Perte de données dans un environnement virtuel : un problème émergeant

Des solutions pour répondre aux impératifs de continuité des activités

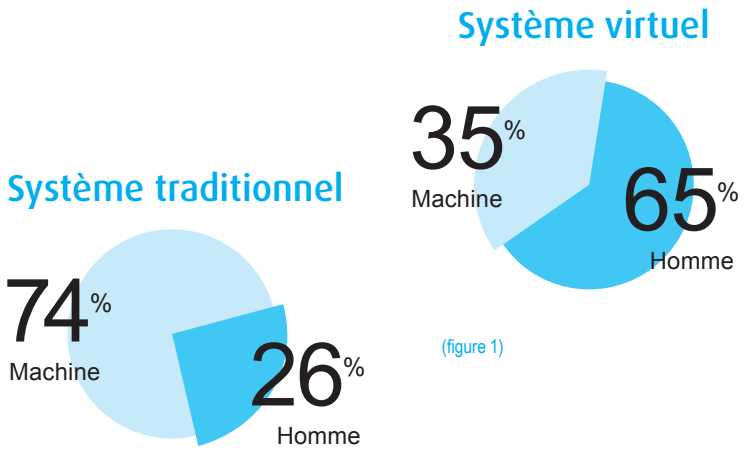
Ontrack®

Introduction

Les termes “continuité des activités” et “reprise après sinistre” ont été galvaudés au fil des ans, ce qui a créé de la confusion au sein des entreprises qui s’efforcent de protéger leurs opérations. Un plan de continuité des activités est une politique exhaustive qui garantit que tous les départements d’une entreprise peuvent fonctionner efficacement en cas d’événement perturbateur[1], avec des conséquences limitées ou minimes. Le plan de reprise après sinistre et les procédures de réponse d’urgence font généralement partie d’un plan plus vaste de continuité des activités.

L’avènement des technologies de virtualisation a permis d’accéder à la planification et la bonne exécution des plans de continuité des activités pour bon nombre d’entreprises. Toutefois, la virtualisation est complexe et nécessite des compétences et connaissances spécialisées aussi bien de la part des équipes informatiques que des dirigeants. En effet, si elle est déployée ou gérée de façon négligée, la virtualisation peut elle-même provoquer des interruptions d’activités ou des sinistres sur les données, comme le montrent les graphiques ci-dessous.

Si elle est déployée ou gérée de façon négligée, la virtualisation peut elle-même provoquer des perturbations des activités ou des sinistres sur les données.



(figure 1)

Source de défaillance : Homme vs Machine

- Erreur humaine,
- Manque de formation,
- Défaillance système ou matérielle,
- Environnement (coupure de courant, surtension...)

D’après le rapport de Forrester Research sur l’état de préparation des entreprises à la reprise après sinistre, établi en collaboration avec le ***Disaster Recovery Journal***[2], bon nombre d’entreprises ont amélioré leurs capacités de reprise après sinistre au cours des dernières années. Malgré le ralentissement de l’économie, les personnes interrogées montrent davantage de confiance quant à leur état de préparation à un sinistre au niveau du centre de traitements ou une défaillance du site.

76 % des personnes interrogées ont ainsi signalé n’avoir subi aucun sinistre ni interruption majeure au cours des cinq dernières années ; néanmoins, Forrester Research prévient que les entreprises ne doivent pas se reposer sur leurs lauriers au vu de ces statistiques. Au contraire, celles-ci doivent servir d’avertissement, dans la mesure où ce sont 25 % des entreprises qui sont susceptibles de déclarer un sinistre. En outre, les perturbations d’activités sont beaucoup plus courantes que les “sinistres déclarés”. Pour qu’une entreprise déclare un sinistre, ce peut être une question de point de vue, indique Don Stewart, directeur des services professionnels à Ongoing Operations, prestataire de services de continuité des activités sans but lucratif pour les caisses d’épargne américaines. “Dans certains cas, les équipes informatiques sont tellement concentrées

[1] Pour les besoins de cet article, une perturbation des activités représente tout ce qui empêche le travail quotidien d’être accompli, notamment une coupure d’électricité, un dérangement des lignes téléphoniques, etc. Un sinistre au niveau des données se produit lorsque des données sont altérées. Il s’agit donc d’un sous-ensemble de la perturbation des activités.

[2] Rapport de 2010 de Forrester Research sur l’état de préparation des entreprises à la reprise après sinistre, en collaboration avec le Disaster Recovery Journal, http://www.drj.com/images/surveys_pdf/forrester/2011Forrester_survey.pdf

sur la résolution du problème qu'elles n'informent pas les dirigeants du sinistre", raconte Don Stewart. Certaines entreprises n'ont pas défini ce qu'est une perturbation des activités, si bien que les dirigeants hésiteront à déclarer un sinistre si l'événement est perçu comme étant mineur ; par exemple, dans le cas de la défaillance du système téléphonique ou de retards dans la messagerie électronique.

15 % des personnes interrogées connaissaient le coût d'un temps d'arrêt de leurs activités : il s'élevait à environ 145 000 dollars par heure en moyenne.

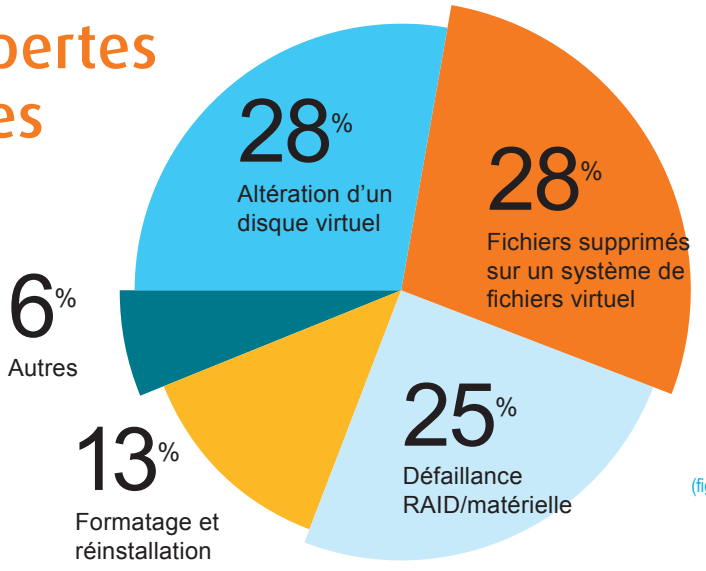
Pour être prêt, il ne suffit pas d'avoir un plan de continuité des activités documenté : cela demande également un travail d'équipe mobilisant toutes les parties prenantes. L'enjeu à ce niveau est de garantir que les opérations de l'entreprise seront maintenues en cas de perturbation. Pour Don Stewart, un plan efficace doit commencer par une analyse de l'impact du risque. De son point de vue, la plupart des entreprises commanderont une évaluation approfondie du risque, mais n'en feront rien : "le rapport reste dans un coin, sans qu'aucune action ultérieure ne soit menée". C'est aussi efficace que de dresser une liste des articles essentiels à regrouper dans un sac en cas d'incendie de sa maison, mais de ne jamais préparer ce sac.

Scénarios courants de perte de données virtuelles

Lorsqu'une perte de données se produit dans un centre de traitements virtuel, elle est généralement due à une erreur humaine. Les autres cas résultent d'une panne matérielle et sont exacerbés par l'absence de plan de reprise après sinistre. Les plans de reprise après sinistre qui sont médiocres ou ne font pas l'objet de tests réguliers obligent le personnel informatique à se concentrer sur des réparations non testées, basées sur un dépannage curatif. Bien évidemment, nul ne veut subir une perte de données ou une interruption d'activités sur les systèmes dont il a la charge. Trop souvent, une grave perte de données ou interruption entraîne un licenciement pour la personne responsable ou jugée responsable.

D'autres scénarios de perte de données sont dus à une trop grande confiance dans la redondance d'un réseau spécialisé de stockage (SAN). Un précieux temps de reprise est perdu lorsque les équipes découvrent au beau milieu d'un sinistre que d'importantes sauvegardes sont altérées ou illisibles. C'est le pire moment possible pour découvrir que les sauvegardes ont échoué ou que le logiciel de sauvegarde n'a pas signalé les erreurs des supports pendant les sessions de sauvegarde. Le diagramme de Kroll Ontrack ci-dessous illustre la répartition des types de pertes de données virtuelles en 2010 :

Types de pertes de données virtuelles



(figure 2)

Études de cas récents sur la perte de données virtuelles

Une perte de données de virtuelles peut être catastrophique pour une entreprise. Il est difficile de déterminer l'impact financier d'une interruption des activités, dans la mesure où elle implique à la fois des facteurs tangibles, comme la perte de productivité, des opportunités de vente ratées et le coût du personnel, et des facteurs moins tangibles liés au temps d'arrêt tels que des sanctions potentielles pour non-conformité, la dégradation de l'image de l'entreprise et la perte de confiance des clients. Selon l'étude Forrester/DRJ précédemment citée, 15 % des personnes interrogées connaissaient le coût d'un temps d'arrêt de leurs activités : il s'élevait à environ 145 000 dollars par heure en moyenne. Avec une telle estimation, tout directeur ou DSI devrait vérifier si son plan de continuité des activités est à jour. La virtualisation peut aggraver ces pertes, comme l'illustrent les cas suivants.

Le cas du serveur reformaté

Une entreprise en Italie a récemment subi une interruption de ses activités lorsque son serveur hôte virtuel de 4 To a perdu l'accès au système de stockage. L'environnement virtuel comportait 40 machines virtuelles dans un environnement de systèmes d'exploitation mixte : systèmes Linux, systèmes Unix hérités et serveurs Microsoft® Windows®. Ces systèmes prenaient en charge des serveurs d'applications, web et de bases de données.

Le serveur hôte virtuel exécutait un hyperviseur Linux auquel étaient rattachées deux unités logiques de 2 To. À un certain moment, les unités logiques de stockage ont été reformatées. La raison de ce reformatage n'a pas été dévoilée, mais les dommages aux structures existantes des systèmes de fichiers étaient graves et étendus. Au cours du processus de reformatage, le gestionnaire de stockage Linux a écrit des métadonnées de système de fichiers EXT dans des zones prédéfinies du volume. Ces métadonnées ne contenaient que quelques milliers d'octets d'informations, mais l'impact sur les fichiers de disques virtuels et le système de fichiers du serveur hôte virtuel a été dévastateur.

Chaque machine virtuelle comportait entre quatre et six fichiers de disque virtuel, soit un total de 70 à 90 fichiers de disque virtuel stockés par le serveur hôte. Certains des serveurs Microsoft Windows virtualisés employaient des configurations de "volume de disque dynamique" entre de multiples fichiers de disques virtuels, ce qui compliquait un peu plus les efforts de récupération.

Une fois que le département informatique de l'entreprise eut épuisé toutes ses ressources internes, une société spécialisée dans la récupération de données a été engagée pour récupérer les données. Malgré les dommages, les fichiers de disques virtuels ont été retrouvés et les données cruciales ont pu être restaurées.

Ce sont les meilleures pratiques de l'industrie combinées à des procédures de gestion informatique qui garantissent la protection des données.

Le cas d'une fusion de données désastreuse

La fusion de deux entreprises aux États-Unis a tourné au désastre lorsque les deux départements informatiques ont fusionné leurs données. Ce désastre est dû à un sabotage par des employés, dont la cause est en cours d'investigation par des enquêteurs spécialisés dans la recherche de preuves informatiques.

Le serveur hôte virtuel de la première entreprise contenait plus de 400 machines virtuelles sur 20 unités logiques de stockage. Au cours de la fusion des données, un individu disposant de droits d'accès administrateur au serveur hôte virtuel a systématiquement supprimé les 400 machines virtuelles et leurs fichiers de disques virtuels, provoquant la perte de plus de 440 disques virtuels et plus d'un millier d'instantanés.

L'entreprise réalisant la fusion a rapidement fait appel à un spécialiste de la récupération de données en urgence et a classé par priorité les serveurs centraux qui fournissaient des services essentiels. En trois jours, ces systèmes étaient de nouveau opérationnels. Pendant les deux semaines suivantes, les efforts de récupération en urgence se sont poursuivis sur le reste du système de stockage. Il a fallu déployer des efforts considérables pour rechercher dans les zones non allouées de l'unité logique de stockage les fichiers potentiels de disques virtuels, identifiables uniquement par leurs attributs du système de fichiers.

Par un effort combiné de restauration des sauvegardes et de récupération des volumes d'origine, les données ont pu être récupérées. La plupart des disques virtuels étaient complets, tandis que les autres disques virtuels ont nécessité que les contenus des fichiers soient extraits en raison de l'altération du système de fichiers.

Le cas du reformatage du SAN hors site

Les efforts de reprise après sinistre sont allés de mal en pis pour une entreprise du Luxembourg. Au cours d'une maintenance de routine sur son système de stockage SAN hébergeant ses machines virtuelles, le SAN a été présenté par accident sur un serveur physique différent. Lorsque le stockage SAN a été identifié comme "inconnu", le volume a été automatiquement reformaté. Dans un premier temps, certains membres du personnel ont paniqué en raison du risque de perte de données. Ils ont été soulagés lorsqu'on leur a rappelé l'existence du système de stockage SAN à l'identique, situé hors du site et qui employait la technologie de réplication automatisée du site intégrée au système SAN. Tous ont alors pensé qu'il s'agirait d'une perturbation mineure des activités.

Toutefois, en se connectant au SAN à distance, l'équipe informatique a découvert qu'il était une copie identique du site primaire ; la technologie de réplication automatisée du site du système SAN n'avait pas été désactivée avant la maintenance. En conséquence, lorsque le reformatage s'est produit sur le site primaire, le SAN secondaire a lui aussi été reformaté. Grâce aux efforts d'ingénieurs expérimentés dans la récupération de données, les machines virtuelles et les fichiers de disques virtuels ont pu être récupérés.

Cette entreprise n'avait aucune sauvegarde, car elle supposait que la double architecture de stockage et les mécanismes de réplication du site lui assuraient une redondance complète des données et systèmes. Ce cas est particulièrement intéressant du fait que les caractéristiques de l'équipement de stockage ont donné un faux sens de sécurité. En réalité, ce sont les meilleures pratiques de l'industrie combinées à des procédures de gestion informatique qui garantissent la protection des données.

Conséquences d'une perte de données

La virtualisation a révolutionné l'industrie informatique et a concrétisé la promesse d'une réduction des dépenses d'installation et des coûts d'équipement. D'après le suivi mondial d'IDC sur les systèmes de stockage sur disques externes, la capacité totale de stockage sur disques dépassait les 5 100 pétaoctets, soit une hausse de 55,7 % par rapport à l'année précédente[3]. Cette croissance continue nécessite une bonne gestion informatique pour tenir à jour la documentation relative à la reprise après sinistre et tester régulièrement les plans de récupération. Cette pratique minimisera voire éliminera les interruptions d'activités dues à la perte de données au sein d'environnements virtualisés.

À mesure que la virtualisation consomme de plus en plus de stockage, une attention supplémentaire doit être portée à la gestion et la protection des ressources virtuelles. Il est essentiel de préserver la continuité des activités en ayant des plans de reprise après sinistre bien planifiés et testés.

Les entreprises performantes ont conscience que toute perturbation dans l'infrastructure, aussi infime soit-elle, aura un impact amplifié sur l'entreprise dans son ensemble. Cela a conduit les responsables informatiques et les planificateurs de la continuité des activités à inclure de façon proactive des services de récupération de données dans leurs plans de secours. En choisissant un prestataire de services de récupération de données avant qu'un sinistre ne se produise, l'équipe informatique est prête à affronter avec succès une interruption des activités provoquée par un sinistre sur les données.

[3] "Les systèmes mondiaux de stockage sur disque terminent l'année 2010 avec une croissance à deux chiffres, sur des résultats probants au quatrième trimestre", IDC, mars 2011

Ontrack®

France : 0 800 10 12 13
www.ontrack.fr

Belgique : +32 (0)2 512 30 22
www.ontrackdatarecovery.be