

BONNES PRATIQUES POUR BIEN DÉMARRER AVEC LE STOCKAGE LOGICIEL

SYNTHÈSE

Aujourd'hui, pour de nombreux spécialistes en informatique, les « Big Data » ou gros volumes de données sont bien plus qu'une simple expression à la mode. Il est impossible de les ignorer, car ils sont à l'origine d'une véritable crise dont la raison est toute simple : les gros volumes de données ne cessent de s'élargir. Dans beaucoup de cas, surtout dans les secteurs très gourmands en données (la distribution, par exemple), les entreprises qui resteront dans le peloton de tête au cours des prochaines années seront celles qui auront trouvé un moyen rentable de valoriser les énormes quantités d'informations dont elles disposent.

Heureusement, il existe des solutions concrètes pour atteindre cet objectif. Les entreprises leaders se tournent vers le stockage logiciel ouvert, car il s'agit d'une méthode viable pour déployer des architectures informatiques à l'échelle du Web. Elle leur laisse une plus grande liberté dans la gestion de toutes sortes de données, même à très grande échelle. Comme c'est souvent le cas avec les nouvelles technologies, les définitions de cette approche novatrice sont variables. La plupart des experts semblent cependant s'accorder sur les caractéristiques des environnements de stockage logiciel : indépendance par rapport au matériel, architectures distribuées, convergence du stockage et prise en charge native des protocoles de données standard.

Mais par où commencer ? Nous proposons dans ce livre blanc un guide détaillé pour l'adoption d'un système de stockage logiciel dans le cadre d'une stratégie pratique et proactive de gestion des données d'entreprise. Le processus variera toujours d'une entreprise à l'autre et d'une application à l'autre. La flexibilité est inhérente à cette technologie. D'ailleurs, une architecture informatique flexible représente peut-être plus que jamais la solution aux problèmes autrefois inconcevables posés par les gros volumes de données, sans oublier les problèmes futurs que nous ne soupçonnons pas encore.

PROBLÈME : DES QUANTITÉS GIGANTESQUES DE DONNÉES

La croissance sans précédent des volumes de données est l'obstacle le plus important auquel de nombreux services informatiques sont confrontés, quel que soit le secteur d'activité. Et les difficultés ne font que s'accumuler. Les entreprises de toutes tailles peinent à enregistrer et à analyser les énormes quantités d'informations qu'elles tirent d'un nombre de plus en plus important de sources. Elles s'adaptent tant bien que mal à des augmentations du volume des données qui leur auraient semblé inimaginables il y a dix ans à peine. Dans ces circonstances, on ne peut guère s'étonner des prévisions des analystes du cabinet de recherche informatique Gartner, selon lesquelles, dans 40 % des entreprises, la taille de l'infrastructure de stockage sur site aura doublé d'ici 2016.¹

Cependant, malgré l'explosion des coûts, les budgets informatiques de nombreuses entreprises continuent de fondre et les équipes internes se voient dans l'incapacité d'innover en matière de services, ce qui est pourtant essentiel à la survie sur un marché mondial encombré. Alors que l'inflation des données promet de devenir encore plus ingérable au cours des prochaines années, certaines entreprises choisiront peut-être de rester fidèles aux solutions existantes, offrant des performances de plus en plus limitées. Ces entreprises ne sont pas celles qui réussiront à long terme. Toute société désireuse d'entrer véritablement dans l'ère des données et de se positionner parmi les leaders du secteur, d'ici 2020 et



facebook.com/redhatinc
@redhatnews

linkedin.com/company/red-hat

fr.redhat.com

¹ « Making the hybrid cloud storage work » (Réussir la mise en place d'un système de stockage dans le cloud hybride), Gartner, 2014.

RED HAT ET CEPH : L'UNION FAIT LA FORCE

En mai 2014, Red Hat a racheté Inktank, éditeur de Ceph et d'Inktank Ceph Enterprise. Ensemble, ces technologies leaders permettent de stocker et de gérer une variété sans précédent de données, allant des données stratégiques actives aux données d'archive inactives.

LA MÉTHODE RED HAT : L'OPEN SOURCE

Chez Red Hat, lorsque nous utilisons le terme « Open Source », nous faisons référence à une méthode de collaboration éprouvée pour la création de technologies, qui laisse à tous la liberté de consulter le code, de se former, de poser des questions et de suggérer des améliorations. Nous préférons travailler et participer au sein de communautés régies par le principe strict de la méritocratie, où les procédures sont ouvertes et où tous disposent des mêmes informations et sont libres d'apporter des améliorations. Dans cet univers communautaire, ce sont toujours les meilleures idées qui l'emportent.

Pour nous, l'Open Source n'est pas vraiment un mouvement doté d'un manifeste, mais tout simplement la meilleure méthode qui existe. Nous pensons également qu'elle est capable de procurer un maximum d'agilité aux entreprises d'aujourd'hui.

au-delà, devra investir massivement dans une infrastructure de stockage modernisée. Sur le court terme, le prix à payer risque donc d'être considérable, mais qu'on ne s'y trompe pas : le prix de l'inaction serait sans aucun doute beaucoup plus élevé.

SOLUTIONS TECHNOLOGIQUES : UN DATACENTER PLUS AGILE

Lorsqu'une technologie se heurte à ses propres limites, il est probablement temps d'en créer une meilleure. Au lieu de solutions anciennes et peu flexibles, basées sur des scénarios obsolètes, les entreprises ont aujourd'hui le choix entre de nombreuses architectures informatiques déployées à l'échelle du Web pour faire le lien entre cloud privé et cloud public. Du point de vue du stockage, cela implique de séparer le matériel de stockage du logiciel qui en assure la gestion. Cette approche est connue sous l'appellation de « stockage logiciel ». Si son adoption reste limitée à l'heure actuelle, il ne faut pas s'attendre à ce que cela soit toujours le cas. Gartner estime en effet que le stockage Open Source représentera d'ici 2018 plus de 20 % du marché du stockage d'entreprise.²

S'il semble que personne ne soit d'accord sur la définition exacte du terme « stockage logiciel », cela n'a rien d'étonnant puisqu'il s'agit d'une technologie émergente. D'ailleurs, les définitions proposées par les fournisseurs pour se mettre en valeur commencent déjà à céder le pas à un ensemble de caractéristiques plus largement acceptées :

Indépendance vis-à-vis du matériel : une solution de stockage logiciel doit pouvoir s'exécuter sur n'importe quelle plateforme de serveur standard et ne pas dépendre de telle ou telle plateforme matérielle. Cela explique, en partie du moins, l'émergence des serveurs de stockage x86 à disques en attachement direct, conséquence évidente des tendances que sont le cloud et l'informatique à l'échelle du Web.

Architecture distribuée : cette caractéristique est essentielle, car l'un des principaux objectifs du stockage logiciel est de dépasser les limites inhérentes aux architectures évolutives plus traditionnelles (serveurs de stockage en réseau et SAN).

Prise en charge des protocoles de données standard : dans notre secteur, on s'accorde de plus en plus à penser que toute solution complète doit inclure la prise en charge de services de données en modes bloc, fichier et objet.

Convergence des ressources de calcul et de stockage : alors qu'un nombre croissant d'entreprises se tournent vers l'informatique unifiée, les solutions de stockage logiciel doivent pouvoir exécuter des charges de travail applicatives sur des nœuds de stockage.

Plan de contrôle de gestion : comme pour toute technologie chargée d'acheminer ou de transférer des informations sur un réseau complexe, la réussite des solutions de stockage logiciel dépendra du niveau de sophistication du plan de contrôle permettant de rationaliser et de simplifier l'accès aux données.

AVANTAGE CONCURRENTIEL : HUIT ÉTAPES POUR BIEN DÉMARRER

L'évolution vers le stockage logiciel ne suit pas une voie toute tracée. Elle est flexible de nature et, dans la pratique, souvent répétitive. Lorsque vous lirez les étapes suivantes, essayez d'imaginer un processus non pas linéaire, mais circulaire : commencez par vous concentrer sur un seul ensemble d'applications, puis parcourez chaque étape l'une après l'autre. Une fois ces étapes terminées, passez à l'ensemble d'applications suivant, et ainsi de suite.

- 1. Définissez vos critères essentiels** : quels sont actuellement vos coûts de stockage ? Quels sont vos besoins et vos contraintes du point de vue de la capacité ? Veillez à tenir compte du coût, ainsi que des critères de flexibilité, de disponibilité et d'agilité.
- 2. Définissez vos charges de travail en fonction de ces critères** : certaines des charges de travail sont-elles basées sur des données non structurées ? Si oui, il se peut que le stockage logiciel soit un excellent choix. Si, en revanche, vous gérez des charges de travail relativement peu volumineuses et basées sur des données structurées, le stockage logiciel n'est peut-être pas la bonne solution pour vous.

² « The five-year storage scenario—Why storage in 2019 won't look like what's on the floor today » (Le stockage à l'horizon 2019 : pourquoi la situation va changer), Gartner, 2014.

EXEMPLES CONCRETS : ENTREPRISES POUR LESQUELLES LE STOCKAGE LOGICIEL EST UNE RÉUSSITE

Metro de Madrid : le métro de Madrid, dont le réseau est le huitième plus grand au monde, possédait un système de stockage réparti entre deux centres de contrôle distincts. Malheureusement, ceux-ci n'ayant pas la même adresse réseau, le personnel devait souvent employer des méthodes chronophages pour parvenir à satisfaire les besoins en matière de stockage. En ayant recours à Red Hat® Storage, le métro de Madrid a réussi à optimiser l'utilisation des ressources existantes et, au final, à résoudre ses problèmes techniques, en dépit de contraintes budgétaires considérables.

Intuit : ce géant du logiciel était « enchaîné » à de grands systèmes propriétaires dont dépendait la compatibilité de tous les composants de stockage. Cette situation faisait qu'il était difficile d'atteindre des niveaux de performance suffisants en période de pointe et limitait la capacité d'Intuit à se protéger des pannes, qu'elles soient localisées ou qu'elles affectent le site entier, avec une architecture adaptée à la réplication. L'entreprise avait donc besoin d'une solution facile à mettre à l'échelle pour s'adapter au rythme toujours plus rapide de la croissance. En choisissant Red Hat Storage, c'est exactement ce qu'elle a obtenu, et à un coût bien inférieur à celui des systèmes de stockage propriétaires.

3. **Déterminez le nombre d'applications concernées :** si vous utilisez surtout une seule application monolithique, ce n'est pas, en règle générale, une utilisation adaptée au stockage logiciel. Si vous utilisez au contraire plusieurs applications variées, le stockage logiciel pourrait vous être très utile.
4. **Migrez d'abord les charges de travail non critiques ou les nouvelles applications :** pour vous habituer au processus et à votre nouvelle plateforme de stockage, et au cas où vous subiriez des temps d'arrêt, commencez par y transférer quelques-unes de vos applications non critiques. Ne vous inquiétez pas si vous rencontrez des difficultés au début : les premières migrations sont toujours les plus problématiques. Une fois que vous aurez affiné et normalisé votre stratégie et vos processus de migration, les transferts suivants, même pour les applications les plus critiques, seront beaucoup plus faciles.
5. **Choisissez entre des charges de travail virtualisées ou dans le cloud :** comptez-vous exécuter vos applications sur des serveurs physiques, dans un environnement virtualisé ou dans le cloud public ? Si vous prévoyez d'utiliser plusieurs de ces modèles de déploiement, vous avez tout intérêt à choisir une plateforme de stockage flexible qui les prenne en charge. Vous éviterez ainsi les complications inutiles liées à la gestion d'îlots de stockage séparés, basés sur des technologies incompatibles.
6. **Déterminez le type d'analyse dont vous avez besoin :** si la progression des gros volumes de données est un problème particulièrement pressant pour votre entreprise, envisagez d'utiliser une solution de stockage logiciel basée sur des technologies éprouvées d'analyse des données, par exemple la version Apache Hadoop de MapReduce. Avec ces outils, vous pourrez stocker et partager directement des données provenant d'applications existantes, sans avoir à transférer d'informations entre les différents silos. Cela vous permettra d'extraire plus rapidement de réservoirs géants de données des informations qui vous seront également plus utiles. D'autre part, vous aurez la possibilité de réduire considérablement vos coûts informatiques en valorisant au maximum l'infrastructure existante.
7. **Déterminez le niveau de protection et de réplication des données adapté à vos besoins :** quels sont les scénarios de reprise d'activité qui doivent être couverts ? Quelle est l'issue probable de chacun des scénarios, notamment du point de vue du coût ? Est-il possible de protéger les données les plus importantes de manière optimale sans pour autant surinvestir dans la sécurité des informations qui ne sont ni sensibles, ni critiques ?
8. **Déterminez la durée de conservation de vos données :** existe-t-il des réglementations qui vous obligent à conserver vos données pour une durée supérieure à celle que vous choisiriez vous-même ? Si ce n'est pas le cas, cherchez un moyen d'effacer les données dont vous n'avez plus besoin. En effet, pourquoi les stocker si elles ne sont pas nécessaires ?

PRIORITÉ AUX SERVICES DE DONNÉES : LA GESTION DES CHARGES DE TRAVAIL INDÉPENDAMMENT DE L'ÉCHELLE

Le stockage logiciel, même dans sa définition la plus étroite, peut revêtir différentes formes en fonction de l'environnement de l'entreprise, de ses besoins métier, de son budget, etc. Cela dit, pour une approche vraiment évolutive du stockage logiciel, les services de données doivent être détachés de la structure de données sous-jacente, ce qui permet d'utiliser des services plus nombreux sur des charges de travail plus diverses. Il s'agit notamment des services suivants :

- **Services de fichiers :** les systèmes de fichiers classiques pâtissent d'une évolutivité limitée en raison de leur structure hiérarchique, ce qui complique les tâches comme la protection des données ou l'optimisation de la capacité.
- **Services d'objets :** les systèmes de stockage d'objets représentent désormais une alternative très répandue aux systèmes en mode fichier et bloc. Ils produisent souvent suffisamment de métadonnées, tout en simplifiant l'indexation des informations non structurées. Le service fourni est basé sur l'architecture REST (REpresentational State Transfer) en tant qu'élément fonctionnel lors de l'abstraction du stockage en dehors de l'application. Cependant, les systèmes en mode objet sont souvent incompatibles avec les systèmes en mode fichier, ce qui oblige les entreprises à passer par un processus coûteux de réécriture des applications.

LIVRE BLANC Bonnes pratiques pour une solution Open Source aux difficultés posées par les gros volumes de données

Casio : cette multinationale spécialisée dans l'électronique, riche d'une longue tradition en matière d'innovation, s'était donné pour mission de faire du stockage à l'échelle du Web une réalité. Cependant, malgré ses efforts pour virtualiser son environnement de stockage, elle ne parvenait pas à s'affranchir de la dépendance qui la liait à un fournisseur particulier. Elle a fini par se tourner vers Red Hat Storage pour résoudre les difficultés liées au stockage, ce qui lui a permis d'intégrer les disques internes de plusieurs serveurs standard et de les utiliser tel un seul pool de stockage.

Cisco : les solutions réseau de cette entreprise sont pensées pour relier, de façon intelligente, les individus, les processus, les données et les objets. Cisco avait besoin d'une plateforme de stockage hautes performances qui prenne en charge au sein d'une même solution unifiée les principaux cas d'utilisation liés au stockage de blocs, d'objets et de fichiers.



facebook.com/redhatinc
@redhatnews

linkedin.com/company/red-hat

fr.redhat.com
INC0204410_V2_0115

- **Services partagés en mode fichier et objet :** ces services permettent aux entreprises de continuer à tirer parti des applications basées sur fichiers déjà utilisées. En même temps, ils mettent les données à la disposition des applications en mode objet, souvent par une méthode de type REST. Cette approche introduit un maximum de souplesse au niveau des normes.
- **Services en mode bloc :** le stockage en mode bloc, souvent utilisé par les SAN, gère les données sous forme de blocs, au sein de « secteurs » et de « pistes ». Le projet OpenStack® Cinder est à l'heure actuelle l'un des services de stockage en mode bloc les plus répandus.

Grâce à un ensemble complet de services de données capables d'interagir librement, les entreprises peuvent mettre fin à la ségrégation des données et acquérir l'agilité nécessaire pour gérer les charges de travail plus complexes, quelle que soit l'échelle.

CONCLUSION

Au cours des prochaines années, sous l'impulsion des spécialistes informatiques qui seront toujours plus nombreux à explorer le potentiel du stockage logiciel, notre façon de concevoir cette technologie va forcément changer. Et c'est très bien comme cela. Nous-mêmes, chez Red Hat, continuons de la découvrir. Nous faisons tout notre possible pour tirer des enseignements des premiers déploiements réalisés dans des entreprises à la pointe de l'innovation. Mais ces efforts ne suffiront pas : la technologie progresse grâce à l'innovation, et les idées les plus novatrices dépassent toujours les attentes qu'elles suscitaient au départ.

Nous sommes favorables à l'exploration active de nouvelles idées. Il semble d'ailleurs que nos clients le soient aussi. Cette confiance vient du fait que nous travaillons au sein d'une communauté. L'expérience nous a appris que notre conception du stockage logiciel ouvert ne nous appartient pas en exclusivité. Il se trouve que nous croyons depuis longtemps au pouvoir de la propriété : pas seulement la propriété de celui qui enferme à double tour une boîte en métal dans un placard, mais aussi celle, plus intéressante, qui naît de la contribution active à la création d'un concept entièrement nouveau.

Nous sommes heureux de pouvoir compter sur votre soutien alors que nous préparons l'avenir du stockage logiciel ouvert. Pour en savoir plus, rendez-vous sur www.redhat.com/fr/technologies/storage.

À PROPOS DE RED HAT

Red Hat est le premier fournisseur mondial de solutions Open Source, s'appuyant sur une approche communautaire pour fournir des technologies cloud, Linux, middleware, de stockage et de virtualisation fiables et performantes. Red Hat propose également des services d'assistance, de formation et de consulting reconnus. Situé au cœur d'un réseau mondial d'entreprises, partenaires et communautés open source, Red Hat participe à la création de technologies innovantes qui permettent de libérer des ressources pour la croissance et de préparer ses clients au futur de l'informatique.

EUROPE, MOYEN-ORIENT
ET AFRIQUE (EMEA)

00800 7334 2835

fr.redhat.com

europe@redhat.com

TURQUIE

00800-448820640

ISRAËL

1-809 449548

ÉAU

8000-4449549

Copyright © 2015 Red Hat, Inc. Red Hat, Red Hat Enterprise Linux, le logo Shadowman et JBoss sont des marques de Red Hat, Inc., déposées aux États-Unis et dans d'autres pays. Linux® est la marque déposée de Linus Torvalds aux États-Unis et dans d'autres pays.

La marque verbale OpenStack® et le logo OpenStack sont soit des marques commerciales/de service déposées, soit des marques commerciales/de service de l'OpenStack Foundation, aux États-Unis et dans d'autres pays. Elles sont utilisées avec l'autorisation de l'OpenStack Foundation. Nous ne sommes ni affiliés, ni approuvés ou sponsorisés par l'OpenStack Foundation ou la communauté OpenStack.