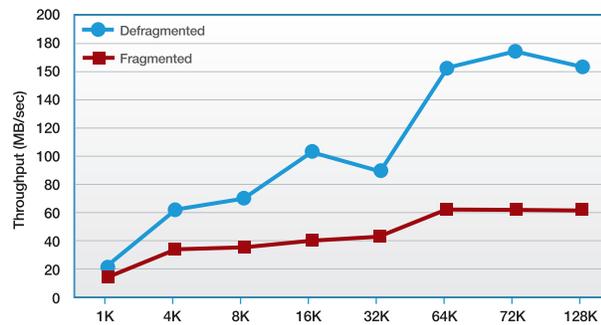
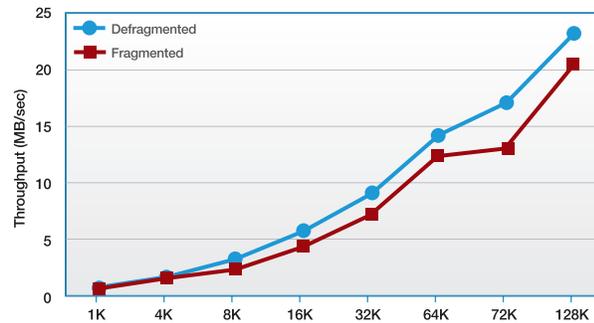
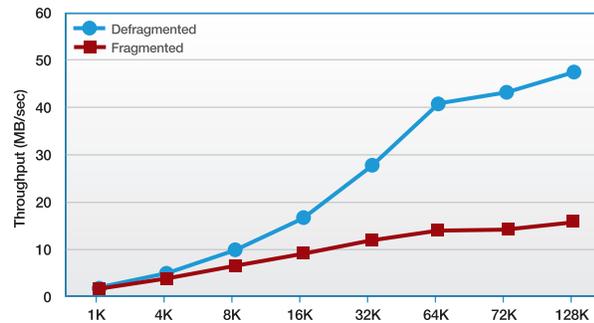
Les informations suivantes, extraites d'un livre blanc de VMware4, expliquent l'utilisation des paramètres lometer.

Les serveurs exécutent généralement un mélange de charges de travail comprenant différents modèles d'accès et tailles de données E/S.

Une charge de travail peut comprendre plusieurs tailles de transferts de données et plus d'un modèle d'accès.

Il existe peu d'applications où l'accès est strictement séquentiel ou purement aléatoire. Par exemple, les journaux de base de données sont écrits de manière séquentielle. Une opération de lecture séquentielle permet de lire à nouveau ces données pendant la phase de restauration de la base de données. Habituellement, l'accès aux bases de données OLTP est de nature majoritairement aléatoire.

La taille du transfert de données dépend de l'application et se mesure souvent en précisant une plage de valeurs plutôt qu'une valeur unique. Pour Microsoft Exchange, la taille des E/S est généralement faible (de 4 à 16 Ko), les accès aléatoires en lecture et en écriture à la base de données Microsoft SQL Server s'élèvent à 8 Ko, et les accès Oracle en général à 8 Ko, tandis que Lotus Domino utilise 4 Ko. Sur la plate-forme Windows, il est possible de déterminer la taille de transfert d'E/S d'une application à l'aide de l'outil Perfmon.



Pour résumer, les caractéristiques d'E/S d'une charge de travail se définissent en termes de rapport entre les opérations de lecture et les opérations d'écriture, de rapport entre les accès séquentiels et les accès aléatoires et de taille de transfert de données. Il est souvent possible de spécifier une plage de tailles de transfert de données plutôt qu'une valeur unique.

Création de la fragmentation

Nous avons utilisé l'outil FragmentFile.exe pour fragmenter le fichier de test d'Iometer (iobw.tst) en 568 572 fragments, soit une quantité de fragmentation moyenne pour un serveur de production.

L'outil V-locity a permis de générer les statistiques recueillies après analyse du volume (voir ci-dessous).

Procédure de test

L'objectif principal était la comparaison des performances des machines virtuelles fragmentées et défragmentées, avec plusieurs tailles de données et plusieurs modèles d'accès. Les tailles de données sélectionnées étaient de 1 Ko, 4 Ko, 8 Ko, 16 Ko, 32 Ko, 64 Ko, 72 Ko et 128 Ko. Les modèles d'accès étaient limités à une combinaison de 100 % en lecture ou en écriture et 100 % en aléatoire ou séquentiel. Chacune de ces quatre charges de travail a été testée avec huit tailles de données, pour un total de 32 points de données par charge de travail.

Afin d'isoler l'impact de la fragmentation, seule la machine virtuelle de test était sous tension et active pendant la durée des tests.

Pour l'exécution initiale, Iometer a créé un fichier non fragmenté et des données sur les performances ont été recueillies. Ensuite, nous avons utilisé l'outil FragmentFile.exe pour fragmenter le fichier de test Iometer, nous avons redémarré la machine virtuelle et relancé la procédure de test. Nous avons obtenu des ensembles de données pour les deux scénarios, non fragmenté et fragmenté. Les résultats sont indiqués dans le graphique ci-dessous.

Résultats en termes de performances :

Comme l'indiquent les graphiques, toutes les charges de travail présentent une augmentation du débit lorsque le volume [fichier] est défragmenté (c'est-à-dire non fragmenté). Il devient également évident que les augmentations des tailles d'E/S en écriture/lecture s'accompagnent d'une hausse considérable du temps de latence des E/S induit par la fragmentation. C'est au niveau des lectures de fichiers, aléatoires et séquentielles, qu'un fichier contigu produit les meilleures améliorations.

Conclusion

Il est évident que la fragmentation ralentit les performances des systèmes d'exploitation Windows invités.

Alors que les essais décrits ont été réalisés sur une seule machine virtuelle, le problème s'aggrave de manière exponentielle dans un environnement comptant plusieurs VM soumises chacune à une fragmentation des fichiers.

Parce que la virtualisation des serveurs instaure une relation symbiotique, il faut tenir compte du fait que la génération d'E/S disque sur une seule machine virtuelle affecte les demandes E/S adressées par les autres systèmes virtuels.

Les temps de latence sur une seule machine virtuelle vont donc accroître de manière artificielle ceux des machines virtuelles co-hébergées (se partageant une plate-forme commune).

La fragmentation accroît de manière artificielle la quantité de demandes d'E/S disque qui, sur une plate-forme de machine virtuelle, accentue bien plus la congestion du disque que sur des systèmes conventionnels.

L'élimination de la fragmentation sur les machines virtuelles et du trafic d'E/S disque inutile correspondant, s'avère essentielle pour les performances de la plate-forme et étend la capacité à accueillir davantage de machines virtuelles dans une infrastructure partagée.



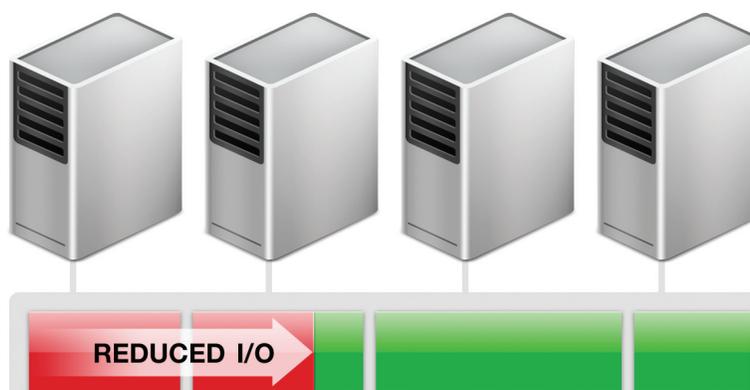
Protocoles de Meilleures Pratiques: Optimisation des systèmes sur VMware



Gérer les E/S des disques de manière optimum est un enjeu essentiel pour les Administrateurs de la sauvegarde et de la virtualisation en maintenant un service de grande qualité comme souhaité. En dépit des bénéfices innés en terme d'efficacité lors d'une migration depuis une sauvegarde physique à accès direct à un mode virtuel, se présentent alors de nouveaux défis et problèmes. Lors de l'apparition de goulots d'étranglement dans une infrastructure de sauvegarde partagée, toute la plateforme peut être affectée et pas simplement la source OS. L'optimisation des disques est devenue plus importante que jamais afin de prévenir les chutes du service et de maintenir une ressource disponible en permanence.

Dans un environnement virtuel, sensible à la ressource, la capacité à assurer des fichiers contigus et à éliminer le trafic d'E/S non-nécessaire est essentiel. Avec l'arrivée de la technologie IntelliWrite®, une solution unique a présent disponible au sein de V-locity® optimiseur de disque sur plateforme virtuelle, l'écriture de fichiers est optimisée afin que l'accès à la donnée puisse s'opérer avec un minimum de besoins en E/S.

“Défragmenter les disques est obligatoire pour solutionner les problèmes rencontrés avec un OS résultant d'une fragmentation du système de fichiers. Les problèmes de fragmentation ralentissent les performances de l'OS.”



– VMware KB article,
[Defragmenting a Disk](#)

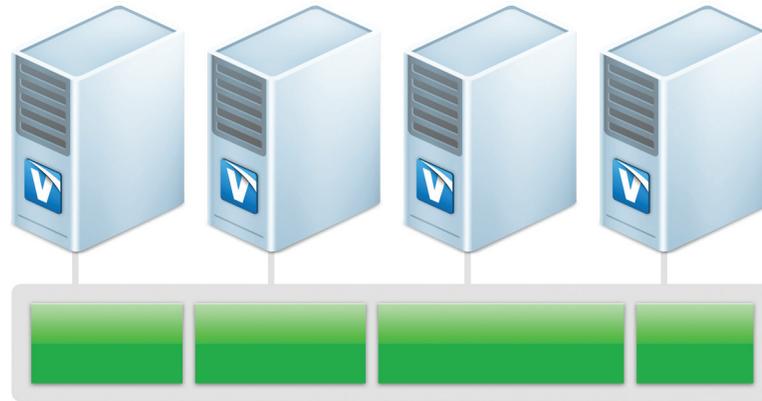
La distribution de la ressource dans une infrastructure de sauvegarde partagée peut varier d'un moment à l'autre et de part sa nature intrinsèque, un goulot d'étranglement à un endroit peut impacter la ressource disponible ailleurs.

V-locity fournit la seule solution hiérarchique afin de traiter le gachis d'E/S des disques à travers les multiples couches d'une infrastructure virtuelle, la ou les E/S doivent être aussi efficaces que possible. L'installation au niveau du guest n'assure pas seulement l'augmentation de l'efficacité des E/S à chaque couche. Les bénéfices se répandent sur chaque composant virtuel à travers toute la plateforme partagée sous la forme d'E/S économisées pour une meilleure allocation de la ressource.

Un guest virtuel est naturellement inconscient de sa position ou de l'allocation de la ressource dans une infrastructure partagée, menant à la possibilité de procédés sur un OS individuel affectant négativement l'infrastructure toute entière. La nouvelle technologie InvisiTasking® permet à V-locity de communiquer entre les couches, permettant une utilisation synchronisée de la ressource. Ainsi, chaque composante V-locity dans chaque OS a zéro impact sur n'importe quel autre hôte virtuel sur le système, améliorant la performance et permettant une densité des machines virtuelles plus importante.

V-locity est 100% compatible avec les disques fins / mous dynamiques, Continuous Data Protection systems (Snapshots), disques différentiels, Linked Clones, de même que les Storage Area Networks (SANs). La prévention de la fragmentation avec IntelliWrite garantit que l'optimisation de la plateforme s'opère sans conflits de ressources

Protocoles de Meilleures Pratiques: Optimisation des systèmes sur VMware



Les Guests de V-locity: Ils optimisent les E/S des disques, ce qui est indispensable au niveau virtuel en terme d'efficacité de la ressource, pas seulement individuellement mais globalement pour toute la plateforme

L'optimisation de l'OS Guest est une Meilleure Pratique recommandée par VMware pour des performances optimum: "Défragmenter les disques est important afin de gérer les problèmes rencontrés avec un OS en conséquence d'une fragmentation des fichiers. Les problèmes de fragmentation entraînent un ralentissement des performances de l'OS" – VMware KB article, Defragmenting a Disk

V-locity a été la première solution certifiée VMware-ready disponible sur le marché.

Démarrer très simplement.

L'installation est très simple:

- Déploiement de la composante Guest de V-locity Guest sur chaque Machine Virtuelle (VM)

La configuration est encore plus simple:

- C'est terminé. Une fois l'installation achevée, tous les composants se synchronisent immédiatement afin de maximiser l'efficacité de l'utilisation de la ressource et des E/S.
- Provisioning note: lors des changements de votre environnement, V-locity ajuste sa configuration de manière dynamique afin de garantir une compatibilité totale et des performances E/S optimales.

La maintenance est intuitive:

- Utiliser l'option de compaction des disques virtuels disponible dans V-locity, fournit un simple GUI montrant le total d'espace pouvant être récupéré, éliminant les séparations entre les blocs d'espace libre sur n'importe quel disque fixe / mou dynamique. Le processus de compaction commence en un clic.