

Livre blanc

Moderniser la protection des données pour plus de performance et de simplicité

Sponsorisé par : Pure Storage

Phil Goodwin
Octobre 2018

INTRODUCTION

La transformation numérique change profondément la façon dont les entreprises perçoivent et gèrent leurs données. Les recherches d'IDC indiquent que 60 % des entreprises se seront investies dans des initiatives de transformation numérique d'ici 2020. Pour nombre d'entre elles, l'objectif consiste à devenir une entreprise centrée sur les données. Les entreprises centrées sur les données sont celles capables de générer des revenus et de créer un avantage concurrentiel en utilisant les données. Mais avant tout, cette transformation numérique requiert que les données soient précises et disponibles pour celui qui les consomme. Ainsi, nous constatons que pour beaucoup d'entreprises, les mesures en nombre d'heures associées aux notions traditionnelles d'objectif de point de rétablissement (RPO) et d'objectif de délai de rétablissement (RTO) ne sont pas acceptables. Les entreprises s'efforcent d'assurer des services de données sans perte de données et ni de temps d'arrêt. Nous pensons qu'avec le temps, les entreprises considéreront que leur principale tâche sera d'assurer la disponibilité des données en y intégrant une composante liée à la protection de ces dernières. La protection des données (c.-à-d., les sauvegardes et les récupérations) n'impliquera pas de tâches distinctes, mais des tâches qui feront partie du plan de disponibilité des données.

Pour de nombreuses entreprises, l'appliance dédiée à la sauvegarde (PBBA ou Purpose-Built Backup Appliance) incarne l'épine dorsale des plans de protection des données. Les PBBA ont été développées pour répondre au besoin de sauvegardes directes sur disque plutôt que sur bande magnétique. Ces systèmes offrent généralement des fonctionnalités de déduplication, ce qui les rend particulièrement efficaces pour les sauvegardes, mais ils exigent que les données soient « réhydratées » pour pouvoir être utilisées. Le processus de « réhydratation » associé à la récupération des données est susceptible d'induire des délais considérablement plus longs que ceux nécessaires aux applications de sauvegarde. Les PBBA ont tendance à être extrêmement efficaces pour la rationalisation des sauvegardes et la réduction du degré de dépendance à l'égard des bandes magnétiques, mais elles n'ont pas été conçues pour faciliter d'autres types d'applications.

En outre, les entreprises centrées sur les données sont en train de réaliser que leurs données de sauvegarde ont une valeur considérable. Ces données ne sont plus destinées à dormir en guise de simple copie supplémentaire. Elles peuvent être utilisées pour des analyses, des tests/développements, du stockage tampon, des procédures d'e-Discovery ou à d'autres fins - à condition de pouvoir les rendre disponibles. Nous qualifions ces cas d'utilisation de secondaires dans la mesure où ils emploient la copie secondaire des données pour des applications supplémentaires. Bien que les PBBA continueront d'apporter de la valeur ajoutée pour les sauvegardes/récupérations,

nous constatons que de nombreuses entreprises se tournent vers les systèmes de stockage secondaire. Le stockage secondaire offre les avantages des PBBA en tant que cible pour la protection des données, mais il libère également la valeur des données en les rendant disponibles pour d'autres applications. Nous pensons que les entreprises tireront parti du stockage secondaire au moins pour les cas d'utilisation suivants :

- Protection des données grâce à une récupération rapide des données et des machines virtuelles (VM)
- Tests/développements
- Analyses
- Intelligence artificielle

CE QUE VOUS TROUVEREZ DANS CE LIVRE BLANC

Ce livre blanc analyse les cas d'utilisation pour les systèmes de stockage secondaire, ainsi que les avantages de ces systèmes par rapport à l'architecture de stockage traditionnelle. En outre, ce document examine la solution FlashBlade de Pure Storage, un nouveau produit de stockage de fichiers et d'objets avec des capacités d'évolution horizontale (scale-out) construit à l'aide d'une architecture de hub de données permettant à une seule plateforme de consolider de nombreux cas d'utilisation secondaires. Dans ce document, nous utilisons indifféremment les termes *systèmes de stockage secondaire* et *hub de données*.

VUE D'ENSEMBLE DE LA SITUATION

La protection traditionnelle des données organisée en silos devient obsolète alors que les départements informatiques cherchent des moyens de protéger leurs données tout en utilisant les données de sauvegarde pour des usages secondaires. Par conséquent, nous voyons apparaître des systèmes de stockage secondaire prévus pour répondre à la fois aux cas d'utilisation de protection des données et à ceux du stockage secondaire. Dans de nombreux cas de figure où les données doivent être sollicitées, le stockage secondaire flash peut offrir un meilleur profil de coût, ainsi que la qualité de service (QoS) traditionnellement associée à la technologie flash.

Le stockage secondaire suscite rapidement de l'intérêt en raison de sa capacité à résoudre simultanément plusieurs problèmes de gestion des données. La vision traditionnelle des activités du datacenter repose sur une séparation des tâches liées aux sauvegardes/récupérations, aux tests/développements, au stockage tampon et aux analyses de données, qui sont généralement déployées à l'aide d'infrastructures, de personnel et de processus métiers distincts. Pour cette raison, les PBBA - placées sous la responsabilité des équipes chargées des sauvegardes - n'ont servi qu'à répondre aux besoins des sauvegardes de données. Lorsque les administrateurs de bases de données ou les développeurs d'applications ont eux-mêmes besoin d'ensembles de données, un processus d'allocation des ressources complètement séparé, qui dure souvent plusieurs jours ou plusieurs semaines, est nécessaire pour mettre à disposition le stockage et effectuer la copie des données. Cette copie de données est généralement réalisée sur des systèmes plus lents et plus anciens.

Les systèmes modernes de stockage secondaire sont susceptibles d'apporter plusieurs avantages au-delà des architectures traditionnelles et peuvent notamment :

- Servir de cible pour des sauvegardes rapides ;
- Restaurer rapidement les données dans des situations de reprise après sinistre ;
- Être exploités simultanément pour plusieurs cas d'utilisation en mettant des copies de données à la disposition de multiples applications ;
- Être adaptés aux services de données non structurées basées sur des fichiers ;
- Fournir une plus grande qualité de service en répondant de la même façon aux requêtes d'accès séquentiel et d'accès aléatoire aux données ;
- Accélérer la mise à disposition des applications en réduisant la durée du processus de l'entreprise nécessaire pour la réalisation des copies de données qui peut s'étendre sur plusieurs jours ou plusieurs semaines grâce à un système à la demande en libre-service ;
- Améliorer la qualité des copies de données en faisant en sorte que les données mises à jour soient immédiatement disponibles pour les tests/développements ou les analyses ;
- Réduire les coûts en consolidant les systèmes de stockage dédiés et réservés à seul usage ;
- Apporter plus de simplicité en réduisant le nombre de systèmes devant être déployés et gérés ; et
- Apporter des capacités de réplication au cloud.

Il est important de noter que le terme « stockage secondaire » n'implique aucunement un niveau de performance également secondaire. Les fournisseurs intègrent de plus en plus le stockage flash à leurs baies de stockage secondaire afin de fournir la qualité de service exigée par diverses applications utilisant le stockage. Dans certains cas, le stockage secondaire est étroitement lié au stockage principal pour permettre une gestion harmonieuse, une mise à disposition rapide et une récupération sûre des données.

Nous pensons qu'une majorité d'applications seront dotées d'une composante cloud d'ici 2020, y compris celles actuellement déployées sur site ; l'architecture la plus courante sera celle du cloud hybride. Pour de nombreuses applications, la principale composante cloud peut être liée à la protection des données. Dans ces cas de figure, les entreprises utilisent le cloud pour le stockage hors site en tant que moyen de protection des données pour les reprises après sinistre ou l'archivage à long terme. Ce transfert des données vers le cloud est davantage considéré comme une composante du processus de déplacement des données de l'entreprise vers le cloud qu'une fonction distincte. À ce titre, les responsables informatiques s'attendent à ce qu'il soit automatisé et s'effectue de façon fluide.

Outre l'environnement de cloud hybride décrit précédemment, certaines entreprises préféreront que les données soient copiées sur plusieurs clouds au sein d'une architecture multicloud. Il pourra s'agir d'ajouter un niveau de protection, d'éviter l'aliénation à un fournisseur ou de rendre les données disponibles pour les applications dans différents clouds. Ainsi, les systèmes de stockage secondaire doivent être capables de fournir leurs services simultanément à plusieurs clouds publics.

ENVISAGER UN HUB DE DONNÉES DE PURE STORAGE

Pure Storage est reconnue dans l'industrie pour ses systèmes de stockage principal 100 % flash extrêmement performants. Récemment, l'entreprise est arrivée sur le marché du stockage non structuré avec sa solution FlashBlade. FlashBlade est une solution 100 % flash optimisée pour les services de stockage de fichiers et d'objets, et elle est construite comme un hub de données pour pouvoir offrir des niveaux de performance élevés à de nombreux cas d'utilisation secondaires au moyen d'une plateforme unique.

Flashblade a été conçue pour contourner le besoin de systèmes dédiés à une seule utilisation tout en prenant en charge des services de données garantissant une qualité de service équivalente à celle du stockage principal. En général, les responsables informatiques doivent faire un compromis entre le stockage hautement performant à coût élevé et le stockage à moindre coût, mais peu performant. La vocation de FlashBlade est d'éviter un tel compromis.

En tant que hub de données, Flashblade a été conçue pour répondre au besoin de plusieurs cas d'utilisation, notamment :

- **La protection des données et les restaurations rapides.** Les données peuvent être copiées depuis une baie Pure Storage de stockage principal vers la plateforme FlashBlade. Compte tenu de l'architecture à évolutivité horizontale 100 % flash de Flashblade, les données et les images de machine virtuelle peuvent être restaurées dans des délais considérablement réduits comparativement aux baies constituées de disques durs. Tandis que certains systèmes dédiés sont optimisés pour les sauvegardes, FlashBlade est optimisée pour les restaurations. Du point de vue de la fourniture des niveaux de service, il en découle une amélioration du RPO et du RTO grâce à des copies de données plus fréquentes et à des restaurations plus rapides.
- **Les tests/développements.** Les développeurs d'applications peuvent avoir accès à la copie secondaire des données pour effectuer des tests par rapport aux données courantes, ce qui évite de devoir copier les données vers un système de stockage plus ancien dédié aux applications de tests/développements. L'accès est en libre-service, ce qui permet de réduire les délais et d'alléger les processus métiers nécessaires pour allouer les ressources et créer une copie. Cette réplication peut inclure le développement d'une sandbox ou d'autres environnements où un accès fréquent aux données courantes est nécessaire.
- **Les NAS et le stockage d'objet.** Flashblade est spécifiquement conçue pour optimiser les services de données basées sur des fichiers et de stockage d'objets, y compris les conteneurs.
- **Les analyses de données.** Les applications analytiques, les Big Data, etc. ont directement accès à l'entrepôt de données secondaires. Les données sont immédiatement accessibles dès leur mise à jour depuis le stockage principal.
- **La réplication cloud.** La plateforme FlashBlade peut être intégrée aux architectures cloud hybride et multicloud en répliquant les données dans différents entrepôts du cloud public.
- **La rentabilité.** Les exigences de Flashblade en termes d'espace et sur le plan environnemental sont nettement moindres que celles des systèmes traditionnels à base de disques durs, ce qui permet de réduire les coûts des déploiements et de la maintenance.

FlashBlade s'appuie sur une conception unique permettant d'offrir des performances équivalentes à celles du stockage flash principal au même niveau de prix et avec les mêmes capacités d'extension que les baies à évolutivité horizontale. Pure Storage a également opté pour une nouvelle orientation intéressante concernant son architecture pour la protection des données. Tandis que de nombreux systèmes cherchent à réduire la fenêtre de sauvegarde et à optimiser le stockage dans une optique de sauvegarde, FlashBlade est conçue pour optimiser la restauration des données, un domaine pour lequel les entreprises ont vraiment besoin des meilleures performances.

DÉFIS ET OPPORTUNITÉS

Au cours des enquêtes conduites par IDC, les entreprises ayant choisi de ne pas utiliser la technologie flash ou de mettre fin à son utilisation pour le stockage ont principalement évoqué les coûts, l'absence d'avantages perçus et les problèmes de fiabilité pour justifier leur décision. Bien que les enquêtes en cours indiquent que ces problèmes ont été atténués, Pure Storage et les autres fournisseurs de ce type de produits doivent encore communiquer clairement sur les avantages procurés par leurs produits de stockage flash. En tant que nouvel entrant sur le marché du stockage secondaire, la plateforme FlashBlade aura besoin de faire ses preuves quant à sa capacité à offrir des gains de performances et d'efficacité dans une optique de réduction des coûts.

Pourtant, les volumes de données ne cessent de croître. Quelles soient structurées ou semi-structurées, les données constituent le nouveau fondement sur lequel repose la notion d'avantage concurrentiel. Le stockage fait partie intégrante de la stratégie des données de l'entreprise dans la mesure où il participe activement au processus de réduction des coûts et de la complexité associés aux immenses volumes de données. Mais surtout, la disponibilité des données est essentielle dans le cadre de la transformation numérique et pour évoluer vers le statut de l'entreprise fondée sur les données. Par conséquent, le besoin d'une protection des données intégrée à divers modèles de déploiements des applications conduit les départements informatiques à acquérir les meilleures solutions de protection des données de leur catégorie pour des environnements spécifiques.

CONCLUSION

Les entreprises centrées sur les données s'éloignent des systèmes de sauvegarde/récupération cloisonnés et réalisent que la récupération rapide et la sollicitation des données constituent des objectifs clés. Les entreprises réalisent également que l'achat d'appliances dédiées à un usage unique, dont le rôle consiste simplement à héberger des données dans l'attente d'un incident éventuel, conduit à une consommation inefficace de ressources. À l'inverse, les systèmes susceptibles d'aider simultanément l'entreprise à protéger et à utiliser leurs données apportent plus de valeur ajoutée par rapport aux sommes investies.

Le stockage secondaire est un domaine qui retient de plus en plus l'attention et suscite un intérêt croissant sur le marché compte tenu de sa capacité à offrir à la fois des fonctionnalités de protection des données et d'utilisation secondaire. Néanmoins, certains systèmes utilisent des disques durs plus lents pour les systèmes secondaires. Ces systèmes occupent plus d'espace, ils consomment plus d'énergie et de ressources de refroidissement que les systèmes 100 % flash, et ils ne sont pas en mesure d'apporter les garanties de qualité de service associées à ces derniers. Nous pensons que les entreprises centrées sur les données seront moins disposées à accepter des compromis en termes de performance pour les cas d'utilisation secondaires. Nous estimons également que Pure Storage a positionné sa solution FlashBlade avec pertinence pour tirer parti de cette tendance et que l'entreprise se situe à l'avant-garde des systèmes de stockage secondaire basés sur la technologie flash.

À propos d'IDC

International Data Corporation (IDC) est le premier fournisseur mondial d'informations sur les marchés commerciaux, de services de conseil, et d'événements sur la technologie de l'information, les télécommunications et les marchés des technologies pour le grand public. IDC aide les professionnels de l'informatique, les dirigeants d'entreprise et la communauté des investisseurs à prendre des décisions qui se fondent sur des faits pour les acquisitions technologiques et la stratégie de l'entreprise. Plus de 1 100 analystes d'IDC apportent une expertise mondiale, régionale et locale sur les questions d'opportunités technologiques et sectorielles, ainsi que sur les tendances qui se dégagent dans plus de 110 pays à travers le monde. Depuis 50 ans, IDC fournit des informations approfondies stratégiques afin d'aider ses clients à atteindre leurs objectifs commerciaux clés. IDC est une filiale d'IDG, la principale société en matière de médias, de recherche et d'événements liés à la technologie.

Siège social mondial :

5 Speen Street
Framingham, MA 01701
États-Unis
+1.508.872.8200
Twitter : @IDC
idc-community.com
www.idc.com

Avis de copyright

Publication externe des données et information d'IDC – toute information d'IDC destinée à être utilisée dans le cadre de publicités, de communiqués de presse ou de supports promotionnels doit préalablement faire l'objet du consentement écrit du vice-président ou du directeur national d'IDC concerné. Un projet du document proposé doit accompagner une telle demande. IDC se réserve le droit de refuser l'approbation de toute utilisation externe, quelle qu'en soit la raison.

Copyright 2018 IDC. Toute reproduction sans autorisation écrite est strictement interdite.

