

Silicon Valley : les start-ups Rubrik et Cohesity veulent réinventer la sauvegarde de données

Confronté aux contraintes de volume de données et de temps réel, le stockage a fortement évolué depuis quelques années. Néanmoins, c'est la sauvegarde/restauration, et plus généralement la protection des données, qui représente la part la plus importante au sein des systèmes d'information. Ce qui se traduit par des coûts conséquents.

Une source d'économies et de simplification d'une telle importance attire forcément les talents et donne naissance à des start-ups souvent créées par des vétérans du stockage, décidés à faire bouger les lignes. Des profils qui facilitent les levées de fonds, comme le prouvent les cas de Rubrik et Cohesity, deux jeunes sociétés rencontrées lors de notre périple dans la Silicon Valley.

Rubrik : redéfinir la sauvegarde en mode scale-out unifié

« Notre ambition: rendre la sauvegarde/restauration aussi simple à utiliser qu'un iPhone ! », lance fièrement Bipul Sinha, co-fondateur et CEO de l'entreprise (et investisseur qui avait judicieusement misé sur Nutanix). « Les plans de sauvegarde/restauration reposent sur un logiciel de sauvegarde édité par un éditeur, souvent complété de solutions (comme la déduplication) issues d'autres éditeurs, avec un stockage sur des équipements émanant d'autres constructeurs encore. Avec sa plateforme de gestion de données convergée, Rubrik supprime le besoin d'un logiciel de sauvegarde en intégrant la protection des données, la reprise instantanée, la réplication dans une fabric scale-out unique. »

L'appliance Rubrik Converged Data Management, ou RCDM, (sur serveur x86 standard) propose une protection complète des données et des services de données évolués à l'échelle du Cloud. **Dans la lignée de l'hyper-convergence**, la solution assemble des services et des matériels généralement distincts (logiciel de sauvegarde, réplication, stockage, ou catalogues.) dans une fabric (nœuds en réseau raccordés par des switches) évolutive jusqu'à des milliers de nœuds. *« Nous utilisons les mêmes technologies à l'échelle du Web que Google, Facebook, ou Amazon, permettant à l'utilisateur d'ajouter simplement plus d'appliances sur le cluster selon les besoins »,* précise le CEO. Une forte automatisation, basée sur un moteur de règles, apporte une meilleure productivité et plus de fiabilité.

Tout le contenu étant indexé, l'accès à l'information sauvegardée via une interface de type Web est simplifié et immédiat. La restauration d'applications ou de fichier est quasi instantanée par un simple montage de disque.

Mais surtout, l'entreprise peut concevoir des workflows pour automatiser l'intégration avec d'autres consoles de supervision, via les API de la plateforme Rubrik.

Au cœur de l'appliance, un système de fichiers distribué fait maison et gérant les versions d'information. Tolérant aux pannes (schéma de répliquée distribuée sur le cluster), mixant flash et disque (pour optimiser les performances au juste coût, avec des entrées/sorties optimales), générant des clones de façon incrémentale selon une image de référence, ce serveur scale-out peut exposer ses informations en tant que serveur NFS pour tout snapshot. Pour gérer efficacement l'ensemble des informations, un système de métadonnées distribué a été optimisé pour procurer des performances maximales.

Côté évolutivité, toute nouvelle appliance est automatiquement détectée et intégrée. Les nœuds déclenchent alors un algorithme d'auto-découverte. Et un mécanisme d'optimisation global exécute les tâches sur les nœuds les plus efficaces à tout instant.

Un moteur automatise les politiques de stockage dès la détection de nouvelles VM, et l'utilisateur peut choisir celles qu'il faut protéger et quelles SLA appliquer selon le type de restauration attendu. Plusieurs politiques ont été préconfigurées par Rubrik. Il est possible de définir le moment et l'endroit où sont stockées ou archivées les données sur une appliance Rubrik ou sur le Cloud (Amazon, Google... avec chiffrement AES 256 bits)

Les appliances Rubrik sont déclinées en trois modèles. Architecturé autour de 3 processeurs Intel 8-Core (Haswell à 2,4 GHz) avec 128 Go de Ram DDR4, le modèle r334 contient 9 disques de 4 To et 3 SSD de 400 Go, et divers ports réseau (6 x 10GbE, 6 x 1GBase-T et 3 x 1GBase-T -IPMI). La version r344 possède une puce supplémentaire avec 256 Go de Ram, pour 12 disques de 4 To et 4 SSD de 400, et plus de ports (8 x 10GbE, 8 x 1GBase-T et 4 x 1GBase-T -IPMI). Enfin le modèle r348 gagne en capacité avec 8 disques de 8 To.

Rubrik a été créée en 2014, par Bipul Sinha, Arvind Nithrakashyap (ex-Oracle Cluster), Arvind Jain (ex-Riverbed et Google) et Soham Mazumdar (ex-Google et cofondateur de TagTile, racheté par Facebook). Après deux versions majeures en un an, plus de 65 appliances déployées dans le monde et déjà plus de 40 partenaires, Rubrik réussit plutôt bien son décollage. Elle attire aussi les investissements de connaisseurs avec 51 millions levés en deux tours, en mars et avril 2015. Parmi ces investisseurs, John Thompson (ex-CEO de Symantec et actuellement Chairman chez Microsoft), Franck Sloopman (ex-CEO de DataDomain racheté par EMC et [actuellement CEO de ServiceNow](#)) ou encore Mark Leslie (fondateur et ex-CEO de Veritas, racheté par Symantec). Belle brochette !

Cohesity : hyper-convergence pour "stockage secondaire"

Le mouvement n'est pas passé inaperçu dans la Silicon Valley : le 8 décembre, le président (ex-CEO) de NetApp rejoint le conseil d'administration de Cohesity, et Steve Mullaney, ex CEO de Nicira (solution SDN [rachetée par VMware](#)), accompagnera la start-up dans son développement.

Pourquoi tant d'intérêt ? Cohesity a été créée en juin 2013 par son CEO, Mohit Aron, co-fondateur et ex-CTO de Nutanix, pionnier des systèmes hyper-convergés. Elle sort de son silence en juin 2015 en levant 55 millions de dollars, complétant les 20 millions déjà obtenus en novembre 2013.

Cohesity s'attaque à ce que son dirigeant appelle le stockage secondaire, autrement dit tout ce qui

n'est pas du stockage primaire (ou en production). *« Avec Nutanix, nous avons appliqué l'hyper-convergence au stockage primaire, principalement pour les environnements virtuels. Cependant, la majorité des données se trouve dans le stockage secondaire constitué non seulement des mécanismes de protection des données (sauvegarde, réplication...), mais de beaucoup plus : données analytiques, développement, tests, etc. Autant d'applications de plus en plus critiques (et des informations aussi utilisées dans les applications majeures), nécessitant divers niveaux de SLA. J'ai estimé qu'il fallait une solution pour résoudre les nombreux problèmes liés au stockage secondaire, avec une autre forme de convergence »*, résume Mohit Aron.

De nombreux problèmes surgissent dans les politiques de sauvegarde, menant à un capharnaüm complexe à gérer et très coûteux. Multiplicité des systèmes et des licences, nombre non maîtrisé de copies de données, variété des interfaces et systèmes de supervision, et surtout dispersion inquiétante des données. *« Une situation créant des îlots de données épars, nuisant aux performances et rendant incertaines les éventuelles restaurations de données. Pire encore, les sauvegardes sont parfois organisées par application, et la déduplication est fragmentée »*, détaille le CEO.

La plateforme Cohesity se matérialise sous la forme d'une appliance 2U (aux composants standards, cela va de soi) déclinée en deux modèles. Le C2300 peut contenir jusqu'à 4 nœuds munis chacun de 2 Intel Xeon E5-2600 (2,4 GHz, 8 cœurs), jusqu'à 12 To de disque et 800 Go de SSD, et plusieurs ports réseau (2x 10 GbE, 2x 1 GbE et 1x IPMI). Les nœuds du C2500 peuvent contenir jusqu'à 24 To de disque et 16 To de SSD. La configuration minimale de départ est de 3 nœuds, avec possibilité d'extension à chaud. Cohesity annonce un prix aux alentours de 100 000 dollars pour 4 nœuds.

La société compte se développer très rapidement en Europe. Confiant, Mohit Aron table sur une nouvelle levée de fonds au cours des deux ans à venir.

A lire aussi :

[Silicon Valley : le logiciel dévore le stockage avec Minio, Flashgrid ou Primary Data](#)

Crédit photo : zimmytws / shutterstock